CWI Syllabus 3

Testing statistical hypotheses: worked solutions

W.C.M. Kallenberg, et.al.

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

1980 Mathematics Subject Classification: 62F03, 62H15, 62G10
ISBN 90 6196 280 3

CONTENTS

The chapters of this syllabus contain the solutions of the problems of the corresponding chapters in Professor Lehmann's book Testing Statistical Hypotheses. References are to be found at the end of each chapter.

PREFACE

It was during the 1978/1979 course that we decided to hold a seminar on Professor Lehmann's fundamental book "Testing Statistical Hypotheses". The objective we set ourselves was to solve all the problems. At some stage we concluded that these solutions might be worth publishing and that this could be done with some extra effort. (We now feel, though, that "some extra effort" is something of an understatement).

The present text is based on the problems as they appear in the first (1959) edition of the book. Professor Lehmann has informed us that he is working on a second edition with extra problems, due to appear shortly, but we have decided to confine ourselves to the first edition. To accommodate readers of the second edition in the matter of changed problem numbers, we shall include a seperate addendum with a cross-reference list as soon as the new problem numbers are available to us.

We thank Professor Lehmann for his support of our project, K. Snel for his excellent typing of the manuscript and the Centre of Mathematics and Computer Science (CWI) for giving us the opportunity to publish this syllabus. Our thanks are due also to all others who have contributed towards its realization. Among the participants it is Wilbert Kallenberg who deserves our special gratitude for doing most of the editorial work.

The members of the Lehmann group:

J. Beirlant
P. van Blokland
J.J. Dik
R.J.M.M. Does
A.J. van Es
R.D. Gill
P.L.J. Janssen
P.J.M. Kallenberg
W.C.M. Kallenberg
C.A.J. Klaassen
E. Opperdoes
C. van Putten
B.F. Schriever
M. Vandemaele
C.J. Warmer

October 1984

CHAPTER 1

Section 2

Problem 1.

(i) For all $x = 0,1,2,\ldots$ we have

$$\begin{aligned} P\{X = x\} &= P[m-1 \text{ successes in the first } x+m-1 \text{ trials}, \\ &\qquad \text{the } (x+m)^{\text{th}} \text{ trial is successful}] \\ &= P[m-1 \text{ successes in the first } x+m-1 \text{ trials}] \\ &\quad P[\text{the } (x+m)^{\text{th}} \text{ trial is successful}] \\ &= \binom{x+m-1}{m-1} p^{m-1}(1-p)^{x+m-1-(m-1)} p = \binom{m+x-1}{x} p^m (1-p)^x. \end{aligned}$$

(ii) Let X_t denote the number of events occurring in any time interval of length t. Then for $t > 0$

$$P\{T > t\} = P\{X_t = 0\} = e^{-\lambda t}$$

and hence

$$p(t) = \frac{d}{dt} P\{T \le t\} = \lambda e^{-\lambda t}.$$

(iii) Let $F(x) = P\{X \le x\}$ for all $x \in \mathbb{R}$; then $F(a) = 0$, $F(b) = 1$ and F is right-continuous, that is

$$\lim_{\varepsilon \downarrow 0} F(x+\varepsilon) = F(x), \qquad x \in \mathbb{R}.$$

Define for all $x \in [0,b-a]$

$$f(x) = F(x+a) = F(x+a) - F(a).$$

The assumption that the probability of X falling in any subinterval of (a,b) depends only on the length of the subinterval, implies that

$$f(x) = F(x+y) - F(y)$$

for all $y \in [a,b]$ and $x \in [0,b-y]$. Moreover, by this assumption it follows that $P\{X = x\} = 0$ for all $x \in (a,b)$ and hence f is continuous on $[0,b-a]$. For all $x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0$ with $x_1 + x_2 + \ldots + x_n \leq b-a$

$$f\left(\sum_{i=1}^{n} x_i\right) = F\left(\sum_{i=1}^{n} x_i + a\right)$$

$$= \sum_{i=2}^{n} \left[F(x_i + \sum_{j=1}^{i-1} x_j + a) - F\left(\sum_{j=1}^{i-1} x_j + a\right)\right] + F(x_1+a)$$

$$= \sum_{i=2}^{n} f(x_i) + f(x_1) = \sum_{i=1}^{n} f(x_i).$$

Hence $f(b-a) = f(\Sigma_{i=1}^{n} n^{-1}(b-a)) = nf(n^{-1}(b-a))$ and thus $f(n^{-1}(b-a)) = n^{-1}f(b-a) = n^{-1}$, since $f(b-a) = 1$ by definition. Therefore, $f(kn^{-1}(b-a)) = kn^{-1}$ for all $k,n \in \mathbb{N}$. By the continuity of f it follows that $f(y(b-a)) = y$ for all $y \in [0,1]$ and thus $f(x) = x(b-a)^{-1}$, $0 \leq x \leq b-a$. Since

$$p(x) = \frac{d}{dx} F(x) = \frac{d}{dx} f(x-a) = (b-a)^{-1}$$

for all $a < x < b$, X has the rectangular distribution $R(a,b)$. (cf. also RUDIN (1970) pp. 50-52.)

Section 5

Problem 2.

For L the square of the error, unbiasedness of δ is equivalent to

$$(1) \qquad \gamma^2(\theta') - \gamma^2(\theta) \geq 2h(\theta)\{\gamma(\theta') - \gamma(\theta)\}$$

for all θ and θ', cf. Example 12.

If $h(\theta) = \gamma(\theta)$ for all θ, it therefore immediately follows that δ is unbiased.

Let δ be unbiased. Assume that $\gamma(\theta_0) > h(\theta_0)$ for some $\theta_0 \in \Omega$. Since $A = \{\theta \in \Omega : \gamma(\theta) > h(\theta_0), \gamma(\theta_0) > h(\theta)\}$ is a non-empty ($\theta_0 \in A$), open subset of Ω (continuity of γ and h) and γ is not constant in any open subset of Ω, there exists $\theta_1 \in A$ satisfying $\gamma(\theta_1) \neq \gamma(\theta_0)$. If $\gamma(\theta_1) > \gamma(\theta_0)$ then $\gamma(\theta_1) > \gamma(\theta_0) > h(\theta_1) \quad (\theta_1 \in A)$, and hence $\gamma^2(\theta_0) - \gamma^2(\theta_1) = \{\gamma(\theta_0) + \gamma(\theta_1)\}\{\gamma(\theta_0) - \gamma(\theta_1)\} < 2h(\theta_1)\{\gamma(\theta_0) - \gamma(\theta_1)\}$, in contradiction with (1). If $\gamma(\theta_1) < \gamma(\theta_0)$ then $h(\theta_0) < \gamma(\theta_1) < \gamma(\theta_0) \quad (\theta_1 \in A)$, and hence $\gamma^2(\theta_1) - \gamma^2(\theta_0) < 2h(\theta_0)\{\gamma(\theta_1) - \gamma(\theta_0)\}$, again in contradiction with (1). In a similar way the assumption $\gamma(\theta_0) < h(\theta_0)$ for some $\theta_0 \in \Omega$ leads to a contradiction, implying that $\gamma(\theta) = h(\theta)$ for all $\theta \in \Omega$.

Note that in this solution the connectedness of the parameter space Ω has not been used!

(LEHMANN (1951))

Problem 3.

(i) If $a_1, a_2 \in \mathbb{R}$ satisfy $a_1 \le a_2$ then

$$|y-a_2| - |y-a_1| = a_2 - a_1 \qquad \text{for all } y \le a_1$$

and

$$|y-a_2| - |y-a_1| \ge a_1 - a_2 \qquad \text{for all } y > a_1.$$

Let m be a median of the random variable Y; then for $m \le a_1 \le a_2$

$$E\{|Y-a_2| - |Y-a_1|\} \ge (a_2-a_1)[P\{Y \le a_1\} - P\{Y > a_1\}]$$
$$= (a_2-a_1)[2P\{Y \le a_1\} - 1] \ge 0$$

with strict inequality if $a_1 < a_2$ and a_1 is not a median of Y. By symmetry this inequality is also valid in the case $a_2 \le a_1 \le m$.

(ii) From formula (9) on p. 12 we see that the estimate $\delta(X)$ of $\gamma(\theta)$ is unbiased with respect to the loss function $L(\theta,d) = |\gamma(\theta) - d|$ iff for all $\theta, \theta' \in \Omega$

$$E_\theta|\gamma(\theta') - \delta(X)| \ge E_\theta|\gamma(\theta) - \delta(X)|. \tag{2}$$

If $\gamma(\theta)$ is a median of $\delta(X)$ for each θ, then (2) holds in view of (i). If for some $\theta \in \Omega$, say θ_0, $\gamma(\theta_0)$ is not a median of $\delta(X)$, then $\gamma(\theta_0) \notin [m^-(\theta_0), m^+(\theta_0)]$. (Note that $m^-(\theta_0)$ and $m^+(\theta_0)$ are medians.) Assume

$\gamma(\theta_0) > m^+(\theta_0)$. The set $A = \{\theta_0 : \gamma(\theta) - m^+(\theta_0) > 0, \gamma(\theta_0) - m^+(\theta) > 0\}$ is a non-empty ($\theta_0 \in A$) and open subset of Ω (continuity of γ and m^+). Since γ is not constant on the open set A, there exists $\theta_1 \in A$ with $\gamma(\theta_1) \neq \gamma(\theta_0)$. If $\gamma(\theta_1) < \gamma(\theta_0)$ then $m^+(\theta_0) < \gamma(\theta_1) < \gamma(\theta_0)$. In view of (i) this implies $E_{\theta_0}\{|\delta(X) - \gamma(\theta_0)| - |\delta(X) - \gamma(\theta_1)|\} > 0$, in contradiction with (2). Interchanging the role of θ_0 and θ_1 a contradiction in the case $\gamma(\theta_1) > \gamma(\theta_0)$ is obtained.
In a similar way it can be shown that the assumption $\gamma(\theta_0) < m^-(\theta_0)$ leads to a contradiction. This completes the proof of (ii).

Note that in this solution the connectedness of the parameter space Ω has not been used!

(LEHMANN (1951))

Problem 4.

The assertion stated in Problem 4 is not correct as is shown by the following example.
Let $\Theta = D = \{0,1\}$, $\omega_0 = \Theta$, $V(0,0) = 0$, $V(0,1) = 1$, $h(0) = 1$ and $h(1) = 2$. The rule δ, given by $P_0\{\delta(X) = 1\} = 1$ and $P_1\{\delta(X) = 0\} = 1$, is unbiased and $R(0,\delta) = 1$.

To avoid this kind of examples we impose an extra condition that *the infimum of* h *on* ω_d *is not attained.*

Let δ be an unbiased procedure and let $R(\theta,\delta)$ be the risk function. We suppose that $h \geq 0$ and $V \geq 0$. Let $\theta \in \omega_d$, that is, d is the unique correct ducision for θ. Then we have

$$R(\theta,\delta) = E_\theta L(\theta,\delta(X)) = h(\theta)E_\theta V(d,\delta(X)).$$

There exists $\theta' \in \omega_d$ with $h(\theta') < h(\theta)$ because of our extra condition. Since δ is unbiased we have

$$E_\theta L(\theta',\delta(X)) \geq E_\theta L(\theta,\delta(X))$$

and hence

$$h(\theta')E_\theta V(d,\delta(X)) \geq h(\theta)E_\theta V(d,\delta(X)).$$

With $h(\theta') < h(\theta)$ and $V \geq 0$ it follows that $E_\theta V(d,\delta(X)) = 0$ and therefore $R(\theta,\delta) = 0$.

Since θ was arbitrarily chosen we have $R(\theta,\delta) = 0$ for all $\theta \in \Omega$.

Take in the example, $\omega_d = \{\theta : \theta = (d,a)\}$, $h(\theta) = a^{-2}$ and $V(d,d') = (d-d')^2$. Note that the infimum of h on ω_d is not attained and the result follows.

The extra condition above can be replaced by the condition: $\inf_{\theta\in\omega_d} h(\theta) = 0$. This condition is also satisfied in the example mentioned in Problem 4.

(LEHMANN (1951))

Problem 5.

Suppose G is a group of transformations that leaves the decision problem invariant. Let $g \in G$ and $\delta_1 \in C$, then $(g^{-1})^*\delta_1 g \in C$ and

$$\begin{aligned} R(\theta,g^*\delta_0 g^{-1}) &= E_\theta L(\theta,g^*\delta_0 g^{-1}(X)) = E_\theta L((\bar{g})^{-1}\theta,\delta_0 g^{-1}(X)) = \\ &= E_{(\bar{g})^{-1}\theta}\, L((\bar{g})^{-1}\theta,\delta_0(X)) = R((\bar{g})^{-1}\theta,\delta_0) \leq \\ &\leq R((\bar{g})^{-1}\theta,(g^{-1})^*\delta_1 g) = R(\theta,\delta_1). \end{aligned}$$

Since δ_1 was arbitrarily chosen $g^*\delta_0 g^{-1}$ uniformly minimizes the risk within the class C. But δ_0 is unique, so $\delta_0 = g^*\delta_0 g^{-1}$ for all $g \in G$, implying that δ_0 is invariant.

If δ_0 is unique only up to sets of measure 0 (i.e. if δ_2 also uniformly minimizes the risk within the class C then $P_\theta\{\delta_0(X) \neq \delta_2(X)\} = 0$ for all $\theta \in \Omega$), then $\delta_0 = g^*\delta_0 g^{-1}$ except on a set N_g of measure 0 and hence δ_0 is almost invariant.

Problem 6.

(i) Let C be the class of unbiased procedures. Let $\delta \in C$. Since $g^{-1}(X)$ has distribution $P_{(\bar{g})^{-1}\theta}$ and $L(\bar{g}\theta,g^*d) = L(\theta,d)$ for all $\theta \in \Omega$, $d \in D$ and $g \in G$, we have

$$\begin{aligned} E_\theta L(\theta',g^*\delta g^{-1}(X)) &= E_{(\bar{g})^{-1}\theta}\, L(\theta',g^*\delta(X)) = \\ &= E_{(\bar{g})^{-1}\theta}\, L((\bar{g})^{-1}\theta',\delta(X)) \geq E_{(\bar{g})^{-1}\theta}\, L((\bar{g})^{-1}\theta,\delta(X)) = \end{aligned}$$

$$= E_\theta L(\theta, g^\star \delta g^{-1}(X)).$$

Hence $g^\star \delta g^{-1} \in \mathcal{C}$ and by Problem 5 δ_0 is almost invariant.

(ii) For any θ, θ' there exists $\bar{g}$ (by transitivity of $\bar{G}$) such that

$$(3) \qquad E_\theta L(\theta', \delta_0(X)) = E_\theta L(\bar{g}\theta, \delta_0(X)) = E_\theta L(\theta, (g^\star)^{-1}\delta_0(X)) = \\ = R(\theta, (g^\star)^{-1}\delta_0).$$

Since δ_0 is (almost) invariant and $G^\star$ is commutative, we have for all $g, h \in G$ (except on a set of measure 0)

$$(g^\star)^{-1}\delta_0 h = (g^\star)^{-1} h^\star \delta_0 = h^\star (g^\star)^{-1}\delta_0$$

and hence $(g^\star)^{-1}\delta_0$ is also (almost) invariant. Therefore,

$$(4) \qquad R(\theta, (g^\star)^{-1}\delta_0) \geq R(\theta, \delta_0) = E_\theta L(\theta, \delta_0(X)).$$

Combination of (3) and (4) completes the proof.

(LEHMANN (1951))

Problem 7.

Let $\theta = (\xi, \sigma^2)$ and $L(\theta, d) = (\xi - d)^2 \sigma^{-2}$.
First consider G_1. Let $g \in G_1$, that is $g(x) = x+b$. Then $\bar{g}$ is given by $\bar{g}\theta = (\xi+b, \sigma^2)$ and $g^\star$ by $g^\star d = d+b$. Note that $L(\bar{g}\theta, g^\star d) = L(\theta, d)$. Since $g_1^\star g_2^\star d = g_1^\star(d+b_2) = d+b_2+b_1 = g_2^\star(d+b_1) = g_2^\star g_1^\star d$, $G_1^\star$ is commutative. $\bar{G}_1$ is not transitive, e.g. there does not exist $\bar{g} \in \bar{G}_1$ satisfying $\bar{g}(0,1) = (0,2)$. An estimate $\delta(x)$ is invariant iff $\delta(x+b) = \delta(x) + b$ for all $b, x \in \mathbb{R}$, implying that the set of invariant estimates is $\{\delta : \delta(x) = x+c,\ c \in \mathbb{R}\}$. Hence, if δ is invariant $R(\theta, \delta) = E_\theta(\xi - X - c)^2 \sigma^{-2} = 1 + c^2\sigma^{-2}$, which is minimized by choosing $c = 0$. Therefore, the best invariant estimate relative to G_1 is X. By Problem 4 it is seen that $\delta(x) = x$ is biased.
Next consider G_2. Let $g \in G_2$, that is $g(x) = ax+b$. Then $\bar{g}$ is given by $\bar{g}\theta = (a\xi+b, a^2\sigma^2)$ and $g^\star$ by $g^\star d = ad+b$. Note that $L(\bar{g}\theta, g^\star d) = L(\theta, d)$. Let (ξ_1, σ_1^2) and $(\xi_2, \sigma_2^2) \in \Omega$. Taking $g(x) = \sigma_2\sigma_1^{-1}x + \xi_2 - \sigma_2\sigma_1^{-1}\xi_1$ we obtain $\bar{g}(\xi_1 \sigma_1^2) = (\xi_2, \sigma_2^2)$ and hence $\bar{G}_2$ is transitive. $G_2^\star$ is not commutative, for let $g_1^\star d = 2d$ and $g_2^\star d = d+1$; then $g_1^\star g_2^\star d = 2d+2$ and $g_2^\star g_1^\star d = 2d+1$. An estimate

$\delta(x)$ is invariant iff $\delta(ax+b) = a\delta(x) + b$ for all $a \in \mathbb{R}^+$, $b \in \mathbb{R}$ and $x \in \mathbb{R}$, implying that $\delta(x) = x$ is the only and therefore also the best invariant estimate relative to G_2. By Problem 4 it is seen that $\delta(x) = x$ is biased.

(LEHMANN (1951))

Section 6

Problem 8.

(i) First it is proved that the conditional probability density of Θ given $X = x$ equals $\pi(\theta|x) = \rho(\theta)p_\theta(x) / \int_\Omega \rho(\theta')p_{\theta'}(x)d\theta'$. For every Borel set B and every Borel set ω' we have by Fubini's theorem

$$P\{X \in B, \Theta \in \omega'\} = \int_{\omega'} \int_B p_\theta(x)dx\rho(\theta)d\theta = \int_B \int_{\omega'} p_\theta(x)\rho(\theta)d\theta dx$$

$$= \int_B \int_{\omega'} p_\theta(x)\rho(\theta)\left\{\int_\Omega \rho(\theta')p_{\theta'}(x)d\theta'\right\}^{-1} d\theta \int_\Omega \rho(\theta)p_\theta(x)d\theta dx$$

$$= \int_B P\{\Theta \in \omega' | X = x\} f(x)dx,$$

where $f(x) = \int_\Omega \rho(\theta)p_\theta(x)d\theta$ is the probability density of X. Next it will be shown that if there exists a rule δ_0, which minimizes $\int L(\theta,\delta(x))\pi(\theta|x)d\theta$ for each x, then δ is a Bayes solution iff $\delta = \delta_0$ a.e. μ^X. Here μ^X denotes the probability measure induced by X. So, suppose that such a rule δ_0 exists. Since by Fubini's theorem

$$\int E_\theta L(\theta,\delta_0(X))\rho(\theta)d\theta = \iint L(\theta,\delta_0(x))p_\theta(x)\rho(\theta)d\theta dx$$

$$= \iint L(\theta,\delta_0(x))\pi(\theta|x)d\theta f(x)dx$$

$$\leq \iint L(\theta,\delta(x))\pi(\theta|x)d\theta f(x)dx = \int E_\theta L(\theta,\delta(X))\rho(\theta)d\theta,$$

δ_0 is a Bayes solution.

Suppose δ_1 is also a Bayes solution. Define $\varphi_i(x) = \int L(\theta,\delta_i(x))\pi(\theta|x)d\theta$ $(i = 0,1)$, then $\int \varphi_0(x)f(x)dx = \int \varphi_1(x)f(x)dx$ and $\varphi_0(x) \leq \varphi_1(x)$ for all x. This implies $\varphi_0(x) = \varphi_1(x)$ a.e. μ^X.

(ii) The loss function in Example 11 is defined to be

$L(.,.)$	d_0	d_1
$\theta \in \omega_0$	0	b
$\theta \in \omega_1$	a	0

Hence

$$\int L(\theta,\delta(x))\pi(\theta|x)d\theta = \begin{cases} a\int_{\omega_1} \pi(\theta|x)d\theta = aP\{\Theta \in \omega_1|x\} & \text{if } \delta(x) = d_0 \\ b\int_{\omega_0} \pi(\theta|x)d\theta = bP\{\Theta \in \omega_0|x\} & \text{if } \delta(x) = d_1 \end{cases}$$

and

$$\delta_0(x) = \begin{cases} d_0 & \\ d_1 & \text{if } aP\{\Theta \in \omega_1|x\} \begin{matrix} < \\ > \\ = \end{matrix} bP\{\Theta \in \omega_0|x\}. \\ d_0 \text{ or } d_1 & \end{cases}$$

According to (i) a Bayes solution is almost equal to δ_0.

(iii) a) $\int L(\theta,\delta(x))\pi(\theta|x)d\theta = \int (g(\theta) - \delta(x))^2\pi(\theta|x)d\theta = \delta^2(x) - 2\delta(x) \int g(\theta)\pi(\theta|x)d\theta + \int g^2(\theta)\pi(\theta|x)d\theta$, which is minimized by choosing $\delta(x) = \int g(\theta)\pi(\theta|x)d\theta = E[g(\Theta|x)]$.

b) $\int L(\theta,\delta(x))\pi(\theta|x)d\theta = \int |g(\theta) - \delta(x)|\pi(\theta|x)d\theta$, which is minimized by choosing $\delta(x)$ is any median of the conditional distribution of $g(\Theta)$ given x (cf. Problem 3(i) with $m = a_1$).

Problem 9.

(i) In this example we have $\Omega = \{HH,HT\}$, the sample space $X = \{H,T\}$, $D = \{HH,HT\}$ and the randomized procedure Y is given by $P\{Y_H = HT\} = \rho$ and $P\{Y_T = HT\} = 1$, where $0 \le \rho \le 1$.
Then the risk function of Y equals

$$R(HH,Y) = \rho L(HH,HT) = \rho \text{ and}$$
$$R(HT,Y) = \tfrac{1}{2}(1-\rho)L(HT,HH) = \tfrac{1}{2}(1-\rho).$$

The maximum risk is minimized for $\rho = \frac{1}{3}$ and equals $\frac{1}{3}$. Note that the maximum risk of the four nonrandomized decision rules is always greater than or equal to $\frac{1}{2}$; so randomization reduces the maximum risk.

(ii) If we replace HH by AA and HT by Aa, the only difference with (i) is that we assume an a priori probability p for AA. The Bayes risk equals

$$r(\rho,Y) = pR(AA,Y) + (1-p)R(Aa,Y) =$$
$$= p\rho + (1-p)\tfrac{1}{2}(1-\rho) = \tfrac{3}{2}\rho(p-\tfrac{1}{3}) + \tfrac{1}{2}(1-p),$$

which is minimized by choosing $\rho = 0$ if $p > \frac{1}{3}$, $\rho = 1$ if $p < \frac{1}{3}$.

Problem 10.

A randomized procedure Y_x based on an observation x can be characterized by

$$\eta_x = P[\text{decision } d_0 \text{ is taken} \mid X = x] = P\{Y_x = d_o\}.$$

For such a procedure we have

$$R(\theta,Y) = E_\theta L(\theta,Y) = \begin{cases} a_1 E_\theta(1-\eta_X) & \text{if } \theta \in \Omega_0 \\ a_0 E_\theta \eta_X & \text{if } \theta \in \Omega_1 . \end{cases}$$

If $a_0 = a_1 = 0$ all procedures are both minimax and unbiased. Therefore, suppose $a_0 + a_1 > 0$.
Now Y is minimax iff

$$\sup_{\theta\in\Omega} R(\theta,Y) = \max\{\sup_{\theta\in\Omega_0} u_1 E_\theta(1-\eta_X), \sup_{\theta\in\Omega_1} a_0 E_\theta \eta_X\} \tag{5}$$
$$\leq \max\{\sup_{\theta\in\Omega_0} a_1 E_\theta(1-\eta_X'), \sup_{\theta\in\Omega_1} a_0 E_\theta \eta_X'\} = \sup_{\theta\in\Omega} R(\theta,Y')$$

for all Y'; and Y is unbiased iff

$$E_\theta L(\theta',Y) \geq E_\theta L(\theta,Y) \qquad \text{for all } \theta \text{ and } \theta'.$$

Because $L(\theta',\cdot) = L(\theta,\cdot)$ if θ and θ' are both in Ω_0 or both in Ω_1, we only have to consider the case $\theta \in \Omega_0$, $\theta' \in \Omega_1$ (the case $\theta' \in \Omega_0$, $\theta \in \Omega_1$ is treated similarly). This gives us

Y is unbiased iff

$$a_0 E_\theta \eta_X \geq a_1 E_\theta(1-\eta_X) = R(\theta,Y) \text{ for all } \theta \in \Omega_0$$
$$\text{and} \quad a_1 E_\theta(1-\eta_X) \geq a_0 E_\theta \eta_X = R(\theta,Y) \text{ for all } \theta \in \Omega_1,$$

which is equivalent to

$$Y \text{ is unbiased iff } R(\theta,Y) \leq a_0 a_1 (a_0 + a_1)^{-1} \text{ for all } \theta \in \Omega. \tag{6}$$

(i) Let Y' be the procedure defined by

$$\eta'_x = a_1(a_0 + a_1)^{-1} \quad \text{for all } x.$$

If Y is minimax then $\sup_\theta R(\theta,Y) \le \sup_\theta R(\theta,Y') = a_0a_1(a_0+a_1)^{-1}$ and hence Y is unbiased.

(ii) Let Y be a randomized procedure. First it will be shown that the continuity of $P_\theta(A)$ for all subsets A of $\mathcal{X}$ implies the continuity of $E_\theta\eta_X$. Note that $E_\theta\eta_X = P_\theta\{Y = d_0\}$, but $\{Y = d_0\}$ is not a subset of $\mathcal{X}$. Let $\varepsilon > 0$. Define $\alpha_i = i\varepsilon$ and $A_i = \{x : i\varepsilon < \eta_x \le (i+1)\varepsilon\}$, $i = 0,1,2,\ldots,N_\varepsilon = [\varepsilon^{-1}]$, where [t] denotes the integral part of t. Then we have

$$\sum_{i=1}^{N_\varepsilon} \alpha_i I_{A_i}(x) \le \eta_x \le \sum_{i=1}^{N_\varepsilon} \alpha_i I_{A_i}(x) + \varepsilon.$$

So, $E_\theta\eta_X$ is the uniform limit of continuous functions of θ and must itself be continuous.

It is assumed thst Ω_0 and Ω_1 have a boundary point in common, say θ_0. Hence

$$\text{(7)}\qquad \begin{aligned} \sup_{\theta\in\Omega} R(\theta,Y) &= \max\{\sup_{\theta\in\Omega_0} a_1E_\theta(1-\eta_X), \sup_{\theta\in\Omega_1} a_0E_\theta\eta_X\} \\ &\ge \max\{a_1E_{\theta_0}(1-\eta_X), a_0E_{\theta_0}\eta_X\} \\ &\ge \frac{a_0}{a_0+a_1} a_1E_{\theta_0}(1-\eta_X) + \frac{a_1}{a_0+a_1} a_0E_{\theta_0}\eta_X = a_0a_1(a_0+a_1)^{-1}. \end{aligned}$$

From (6) it follows that an unbiased randomized procedure attains the lower bound in (7). Hence Y is minimax.

Problem 11.

(i) Let Y be a randomized procedure. Define the procedure Y' by

$$dP^{Y'_x}(d) = N^{-1} \sum_{i=1}^{N} dP^{Y_{g_i x}}(g_i^* d)$$

for all $x \in \mathcal{X}$. Then we have by the invariance of the problem

$$\text{(8)}\qquad \begin{aligned} R(\theta,Y') = EL(\theta,Y') &= \iint L(\theta,d)dP^{Y'_x}(d)dP_\theta(x) \\ &= N^{-1}\sum_{i=1}^{N} \iint L(\theta,d)dP^{Y_{g_i x}}(g_i^* d)dP\theta(x) \end{aligned}$$

$$= N^{-1} \sum_{i=1}^{N} \iint L(\theta,d)\,dP^{Y_x}(g_i^* d)\,dP_{\bar{g}_i\theta}(x)$$

$$= N^{-1} \sum_{i=1}^{N} \iint L(\bar{g}_i\theta, g_i^* d)\,dP^{Y_x}(g_i^* d)\,dP_{\bar{g}_i\theta}(x) = N^{-1} \sum_{i=1}^{N} R(\bar{g}_i\theta, Y).$$

The fact that Y' is an invariant procedure follows on observing that for any $j = 1,2,\ldots,N$

$$dP^{(g_j^*)^{-1}Y'_{g_j x}}(d) = dP^{Y'_{g_j x}}(g_j^* d) = N^{-1} \sum_{i=1}^{N} dP^{Y_{g_i g_j x}}(g_i^* g_j^* d).$$

Now (ii) on p. 11 implies that $g_i^* g_j^* = (g_i g_j)^*$. Furthermore, since G is a group, $g_i g_j$ takes on the same values as g_i as i runs from 1 up to and including N, whence it follows that

$$N^{-1} \sum_{i=1}^{N} dP^{Y_{g_i g_j x}}(g_i^* g_j^* d) = N^{-1} \sum_{i=1}^{N} dP^{Y_{g_i x}}(g_i^* d) = dP^{Y'_x}(d).$$

Now suppose that Y is a minimax procedure, i.e.

$$\sup_\theta R(\theta,Y) = \min_Z \sup_\theta R(\theta,Z), \tag{9}$$

where Z runs through all randomized procedures. Then by (8) and (9)

$$\sup_\theta R(\theta,Y') \le N^{-1} \sum_{i=1}^{N} \sup_\theta R(\bar{g}_i\theta,Y)$$

$$= N^{-1} \sum_{i=1}^{N} \min_Z \sup_\theta R(\theta,Z) = \min_Z \sup_\theta R(\theta,Z).$$

On the other hand, $\sup_\theta R(\theta,Y') \ge \min_Z \sup_\theta R(\theta,Z)$, implying that Y' is a minimax procedure, thus establishing part (i).

(ii) Let $\theta = \pi_1\pi_2\ldots\pi_n$ for some n, and let $X = \theta Y$, where $P\{Y = a\} = P\{Y = a^{-1}\} = P\{Y = b\} = P\{Y = b^{-1}\} = \frac{1}{4}$. Let U be any (possibly randomized) invariant procedure. Then in view of the invariance

$$P\{U = \theta\} = \sum_{y\in\{a,a^{-1},b,b^{-1}\}} P\{U_{\theta y} = \theta\}P\{Y = y\}$$

$$= \sum_{y\in\{a,a^{-1},b,b^{-1}\}} P\{U_e = y^{-1}\}\tfrac{1}{4} = \tfrac{1}{4}P\{U \in \{a,a^{-1},b,b^{-1}\}\} \le \tfrac{1}{4},$$

that is $R(\theta,U) = P\{U \ne \theta\} \ge \frac{3}{4}$. So, any invariant procedure U with

$P\{U_e \in \{a,a^{-1},b,b^{-1}\}\} = 1$ minimizes the maximum risk (for instance, the nonramdomized decision rule $\delta(x) = xa$). In such a case the risk function $R(\theta,U)$ equals $\frac{3}{4}$.
Now let V be the procedure described in the hint. Then

$$P\{V = \theta\} = P\{V=\pi_1\ldots\pi_n\} \geq \sum_{\substack{y\in\{a,a^{-1},b,b^{-1}\} \\ y\neq\pi_n^{-1}}} P\{V_{\theta y} = \theta\}P\{Y = y\} = \tfrac{3}{4},$$

that is $\max_\theta R(\theta,V) \leq \frac{1}{4}$.

(PEISAKOFF (1951), KIEFER (1957), KUDO (1955))

Section 7

Problem 12.

(i) To obtain the Bayes solution we have to maximize the expression $n^{-1}\sum_{i=1}^n E_{\theta_i}G(\theta_i,\delta(X))$, where δ is a decision rule and

$$G(\theta_i,d_j) = \begin{cases} a(\theta_i) & \text{if } i = j \\ 0 & \text{otherwise.}\end{cases}$$

Since

$$n^{-1}\sum_{i=1}^n E_{\theta_i}G(\theta_i,\delta(X)) = n^{-1}\sum_{i=1}^n \int_{\{x:\delta(x)=\theta_i\}} a(\theta_i)L_x(\theta_i)dx$$
$$\leq n^{-1}\sum_{i=1}^n \int_{\{x:\delta(x)=\theta_i\}} \max_i \{a(\theta_i)L_x(\theta_i)\}dx$$
$$= n^{-1}\int \max_i \{a(\theta_i)L_x(\theta_i)\}dx$$

with equality iff, for a.e. x, we choose $\delta(x) = \theta_k$, where $\theta_k = \theta_{k(x)}$ satisfies $\max_i a(\theta_i)L_x(\theta_i) = a(\theta_k)L_x(\theta_k)$, i.e. if δ is the maximum likelihood procedure. So we see that the Bayes solution coincides with the maximum likelihood procedure (a.e.).

(ii) In Problem 8 the a posteriori density is defined as $\pi(\theta|x) = \rho(\theta)p_\theta(x)\{\int \rho(\theta')p_{\theta'}(x)d\theta'\}^{-1}$. If Θ is uniformly distributed on (0,1) then

$$\max_\theta \pi(\theta|x) = \max_\theta \{p_\theta(x) / \int p_{\theta'}(x)d\theta'\}. \tag{10}$$

Let $\hat\theta$ be the maximum likelihood estimate, i.e. $p_{\hat\theta}(x) = \max_\theta p_\theta(x)$. In view of (10)

$$\pi(\hat\theta|x) = p_{\hat\theta}(x)\{\int p_{\theta'}(x)d\theta'\}^{-1} = \max_\theta \pi(\theta|x),$$

that is, $\hat\theta$ is the mode of the a posteriori density of Θ given x.

Problem 13.

By formula 14 on page 15 the likelihood ratio procedure takes decision d_0 or d_1 according to whether $\sup_{\theta\in\omega_0} L_x(\theta) / \sup_{\theta\in\omega_1} L_x(\theta) > a_1a_0^{-1}$ or $< a_1a_0^{-1}$.
(i) Here $\omega_0 = \{(\xi,\sigma^2) : \xi < 0\}$, $\omega_1 = \{(\xi,\sigma^2) : \xi \geq 0\}$ and $\log L_x(\theta) = -\frac{1}{2}n\log 2\pi - \frac{1}{2}n\log\sigma^2 - \frac{1}{2}\Sigma_{i=1}^n (x_i-\xi)^2\sigma^{-2}$.
Differentiating with respect to σ^2 we see that

$$(11) \qquad \sup_{\sigma^2>0} L_x(\theta) = L_x(\xi, n^{-1}\sum_{i=1}^n (x_i-\xi)^2)$$

$$= (2\pi)^{-\frac{1}{2}n}\{n^{-1}\sum_{i=1}^n (x_i-\xi)^2\}^{-\frac{1}{2}n}e^{-\frac{1}{2}n}$$

holds. Define

$$a \wedge b = \min(a,b) \text{ and } a \vee b = \max(a,b).$$

In view of (11) the likelihood ratio procedure takes decision d_0 when

$$\frac{(2\pi)^{-\frac{1}{2}n}\{n^{-1}\Sigma_{i=1}^n (x_i - (\bar x \wedge 0))^2\}^{-\frac{1}{2}n}e^{-\frac{1}{2}n}}{(2\pi)^{-\frac{1}{2}n}\{n^{-1}\Sigma_{i=1}^n (x_i - (\bar x \vee 0))^2\}^{-\frac{1}{2}n}e^{-\frac{1}{2}n}} > \frac{a_1}{a_0},$$

i.e. when

$$\sum_{i=1}^n (x_i - (\bar x \wedge 0))^2 \Big/ \sum_{i=1}^n (x_i - (\bar x \vee 0))^2 < c^{-1}.$$

Using $\Sigma_{i=1}^n x_i^2 = \Sigma_{i=1}^n (x_i - \bar x)^2 + n\bar x^2$ this may be written in the required form.
(ii) In this case $\omega_0 = \{(\xi,\sigma^2) : \sigma^2 < \sigma_0^2\}$, $\omega_1 = \{(\xi,\sigma^2) : \sigma^2 \geq \sigma_0^2\}$.
Differentiating $\log L_x(\xi,\sigma^2)$ with respect to ξ we obtain

$$\sup_{\xi\in\mathbb{R}} \log L_x(\xi,\sigma^2) = \log L_x(\bar x,\sigma^2)$$

$$= -\tfrac{1}{2}n\log 2\pi - \tfrac{1}{2}n\log\sigma^2 - \tfrac{1}{2}\sigma^{-2}\sum_{i=1}^n (x_i - \bar x)^2.$$

Since $\frac{d}{d\sigma^2} \log L_x(\bar{x},\sigma^2) = \frac{1}{2}n\sigma^{-4}(s_n^2 - \sigma^2)$ with $s_n^2 = n^{-1}\,\Sigma_{i=1}^n (x_i - \bar{x})^2$, the likelihood ratio procedure takes decision d_0 when

$$\frac{(2\pi)^{-\frac{1}{2}n}(s_n^2 \wedge \sigma_0^2)^{-\frac{1}{2}n} \exp\{-\frac{1}{2}ns_n^2 / (s_n^2 \wedge \sigma_0^2)\}}{(2\pi)^{-\frac{1}{2}n}(s_n^2 \vee \sigma_0^2)^{-\frac{1}{2}n} \exp\{-\frac{1}{2}ns_n^2 / (s_n^2 \vee \sigma_0^2)\}} > \frac{a_1}{a_0},$$

i.e. when

$$(q_n \vee q_n^{-1}) \exp\{(1-q_n) \wedge (q_n-1)\} > c,$$

where $q_n = s_n^2\sigma_0^{-2}$. By discerning the four cases $q_n \gtrless 1$, $c \gtrless 1$, (12) may be written in the required form.

Section 8

Problem 14.

(i) In Problem 10 we have seen (cf. (6)) that

(13) $\quad \delta$ is unbiased iff $R(\theta,\delta) \le a_0a_1(a_0 + a_1)^{-1}$ for all $\theta \in \Omega$.

Let δ_0 be an unbiased procedure with uniformly minimum risk and suppose that δ_0 is inadmissible. Then there exists a procedure δ_1 with

(14) $\quad R(\theta,\delta_1) \le R(\theta,\delta_0) \quad$ for all $\theta \in \Omega$

$\quad R(\theta,\delta_1) < R(\theta,\delta_0) \quad$ for some $\theta \in \Omega$.

Since δ_0 is unbiased we have $R(\theta,\delta_0) \le a_0a_1(a_0 + a_1)^{-1}$ for all $\theta \in \Omega$ and hence $R(\theta,\delta_1) \le a_0a_1(a_0 + a_1)^{-1}$ for all $\theta \in \Omega$. By (13) δ_1 is unbiased. Because δ_0 minimizes the risk uniformly among all unbiased procedures, a contradiction with (14) is obtained. So δ_0 is admissible.

(ii) In the text of the problem the loss function is defined as $L(\theta,d) = (d - \theta)^2$. This seems to be a misprint. In this solution we take the natural loss function $L(\theta,d) = (d - e^{-\theta})^2$.
By Problem 2 we have

$$\delta \text{ is unbiased} \Leftrightarrow E_\theta\delta(X) = e^{-\theta} \qquad \text{for all } \theta \in \Omega$$

$$\Leftrightarrow \sum_{x=1}^{\infty} \delta(x)\theta^x e^{-\theta}\{x!(1-e^{-\theta})\}^{-1} = e^{-\theta} \qquad \text{for all } \theta \in \Omega$$

$$\Leftrightarrow \sum_{x=1}^{\infty} \delta(x)\theta^x (x!)^{-1} = 1 - e^{-\theta} = \sum_{x=1}^{\infty} (-1)^{x+1}\theta^x (x!)^{-1} \quad \text{for all } \theta \in \Omega$$

$$\Leftrightarrow \delta(x) = (-1)^{x+1} \qquad x = 1,2,\ldots$$

Therefore, $\delta_0(X) = (-1)^{X+1}$ is the unique unbiased estimate. Define

$$\delta_1(x) = \begin{cases} 0 & \text{if } x \text{ is even} \\ 1 & \text{if } x \text{ is odd.} \end{cases}$$

Then for all $\theta \in \Omega$ we have $R(\theta,\delta_1) = \frac{1}{2}(1-e^{-\theta}) < (1+e^{-\theta})(1-e^{-\theta}) = 1 - e^{-2\theta} = R(\theta,\delta_0)$. So δ_0 is inadmissible.

(LEHMANN (1951))

Problem 15.

We use randomized procedures in this problem, which are denoted by capitals (this differs from the notation in the statement of the problem).
Let $g_1,\ldots,g_N$ be the N different elements of the finite group. Let $Y^{(0)}$ be a procedure that uniformly minimizes the risk among all invariant procedures. The invariance of $Y^{(0)}$ implies

$$R(\theta,Y^{(0)}) = \iint L(\theta,d)dP^{Y_x^{(0)}}(d)dP_\theta(x) =$$

$$= \iint L(\bar{g}_i\theta, g_i^* d)dP^{Y_x^{(0)}}(d)dP_\theta(x) = \iint L(\bar{g}_i\theta,d)dP^{Y_{g_i x}^{(0)}}(d)dP_\theta(x)$$

$$= \iint L(\bar{g}_i\theta,d)dP^{Y_x^{(0)}}(d)dP_{\bar{g}_i\theta}(x) = R(\bar{g}_i\theta,Y^{(0)})$$

for all $\theta \in \Omega$ and $i \in \{1,\ldots,N\}$.
Suppose that $Y^{(0)}$ is inadmissible. Then there exists a procedure $Y^{(1)}$ that dominates $Y^{(0)}$. Define the procedure Y by

$$dP^{Y_x}(d) = N^{-1} \sum_{i=1}^{N} dP^{Y_{g_i x}^{(1)}}(d).$$

In Problem 11 it is shown that Y is invariant and $R(\theta,Y) = N^{-1} \Sigma_{i=1}^N R(\bar{g}_i\theta,Y^{(1)})$ for all θ. Since

$$N^{-1}\sum_{i=1}^{N} R(\bar{g}_i\theta, Y^{(1)}) \le N^{-1}\sum_{i=1}^{N} R(\bar{g}_i\theta, Y^{(0)}) = R(\theta, Y^{(0)})$$

for all θ with strict inequality for some θ,Y dominates $Y^{(0)}$ in contradiction with the definition of $Y^{(0)}$. So $Y^{(0)}$ is inadmissible.

Problem 16.

(i) We first show that the problem is invariant. We verify Lehmann's conditions on p. 11.

Let $\mathcal{P}$ be the set of distributions of the form $P\{X = \theta-1\} = P\{X = \theta+1\} = \frac{1}{2}$ for $\theta \in \mathbb{R}$.

1° $\left.\begin{array}{l}\frac{1}{2} = P\{X = \theta-1\} = P\{X+c = \theta+c-1\} = P\{gX = \theta+c-1\}\\ \frac{1}{2} = P\{X = \theta+1\} = P\{X+c = \theta+c+1\} = P\{gX = \theta+c+1\}\end{array}\right\}$ distribution of gX also in $\mathcal{P}$.

2° $\bar{g}\theta = \theta+c$ is 1 : 1 and $\bar{g}\Omega = \Omega$.

3° $g^*d = d+c$ is a homomorphism by linearity.

4° $L(\bar{g}\theta, g^*d) = \min(|\theta+c-d-c|, 1) = L(\theta,d)$.

Next we derive the invariant estimators which uniformly minimize the risk. An estimator Y is invariant iff $dP^{Y_x}(d) = dP^{Y_{x+c}}(d+c)$ for all x and c, or iff $dP^{Y_x}(d) = dP^{Y_0}(d-x)$ for all x. Furthermore,

$$R(\theta,Y) = \iint L(\theta,d)dP^{Y_x}(d)dP_\theta(x) = \iint L(\theta,d+x)dP_\theta(x)dP^{Y_0}(d)$$

$$= \tfrac{1}{2}\int (|d-1| \wedge 1) + (|d+1| \wedge 1)dP^{Y_0}(d)$$

$$\ge \tfrac{1}{2} P\{Y_0 = 1\} + \tfrac{1}{2}P\{Y_0 = -1\} + \tfrac{1}{2}P\{Y_0 \notin \{-1,1\}\} = \tfrac{1}{2}$$

with equality iff $P\{Y_0 \in \{-1,1\}\} = 1$. Therefore, $R(\theta,Y)$ is minimized by taking $P\{Y_0 \in \{-1,1\}\} = 1$, i.e. by taking the values X-1 and X+1 with probabilities p and q (independent of X). Note that for invariant estimators the risk does not depend on θ.

(ii) For the (nonrandomized) rule δ_1 the risk does depend on θ. Let $\theta < -1$, then $\theta-1$ and $\theta+1$ are both negative, so $P\{X < 0\} = 1$. Hence $R(\theta,\delta_1) = E_\theta L(\theta,X+1) = \frac{1}{2}$.

Similarly, $R(\theta,\delta_1) = \frac{1}{2}$ for $\theta \ge 1$. But if $-1 \le \theta < 1$, then $\theta-1 < 0$ and $\theta+1 \ge 0$, so $R(\theta,\delta_1) = 0$.

The rule δ_1 dominates any invariant estimator which uniformly minimizes the risk. This implies that the conclusion of Problem 15 need not hold when G is infinite.

Section 9

Problem 17.

Suppose that T is not minimal sufficient. Then there exists a function f such that U = f(T) is sufficient and $k_1, k_2, \ldots, k_r$ are the solutions of the equation f(t) = u for some u and some $r \geq 2$, where $k_i \in \{0,1,\ldots,n\}$ for all $i = 1,2,\ldots,r$ and $k_i \neq k_j$ if $i \neq j$. This means that for any $j \in \{1,2,\ldots,r\}$,

$$P\{T = k_j \mid U = u\} = P\{T = k_j, f(T) = u\} / P\{U = u\}$$
$$= \frac{P\{T = k_j\}}{\Sigma_{i=1}^{r} P\{T = k_i\}} = \frac{\binom{n}{k_j} p^{k_j}(1-p)^{n-k_j}}{\Sigma_{i=1}^{r} \binom{n}{k_i} p^{k_i}(1-p)^{n-k_i}}.$$

This expression depends on p, since $r \geq 2$, and hence U is not sufficient, implying that T is minimal sufficient.

Problem 18.

(i) Since $T = \max(X_1,\ldots,X_n) \in (0,\theta)$, we only have to consider the conditional distribution of $X_1,\ldots,X_n$ given T = t for $t \in (0,\theta)$. For all $t \in (0,\theta)$ and $x_i > 0$ $(i = 1,\ldots,n)$ $P\{X_1 \leq x_1,\ldots,X_n \leq x_n, T \leq t\} = \theta^{-n} \Pi_{i=1}^{n} (t \wedge x_i)$. On the other hand

$$P\{X_1 \leq x_1,\ldots,X_n \leq x_n, T \leq t\} =$$
$$= \int_{(0,t)} P\{X_1 \leq x_1,\ldots,X_n \leq x_n \mid T = u\} n u^{n-1}\theta^{-n} du,$$

implying that $P\{X_1 \leq x_1,\ldots,X_n \leq x_n \mid T = u\}$ is independent of θ. Therefore, T is sufficient by the definition of sufficiency.
With

$$g_\theta(T(x)) = \begin{cases} \theta^{-n} & \text{if } T \leq \theta \\ 0 & \text{if } T > \theta \end{cases} \text{ and } h(x_1,\ldots,x_n) = \begin{cases} 1 & \text{if } \min_{1\leq i\leq n} x_i \geq 0 \\ 0 & \text{if } \min_{1\leq i\leq n} x_i < 0 \end{cases}$$

in formula 20 on p. 20 it follows by the factorization criterion that T is sufficient.

(ii) With $T(x_1,\ldots,x_n) = (\min(x_1,\ldots,x_n), \Sigma_{i=1}^{n} x_i)$,

$$g_{(a,b)}(t_1,t_2) = \begin{cases} a^n \exp(nab - at_2) \\ 0 \end{cases} \text{if } t_1 \gtrless b,\ h(x_1,\ldots,x_n) = 1$$

in formula 20 on p. 20 it follows by the factorization criterion that T is sufficient.

Problem 19.

Suppose T satisfies formula 20 on p. 20, then

$$p_\theta^{T,Y}(t,y) = |J|^{-1} g_\theta(t)h(x),$$

where x and (t,y) correspond to one another with respect to the mapping given by formula 17 on p. 19. Hence

$$\int p_\theta^{T,Y}(t,y)dy = g_\theta(t) \int |J(x(t,y))|^{-1} h(x(t,y))dy.$$

Therefore, by formula 19 on p. 19

$$p_\theta^{Y|t}(y) = |J(x(t,y))|^{-1}h(x(t,y)) / \int |J(x(t,y'))|^{-1}h(x(t,y'))dy',$$

and thus is independent of θ, implying that T is sufficient for θ. Suppose T is sufficient, the $p_\theta^{Y|t}(y)$ is independent of θ and we may delete the subscript θ. From formulae 18 and 19 on p. 19 we get

$$p_\theta^X(x) = p_\theta^{T,Y}(T(x),Y(x))|J| = p^{Y|T(x)}(Y(x)) \int p_\theta^{T,Y}(T(x),y')dy'|J|.$$

Thus functions g_θ and h can be defined to satisfy formula 20 on p. 20 by

$$g_\theta(t) = \int p_\theta^{T,Y}(t,y')dy' \text{ and } h(x) = p^{Y|T(x)}(Y(x))|J|.$$

REFERENCES

KIEFER, J. (1957). Invariance, minimax sequential estimation, and continuous time processes. *Ann. Math. Statist.* 28, 573-601.

KUDO, H. (1955). On minimax invariant estimates of the transformations parameter. *Nat. Sci. Rept. Ochanomizu Univ., Tokyo*, 6, 31-73.

LEHMANN, E.L. (1951). A general concept of unbiasedness. *Ann. Math. Statist.* 22, 587-597.

PEISAKOFF, M. (1951). Transformation of parameters. Unpublished thesis, Princeton.

RUDIN, W.R. (1970). *Real and Complex Analysis*. International Student Edition. Mc Graw-Hill.

CHAPTER 2

Section 1

Problem 1.

This problem can be found as Theorem (21.6) in HEWITT and STROMBERG (1965) pp. 380-381.

Section 2

Problem 2.

(i) This follows from Theorem (19.24) and definition (19.43) in the above book, pp. 315, 316, 328.

(ii) By (19.44) of the same book (p. 328) the result is established.

(iii) Since

$$\frac{d\nu}{d\nu} = 1 \qquad \text{a.e. } \nu,$$

application of (ii) yields

$$1 = \frac{d\nu}{d\nu} = \frac{d\nu}{d\mu}\frac{d\mu}{d\nu} \qquad \text{a.e. } \nu$$

and hence

$$\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1} \qquad \text{a.e. } \mu,\nu.$$

(iv) For each $A \in \mathcal{A}$ we have

$$\sum_{k=1}^{n} \mu_k(A) = \sum_{k=1}^{n} \int_A \frac{d\mu_k}{d\lambda}\, d\lambda = \int_A \sum_{k=1}^{n} \frac{d\mu_k}{d\lambda}\, d\lambda$$

and therefore

$$\frac{d\,\Sigma_{k=1}^{n}\mu_k}{d\lambda} = \sum_{k=1}^{n}\frac{d\mu_k}{d\lambda} \qquad \text{a.e. } \lambda.$$

By Lebesgue's monotone convergence theorem and the definition of the finite measure μ we have

$$\mu(A) = \sum_{k=1}^{\infty}\mu_k(A) = \lim_{n\to\infty}\sum_{k=1}^{n}\int_A \frac{d\mu_k}{d\lambda}\,d\lambda = \lim_{n\to\infty}\int_A \sum_{k=1}^{n}\frac{d\mu_k}{d\lambda}\,d\lambda =$$

$$= \int_A \lim_{n\to\infty}\sum_{k=1}^{n}\frac{d\mu_k}{d\lambda}\,d\lambda = \int_A \lim_{n\to\infty}\frac{d\,\Sigma_{k=1}^{n}\mu_k}{d\lambda}\,d\lambda,$$

implying that μ is absolutely continuous with respect to λ, and that

$$\lim_{n\to\infty}\frac{d\,\Sigma_{k=1}^{n}\mu_k}{d\lambda} = \frac{d\mu}{d\lambda} \qquad \text{a.e. } \lambda.$$

Section 3

Problem 3.

Let $X = \mathbb{R}$, let $\mathcal{A}$ be the Borel sets of $\mathbb{R}$ and let $\mathcal{A}_0 = \{A : A \text{ or } \tilde{A} \text{ is a countable subset of } \mathbb{R}\}$. Then $\mathcal{A}_0$ is a σ-field, $\mathcal{A}_0 \subset \mathcal{A}$ and $\mathcal{A}_0 \neq \mathcal{A}$. Suppose there exists a function T such that $T^{-1}(\mathcal{B}) = \mathcal{A}_0$.
Let $A \in \mathcal{A}$ and $T(x) \in T(A)$; then $T(x) = T(a)$ for some $a \in A$. Since $\{a\} \in \mathcal{A}_0$ and $\mathcal{A}_0 = T^{-1}(\mathcal{B})$, there exists a set $B \in \mathcal{B}$ such that $\{a\} = T^{-1}(B)$. Because $T(x) = T(a)$ this implies $x \in T^{-1}(B) = \{a\}$ and hence $x = a$. So, $T^{-1}(T(A)) \subset A$. We always have $A \subset T^{-1}(T(A))$ and therefore $T^{-1}(T(A)) = A \in \mathcal{A}$, implying, by definition of $\mathcal{B}$, $T(A) \in \mathcal{B}$ and hence $A = T^{-1}(T(A)) \in \mathcal{A}_0$, in contradiction with $\mathcal{A}_0 \neq \mathcal{A}$.
So there does not exist a function T such that $T^{-1}(\mathcal{B}) = \mathcal{A}_0$.

(BAHADUR and LEHMANN (1955))

Section 4

Problem 4.

First (iii) will be proved. Then it will be shown that (i) and (ii) are special cases of (iii).
(iii) Let $B \in \mathcal{B}^n$, the Borel sets of $\mathbb{R}^n$, and let f be a $\mathcal{B}^n$-measurable

and P-integrable function. Then we have

$$\int_{T^{-1}(B)} \frac{\Sigma_{k=1}^{r} f(g_k T(x))h(g_k T(x))}{\Sigma_{i=1}^{r} h(g_i T(x))} dP(x)$$

$$= \sum_{k=1}^{r} \sum_{j=1}^{r} \int_{T^{-1}(B)} \frac{f(g_k g_j x)h(g_k g_j x)}{\Sigma_{i=1}^{r} h(g_i g_j x)} I_{\{x:T(x)=g_j x\}}(x) h(x)d\mu(x)$$

$$= \sum_{k=1}^{r} \sum_{j=1}^{r} \int_{T^{-1}(B)} \frac{f(x)h(x)}{\Sigma_{i=1}^{r} h(g_i g_k^{-1}x)} I_{\{x:T(x)=g_k^{-1}x\}}(x) h(g_j^{-1} g_k^{-1} x)d\mu(x)$$

$$= \int_{T^{-1}(B)} f(x)dP(x),$$

where we have used that $T(gx) = T(x)$ for all $g \in G$, implying $g_j^{-1}g_k^{-1}x \in \{x : T(x) = g_j x\} \Leftrightarrow x \in \{x : T(x) = g_k^{-1}x\}$. By definition of conditional expectation the proof of (iii) is complete.

(ii) follows from (iii) by taking $h(x) = 1$.

(i) follows from (ii) by taking $r = n$, $g_k(x_1,\ldots,x_n) = (x_k,x_{k+1},\ldots, x_n, x_1,\ldots,x_{k-1})$, $k = 1,\ldots,n$, and $T(x_1,\ldots,x_n) = (y_1,\ldots,y_n)$.

(It should be noted that for each sample point $(x_1,\ldots,x_n)$ the index i such that $x_i = x^{(1)} = \min (x_1,\ldots,x_n)$ has to be uniquely defined; e.g. by choosing the smallest such index in case of ties. Otherwise $(y_1,\ldots,y_n)$ is not defined uniquely.)

Section 5

Problem 5.

Theorem 4 of Chapter 2 holds true if X is a Borel space, so in particular if $X = \mathbb{R}^n$. For a proof see BREIMAN (1968) pp. 79,401 or ASH (1972) p. 265.

Problem 6.

The independence of Y and T under P_0 implies

$$E_0\left[\frac{h(y,T)}{f(y)g(T)}\right] = E_0\left[\frac{h(y,T)}{f(y)g(T)} \,\middle|\, Y = y\right]$$

and hence

$$p_1^Y(y) = \int_T h(y,t)d\nu(t) = f(y) \int_T \frac{h(y,t)}{f(y)g(t)} g(t)d\nu(t)$$

$$= f(y)E_0\left[\frac{h(y,T)}{f(y)g(T)}\right] = f(y)E_0\left[\frac{h(y,T)}{f(y)g(T)} \,\middle|\, Y = y\right].$$

Section 6

Problem 7

(i) Application of Problem 4 (iii) with G the group of all n! permutations, $h(x_1,\ldots,x_n) = 1$ and $T(x_1,\ldots,x_n) = (x^{(1)},\ldots,x^{(n)})$ yields for any function f which is P-integrable for all $P \in \mathcal{P}$ $E[f(X) \mid T(x)] = f_0(x)$, independently of $P \in \mathcal{P}$, implying that T is sufficient for $\mathcal{P}$.
(ii) From Problem 4 we know that $E[f(X) \mid Y = y] = f_0(y)$, where we use the notation introduced there. This means that Y is sufficient for $\mathcal{P}$.
(iii) Let G be the group of $2^n n!$ transformations given by

$$g(x_1,\ldots,x_n) = (c_1 x_{i_1},\ldots,c_n x_{i_n}),$$

where $(i_1,\ldots,i_n)$ is a permutation of $(1,\ldots,n)$ and $c_i = 1$ or -1, $i = 1,\ldots,n$. Defining

$$f_0(x) = (2^n n!)^{-1} \sum_{g \in G} f(gx),$$

we obtain by application of (the solution of) Problem 4 (ii) with $T(x_1,\ldots,x_n) = (W_1,\ldots,W_n)$ that for any integrable function f $E[f(X) \mid W(x)] = f_0(x)$, independently of $P \in \mathcal{P}$, inplying that W is sufficient for $\mathcal{P}$.
(Note that in the solution of Problem 4 (ii) the group G may by any finite group of transformations of $x \in \mathbb{R}^n$, not necessarily a group of transformations corresponding to permutations of coordinates.)

Problem 8.

Defining $T(x)$, $g_0(t)$, $g_1(t)$ and $h(x)$ by

$$T(x) = \begin{cases} p_1(x)/p_0(x) & \text{if } p_0(x) > 0 \\ \infty & \text{if } p_0(x) = 0 \end{cases},$$

$$g_0(t) = \begin{cases} 1 & \text{if } t \in [0,\infty) \\ 0 & \text{if } t = \infty \end{cases},$$

$$g_1(t) = \begin{cases} t & \text{if } t \in [0,\infty) \\ 1 & \text{if } t = \infty \end{cases},$$

$$h(x) = \begin{cases} p_0(x) & \text{if } p_0(x) > 0 \\ p_1(x) & \text{if } p_0(x) = 0 \end{cases},$$

if follows that $p_i(x) = g_i(T(x))h(x)$ for $i = 0,1$. In view of (31), T is sufficient for $\mathcal{P}$.

Problem 9.

(i) Following the hint we see that $P_0(S) = 0$; on $\mathcal{X} - S$ λ_j is absolutely continuous with respect to P_0 and hence

$$\frac{dP_0}{d\,\Sigma_{j=0}^n c_jP_j} = \frac{dP_0}{d\,\Sigma_{j=1}^n ((c_0/n)P_0+c_jP_j)} = \left(\frac{\Sigma_{j=1}^n ((c_0/n)dP_0+c_jdP_j)}{dP_0}\right)^{-1}$$

$$= \left(\sum_{j=1}^n f_j^{-1}\right)^{-1},$$

which is $\mathcal{A}_0$-measurable.
By Problem 2 (ii) it follows that

$$\frac{dP_0}{d\lambda} = \frac{dP_0}{d\,\Sigma_{j=0}^n c_jP_j} \frac{d\,\Sigma_{j=0}^n c_jP_j}{d\lambda}.$$

By Problem 2 (iv) we have

$$\lim_{n\to\infty} \frac{d\,\Sigma_{j=0}^n c_jP_j}{d\lambda} = 1.$$

Therefore

$$\frac{dP_0}{d\lambda} = \lim_{n\to\infty} \frac{dP_0}{d\,\Sigma_{j=0}^n c_jP_j};$$

since the limit of $\mathcal{A}_0$-measurable functions is $\mathcal{A}_0$-measurable, $\frac{dP_0}{d\lambda}$ is also $\mathcal{A}_0$-measurable.
Because P_0 is arbitrarily chosen, we have by Lemma 1 and Corollary 1 that T is sufficient for $\mathcal{P}$.
(ii) Define λ as in the hint for part (ii). By what we have just proved, $dP_{\theta_0}/(dP_{\theta_0} + d\lambda)$ is $\mathcal{A}_0$-measurable. Since

$$\frac{dP_{\theta_0}}{d\lambda} = \frac{dP_{\theta_0}}{dP_{\theta_0} + d\lambda} \left(1 - \frac{dP_{\theta_0}}{dP_{\theta_0} + d\lambda}\right)^{-1},$$

the result is established.

There are some misprints in the hint of (i): $\Sigma_{j=1}^{n} 1/f_j$ should be $(\Sigma_{j=1}^{n} 1/f_j)^{-1}$, and in the displayed formula the summation should be from 0 to n and not from 1 to n.

(HALMOS and SAVAGE (1949))

Problem 10.

The solution is given in the hint. There is one misprint: Lemma 3 (ii) should be Lemma 3 (i).

Problem 11.

Let $Y_x = (Y_x^{(1)},\dots,Y_x^{(m)})$ be a decision procedure based on x, i.e. $Y_x^{(i)}$ is the probability that decision d_i is taken when $X = x$. By Problem 10 there exists a procedure Z based on T such that

$$Z_t^{(i)} = E[Y_X^{(i)} \mid t].$$

Note that by sufficiency this conditional expectation does not depend on $P \in \mathcal{P}$, the family of distributions of X. For each $P \in \mathcal{P}$ we have

$$R(P,Y) = \int \sum_{i=1}^{m} L(P,d_i)Y_x^{(i)}dP(x) = \sum_{i=1}^{m} L(P,d_i)EY_X^{(i)} =$$

$$= \sum_{i=1}^{m} L(P,d_i)EE[Y_x^{(i)} \mid T] = \sum_{i=1}^{m} L(P,d_i)EZ_T^{(i)} = R(P,Z),$$

and hence the class of procedures based on T is essentially complete.

Section 7

Problem 12.

The first statement is proved by induction w.r.t. s. For $s = 1$ the statement is obviously true. Suppose the statement holds for s-1. Then for all integers $x \geq 0$

$$P\left\{\sum_{j=1}^{s} X_j = x\right\} = \sum_{y=0}^{x} P\left\{\sum_{j=1}^{s-1} X_j = y,\ X_s = x-y\right\}$$

$$= \sum_{y=0}^{x} \exp\left(-\sum_{j=1}^{s-1} \lambda_j\right)\left(\sum_{j=1}^{s-1} \lambda_j\right)^y (y!)^{-1} \exp(-\lambda_s)\lambda_s^{x-y}\{(x-y)!\}^{-1}$$

$$= \exp(-\lambda) \sum_{y=0}^{x} \binom{x}{y}(\lambda - \lambda_s)^y \lambda_s^{x-y}(x!)^{-1} = e^{-\lambda}\lambda^x/x!,$$

where $\lambda = \Sigma_{j=1}^{s} \lambda_j$. This completes the proof of the first statement. For all non-negative integers $t_1,\ldots,t_{s-1}$ satisfying $\Sigma_{i=1}^{s-1} t_i \le t_0$ we have with $t_s = t_0 - \Sigma_{i=1}^{s-1} t_i$

$$P\{T_1 = t_1,\ldots,T_{s-1} = t_{s-1} \mid T_0 = t_0\}$$

$$= P\{X_1 = t_1,\ldots,X_{s-1} = t_{s-1}, X_s = t_s\} / P\left\{\sum_{i=1}^{s} X_i = t_0\right\}$$

$$= \left\{\prod_{i=1}^{s} e^{-\lambda_i}\lambda_i^{t_i}(t_i!)^{-1}\right\} \Big/ \{e^{-\lambda}\lambda^{t_0}(t_0!)^{-1}\} = t_0! \prod_{i=1}^{s} p_i^{t_i}(t_i!)^{-1},$$

where $\lambda = \Sigma_{i=1}^{s} \lambda_i$ and $p_i = \lambda_i/\lambda$. This completes the proof of the second statement.

Problem 13.

(i) The density of $Y_1,\ldots,Y_n$ is given by

$$f_n(y_1,\ldots,y_n) = n! \prod_{i=1}^{n} (2\theta)^{-1} \exp\{-y_i/(2\theta)\}, \quad 0 \le y_1 \le \ldots \le y_n.$$

This implies that the density $g_n(y_1,\ldots,y_r)$ of $Y_1,\ldots,Y_r$ equals

$$\int_{y_r \le y_{r+1} \le \ldots \le y_n} \cdots \int f_n(y_1,\ldots,y_n)dy_{r+1}\ldots dy_n$$

$$= \frac{1}{(2\theta)^r} \frac{n!}{(n-r)!} \exp\left[-\frac{\Sigma_{i=1}^{r} y_i + (n-r)y_r}{2\theta}\right], \qquad 0 \le y_1 \le \ldots \le y_r,$$

cf. ROHATGI (1976) pp. 150-152.

(ii) Defining $z_i(y_1,\ldots,y_r) = (n-i+1)(y_i - y_{i-1})\theta^{-1}$, $i = 2,\ldots,r$, and $z_1(y_1,\ldots,y_r) = ny_1\theta^{-1}$ it follows that $\Sigma_{i=1}^{r} z_i(y_1,\ldots,y_r) = [\Sigma_{i=1}^{r} y_i + (n-r)y_r]\theta^{-1}$. For all Borel sets $B_1,\ldots,B_r$ we have

$$P\{Z_1 \in B_1,\ldots,Z_r \in B_r\} =$$

$$= \int_{\{(y_1,\ldots,y_r):z_i(y_1,\ldots,y_r)\in B_r, i=1,\ldots,r\}} \cdots \int g_n(y_1,\ldots,y_r)dy_1,\ldots,dy_r$$

$$= \int_{\{(z_1,\ldots,z_r):z_i\in B_i, i=1,\ldots,r\}} \cdots \int 2^{-r} \exp\left(-\tfrac{1}{2}\sum_{i=1}^{r} z_i\right)dz_1,\ldots,dz_r$$

$$= \prod_{i=1}^{r} \int_{B_i} \tfrac{1}{2}e^{-\frac{1}{2}t}dt,$$

and hence Z_i has a chi-square distribution with 2 degrees of freedom and $Z_1,\ldots,Z_r$ are independent. Since $[\Sigma_{i=1}^{r} Y_i + (n-r)Y_r]\theta^{-1} = \Sigma_{i=1}^{r} Z_i$, it has a chi-square distribution with 2r degrees of freedom. Cf. also EPSTEIN and SOBEL (1954), Corollary 2.

(iii) By Problem 1 of Chapter 1 it follows that $Z_1,\ldots,Z_r$ are independent and that Z_i $(i = 1,\ldots,r)$ has an exponential ditribution with parameter $\frac{1}{2}$, which is the same as a chi-square distribution with 2 degrees of freedom. Since $Y_i = \theta' \Sigma_{j=1}^{i} Z_j$ we have for all Borel sets $B_1,\ldots,B_r$

$$P\{Y_1 \in B_1,\ldots,Y_r \in B_r\}$$

$$= \int_{\{(z_1,..,z_r):\theta'\Sigma_{j=1}^{i} z_j\in B_i, z_i\geq 0, i=1,..,r\}} \cdots \int 2^{-r} \exp\left(-\tfrac{1}{2}\Sigma_{i=1}^{r} z_i\right)dz_1..dz_r$$

$$= \int_{\{(y_1,..,y_r):0\leq y_1\leq..\leq y_r, y_i\in B_i, i=1,..,r\}} \cdots \int (2\theta')^{-r}\exp(-y_r/(2\theta'))dy_1..dy_r$$

and hence the density of $Y_1,\ldots,Y_r$ equals

$$(2\theta')^{-r} \exp(-y_r/(2\theta')), \quad 0 \leq y_1 \leq \ldots \leq y_r.$$

Note that $Y_r/\theta' = \Sigma_{j=1}^{r} Z_j$. This implies that Y_r/θ' has a chi-square distribution with 2r degrees of freedom.

(iv) A sample of n units (tubes) is randomly selected from a population having an exponential lifetime distribution with parameter $(2\theta)^{-1}$, i.e. P[length of life of the unit $\leq x] = 1 - \exp(-(2\theta)^{-1}x)$. Each time a failure occurs on the i^{th} place $(i = 1,\ldots,n)$, the unit is replaced by a new, randomly selected unit. Let $N_i(t)$ denote the number of failures at the i^{th} location. Since the times elapsed between consecutive failures are independent and exponentially distributed with parameter $(2\theta)^{-1}$, N_i is a Poisson process with intensity $(2\theta)^{-1}$ (cf. Problem 1 of Chapter 1). Let $N(t) = \Sigma_{i=1}^{n} N_i(t)$. The sum of independent Poisson variables is also Poisson distributed. So we see that the total numbers of events (i.e. adding over $i = 1,\ldots,n$) in nonoverlapping time intervals are independently

Poisson distributed, the number in an interval of length τ having expectation $n(2\theta)^{-1}\tau$. Thus N is a Poisson proces with intensity $n(2\theta)^{-1}$.

(EPSTEIN and SOBEL (1954))

Problem 14.

It is assumed that the statements have to be proved for $\theta \in \text{int}\,\Theta$. Using Theorem 9 (ii) the proof is straightforward.

Problem 15.

We write (35) in the following form

$$dP^{X}_{\theta,\vartheta}(x) = C(\theta,\vartheta)\exp\left[\sum_{i=1}^{r}\theta_i U_i(x) + \sum_{j=1}^{k-r}\vartheta_j T_j(x)\right]d\mu(x).$$

Let $(\theta^0,\vartheta^0) \in \Omega$. For any fixed $t = t_1,\dots,t_{k-r}$ define the probability measure ν_t by

$$d\nu_t(u) = C_t(\theta^0,\vartheta^0)\exp\left(-\sum_{i=1}^{r}\theta_i^0 u_i\right)dP^{U/t}_{\theta^0,\vartheta^0}(u),$$

where $C_t(\theta^0,\vartheta^0)$ is a normalizing constant. Then we have for all $(\theta,\vartheta) \in \Omega$

$$dP^{U/t}_{\theta}(u) = C_t(\theta)\exp\left(\sum_{i=1}^{r}\theta_i u_i\right)d\nu_t(u),$$

cf. (37). Let

$$\Omega' = \Omega'(t) = \{\theta : \int \exp\left(\sum_{i=1}^{r}\theta_i u_i\right)d\nu_t(u) < \infty\}.$$

Define

$$A(\theta) = \{t : \int \exp\left(\sum_{i=1}^{r}\theta_i u_i\right)d\nu_t(u) = \infty\}.$$

We will prove (i) in the following sense:

$$\text{If } (\theta,\vartheta) \in \Omega \text{ then } P^{T}_{\theta,\vartheta}(A(\theta)) = 0.$$

Note that if $t \notin A(\theta)$ then $\theta \in \Omega'$.

So, suppose $(\theta,\vartheta) \in \Omega$, then

$$\iint c_t(\theta)\ \exp\left(\sum_{i=1}^{r}\ \theta_i u_i\right) d\nu_t(u) dP^T_{(\theta,\vartheta)}(t) < \infty$$

and hence $P^T_{\theta,\vartheta}(A(\theta)) = 0$. (Note that by equivalence of the measures this implies that $P^T_{\theta',\vartheta'}(A(\theta)) = 0$ for all $(\theta',\vartheta') \in \Omega$.)

(ii) By easy calculations it follows that $\Omega = \{(\theta_1,\theta_2) : \theta_1 < 0, \theta_2 < 0\}$. So, the projection of Ω onto θ_1 is $\{\theta_1 : \theta_1 < 0\}$. The conditional density of X given $Y = y_0 > 0$ is given by

$$p_{\theta_1}^{X|y_0}(x) = (y_0 - \theta_1)\exp\{-x(y_0 - \theta_1)\}, \quad x > 0.$$

Therefore $\Omega' = \{\theta_1 : \theta_1 < y_0\}$, implying that the projection of Ω onto θ_1 is a proper subset of Ω'.

REFERENCES

ASH, R.B. (1972). *Real Analysis and Probability*. Acad. Press.

BAHADUR, R.R. & E.L. LEHMANN (1955). Two comments on 'sufficiency and statistical decision functions'. *Ann. Math. Stat.* 26, 139-142.

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading (Mass.).

EPSTEIN, B. & M. SOBEL (1954). Some theorems relevant to life testing from an exponential distribution. *Ann. Math. Stat.* 25, 373-381.

HALMOS, P.R. & L.J. SAVAGE (1949). Application of the Radon-Nikodym Theorem to the theory of sufficient statistics. *Ann. Math. Stat.* 20, 225-241.

HEWITT, E. & K. STROMBERG (1965). *Real and Abstract Analysis*. Springer-Verlag Berlin.

ROHATGI, V.K. (1976). *An Introduction to Probability Theory and Mathematical Statistics*. Wiley, New York.

CHAPTER 3

Section 2

Problem 1.

(i) We denote max $(x_1,\ldots,x_n)$ by $x_{(n)}$. Then the density p_θ with respect to Lebesgue measure λ of a sample of size n from a uniform distribution on $(0,\theta)$ satisfies

$$p_\theta(x) = \theta^{-n} I_{(0,\theta)}(x_{(n)}), \qquad x \in \mathbb{R}^n.$$

If φ is defined as in the problem, then for any $\theta_1 > \theta_0$ the inequality

$$p_{\theta_1}(x) > \theta_0^n \theta_1^{-n} p_{\theta_0}(x)$$

implies $\varphi(x) = 1$. Because the set

$$\{x : x \in \mathbb{R}^n, p_{\theta_1}(x) < \theta_0^n \theta_1^{-n} p_{\theta_0}(x)\}$$

is empty it follows from Theorem 1 (ii) that φ is MP for testing $\theta = \theta_0$ against $\theta = \theta_1$. Because φ has been defined independently of θ_1, φ is UMP for testing $\theta = \theta_0$ against $K : \theta > \theta_0$. In view of $E_\theta\varphi(X) \le \alpha$, $\theta \le \theta_0$, the test φ is also UMP for testing $H : \theta \le \theta_0$ against $K : \theta > \theta_0$.

(ii) With φ as defined in the problem

$$E_{\theta_0}\varphi(X) = P_{\theta_0}\{X_{(n)} \le \theta_0\alpha^{\frac{1}{n}}\} = \alpha.$$

For the cases $\theta_1 > \theta_0$, $\theta_0\alpha^{\frac{1}{n}} < \theta_1 < \theta_0$ and $\theta_1 \le \theta_0\alpha^{\frac{1}{n}}$ Theorem 1 (ii) with $k = \theta_0^n\theta_1^{-n}$, $k = \theta_0^n\theta_1^{-n}$ respectively $k = 0$ shows that φ is MP for testing $\theta = \theta_0$ against $\theta = \theta_1$. Consequently φ is UMP for testing $H : \theta = \theta_0$ against $K : \theta \ne \theta_0$.

To prove uniqueness, let φ^* be any UMP test for the problem. Define $D = \{x : \varphi(x) \ne \varphi^*(x)\}$, $D_1 = D \cap \{x : x_{(n)} \le \theta_0\alpha^{\frac{1}{n}}\}$,

$D_2 = D \cap \{x : \theta_0\alpha^{\frac{1}{n}} < x_{(n)} \leq \theta_0\}$ and $D_3 = D \cap \{x : \theta_0 < x_{(n)} < K\}$ for some $K > \theta_0$. Because φ and φ^* are both UMP $E_\theta\varphi(X) = E_\theta\varphi^*(X)$ for all $\theta > 0$, which implies

$$\int_{x_{(n)}<\theta} [\varphi(x) - \varphi^*(x)]d\lambda(x) = 0, \quad \theta > 0.$$

Consequently

$$\int_{D_1} [1 - \varphi^*(x)]d\lambda(x) = \int_{x_{(n)}<\theta_0\alpha^{\frac{1}{n}}} [\varphi(x) - \varphi^*(x)]d\lambda(x) = 0 \Rightarrow \lambda(D_1) = 0,$$

$$\int_{D_2} [-\varphi^*(x)]d\lambda(x) = \int_{x_{(n)}<\theta_0} [\varphi(x) - \varphi^*(x)]d\lambda(x) +$$

$$- \int_{x_{(n)}<\theta_0\alpha^{\frac{1}{n}}} [\varphi(x) - \varphi^*(x)]d\lambda(x) = 0 \Rightarrow \lambda(D_2) = 0$$

and

$$\int_{D_3} [1 - \varphi^*(x)]d\lambda(x) = \int_{x_{(n)}<K} [\varphi(x) - \varphi^*(x)]d\lambda(x) +$$

$$- \int_{x_{(n)}<\theta_0} [\varphi(x) - \varphi^*(x)]d\lambda(x) = 0 \Rightarrow \lambda(D_3) = 0.$$

Hence, for all $K > \theta_0$, we have $\lambda(D \cap \{x : x_{(n)} < K\}) = 0$ which implies $\lambda(D) = 0$.

(NEYMAN and PEARSON (1933)).

<u>Problem 2</u>

(i) The variables $Y_i = e^{-aX_i}$, $i = 1,2,\ldots,n$, constitute a sample from the uniform distribution on $(0,e^{-ab})$. The testing problem reduces to the problem of testing $H^* : \theta = e^{-ab_0}$ against $K^* : \theta \neq e^{-ab_0}$ on the basis of $Y = (Y_1,\ldots,Y_n)$. By Problem 3.1 (ii) the UMP level α test for the original problem is given by $\varphi(x) = 1$ if $\min(x_1,\ldots,x_n) < b_0$ or $\min(x_1,\ldots,x_n) \geq b_0 - (an)^{-1}\log\alpha$ and $\varphi(x) = 0$ otherwise.

(ii) First consider the problem of testing $H : a = a_0, b = b_0$ against $K' : a = a_1, b = b_1$, where $a_1 > a_0$ and $b_1 > b_0$. By Neyman-Pearson's

fundamental lemma it follows that the test φ, given by $\varphi(x) = 1$ if $\min(x_1,\dots,x_n) < b_0$ or $\Sigma_{i=1}^n x_i < (2a_0)^{-1}\chi^2_{2n;\alpha} + nb_0$ and $\varphi(x) = 0$ otherwise, is MP at level of significance α. Here $\chi^2_{2n;\alpha}$ is the lower α-quantile of the chi-square distribution with 2n degrees of freedom.
Since the test does not depend on the particular alternative (a_1,b_1) chosen it is UMP against alternatives $a > a_0$, $b < b_0$.

The (very unusual) existence in this case of a UMP test on a two-parameter problem can be explained in the following way:
Since $P_{(a_0,b_0)}\{\min(X_1,\dots,X_n) < b_0\} = 0$ we obviously reject H if $\min(X_1,\dots,X_n) < b_0$. On the set $\{(x_1,\dots,x_n) : \min(x_1,\dots,x_n) \geq b_0\}$ the densities are strictly positive both under H and K and the role of the second parameter b is played out. What remains is a one-sided testing problem for exponential distributions.
In that sense the problem can be generalized:
Let $X_1,\dots,X_n$ be a sample from a distribution with probability density (with respect to some measure μ)

$$p_{\theta,b}(x) = C(\theta,b)e^{Q(\theta)T(x)}h(x)\cdot I_{[b,\infty)}(x),$$

where Q is strictly monotone. Then there exists a UMP test φ for testing $H : \theta = \theta_0, b = b_0$ against $K : \theta > \theta_0, b < b_0$. If Q is increasing (decreasing)

$$\varphi(x_1,\dots,x_n) = \begin{cases} 1 & \text{if } \min(x_1,\dots,x_n) < b_0 \\ 1 & \left.\begin{matrix} & >(<) \\ \text{if } \Sigma_{i=1}^n T(x_i) & = C \\ & <(>) \end{matrix}\right\} \text{ and } \min(x_1,\dots,x_n) \geq b_0 \\ \gamma & \\ 0 & \end{cases}$$

where C and γ satisfy $E_{(\theta_0,b_0)}\varphi(X_1,\dots,X_n) = \alpha$. A similar assertion can be made about families with monotone likelihood ratio and truncation.
Note that for testing H against $K^* : \theta < \theta_0, b < b_0$ the same argument holds, but that it does not work for testing H against $K' : \theta > \theta_0, b > b_0$.

(NEYMAN and PEARSON (1936))

Problem 3.

If $\alpha = 0$ or $\alpha = 1$ a nonrandomized most powerful level α test trivially exist. Therefore let $0 < \alpha < 1$ and let φ be a most powerful test for testing P_0 against P_1 at level α. Then, by Theorem 1 (iii), φ satisfies

$$\varphi(x) = \begin{cases} 1 & \text{when } p_1(x) > kp_0(x) \\ 0 & \text{when } p_1(x) < kp_0(x) \end{cases} \qquad \text{(Lebesgue a.e.)}$$

where k is some constant and p_0 and p_1 are densities (with respect to Lebesgue-measure) of P_0 and P_1 respectively.
Define the measurable set A by $A = \{x : p_1(x) = kp_0(x)\}$.
An application of the lemma with $f(x) = p_0(x)$, $a = P_0(A)$ and $b = \int_A \varphi(x)p_0(x)dx$ yields the existence of a subset B of A that satisfies $P_0(B) = \int_B p_0(x)dx = \int_A \varphi(x)p_0(x)dx$. Therefore the nonrandomized test defined by

$$\tilde{\varphi}(x) = \begin{cases} 1 & \text{when } p_1(x) > kp_0(x) \text{ or } x \in B \\ 0 & \text{when } p_1(x) < kp_0(x) \text{ or } x \in A-B \end{cases}$$

has the same level and power as φ and consequently it is a most powerful level α test.

Problem 4.

Let P_0 and P_1 with densities p_0 and p_1 belong to $\mathcal{P}$ and note that for nonnegative k

$$\{x : p_1(x) \geq kp_0(x)\} = \{x : g_1(x) \geq k/(k+1)\},$$

where $g_1 = p_1/(p_0+p_1)$ is a density of P_1 with respect to P_0+P_1. Let $\varphi_c(\cdot)$ be the critical function of the most powerful test based on T with level of significance $P_0\{g_1(x) \geq c\}$, $0 < c \leq 1$. Since T is fully informative, the tests based only on T form an essentially complete class and Theorem 1 (iii) implies that there exists a $k \in [0,1]$ such that

$$(1) \qquad \{x : g_1(x) > k\} \subset \{x : \varphi_c(T(x)) = 1\} \subset \{x : g_1(x) \geq k\}$$

holds up to sets of (P_0+P_1)-measure zero. If c is such that $P_1\{g_1(X) \geq c\} < 1$, then Theorem 1 (iii) also yields $E_0\varphi_c(T) = P_0\{g_1(X) \geq c\}$ and hence we have

$$P_0\{g_1(X) > k\} \leq E_0\varphi_c(T) = P_0\{g_1(X) \geq c\} \leq P_0\{g_1(X) \geq k\}$$

and

$$P_1\{g_1(X) > k\} \leq E_1\varphi_c(T) = P_1\{g_1(X) \geq c\}.$$

Consequently $\{x : g_1(x) > k\} \subset \{x : g_1(x) \geq c\} \subset \{x : g_1(x) \geq k\}$ holds up to sets of $(P_0 + P_1)$-measure zero. (Note that $P_0\{c \leq g_1(X) < k\} = 0$ implies $P_1\{c \leq g_1(X) < k\} = 0$). This implies that k in (1) can be chosen equal to c and hence that

(2) $$\{x : g_1(x) > c\} \subset \{x : \varphi_c(T(x)) = 1\} \subset \{x : g_1(x) \geq c\}$$

holds up to sets of $(P_0 + P_1)$-measure zero.
If c is such that $P_1\{g_1(X) \geq c\} = 1$ then

(3) $$P_0\{g_1(X) > k\} \leq P_0\{\varphi_c(T) = 1\} \leq E_0\varphi_c(T) \leq P_0\{g_1(X) \geq c\},$$

(4) $$P_1\{g_1(X) > k\} \leq P_1\{\varphi_c(T) = 1\} = E_1\varphi_c(T) = P_1\{g_1(X) \geq c\} = 1$$

and k may be chosen $\geq c$. So the second inclusion in (2) holds again because of (1). In view of (4) we also have

$$P_0\{\varphi_c(T(X)) < 1 \text{ and } g_1(X) > c\} \leq \frac{1-c}{c} P_1\{\varphi_c(T(X)) < 1\} =$$

$$= \frac{1-c}{c}[1 - P_1\{\varphi_c(T) = 1\}] = 0$$

and we conclude that (2) holds in general. But this implies that

$$\{x : g_1(x) \geq c\} = \bigcap_{n=\left[\frac{1}{c}\right]+1}^{\infty} \{x : \varphi_{c-\frac{1}{n}}(T(x)) = 1\}$$

up to a set of $(P_0 + P_1)$-measure zero. Consequently the sets $\{x : g_1(x) \geq c\}$, $0 < c \leq 1$, are contained, up to $(P_0 + P_1)$-nullsets, in the σ-field induced by T and obviously the same holds with g_1 replaced by $g_0 = p_0/(p_0 + p_1)$.
Since $(g_0(X), g_1(X))$ is clearly sufficient for (P_0, P_1) (cf. Problem 8 of Chapter 2), this implies that the statistic T is sufficient for (P_0, P_1). Problem 9 of Chapter 2 completes the argument.

Section 3

Problem 5.

(i) The number of successes X has a binomial distribution with density (with respect to counting measure) $p_p(x) = \binom{n}{x} p^x (1 - p)^{n-x}$.
For $p < p'$

$$\frac{p_{p'}(x)}{p_p(x)} = \left(\frac{p'}{p}\right)^x \left(\frac{1-p'}{1-p}\right)^{n-x}$$

is a nondecreasing function of x.
So we have to find C and γ, $0 \le \gamma \le 1$, such that φ, defined by $\varphi(x) = 1$ if $x > C$, $\varphi(x) = \gamma$ if $x = C$ and $\varphi(x) = 0$ if $x < C$, satisfies

$$E_{p_0}\varphi(X) = \gamma \binom{n}{C} p_0^C (1-p_0)^{n-C} + \sum_{x=C+1}^{n} \binom{n}{x} p_0^x (1-p_0)^{n-x} = \alpha.$$

With the help of a programmable calculator one may find

$\alpha = .05 : C = 3, \gamma = .094$
$\alpha = .10 : C = 3, \gamma = .473$
$\alpha = .20 : C = 2, \gamma = .440$

The power of the test against p_1 is given by

$$\beta(p_1) = \gamma \binom{n}{C} p_1^C (1-p_1)^{n-C} + \sum_{x=C+1}^{n} \binom{n}{x} p_1^x (1-p_1)^{n-x}.$$

This gives the following table:

p_1 \ α	.05	.10	.20
.3	.088	.158	.398
.4	.205	.310	.593
.5	.373	.492	.759
.6	.570	.675	.882
.7	.762	.835	.956

(ii) (b) In order to estimate the answer of (a) we use the normal approximation first (X is approximately distributed as N(np,np(1-p))).
If U has a N(0,1)-distribution then

$$E_{p_0}\varphi(X) = P_{p_0}\{X > C\} + \gamma P_{p_0}\{X = C\} \approx P\left\{U > \frac{C - np_0}{\sqrt{np_0(1-p_0)}}\right\} =$$

$$= \alpha = .05$$

Hence $(C - np_0)/\sqrt{np_0(1-p_0)} \approx 1.65$ ("$\approx$" means "is approximately equal to").
In the same way we find $E_{p_1}\varphi(X) = \beta = .90 \Rightarrow (C - np_1)/\sqrt{np_1(1-p_1)} \approx -1.28$.
With $p_0 = .2$ and $p_1 = .4$ this gives $n \approx 42.8$.

So the minimum sample size required in order to have $\beta(.4) \geq .90$ is approximately $n = 43$.

(a) Using binomial tables we get
for $n = 43$: $C = 13$, $\gamma = .3713$ and $\beta(.4) = .896 < .90$;
for $n = 44$: $C = 13$, $\gamma = .1508$ and $\beta(.4) = .905 > .90$.
So the required sample size is $n = 44$.

(iii) Solving $(C - np_0)/\sqrt{np_0(1-p_0)} \approx 1.65$ and $(C - np_1)/\sqrt{np_1(1-p_1)} \approx -1.28$ for $p_0 = .01$ and $p_1 = .02$ we get $n \approx 1179.05$. Hence $n = 1180$.

Problem 6.

(i) We assume that the mixed second derivative

$$\frac{\partial^2}{\partial\theta\partial x} \log p_\theta(x)$$

exists. This derivative is nonnegative for all θ and x, iff $\frac{\partial}{\partial x} \log [p_{\theta'}(x)/p_\theta(x)]$ is nonnegative for all $\theta' > \theta$ and all x; or, iff $\log [p_{\theta'}(x)/p_\theta(x)]$ is nondecreasing in x for all $\theta' > \theta$. Since the logarithm is increasing this is equivalent to the monotone likelihood ratio property.
Here we assumed also that for any $\theta' > \theta$ the densities p_θ and $p_{\theta'}$ are distinct, which may be phrased as follows. There exist θ'' and x' such that $\theta < \theta'' < \theta'$ and

$$\frac{\partial^2}{\partial\theta_0\partial x} \log p_{\theta_0}(x) > 0 \quad \text{for } \theta_0 = \theta'' \text{ and } x = x'.$$

(ii) The result follows by observing

$$\frac{\partial^2 \log p_\theta(x)}{\partial\theta\partial x} = \frac{\partial}{\partial\theta}\left(\frac{1}{p_\theta(x)} \cdot \frac{\partial p_\theta(x)}{\partial x}\right) = \frac{1}{p_\theta(x)} \frac{\partial^2 p_\theta(x)}{\partial\theta\partial x} + - \frac{1}{p_\theta(x)^2} \frac{\partial p_\theta(x)}{\partial\theta} \cdot \frac{\partial p_\theta(x)}{\partial x} .$$

Problem 7.

Consider the UMP tests φ_α based on T and define $\hat{\alpha} = \inf \{\alpha : \varphi_\alpha(x) = 1\}$. By Theorem 2 a UMP test is given by

$$\varphi_\alpha(x) = \begin{cases} 1 & > \\ \gamma_\alpha \quad \text{when } T(x) & = c_\alpha, \\ 0 & < \end{cases}$$

where c_α and γ_α satisfy

$$E_{\theta_0}\varphi_\alpha(X) = P_{\theta_0}\{T > c_\alpha\} + \gamma_\alpha P_{\theta_0}\{T = c_\alpha\} = \alpha.$$

Now we note that

$$\varphi_\alpha(x) < 1 \Rightarrow T(x) \le c_\alpha \Rightarrow P_{\theta_0}\{T \ge T(x)\} \ge P_{\theta_0}\{T \ge c_\alpha\} \ge \alpha.$$

On the other hand, if $\gamma_\alpha = 1$ then

$$\varphi_\alpha(x) = 1 \Rightarrow T(x) \ge c_\alpha \Rightarrow P_{\theta_0}\{T \ge T(x)\} \le P_{\theta_0}\{T \ge c_\alpha\} = \alpha$$

and if $\gamma_\alpha < 1$ then

$$\varphi_\alpha(x) = 1 \Rightarrow T(x) > c_\alpha \Rightarrow P_{\theta_0}\{T \ge T(x)\} \le P_{\theta_0}\{T > c_\alpha\} \le \alpha.$$

Summarizing we see that

$$(5) \qquad P_{\theta_0}\{T \ge T(x)\} < \alpha \Rightarrow \varphi_\alpha(x) = 1 \Rightarrow P_{\theta_0}\{T \ge T(x)\} \le \alpha.$$

But this implies, with $t = T(x)$,

$$(6) \qquad \begin{aligned} P_{\theta_0}\{T \ge t\} &= \inf\{\alpha : P_{\theta_0}\{T \ge T(x)\} \le \alpha\} \\ &\le \inf\{\alpha : \varphi_\alpha(x) = 1\} (= \hat{\alpha}) \\ &\le \inf\{\alpha : P_{\theta_0}\{T \ge T(x)\} < \alpha\} = P_{\theta_0}\{T \ge t\}. \end{aligned}$$

Problem 8.

(i) By Problem 13, Chapter 2, the joint distribution of $Y = (Y_1, \ldots, Y_r)$ is an exponential family with density (with respect to Lebesgue-measure)

$$p_\theta(y) = \frac{1}{(2\theta)^r} \cdot \frac{n!}{(n-r)!} \exp\left(-\frac{T(y)}{2\theta}\right), \quad 0 \le y_1 \le \ldots \le y_r,$$

where $T(y) = \Sigma_{i=1}^r y_i + (n-r)y_r$,
Because, by Theorem 9, Chapter 2, the power function $\beta(\theta) = E_\theta\varphi(Y)$ of any test φ is continuous in θ (> 0), a UMP test for testing $H^* : \theta \ge \theta_0 = 1000$ against $K : \theta < \theta_0$ is also UMP for testing $H : \theta > \theta_0$ against K and vice

versa. By reversing inequalities in Corollary 2 we have that the test φ, given by $\varphi(y) = 1$ if $T(y) \le C$ and $\varphi(y) = 0$ if $T(y) > C$, is UMP for testing H against K.

Also by Problem 13, Chapter 2, we see that the statistic $T(Y)/\theta$ has a χ^2-distribution with 2r degrees of freedom and distribution function F_r. Hence C is determined by $C/1000 = \chi^2_{2r;.05}$ (cf. Problem 2). For r = 4 we find C = 2733.0.

The power of this test against $\theta_1 = 500$ is $P_{500}\{T(Y) \le 2733.0\} = F_4\left(\frac{2733.0}{500}\right) = F_4(5.466) = .29$.

(ii) We have to find values of r such that $P_{500}\{X \le C\} \ge .95$.
Since $C = 1000 \cdot \chi^2_{2r;.05}$ we have

$$P_{500}\{T(Y) \le C\} = P_{500}\{T(Y)/500 \le \frac{1000}{500}\chi^2_{2r;.05}\} = F_r(2\chi^2_{2r;.05})$$

In order that this expression is at least .95 r must satisfy $r \ge 23$.

(EPSTEIN and SOBEL (1953))

<u>Problem 9</u>.

That X has a Poisson distribution with parameter $\lambda\tau$ and that $2\lambda T$ has a χ^2-distribution with 2r degrees of freedom can be seen for example from Section I.4 of FELLER (1971).

We consider the UMP tests for testing $H : \lambda \le \lambda_0$ at level α based on X respectively T. For each choice of r the test based on T has the power $\beta(r)$, say, at the alternative λ_1. Hence, since r is a natural number, one can *not* obtain every prespecified power β at λ_1. Thanks to the fact that τ is real, this phenomenon does not occur for the tests based on X.

Let r be fixed and let us consider now the UMP test based on T for testing H at level α, which has power $\beta(r)$ at λ_1. This test rejects for realizations of T less than τ, say. If X is the number of events occuring in a time interval of length τ then (cf. again Section I.4 of Feller (1971))

$$P_{\lambda,r}\{T \le \tau\} = P_{\lambda,\tau}\{X \ge r\} \tag{7}$$

and we see that the nonrandomized UMP tests based on X respectively T have the same power functions and hence are equivalent.

(i) Since $ET = (2\lambda)^{-1}E(2\lambda T) = r/\lambda$, the desired ratio equals $\lambda\tau/r$.

(ii) Using the χ^2-distribution we see that we have to choose here r = 19 and $\tau \approx 12.44$. Consequently the first design has a smaller (expected) time of observation than the second one iff $\lambda < r/\tau \approx 1.53$.

Problem 10.

(i) If g is a function with $p_1(x)/p_0(x) = g(T(x))$, then by Lemma 2.2 we have for every Borel set B

$$\int_B p_1'(t)d\nu(t) = \int_{T^{-1}(B)} p_1(x)d\mu(x) \geq \int_{T^{-1}(B)} g(T(x))p_0(x)d\mu(x)$$
$$= \int_B g(t)p_0'(t)d\nu(t)$$

with equality if g(t) is finite on B. In view of this and

$$P_0\{g(T) = \infty\} = P_0\{g(T(X)) = \infty\} = P_0\{p_0(X) = 0\} = 0$$

we see that $p_1'(t)/p_0'(t) = g(t)$ holds for almost all t.

(ii) This is an immediate consequence of Lemma 2.(i) in the proof of which it is implicitly assumed that $E_1\psi(T) \geq E_0\psi(T)$ is implied by $\int \psi(t)[p_1'(t) - p_0'(t)]d\nu(t) \geq 0$. This implication is valid under the convention that $E_1\psi(T) \geq E_0\psi(T)$ holds if $E_1\psi(T)$ and $E_0\psi(T)$ both exist and satisfy the inequality or if $E_1\psi(T)$ does not exist and $E_0\psi(T) = -\infty$, or if $E_1\psi(T) = \infty$ and $E_0\psi(T)$ does not exist or if $E_1\psi(T)$ and $E_0\psi(T)$ do not exist.

(iii) The first part follows from (i) and the version of the second part we will show is the following one: $E_0\psi(T) < E_1\psi(T)$ unless $\psi(T(x))$ is constant where $p_0 \neq p_1$ a.e. $(P_0 + P_1)$ or $E_0\psi(T)$ and $E_1\psi(T)$ are both ∞, $-\infty$ or undefined. Here we agree to say that $E_0\psi(T) < E_1\psi(T)$ also holds if $E_0\psi(T)$ does not exist but $E_1\psi(T) = \infty$ or if $E_0\psi(T) = -\infty$ but $E_1\psi(T)$ does not exist. With this convention we must only show that $E_0\psi(T) < E_1\psi(T)$ when both $E_0\psi(T)$ and $E_1\psi(T)$ are finite and $\psi(T(x))$ is not constant where $p_0 \neq p_1$ a.e. $(P_0 + P_1)$. However in this case, the proof of Lemma 2 (i) can be carried out with the first or second inequality in the displayed chain of inequalities being strict; for either $\psi(T(x))$ is not constant on A a.e. $(P_0 + P_1)$ and $\mu(A) > 0$, or on B a.e. $(P_0 + P_1)$ and $\mu(B) > 0$, or $a < b$ and $\mu(B) > 0$. Note that the extra condition "where $p_0 \neq p_1$" is necessary, for otherwise it is possible that a t exists such that $0 < p_0'(t) = p_1'(t) < 1$, $\nu((-\infty,t)) > 0$, $\nu(\{t\}) > 0$,

$\nu((t,\infty)) = 0$, and such that $\psi = 0$ on $(-\infty,t)$ and $\psi = b > 0$ on $[t,\infty)$.

(iv) By the concavity of the log-function we have

$$-\infty \le E_0 \log \left(\frac{p_1(X)}{p_0(X)}\right) \le \log \left(E_0 \frac{p_1(X)}{p_0(X)}\right) \le 0$$

and similarly

$$\infty \ge E_1 \log \left(\frac{p_1(X)}{p_0(X)}\right) = - E_1 \log \left(\frac{p_0(X)}{p_1(X)}\right) \ge 0.$$

With $T(x) = p_1(x)/p_0(x)$ and $\psi(t) = \log t$ the strict inequality now follows from (iii) once we have verified that $\log (p_1(x)/p_0(x))$ is not constant where $p_0 \neq p_1$, a.e. $(P_0 + P_1)$.
It this were not the case, then we should have either $p_1 \ge p_0$ a.e. $(P_0 + P_1)$ or $p_1 \le p_0$ a.e. $(P_0 + P_1)$; but this is impossible if $P_0 \neq P_1$. We note that it is just as easy to check the conditions for strict inequality in the concavity argument.

Problem 11.

By Lemma 1 there exist two nondecreasing functions f_0 and f_1, and a random variable V, such that $f_0(v) \le f_1(v)$, for all real v, and the cumulative distribution functions of $f_0(V)$ and $f_1(V)$ are F_0 and F_1, respectively. Hence for any nondecreasing function Ψ

$$E_0\Psi(X) = E\Psi[f_0(V)] \le E\Psi[f_1(V)] = E_1\Psi(X)$$

if the expectations exist.

Section 4

Problem 12.

The experiment (f,g) is more informative than (f',g'), that is

(8) $$\sup \{E_{g'}\varphi(X') : E_{f'}\varphi(X') \le \alpha\} \le \sup \{E_g\varphi(X) : E_f\varphi(X) \le \alpha\},$$

for all $\alpha \in [0,1]$.
Consider any $\alpha' \in [0,1]$. By Theorem 1 there exists a test φ' such that

$$\sup \{E_{f'}\varphi(X') : E_{g'}\varphi(X') \le \alpha'\} = E_{f'}\varphi'(X')$$

and $E_{g'}\varphi'(X') = \alpha'$. Using (8) with $\alpha = E_{f'}[1-\varphi'(X')]$ and, again, Theorem 1 we conclude that there exists a test φ^* such that

$$E_f\varphi^*(X) \le \alpha = E_{f'}[1-\varphi'(X')] \text{ and } E_g\varphi^*(X) \ge E_{g'}[1-\varphi'(X')].$$

This implies

$$E_f[1-\varphi^*(X)] \ge E_{f'}\varphi'(X') \text{ and } E_g[1-\varphi^*(X)] \le E_{g'}\varphi'(X') = \alpha'$$

and hence

$$\sup\{E_{f'}\varphi(X') : E_{g'}\varphi(X') \le \alpha'\} = E_{f'}\varphi'(X') \le E_f[1-\varphi^*(X)] \le$$

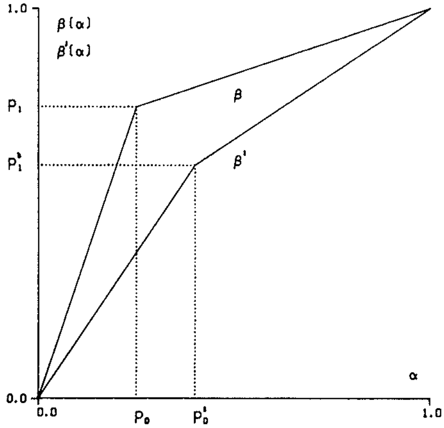
$$\le \sup\{E_f\varphi(X) : E_g\varphi(X) \le \alpha'\}.$$

Since α' was arbitrarily chosen, it follows that (g,f) is more informative than (g',f').

(BLACKWELL (1951, 1953))

Problem 13.

(i) Let X and X' be two random variables taking on the values 1 and 0 and let under H_0 : $P\{X = 1\} = p_0$, $P\{X' = 1\} = p_0'$, and under H_1 : $P\{X = 1\} = p_1$, $P\{X' = 1\} = p_1'$ $(0 < p_0, p_1, p_0', p_1' < 1)$.
Without loss of generality we assume: $p_0 < p_0'$, $p_0' < p_1'$ and $p_0 < p_1$.
Let φ_α be the critical function and $\beta(\alpha)$ the power of the most powerful level α test for testing H_0 against H_1 based on X.
Then for $\alpha \le p_0$ $\varphi_\alpha(1) = \alpha p_0^{-1}$, $\varphi_\alpha(0) = 0$, $\beta(\alpha) = \alpha p_0^{-1} p_1$ and for $p_0 \le \alpha$ $\varphi_\alpha(1) = 1$, $\varphi_\alpha(0) = (\alpha - p_0)(1-p_0)^{-1}$, $\beta(\alpha) = p_1 + (\alpha - p_0)(1-p_0)^{-1}(1-p_1)$.
An analogous expression for the power $\beta'(\alpha)$ of the most powerful level α test for testing H_0 against H_1 based on X' may be derived.
We have the following situation:



Hence X is more informative than X'

iff $\beta'(\alpha) \le \beta(\alpha)$ for all α, $0 \le \alpha \le 1$

iff $p_1' \le \beta(p_0') = p_1 + (p_0' - p_0)(1-p_0)^{-1}(1-p_1)$

iff $1 - p_1' \ge (1-p_1)[1-(p_0'-p_0)(1-p_0)^{-1}]$

iff

(9) $$(1-p_0)(1-p_1') \ge (1-p_0')(1-p_1).$$

(ii) We shall prove that a sample $X_1,\ldots,X_n$ from X is more informative than a sample $X_1',\ldots,X_n'$ from X' iff (9) holds. Sufficiency: let U_0 and U_1 be uniformly distributed over (0,1) and let U_0, U_1 and X be independent under both H_0 and H_1. For any $\gamma_0,\gamma_1 \in [0,1]$, define $Y(\gamma_0,\gamma_1) = 1$ if $X = 1$ and $U_1 \le \gamma_1$ and if $X = 0$ and $U_0 \le \gamma_0$ and $Y(\gamma_0,\gamma_1) = 0$ otherwise. Then $P_{H_0}\{Y(\gamma_0,\gamma_1) = 1\} = \gamma_1 p_0 + \gamma_0(1-p_0)$ and $P_{H_1}\{Y(\gamma_0,\gamma_1) = 1\} = \gamma_1 p_1 + \gamma_0(1-p_1)$. So $Y(\gamma_0,\gamma_1)$ has the same distribution as X' under both H_0 and H_1, for some γ_0 and γ_1, iff the system of equations

(10) $$\begin{cases} \gamma_1 p_0 + \gamma_0(1-p_0) = p_0' \\ \gamma_1 p_1 + \gamma_0(1-p_1) = p_1' \end{cases}$$

has a solution $\gamma_0,\gamma_1 \in [0,1]$. Since the solution without the restriction $\gamma_0,\gamma_1 \in [0,1]$ is

$$\hat{\gamma}_0 = \frac{p_0'p_1 - p_0p_1'}{p_1 - p_0} \le \frac{p_0'p_1 - p_0p_1' + p_1' - p_0'}{p_1 - p_0} = \hat{\gamma}_1$$

and since, in view of (10), $\gamma_1 \le 1$ implies $\gamma_0 \ge 0$, $Y(\hat{\gamma}_0,\hat{\gamma}_1)$ has the same distribution as X' under both H_0 and H_1 iff $p_0'p_1 - p_0p_1' + p_1' - p_0' \le p_1 - p_0$ iff $p_1'(1-p_0) - p_0'(1-p_1) \le (1-p_0) - (1-p_1)$ iff (9) holds. Now the sufficiency follows from the theory of Section 4, p. 76.

Necessity: for $\gamma \in [0,1]$, consider the most powerful level $\alpha(\gamma) = 1-\gamma(1-p_0)^n$ test for testing H_0 against H_1 based on $X_1,\ldots,X_n$. Its critical function is the one that rejects with probability $1-\gamma$ when $\Sigma X_i = 0$ and with probability 1 otherwise. Hence its power equals $\beta[\alpha(\gamma)] = 1-(1-p_1)^n + (1-\gamma)(1-p_1)^n = 1-\gamma(1-p_1)^n$. The most powerful test based on $X_1',\ldots,X_n'$, which rejects with probability $1-\delta$, $0 \le \delta \le 1$, when $\Sigma X_i = 0$ and with probability 1 otherwise, has the same level iff $1-\delta(1-p_0')^n = 1-\gamma(1-p_0)^n$, which is only possible for $0 \le \gamma \le (1-p_0')^n(1-p_0)^{-n} < 1$.
If the sample $X_1,\ldots,X_n$ is more informative than $X_1',\ldots,X_n'$ then $1-\delta(1-p_1')^n \le 1-\gamma(1-p_1)^n$ with $\delta(1-p_0')^n = \gamma(1-p_0)^n$ and $\gamma > 0$ sufficiently small. This implies $\gamma(1-p_1)^n \le \gamma(1-p_0)^n(1-p_0')^{-n}(1-p_1')^n$ for sufficiently small, positive γ, which is equivalent to (9).

(BLACKWELL (1951, 1953))

Problem 14.

Let X and X' be 0-1 random variables and let their probabilities of being equal to 1 be given by the third and fourth row respectively of the second table in Example 4. Let $p \le \pi$ and let U be uniformly distributed on (0,1) and independent of X. We define

$$Y = \begin{cases} 0 & \text{if } X = 0 \text{ and } U \le (1-\pi)/(1-p) \\ 1 & \text{elsewhere.} \end{cases}$$

Now we have $P_H\{Y = 0\} = 1-\pi$ and $P_K\{Y = 0\} = 1-(\pi-\rho)/(1-p)$. Consequently Y and X' have the same distribution, which shows that X is sufficient for X' and hence is the more informative of the two.
In proving that B and $\tilde{B}$ are not comparable we may assume without loss of generality that $p \le \min(\pi, 1-\pi) \le \frac{1}{2}$ and $\rho < p\pi$ as can be seen by studying the first table of Example 4 and interchanging B and $\tilde{B}$ if necessary.
Now let $V = (V_1,\ldots,V_n)$ and $V' = (V_1',\ldots,V_n')$ be samples from $\tilde{B}$ and B, respectively. Put $Y_i = 1$ if $V_i \in A$, $Y_i = 0$ if $V_i \in \tilde{A}$ and $Y_i' = 1$ if $V_i' \in \tilde{A}$,

$Y_i' = 0$ if $V_i' \in A$ $(i = 1,\dots,n)$. Then under H $q_0 \overset{\text{def}}{=} P\{Y_i = 1\} = p$, $q_0' \overset{\text{def}}{=} P\{Y_i' = 1\} = 1 - p$ and under K $q_1 \overset{\text{def}}{=} P\{Y_i = 1\} = (p - \rho)/(1 - \pi)$, $q_1' \overset{\text{def}}{=} P\{Y_i' = 1\} = 1 - \rho/\pi$. Hence $q_0 < q_0'$, $q_0 < q_1$ and $q_0' < q_1'$.
Suppose V is more informative than V'. Then, by Problem 13 (i), $(1 - q_1)(1 - q_0') \le (1 - q_0)(1 - q_1')$, that is

$$\left(1 - \frac{p-\rho}{1-\pi}\right)p \le (1-p)\frac{\rho}{\pi} \text{ iff } (\rho - p\pi)(1 - \pi - p) \ge 0$$

iff $\rho \ge p\pi$, a contradiction.
Furthermore, suppose V' is more informative than V. Then, by Problem 13 (ii), $q_0'q_1 \le q_0q_1'$, that is

$$(1 - p)\frac{p-\rho}{1-\pi} \le p\left(1 - \frac{\rho}{\pi}\right) \text{ iff } (\pi - p)(\rho - \pi p) \ge 0 \text{ iff } \rho \ge p\pi,$$

again a contradiction.
Hence samples from B and $\tilde{B}$ are not comparable.

(BLACKWELL (1951, 1953))

Problem 15.

Consider the problem of testing $H : \lambda = \lambda_0$ against $K : \lambda = \lambda_1$. Let v_1 and v_2 be two different values of v. Without loss of generality we assume $0 < \lambda_0 < \lambda_1$ (Problem 12) and $0 < v_1 < v_2$.
Let X and X' be random variables taking on the values 0 and 1 with $P\{X = 1\} = 1 - e^{-\lambda_0 v_1}$ or $1 - e^{-\lambda_1 v_1}$ and $P\{X' = 1\} = 1 - e^{-\lambda_0 v_2}$ or $1 - e^{-\lambda_1 v_2}$.
Then the conditions of Problem 13 (i) are satisfied. Hence X is more informative than X' iff

$$\begin{aligned} &e^{-\lambda_1 v_1}e^{-\lambda_0 v_2} \le e^{-\lambda_0 v_1}e^{-\lambda_1 v_2} \text{ iff } \lambda_0(v_1 - v_2) \le \lambda_1(v_1 - v_2) \\ &\text{iff } \lambda_0 \ge \lambda_1. \end{aligned} \tag{11}$$

It follows that X is not more informative than X'. By Problem 13 (ii) we have: X' is more informative than X iff $(1 - e^{-\lambda_1 v_1})(1 - e^{-\lambda_0 v_2}) \le (1 - e^{-\lambda_0 v_1})(1 - e^{-\lambda_1 v_2})$. Define $f(v_1,v_2) = (1 - e^{-\lambda_1 v_1})(1 - e^{-\lambda_0 v_2}) + - (1 - e^{-\lambda_0 v_1})(1 - e^{-\lambda_1 v_2})$. Then

$$\frac{\partial}{\partial v_2} f(v_1,v_2) = \lambda_0(1 - e^{-\lambda_1 v_1})e^{-\lambda_0 v_2} - \lambda_1(1 - e^{-\lambda_0 v_1})e^{-\lambda_1 v_2}$$

and

$$\frac{\partial^2}{\partial v_1 \partial v_2} f(v_1,v_2) = \lambda_0\lambda_1 e^{-\lambda_1 v_1}e^{-\lambda_0 v_2} - \lambda_0\lambda_1 e^{-\lambda_0 v_1}e^{-\lambda_1 v_2} > 0,$$

by (11).

So $(\partial/\partial v_2)f(v_1,v_2)$ is strictly increasing in the first coordinate v_1 $(< v_2)$. Furthermore $(\partial/\partial v_2)f(0,v_2) = 0$. Hence $(\partial/\partial v_2)f(v_1,v_2) > 0$ if $0 < v_1 \le v_2$. This means that $f(v_1,v_2)$ is stricly increasing in v_2 for $0 < v_1 < v_2$. Because $f(v_1,v_1) = 0$ it follows that $f(v_1,v_2) > 0$ for $0 < v_1 < v_2$.

Hence X' is not more informative than X.

Concluding we see that X and X' are not comparable.

(BLACKWELL (1951, 1953))

Section 5

Problem 16.

First we note that for $q = 1-p$ and $t > [t]$ we have $[n-t+1] = n-[t]$ and hence

$$\begin{aligned} P_p\{T < t\} &= (t-[t])P_p\{X = [t]\} + P_p\{X \le [t]-1\} \\ &= (t-[t])P_q\{X = n-[t]\} + P_q\{X \ge n-[t]+1\} \\ &= 1 - \{(1+[t]-t)P_q\{X = n-[t]\} + P_q\{X \le n-[t]-1\}\} \\ &= 1 - P_q\{T < n-t+1\}. \end{aligned}$$

By continuity considerations we see that for all t

$$P_p\{T < t\} = 1 - P_q\{T < n-t+1\}.$$

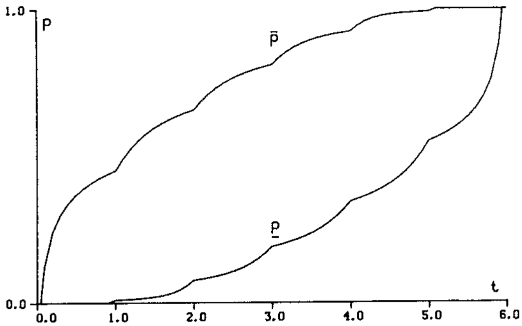
In view of the above $\underline{p}$ is easily obtained from $\bar{p}$ by

$$\underline{p}(t) = 1-\bar{p}(n-t+1).$$

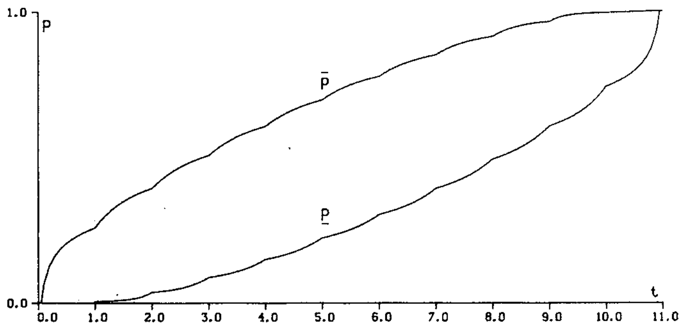
Further we have $\bar{p}^*(t) = \bar{p}(t)$ for $t \ge .05$ and $\bar{p}^*(t) = 0$ for $t \le .05$. We therefore only present graphs of $\bar{p}$ and $\underline{p}$.

Note that for $t \in \mathbb{N}$ the left- and righthand derivatives of $\bar{p}$ are different!

N = 5



N = 10



Problem 17.

The hint is a complete solution. Note that

$$E_\theta L(\theta,\underline{\theta}) = P_\theta\{\underline{\theta}^* \le \theta\} \cdot \int L(\theta,u)dF(u).$$

Section 6

Problem 18.

Absorbing, without loss of generality, the factor h(x) into μ, define

$\Omega = \{\theta : 0 < \int \exp [Q(\theta)T(x)]d\mu(x) < \infty\}$. By Theorem 9, Chapter 2, the integral $\int \psi(x) \exp [\eta T(x)]d\mu(x)$ is differentiable with respect to η in the interior of $Q(\Omega) = \{Q(\theta) : \theta \in \Omega\}$, for any bounded measurable function ψ. Its derivative is $\int \psi(x)T(x) \exp [\eta T(x)]d\mu(x)$.
Since Q is differentiable in int Ω (= interior of Ω) and since Q is strictly monotone, Q (int Ω) $\subset$ int $Q(\Omega)$ and $\int \psi(x) \exp [Q(\theta)T(x)]d\mu(x)$ is differentiable in int Ω with derivative $Q'(\theta) \int \psi(x)T(x) \exp [Q(\theta)T(x)]d\mu(x)$.
Hence, taking $\psi(x) \equiv 1$, $C(\theta) \int \exp [Q(\theta)T(x)]d\mu(x) \equiv 1$ implies

$$C'(\theta) \int \exp [Q(\theta)T(x)]d\mu(x) + C(\theta)Q'(\theta) \int T(x) \exp [Q(\theta)T(x)]d\mu(x) \equiv 0,$$

that is

$$(12) \qquad \frac{C'(\theta)}{C(\theta)} = -Q'(\theta)E_\theta T(X), \quad \text{for all } \theta \in \text{int } \Omega.$$

Also, taking $\psi(x) \equiv \varphi(x)$, the power function $\beta(\theta) = C(\theta) \int \varphi(x) \exp [Q(\theta)T(x)]d\mu(x)$ is differentiable in int Ω with derivative

$$\begin{aligned}\beta'(\theta) &= C'(\theta) \int \varphi(x) \exp [Q(\theta)T(x)]d\mu(x) + \\ &\quad + C(\theta)Q'(\theta) \int \varphi(x)T(x) \exp [Q(\theta)T(x)]d\mu(x) \\ &= \frac{C'(\theta)}{C(\theta)} E_\theta\varphi(X) + Q'(\theta)E_\theta\varphi(X)T(X).\end{aligned}$$

Combining this with (12) we see that

$$\beta'(\theta) = Q'(\theta)\{E_\theta\varphi(X)T(X) - E_\theta\varphi(X)\cdot E_\theta T(X)\}, \qquad \theta \in \text{int } \Omega.$$

Define $\varphi_0(t) = 1, \gamma, 0$ as $t >, =, < C$ then, for $\theta \in$ int Ω,

$$\begin{aligned}&E_\theta\varphi(X)T(X) - E_\theta\varphi(X)\cdot E_\theta T(X) \\ &= E_\theta\varphi_0(T)[T - E_\theta T] = E_\theta[\varphi_0(T) - 1][T - E_\theta T] \\ &= \begin{cases} \int_{(E_\theta T,\infty)} \varphi_0(t)[t - E_\theta T]dP_\theta^T(t) > 0, & \text{when } C \geq E_\theta T \\ \int_{(-\infty,E_\theta T)} [\varphi_0(t) - 1][t - E_\theta T]dP_\theta^T(t) > 0, & \text{when } C < E_\theta T. \end{cases}\end{aligned}$$

Remark: it is implicitely assumed that T(x) is not a constant (a.e. μ).

It follows that $\beta'(\theta) > 0$ for all $\theta \in$ int Ω for which $Q'(\theta) > 0$.

Problem 19.

(i) Preliminary remark: since the formulation of the problem is ambiguous we assume, in view of the hint, that one wishes to find a selection procedure such that the expected proportion of the candidates being selected is α. Let $\varphi(x)$ be the probability of being selected for a member with measurements $x = (x_1,\ldots,x_n)$. Then "the expectation of Y for the selected group" equals $EY\cdot\varphi(X) = \int \varphi(x)E(Y \mid x)dP^X(x)$. So we must maximize $\int \varphi(x)E(Y \mid x)dP^X(x)$ subject to $\int \varphi(x)dP^X(x) = \alpha$. For this we apply theorem 5 (ii) with $m = 1$, $f_1(x) = 1$, $f_2(x) = E(Y \mid x)$, $c_1 = \alpha$ and $\varphi(x) = 1, \gamma, 0$ as $E(Y \mid x) >, =, < C$, where γ and C satisfy $P\{E(Y \mid X) > C\} = \gamma P\{E(Y \mid X) = C\} = \alpha$. Now the desired result follows.

(ii) This follows in the same way when we take $f_2(x) = P\{Y \geq y_0 \mid x\}$.

(BIRNBAUM and CHAPMAN (1950))

Problem 20.

(i) Suppose there existed $c_1, c_2, \ldots$ such that $p_0 = \Sigma\, c_n p_n$ (a.e. with respect to Lebesgue measure). Now up to a set of Lebesgue measure zero, $p_0 = 1$ on $[0,1]$, zero elsewhere; and $p_n = \frac{n}{n+1}$ on $[0,1+\frac{1}{n}]$, zero elsewhere. So on $(1+\frac{1}{n+1}, 1+\frac{1}{n}]$ we have (a.e.) $0 = \Sigma_{i=1}^{n} c_i \frac{i}{i+1}$, $n = 1,2,\ldots$ Successive substitution shows that $c_i = 0$ for all i, a contradiction.

(ii) Suppose φ is a test such that $\int \varphi(x)p_n(x)dx = \alpha$ for all $n \geq 1$. Then we must have

$$\alpha = \frac{n}{n+1}\int_0^{1+\frac{1}{n}} \varphi(x)dx, \text{ for all } n \geq 1.$$

Now dominated convergence implies

$$\alpha = \lim_{n\to\infty} \frac{n}{n+1}\int_0^{1+\frac{1}{n}} \varphi(x)dx = \int_0^1 \varphi(x)dx = \int \varphi(x)p_0(x)dx,$$

as was to be shown.

Problem 21.

Let u satisfy the side conditions $F_i(u) \leq c_i$ $(i = 1,2,\ldots,m)$. Then from

$$k_i[F_i(u_0) - F_i(u)] \geq 0 \quad (i = 1,2,\ldots,m)$$

and

$$F_{m+1}(u) - \sum_{i=1}^{m} k_i F_i(u) \le F_{m+1}(u_0) - \sum_{i=1}^{m} k_i F_i(u_0)$$

it follows that

$$F_{m+1}(u_0) - F_{m+1}(u) \ge \sum_{i=1}^{m} k_i [F_i(u_0 - F_i(u)] \ge 0.$$

Hence, if u_0 satisfies the side conditions, u_0 maximizes F_{m+1} subject to $F_i(u) \le c_i$ $(i = 1,2,\ldots,m)$.

Section 7

Problem 22.

Using the method of p. 90 one obtains $C_1 = 5$, $\gamma_1 = .3969$, $C_2 = 8$ and $\gamma_2 = .6459$. The power of the test against the alternative $p = .4$ equals .5338.

Problem 23.

We have to prove that the exponential family with densities

$$p_\theta(x) = C(\theta)e^{Q(\theta)T(x)}g(x),$$

with $T(x) = x$ and $Q(\theta) = \theta$ is strictly of Pōlya type. It is necessary to assume that the function $g(x)$ is strictly positive, since $\Delta_n = 0$ if there exists an x_i such that $g(x_i) = 0$. This is however no restriction, because it can always be achieved by choosing an appropriate measure μ. The problem now reduces to the proof that for all $x_1 < \ldots < x_n$ and $\theta_1 < \ldots < \theta_n$

$$\Delta_n^* = \begin{vmatrix} e^{\theta_1 x_1} & \ldots\ldots & e^{\theta_1 n_n} \\ \vdots & & \vdots \\ e^{\theta_n x_1} & \ldots\ldots & e^{\theta_n x_n} \end{vmatrix} > 0.$$

For $n = 1$, we have $\Delta_1^* = e^{\theta_1 x_1} > 0$. Now suppose that the assertion holds for $\Delta_1^*, \Delta_2^*, \ldots, \Delta_{n-1}^*$. Divide the i-th column of Δ_n^* by $\exp(\theta_1 x_i)$, $i = 1,\ldots,n$, to obtain

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \exp\{(\theta_2-\theta_1)x_1\} & \exp\{(\theta_2-\theta_1)x_2\} & \cdots & \exp\{(\theta_2-\theta_1)x_n\} \\ \vdots & \vdots & & \vdots \\ \exp\{(\theta_n-\theta_1)x_1\} & \exp\{(\theta_n-\theta_1)x_2\} & \cdots & \exp\{(\theta_n-\theta_1)x_n\} \end{vmatrix}.$$

Subtracting in the resulting determinant column (j-1) from column j, $j = n, n-1, \ldots, 2$, we get

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ e^{(\theta_2-\theta_1)x_1} & e^{(\theta_2-\theta_1)x_2}-e^{(\theta_2-\theta_1)x_1} & \cdots & e^{(\theta_2-\theta_1)x_n}-e^{(\theta_2-\theta_1)x_{n-1}} \\ \vdots & \vdots & & \vdots \\ e^{(\theta_n-\theta_1)x_1} & e^{(\theta_n-\theta_1)x_2}-e^{(\theta_n-\theta_1)x_1} & \cdots & e^{(\theta_n-\theta_1)x_n}-e^{(\theta_n-\theta_1)x_{n-1}} \end{vmatrix}.$$

Hence, by elementary properties of determinants, Δ_n^* has the same sign as

$$\Delta_{n-1}^{(1)} = \begin{vmatrix} e^{\eta_2x_2}-e^{\eta_2x_1} & e^{\eta_2x_3}-e^{\eta_2x_2} & \cdots & e^{\eta_2x_n}-e^{\eta_2x_{n-1}} \\ \vdots & \vdots & & \vdots \\ e^{\eta_nx_2}-e^{\eta_nx_1} & e^{\eta_nx_3}-x^{\eta_nx_2} & \cdots & e^{\eta_nx_n}-e^{\eta_nx_{n-1}} \end{vmatrix}$$

with $\eta_i = \theta_i - \theta_1$. Now

$$\Delta_{n-1}^{(1)} = \begin{vmatrix} e^{\eta_2x_2} & e^{\eta_2x_3}-e^{\eta_2x_2} & \cdots & e^{\eta_2x_n}-e^{\eta_2x_{n-1}} \\ \vdots & \vdots & & \vdots \\ e^{\eta_nx_2} & e^{\eta_nx_3}-e^{\eta_nx_2} & \cdots & e^{\eta_nx_n}-e^{\eta_nx_{n-1}} \end{vmatrix}$$

$$- \begin{vmatrix} e^{\eta_2x_1} & e^{\eta_2x_3}-e^{\eta_2x_2} & \cdots & e^{\eta_2x_n}-e^{\eta_2x_{n-1}} \\ \vdots & \vdots & & \vdots \\ e^{\eta_nx_1} & e^{\eta_nx_3}-e^{\eta_nx_2} & \cdots & e^{\eta_nx_n}-e^{\eta_nx_{n-1}} \end{vmatrix}.$$

Expanding the two determinants by the first column, and using the notation

$$h(x) = \sum_{k=2}^{n} a_k e^{\eta_k x},$$

where a_k denotes the cofactor corresponding to $e^{\eta_k x}$, we get

$$\Delta_{n-1}^{(1)} = h(x_2) - h(x_1) = (x_2 - x_1)h'(y_2),$$

where $y_2 \in (x_1,x_2)$; the second equality follows from the mean value theorem of Lagrange, since $h(x)$ is differentiable. Notice that

$$h'(x) = \sum_{k=2}^{n} a_k \eta_k e^{\eta_k x},$$

implying that

$$\Delta_{n-1}^{(1)} = (x_2 - x_1) \begin{vmatrix} \eta_2 e^{\eta_2 y_2} & e^{\eta_2 x_3} - e^{\eta_2 x_2} & \dots & e^{\eta_2 x_n} - e^{\eta_2 x_{n-1}} \\ \vdots & \vdots & & \vdots \\ \eta_n e^{\eta_n y_2} & e^{\eta_n x_3} - e^{\eta_n x_2} & \dots & e^{\eta_n x_n} - e^{\eta_n x_{n-1}} \end{vmatrix}$$

$$= (x_2 - x_1)\Delta_{n-1}^{(2)} .$$

Now we expand $\Delta_{n-1}^{(2)}$ by the second column and then proceed in the same manner as for $\Delta_{n-1}^{(1)}$ to obtain

$$\Delta_{n-1}^{(2)} = (x_3 - x_2) \begin{vmatrix} \eta_2 e^{\eta_2 y_2} & \eta_2 e^{\eta_2 y_3} & e^{\eta_2 x_4} - e^{\eta_2 x_3} & \dots & e^{\eta_2 x_n} - e^{\eta_2 x_{n-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ \eta_n n^{\eta_n y_2} & \eta_n e^{\eta_n y_3} & e^{\eta_n x_4} - e^{\eta_n x_3} & \dots & e^{\eta_n x_n} - e^{\eta_n x_{n-1}} \end{vmatrix}$$

$$= (x_3 - x_2)\Delta_{n-1}^{(3)} ,$$

where $y_3 \in (x_2,x_3)$. Then we proceed in an analogous way with the third column and we continue this procedure up to the n-th column. So we finally get

$$\Delta_{n-1}^{(1)} = \prod_{k=2}^{n} (x_k - x_{k-1}) \begin{vmatrix} \eta_2 e^{\eta_2 y_2} & \eta_2 e^{\eta_2 y_3} & \dots & \eta_2 e^{\eta_2 y_n} \\ \vdots & \vdots & & \vdots \\ \eta_n e^{\eta_n y_2} & \eta_n e^{\eta_n y_3} & \dots & \eta_n e^{\eta_n y_n} \end{vmatrix},$$

where $y_2 < y_3 < \dots < y_n$, or, since $\eta_i = \theta_i - \theta_1$,

$$\Delta_{n-1}^{(1)} = \prod_{k=2}^{n} (x_k - x_{k-1}) \prod_{k=2}^{n} (\theta_k - \theta_1) \begin{vmatrix} e^{\eta_2 y_2} & \dots & e^{\eta_2 y_n} \\ \vdots & & \vdots \\ e^{\eta_n y_2} & \dots & e^{\eta_n y_n} \end{vmatrix}.$$

Since the obtained determinant is $(n-1) \times (n-1)$, the induction hypothesis implies that $\Delta_{n-1}^{(1)}$, and therefore Δ_n^* is strictly positive.

Remark. The y_k's obtained by application of the mean value theorem are inner points of the interval $[x_{k-1},x_k]$, and not $y_k \in [x_{k-1},x_k]$ as is stated in the hint. This is essential, because otherwise it would be possible that $y_k = y_{k+1} = x_k$ for some k, which would imply that

$$|e^{\eta_i y_j}| \quad i,j = 2,3,\dots,n$$

has two identical columns, that is $|e^{\eta_i y_j}| = 0$.

(KARLIN (1955, 1957))

Problem 24.

First we prove that b) implies a). The determinant Δ_3 is positive for all $\theta_1 < \theta_2 < \theta_3$, $x_1 < x_2 < x_3$. So we have for $k_1,k_2,k_3 > 0$ that

$$\begin{vmatrix} g(x_1) & g(x_2) & g(x_3) \\ p_{\theta_2}(x_1) & p_{\theta_2}(x_2) & p_{\theta_2}(x_3) \\ p_{\theta_3}(x_1) & p_{\theta_3}(x_2) & p_{\theta_3}(x_3) \end{vmatrix} = k_1\Delta_3 > 0,$$

where $g(x) = k_1p_{\theta_1}(x) - k_2p_{\theta_2}(x) + k_3p_{\theta_3}(x)$. The equation $g(x) = 0$ has therefore at most two solutions. If $g(x_1) = g(x_3) = 0$, we have

$$0 < k_1\Delta_3 = -g(x_2)\{p_{\theta_2}(x_1)p_{\theta_3}(x_3) - p_{\theta_2}(x_3)p_{\theta_3}(x_1)\}.$$

Monotonicity of the likelihood ratios implies that $g(x_2) < 0$. If $g(x_2) = g(x_3) = 0$, we have

$$0 < k_1\Delta_3 = g(x_1)\{p_{\theta_2}(x_2)p_{\theta_3}(x_3) - p_{\theta_2}(x_3)p_{\theta_3}(x_2)\}.$$

Again by the monotonicity of the likelihood ratios it follows that $g(x_1) > 0$. Finally if $g(x_1) = g(x_2) = 0$, the same argument yields that $g(x_3) > 0$.

To prove that a) implies b), let $\theta_1 < \theta_2 < \theta_3$ and $x_1 < x_2 < x_3$, and write the following set of equations in λ_2 and λ_3:

$$(13) \qquad \begin{cases} p_{\theta_1}(x_1) = \lambda_2p_{\theta_2}(x_1) + \lambda_3p_{\theta_3}(x_1) \\ p_{\theta_1}(x_3) = \lambda_2p_{\theta_2}(x_3) + \lambda_3p_{\theta_3}(x_3) \end{cases} .$$

Monotonicity of the likelihood ratios implies that the determinant D defined by

$$D = \begin{vmatrix} p_{\theta_2}(x_1) & p_{\theta_3}(x_1) \\ p_{\theta_2}(x_3) & p_{\theta_3}(x_3) \end{vmatrix}$$

is positive. Therefore the set of equations (13) has a solution, $\lambda_2 = k_2$, $\lambda_3 = -k_3$ say, and

$$k_2 = \{p_{\theta_1}(x_1)p_{\theta_3}(x_3) - p_{\theta_1}(x_3)p_{\theta_3}(x_1)\}/D$$

$$-k_3 = -\{p_{\theta_1}(x_1)p_{\theta_2}(x_3) - p_{\theta_1}(x_3)p_{\theta_2}(x_1)\}/D.$$

The monotonicity of the likelihood ratios implies that k_2 and k_3 are positive. Define

$$g(x) = p_{\theta_1}(x) - k_2 p_{\theta_2}(x) + k_3 p_{\theta_3}(x),$$

then we have that $g(x_1) = g(x_3) = 0$. Furthermore $1 = k_1$, k_2 and k_3 are positive. Hence we can apply a) to obtain that $g(x_2) < 0$. Since

$$\Delta_3 = \begin{vmatrix} p_{\theta_1}(x_1) & p_{\theta_1}(x_2) & p_{\theta_1}(x_3) \\ p_{\theta_2}(x_1) & p_{\theta_2}(x_2) & p_{\theta_2}(x_3) \\ p_{\theta_3}(x_1) & p_{\theta_3}(x_2) & p_{\theta_3}(x_3) \end{vmatrix} = \begin{vmatrix} g(x_1) & g(x_2) & g(x_3) \\ p_{\theta_2}(x_1) & p_{\theta_2}(x_2) & p_{\theta_1}(x_3) \\ p_{\theta_3}(x_1) & p_{\theta_3}(x_2) & p_{\theta_3}(x_3) \end{vmatrix}$$

$$= \begin{vmatrix} 0 & g(x_2) & 0 \\ p_{\theta_2}(x_1) & p_{\theta_2}(x_2) & p_{\theta_2}(x_3) \\ p_{\theta_3}(x_1) & p_{\theta_3}(x_3) & p_{\theta_3}(x_3) \end{vmatrix} = -g(x_2)\cdot D,$$

it follows that $\Delta_3 > 0$.

(KARLIN (1955, 1957))

<u>Problem 25</u>.

We prove the results of Theorem 6, Section 3.7 for the family of densities $\{p_\theta(x)\}$ satisfying (a) and (b) of the problem.

Remark: In formula (24) of Theorem 6 we must replace "$C_1 < C_2$" by "$C_1 \leq C_2$".

First note (cf. (46)) that $\Delta_3 > 0$ implies that there do not exist $x_1 < x_2 < x_3$ and θ_1 with $p_{\theta_1}(x_1) = p_{\theta_1}(x_2) = p_{\theta_1}(x_3) = 0$. Consequently the inequality (c) $p_\theta(x) > 0$ is violated for at most two points in $\mathbb{R}$.

(i) Let $\theta_1 < \theta' < \theta_2$. As in the proof of Theorem 6, we use Theorem 5 (iv) to prove the existence of constants k_1 and k_2 and a test φ_0 with

$$E_{\theta_1}\varphi_0(X) = E_{\theta_2}\varphi_0(X) = \alpha$$

and

$$\varphi_0(x) = \begin{cases} 1 \\ 0 \end{cases} \text{when } k_1 p_{\theta_1}(x) + k_2 p_{\theta_2}(x) \begin{matrix} < \\ > \end{matrix} p_{\theta'}(x)$$

or, in view of (c),

$$\varphi_0(x) = \begin{cases} 1 \\ 0 \end{cases} \text{when } k_1(p_{\theta_1}(x)/p_{\theta'}(x)) + k_2(p_{\theta_2}(x)/p_{\theta'}(x)) \begin{matrix} < \\ > \end{matrix} 1$$

k_1 and k_2 can not be both ≤ 0, for then:

$$E_{\theta_1}\varphi(X) = E_{\theta_2}\varphi(X) = 1.$$

If one of the k's is non-positive and the other is positive, then, as in the book, φ_0 has a strictly monotone power function, which is also impossible in view of (25), p. 89, so $k_1 > 0$ and $k_2 > 0$. Let

$$g(x) = k_1 p_{\theta_1}(x) - p_{\theta'}(x) + k_2 p_{\theta_2}(x).$$

Then we have

$$\varphi_0(x) = \begin{cases} 1 \\ 0 \end{cases} \text{when } g(x) \begin{matrix} < \\ > \end{matrix} 0.$$

By the continuity of $p_\theta(x)$ in x (assumption (a)) and the fact that φ_0 does not p_θ-a.e. reject, there exists at least one point C with $g(C) = 0$. It follows from Problem 24 that there exist at most two points C_1 and C_2 with $g(C_1) = g(C_2) = 0$, and that φ_0 is of the form (24) if $C_1 \neq C_2$. If $C_1 = C_2 = 0$, then g(x) must be of equal sign for $x < C$ and $x > C$. Otherwise φ_0 would be a one-sided test, which has a strictly monotone power function. This is however impossible by (25). Observing the determinant $k_1\Delta_3$ in Problem 24 with $x_2 = C$, we see that for $x_1 < C < x_3$

$$0 < k_1\Delta_3 = g(x_1)\{p_{\theta'}(C)p_{\theta_2}(x_3) - p_{\theta'}(x_3)p_{\theta_2}(C)\} +$$
$$+ g(x_3)\{p_{\theta'}(x_1)p_{\theta_2}(C) - p_{\theta'}(C)p_{\theta_2}(x_1)\}.$$

We know already that $g(x_1)$ and $g(x_3)$ have the same sign, and combining this with assumption (b) we see that $g(x) > 0$ for $x \neq C$. This means that also in this case φ_0 is of the form (24) with $C_1 = C_2 = C$. By Theorem 5 (iii), φ_0 maximizes $E_\theta\varphi(X)$ subject to

$$E_{\theta_1}\varphi(X) \le \alpha \quad \text{and } E_{\theta_2}\varphi(X) \le \alpha$$

for $\theta_1 \le \theta \le \theta_2$. From (ii) it follows by comparison with the test $\varphi(x) \equiv \alpha$ that $E_\theta\varphi_0(X) \le \alpha$ for $\theta \le \theta_1$, $\theta \ge \theta_2$, so φ_0 is UMP.

(ii) 1. Suppose that the equation

$$(14) \qquad k_1 p_{\theta_1}(x) + k_2 p_{\theta_2}(x) = p_{\theta'}(x)$$

has two solutions C_1 and C_2. Then we know from part (i) that

$$(15) \qquad \varphi_0(x) = \begin{cases} 1 & \text{when } C_1 < x < C_2 \\ 0 & \text{when } x < C_1 \text{ or } x > C_2. \end{cases}$$

Let $\theta'' < \theta_1 < \theta_2$. The set of equations

$$k_1^* p_{\theta''}(C_i) + k_2^* p_{\theta_2}(C_i) = p_{\theta_1}(C_i), \qquad i = 1,2,$$

has exactly one solution

$$k_1^* = \frac{p_{\theta_2}(C_2)p_{\theta_1}(C_1) - p_{\theta_1}(C_2)p_{\theta_2}(C_1)}{p_{\theta_2}(C_2)p_{\theta''}(C_1) - p_{\theta''}(C_2)p_{\theta_2}(C_1)}$$

$$k_2^* = \frac{p_{\theta''}(C_2)p_{\theta_1}(C_1) - p_{\theta_1}(C_2)p_{\theta''}(C_1)}{p_{\theta''}(C_2)p_{\theta_2}(C_1) - p_{\theta_2}(C_2)p_{\theta''}(C_1)}$$

with k_1^* and k_2^* both positive by assumption (b). Then it follows from (15) and Problem 24 that

$$\varphi_0(x) = \begin{cases} 1 \\ 0 \end{cases} \text{when } k_1^* p_{\theta''}(x) + k_2^* p_{\theta_2}(x) \begin{matrix} < \\ > \end{matrix} p_{\theta_1}(x),$$

or, since $k_1^* > 0$,

$$\varphi_0(x) = \begin{cases} 1 \\ 0 \end{cases} \text{when } p_{\theta_1}(x)/k_1^* - k_2^* p_{\theta_2}(x)/k_1^* \begin{matrix} > \\ < \end{matrix} p_{\theta''}(x).$$

By Theorem 5 (ii) we see that φ_0 minimizes $E_{\theta''}\varphi(X)$ subject to (25). Similarly we can prove that φ_0 minimizes $E_\theta\varphi(X)$ subject to (25) for $\theta > \theta_2$. By comparison with the test $\varphi(x) \equiv \alpha$ we see that $E_\theta\varphi_0(X) \le \alpha$ for $\theta \le \theta_1$, $\theta \ge \theta_2$.

2. Now suppose that the equation (14) has only one solution $x = C$. From (25) it follows that $p_{\theta_1}(C) = p_{\theta_2}(C)$.
Let again $\theta'' < \theta_1 < \theta_2$. We have to prove the existence of constants $\hat{k}_1, \hat{k}_2$ such that

(16) $$p_{\theta''}(x) > \hat{k}_1 p_{\theta_1}(x) + \hat{k}_2 p_{\theta_2}(x) \qquad \text{when } x \neq C$$

(17) $$p_{\theta''}(x) = \hat{k}_1 p_{\theta_1}(x) + \hat{k}_2 p_{\theta_2}(x) \qquad \text{when } x = c.$$

Since $p_{\theta_1}(C) = p_{\theta_2}(C)$, it follows from (17) that

$$\hat{k}_2 = \{p_{\theta''}(C)/p_{\theta_2}(C)\} - \hat{k}_1.$$

Substituting this into (16) gives

$$p_{\theta''}(x) > \hat{k}_1\{p_{\theta_1}(x) - p_{\theta_2}(x)\} + \{p_{\theta''}(C)/p_{\theta_2}(C)\}\cdot p_{\theta_2}(x), \quad x \neq C.$$

Rewriting this inequality gives

$$\hat{k}_1 \;\substack{<\\>}\; \frac{p_{\theta''}(x)p_{\theta_2}(C) - p_{\theta''}(C)p_{\theta_2}(x)}{p_{\theta_1}(x)p_{\theta_2}(C) - p_{\theta_1}(C)p_{\theta_2}(x)} \quad \text{when } x \;\substack{<\\>}\; C.$$

Note that the right hand side of this inequality is positive for all $x \neq C$. Let $x_1 < C < x_3$, then we must find $\hat{k}_1$ such that $0 < f(x_3) < \hat{k}_1 < f(x_1)$ for all $x_1 < C < x_3$, where

$$f(x) = \frac{p_{\theta''}(x)p_{\theta_2}(C) - p_{\theta''}(C)p_{\theta_2}(x)}{p_{\theta_1}(x)p_{\theta_2}(C) - p_{\theta_1}(C)p_{\theta_2}(x)}.$$

To this end we start with proving that $f(x_3) < f(x_1)$ for all $x_1 < C < x_3$. By assumption (b)

$$0 < \Delta_3 = \begin{vmatrix} p_{\theta''}(x_1) & p_{\theta''}(C) & p_{\theta''}(x_3) \\ p_{\theta_1}(x_1) & p_{\theta_1}(C) & p_{\theta_1}(x_3) \\ p_{\theta_2}(x_1) & p_{\theta_2}(C) & p_{\theta_2}(x_3) \end{vmatrix}$$

$$= \begin{vmatrix} p_{\theta''}(x_1) & p_{\theta''}(C) & p_{\theta''}(x_3) \\ p_{\theta_1}(x_1) - p_{\theta_2}(x_1) & 0 & p_{\theta_1}(x_3) - p_{\theta_2}(x_3) \\ p_{\theta_2}(x_1) & p_{\theta_2}(C) & p_{\theta_2}(x_3) \end{vmatrix}$$

$$= \{p_{\theta_2}(x_1) - p_{\theta_1}(x_1)\}\{p_{\theta''}(C)p_{\theta_2}(x_3) - p_{\theta''}(x_3)p_{\theta_2}(C)\} +$$

$$+ \{p_{\theta_2}(x_3) - p_{\theta_1}(x_3)\}\{p_{\theta''}(x_1)p_{\theta_2}(C) - p_{\theta''}(C)p_{\theta_2}(x_1)\},$$

which can be rewritten to $f(x_3) < f(x_1)$ as was to be proved.
So there exists a $\hat{k}_1 > 0$ with $f(x_3) \le \hat{k}_1 \le f(x_1)$ for all $x_1 < C$ and $x_3 > C$, that is there exist $\hat{k}_1$ and $\hat{k}_2$ such that

(18) $$p_{\theta''}(x) \ge \hat{k}_1 p_{\theta_1}(x) + \hat{k}_2 p_{\theta_2}(x) \qquad \text{when } x \ne C$$

(19) $$p_{\theta''}(x) = \hat{k}_1 p_{\theta_1}(x) + \hat{k}_2 p_{\theta_2}(x) \qquad \text{when } x = C.$$

Furthermore, $\hat{k}_2 = \{p_{\theta''}(C)/p_{\theta_2}(C)\} - \hat{k}_1 \le \{p_{\theta''}(C)/p_{\theta_2}(C)\} - f(x_3) < 0$ by assumption (b), and so Problem 24 implies that the strict inequality must hold in (18) by which (16) and (17) are proved.
Application of Theorem 5 (ii) gives that φ_0 minimizes $E_{\theta''}\varphi(X)$ subject to (25). Similarly it follows that φ_0 minimizes $E_\theta\varphi(X)$ subject to (25) for $\theta > \theta_2$. By comparison with the test $\varphi(x) \equiv \alpha$ we see that $E_\theta\varphi_0(X) \le \alpha$ for $\theta \le \theta_1$ and $\theta \ge \theta_2$.

(iii) In the definition given in problem 23 a family of distributions with densities $p_\theta(x)$ is said to be of Pólya type if, among other things, $p_\theta(x)$ is continuous in θ. To avoid serious difficulties we suppose also in this problem that the probability densities $p_\theta(x)$ are continuous in θ. Then we can follow the lines of the proof in the book (p. 90), yielding that if $\beta(\theta)$ does not satisfy (iii), there exist $\theta' < \theta'' < \theta'''$ and x_1, x_2, x_3 such that

$$p_{\theta''}(x_i) = k_1 p_{\theta'}(x_i) + k_2 p_{\theta'''}(x_i), \qquad i = 1,2,3.$$

This is however impossible in view of Problem 24.

(KARLIN (1955, 1957))

<u>Problem 26.</u>

We prove the following (stronger) result:

Let θ be a real parameter and let the random variable X have probability density $p_\theta(x)$ (with respect to some measure μ) with strictly monotone likelihood ratio in $T(x)$ on $S = \{x : p_\theta(x) > 0\}$, where S is independent of θ. Suppose θ_1 and θ_2 are such that $\theta_1 \le \theta_2$ and that there exist

$\theta_1^* < \theta_1$ and $\theta_2^* > \theta_2$. It will be shown that under these conditions a UMP test of $H' : \theta_1 \le \theta \le \theta_2$ against $K' : \theta < \theta_1$ or $\theta > \theta_2$ does not exist.

Let $0 < \alpha < 1$, $\theta_1^* < \theta_1$ and $\theta_2^* > \theta_2$. By the proof of Theorem 2 there exist MP size-α-tests φ_1 and φ_2 for testing $H_1^* : \theta = \theta_1$ against $K_1^* : \theta = \theta_1^*$ and $H_2^* : \theta = \theta_2$ against $K_2^* : \theta = \theta_2^*$ respectively, given by

$$\varphi_1(x) = \begin{cases} 1 & < \\ \gamma_1 \quad \text{when } T(x) & = C_1 \\ 0 & > \end{cases}$$

and

$$\varphi_2(x) = \begin{cases} 1 & > \\ \gamma_2 \quad \text{when } T(x) & = C_2. \\ 0 & < \end{cases}$$

Comparison with the test $\varphi(x) \equiv \alpha$ yields in view of Theorem 2 (iv) that φ_1 and φ_2 are also size-α-tests for testing H' against K'. Suppose that φ_0 is a size-α UMP test for testing H' against K'. Then it follows that

$$E_{\theta_i^*}\varphi_0(X) \ge E_{\theta_i^*}\varphi_i(X), \quad i = 1,2,$$

and hence φ_0 is a level-α MP test both for testing H_1^* against K_1^* and for H_2^* against K_2^*. According to Theorem 1 (iii) there exist constants k_1 and k_2 such that except for a p_θ-null set N

$$(20) \qquad \varphi_0(x) = \begin{cases} 1 & > \\ \quad \text{when } p_{\theta_1^*}(x) & k_1 p_{\theta_1}(x) \\ 0 & < \end{cases}$$

and

$$(21) \qquad \varphi_0(x) = \begin{cases} 1 & > \\ \quad \text{when } p_{\theta_2^*}(x) & k_2 p_{\theta_2}(x). \\ 0 & < \end{cases}$$

Let $A = S \cap \tilde{N}$, then $P_\theta(A) = 1$ for all θ. Suppose

$$p_{\theta_1^*}(x) > k_1 p_{\theta_1}(x)$$

for some $x \in A$, then by the fact that $p_\theta(x)$ has monotone likelihood ratio in $T(x)$, we have for all $y \in A$ with $T(y) \le T(x)$ that $p_{\theta_1^*}(y) > k_1 p_{\theta_1}(y)$, and hence by (20) that $\varphi_0(y) = 1$. Then (21) implies that $p_{\theta_2^*}(x) \ge k_2 p_{\theta_2}(x)$, and therefore for all $y \in A$ with $T(y) > T(x)$, we have that

$p_{\theta_2^*}(y) > k_2 p_{\theta_2}(y)$, implying that $\varphi_0(y) = 1$. So in this case $\varphi_0(y) = 1$ for all $y \in A$, in contradiction with $\alpha < 1$.
Similarly $p_{\theta_1^*}(x) < k_1 p_{\theta_1}(x)$ for some $x \in A$ leads to a contradiction. Therefore $p_{\theta_1^*}(x) = k_1 p_{\theta_1}(x)$ a.e., in contradiction with $\theta_1^* < \theta_1$. The conclusion is that such a UMP test φ_0 does not exist.

Section 8

Problem 27.

$$\varphi(x_1,\ldots,x_s) = \begin{cases} 1 & > \\ \gamma \quad \text{when } \Sigma_{i=1}^r x_i = C, \\ 0 & < \end{cases}$$

with $\gamma \in [0,1]$ and C chosen such that they satisfy $E\varphi(Z_1,\ldots,Z_s) = \alpha$, where Z_i, $i = 1,\ldots,s$ are independently distributed with Poisson distribution $P(\lambda_i^*)$ and $\Sigma_{i=1}^s \lambda_i^* = a$.

The joint density of $(X_1,\ldots,X_s)$ with respect to counting measure on $\mathbb{N}$ is

$$p_{(\lambda_1,\ldots,\lambda_s)}(x) = \prod_{i=1}^s e^{-\lambda_i}\lambda_i^{x_i}/x_i! =$$
$$= \left[\exp\left\{-\sum_{i=1}^s \lambda_i\right\}\right] \cdot \prod_{i=1}^s \lambda_i^{x_i} \cdot \left\{\prod_{i=1}^s x_i!\right\}^{-1}.$$

Let $(\mu_1,\ldots,\mu_s)$ be any alternative, i.e. $\Sigma_{i=1}^s \mu_i > a$, then $(\mu_1^*,\ldots,\mu_s^*) \in H$, where

$$\mu_i^* = a\mu_i \sum_{i=1}^s \mu_i, \qquad i = 1,2,\ldots,s.$$

Since

$$\sum_{i=1}^s x_i \overset{>}{\underset{<}{=}} C \text{ iff } p_{(\mu_1,\ldots,\mu_s)}(x) \overset{>}{\underset{<}{=}} kp_{(\mu_1^*,\ldots,\mu_s^*)}(x),$$

where k is some constant depending on $(\mu_1,\ldots,\mu_s)$, a and C, Theorem 1 (ii) implies that φ is MP for testing $(\mu_1^*,\ldots,\mu_s^*)$ against $(\mu_1,\ldots,\mu_s)$.
Since the distribution of $\Sigma_{i=1}^s X_i$ depends only on $\Sigma_{i=1}^s \lambda_i$, and $\Sigma_{i=1}^s \lambda_i \le \Sigma_{i=1}^s \lambda_i'$ implies that

$$E_{(\lambda_1,\ldots,\lambda_s)}\varphi(X_1,\ldots,X_s) \le E_{(\lambda_1',\ldots,\lambda_s')}\varphi(X_1,\ldots,X_s),$$

φ is MP for testing H against $(\mu_1,\ldots,\mu_s)$ at level α. Since φ does not depend on the particular alternative chosen, φ is UMP.

Problem 28.

(i) In order to determine a uniformly most accurate lower confidence bound, we have to find the acceptance region $A(\xi_0)$ of a UMP test for $H(\xi_0) : \xi = \xi_0$ against $K(\xi_0) : \xi > \xi_0$. Defining $p(\xi)$ by $p(\xi) = P_\xi\{X_1 \le \xi_0\}$, it follows that, just as in Example 8, the joint density of $X_1,\dots,X_n$ at a sample point $x_1,\dots,x_n$ satisfying

$$x_{i_1},\dots,x_{i_m} \le \xi_0 < x_{j_1},\dots,x_{j_{n-m}}$$

is given by

$$p(\xi)^m(1-p(\xi))^{n-m}p_-(x_{i_1})\dots p_-(x_{i_m})p_+(x_{j_1})\dots p_+(x_{j_{n-m}}),$$

where we use the notation of Example 8. This means that the MP test for testing $H_0 : \xi = \xi_0$ against $K : \xi = \xi_1 > \xi_0$ is given by

$$\varphi(x_1,\dots,x_n) = \begin{cases} 1 & \\ 1-\rho & \text{when } \dfrac{p(\xi)^m(1-p(\xi))^{n-m}}{(\frac{1}{2})^n} \; \begin{matrix} > \\ = \\ < \end{matrix} \; C, \\ 0 & \end{cases}$$

or equivalently

$$\varphi(x_1,\dots,x_n) = \begin{cases} 1 & \\ 1-\rho & \text{when } m \; \begin{matrix} < \\ = \\ > \end{matrix} \; k, \\ 0 & \end{cases}$$

where k and ρ satisfy

$$P\{M < k\} + (1-\rho)P\{M = k\} = \alpha \quad \text{for } \xi = \xi_0,$$

and M is the number of X's $\le \xi_0$. Now for $\xi = \xi_0$ it holds that

$$P\{M = m\} = \binom{n}{m}(\tfrac{1}{2})^n,$$

implying that k and ρ satisfy

$$\sum_{j=0}^{k-1}\binom{n}{j}(\tfrac{1}{2})^n + (1-\rho)\binom{n}{k}(\tfrac{1}{2})^n = \alpha,$$

that is

$$\rho\sum_{j=k}^{n}\binom{n}{j}(\tfrac{1}{2})^n + (1-\rho)\sum_{j=k+1}^{n}\binom{n}{j}(\tfrac{1}{2})^n = 1-\alpha.$$

Since the test is independent of the particular alternative chosen, it is UMP. This test can also be written as

$$\varphi(x_1,\ldots,x_n) = \begin{cases} 1 & x^{(k)} > \xi_0 \\ 1-\rho \quad \text{when} & x^{(k)} \le \xi_0 < x^{(k+1)} \\ 0 & x^{(k+1)} \le \xi_0 . \end{cases}$$

Hence Theorem 4 (ii) implies that a uniformly most accurate lower confidence bound is given by

$$\underline{\xi} = \begin{cases} X^{(k)} \\ X^{(k+1)} \end{cases} \text{with probability} \quad \begin{matrix} \rho \\ 1-\rho . \end{matrix}$$

(ii) If ξ is a median of F, then

$$P_\xi\{\underline{\xi}(X_1,\ldots,X_n) \le \xi\} = \rho P_\xi\{X^{(k)} \le \xi\} + (1-\rho)P_\xi\{X^{(k+1)} \le \xi\}$$

$$= \rho \sum_{j=k}^{n} \binom{n}{j}(\tfrac{1}{2})^n + (1-\rho) \sum_{j=k+1}^{n} \binom{n}{j}(\tfrac{1}{2})^n = 1-\alpha .$$

(iii) The only difference with part (i) is that instead of $p(\xi_0) = \frac{1}{2}$, it now holds that $p(\xi_0) = p$, and so for $\xi = \xi_0$

$$P\{M=m\} = \binom{n}{m} p^m (1-p)^{n-m} .$$

Hence we obtain that a uniformly most accurate lower confidence bound is given by

$$\underline{\xi} = \begin{cases} X^{(k)} \\ X^{(k+1)} \end{cases} \text{with probability} \quad \begin{matrix} \rho \\ 1-\rho , \end{matrix}$$

where k and ρ satisfy

$$\rho \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} + (1-\rho) \sum_{j=k+1}^{n} \binom{n}{j} p^j (1-p)^{n-j} = 1-\alpha .$$

(THOMPSON (1936))

<u>Problem 29</u>.

Let p_i, $i = 1,2,3,4$, be the probabilities with which H is rejected when X takes on the value i. To find the most powerful test of H: the distribution of X is P_0 or P_1, against the alternative that it is Q, we must solve the following problem:

maximize $\frac{4}{13}p_1 + \frac{3}{13}p_2 + \frac{2}{13}p_3 + \frac{4}{13}p_4$ subject to

$$\begin{cases} \frac{2}{13}p_1 + \frac{4}{13}p_2 + \frac{3}{13}p_3 + \frac{4}{13}p_4 \le \alpha \\ \frac{4}{13}p_1 + \frac{2}{13}p_2 + \frac{1}{13}p_3 + \frac{6}{13}p_4 \le \alpha \\ 0 \le p_i \le 1, \quad i = 1,2,3,4. \end{cases}$$

With the help of e.g. the simplex method, it may be verified that for $\alpha = 5/13$ the solution of this problem is $p_1 = p_3 = 1$, $p_2 = p_4 = 0$ and for $\alpha = 6/13$ it is $p_1 = p_2 = 1$, $p_3 = p_4 = 0$. So in both cases the most powerful test is non-randomized. We have thus

$$R_{5/13} = \{1,3\} \not\subset \{1,2\} = R_{6/13}.$$

(STEIN (1951))

Problem 30.

(i) Let X and Y be independently distributed as $b(n,p_1)$ and $b(n,p_2)$, respectively. Let $H : p_2 \le p_1$; $K : (p_1', p_2')$ with $p_1' < p_2'$ and $p_1' + p_2' = 1$; $\omega = \{(p_1^*, p_2^*) : p_2^* \le p_1^*\}$; $\omega' = \{(\frac{1}{2},\frac{1}{2})\}$; λ a probability distribution over ω such that $\lambda(\omega') = 1$, $H_\lambda : (p_1,p_2) = (\frac{1}{2},\frac{1}{2})$ and $\alpha \in (0,\frac{1}{2})$.
By the fundamental lemma of Neyman and Pearson the most powerful level α test for testing H_λ against K is

$$\varphi(x,y) = \begin{cases} 1 & \\ \gamma & \text{when } \binom{n}{y}\binom{n}{x}(p_1')^{n-y+x}(p_2')^{y+n-x} \begin{matrix} > \\ = \\ < \end{matrix} k\binom{n}{x}\binom{n}{y}2^{-2n}, \\ 0 & \end{cases}$$

or equivalently, since $p_2' > p_1'$

$$\varphi(x,y) = \begin{cases} 1 & \\ \gamma & \text{when } y - x \begin{matrix} > \\ = \\ < \end{matrix} C, \\ 0 & \end{cases}$$

where C and γ satisfy $P_{(\frac{1}{2},\frac{1}{2})}\{Y - X > C\} + \gamma P_{(\frac{1}{2},\frac{1}{2})}\{Y - X = C\} = \alpha$. For $C < 0$, it holds that $P_{(\frac{1}{2},\frac{1}{2})}\{Y - X > C\} > \frac{1}{2}$, so $\alpha < \frac{1}{2}$ implies $C \ge 0$. Furthermore, since $P_{(\frac{1}{2},\frac{1}{2})}\{Y - X > 0\} + \frac{1}{2}P_{(\frac{1}{2},\frac{1}{2})}\{Y - X = 0\} = \frac{1}{2}$, it follows that $\gamma \in [0,\frac{1}{2})$ if $C = 0$, and since $P_{(\frac{1}{2},\frac{1}{2})}\{Y - X > n\} = 0$, that $\gamma \in (0,1]$ if $C = n$.
By Theorem 7 it is sufficient to prove that φ is of size $\le \alpha$ with respect to $H : p_2 \le p_1$, i.e. to prove that

$$\sup_{(p_1^*,p_2^*)\in\omega} \beta(p_1^*,p_2^*) \le \alpha = \beta(\tfrac{1}{2},\tfrac{1}{2}),$$

where $\beta(p_1^*,p_2^*) = P_{(p_1^*,p_2^*)}\{Y - X > C\} + \gamma P_{(p_1^*,p_2^*)}\{Y - X = C\}$. For $p_2^* \le p_1^*$ it holds that

$$\beta(p_1^*,p_2^*) = E(P_{(p_1^*,p_2^*)}\{Y - X > C \mid X\} + \gamma P_{(p_1^*,p_2^*)}\{Y - X = C \mid X\})$$
$$\le E(P_{(p_1^*,p_1^*)}\{Y - X > C \mid X\} + \gamma P_{(p_1^*,p_1^*)}\{Y - X = C \mid X\}) = \beta(p_1^*,p_1^*),$$

where the inequality follows from Example 2 and Theorem 2 (ii).

Finally we shall prove that

$$(22) \qquad \beta(p_1^*,p_1^*) \le \beta(\tfrac{1}{2},\tfrac{1}{2}), \quad p_1^* \in [0,1].$$

To this end it suffices to prove that for all nonnegative integers c and all $p \in [0,1]$

$$(23) \qquad P_p\{|S_n| \ge c\} \le P_{\frac{1}{2}}\{|S_n| \ge c\},$$

where $S_n = \Sigma_1^n (Y_i - X_i)$ (see the hint). First we show that S_n has a unimodal distribution under $p = \frac{1}{2}$, i.e. for all nonnegative integers c

$$(24) \qquad P_{\frac{1}{2}}\{S_n = c\} \ge P_{\frac{1}{2}}\{S_n = c+1\}.$$

Indeed, since $P_{\frac{1}{2}}\{Y_n - X_n = 0\} = \frac{1}{2}$ and $P_{\frac{1}{2}}\{Y_n - X_n = 1\} = P_{\frac{1}{2}}\{Y_n - X_n = -1\} = \frac{1}{4}$ hold, we have

$$P_{\frac{1}{2}}\{S_n = c\} - P_{\frac{1}{2}}\{S_n = c+1\} = \tfrac{1}{4}[P_{\frac{1}{2}}\{S_{n-1} = c-1\} + P_{\frac{1}{2}}\{S_{n-1} = c\}$$
$$- P_{\frac{1}{2}}\{S_{n-1} = c+1\} - P_{\frac{1}{2}}\{S_{n-1} = c+2\}],$$

which is nonnegative by the induction hypothesis; note that for $c = 0$ the first and third summand in the last expression cancel because of symmetry. Since (24) clearly holds for $n = 1$ the unimodality has been proved.

For $c = 0$ inequality (23) is a trivial equality. For $c \ge 1$ we have

$$
\begin{aligned}
P_p\{|S_n| \geq c\} &= P_p\{|Y_n - X_n + S_{n-1}| \geq c\} \\
&= p(1-p)[P_p\{|S_{n-1}| \geq c-1\} + P_p\{|S_{n-1}| \geq c+1\}] \\
&\quad + \{p^2 + (1-p)^2\}P_p\{|S_{n-1}| \geq c\} \\
&\leq p(1-p)[P_{\frac{1}{2}}\{|S_{n-1}| \geq c-1\} + P_{\frac{1}{2}}\{|S_{n-1}| \geq c+1\}] \\
&\quad + \{p^2 + (1-p)^2\}P_{\frac{1}{2}}\{|S_{n-1}| \geq c\} \qquad (25) \\
&= P_{\frac{1}{2}}\{|S_{n-1}| \geq c\} + p(1-p)[P_{\frac{1}{2}}\{|S_{n-1}| = c-1\} - P_{\frac{1}{2}}\{|S_{n-1}| = c\}] \\
&\leq P_{\frac{1}{2}}\{|S_{n-1}| \geq c\} + \tfrac{1}{4}[P_{\frac{1}{2}}\{|S_{n-1}| = c-1\} - P_{\frac{1}{2}}\{|S_{n-1}| = c\}] \\
&= \tfrac{1}{2}P_{\frac{1}{2}}\{|S_{n-1}| \geq c\} + \tfrac{1}{4}[P_{\frac{1}{2}}\{|S_{n-1}| \geq c-1\} + P_{\frac{1}{2}}\{|S_{n-1}| \geq c+1\}] \\
&= P_{\frac{1}{2}}\{|S_n| \geq c\},
\end{aligned}
$$

where the first inequality follows by induction and the second one from the unimodality (24).

(ii) Since $\beta(p,p) < \alpha$ for $p \in (0,\frac{1}{2}) \cup (\frac{1}{2},1)$, it holds that $\beta(p_1,p_2) < \alpha$ for alternatives $p_1 < p_2$ sufficiently close to the line $p_1 = p_2$. Against these alternatives the level α test $\varphi(x,y) \equiv \alpha$ has power α. So the test described in (i) is not UMP against the alternatives $p_1 < p_0$.

Problem 31.

(i) Let Γ and Δ denote the sets of possible θ- and η-values, respectively. In view of theorem 1 and (b) there exists a level α test ψ_0 which is most powerful within the class of tests based on T, satisfying

$$
\psi_0(t) = \begin{cases} 1 & > \\ \gamma & \text{when } p_1(t) = Cp_0(t), \\ 0 & < \end{cases}
$$

for some C and γ, where

$$
p_i = \frac{dP^T_{\theta_i}}{d(P^T_{\theta_0} + P^T_{\theta_1})}, \qquad i = 0,1,
$$

and

(26) $$E_{\theta_0}\psi_0(T) = \alpha.$$

Condition (b) of the problem implies that for all $\eta \in \Delta$

$$E_{\theta_0,\eta}\psi_0[T(X)] = E_{\theta_0}\psi_0(T) = \alpha,$$

in view of (26). Let φ^* be a test which satisfies

$$E_{\theta_0,\eta}\varphi^*(X) \le \alpha$$

for all $\eta \in \Delta$. Consider for any fixed $\eta \in \Delta$, a version of $E_{\theta_0,\eta}[\varphi^*(X) \mid t]$ that does not depend on θ_0. It follows from Lemma 2.3. (iii) that $P_\theta^T(N_\eta) = 0$ for all $\theta \in \Gamma$, where

$$N_\eta = \{t;\ E_\eta[\varphi^*(X) \mid t] \notin [0,1]\}.$$

Define ψ_η by

$$\psi_\eta(t) = \begin{cases} E_\eta[\varphi^*(X) \mid t] & \text{when } t \notin N_\eta \\ 0 & \text{when } t \in N_\eta, \end{cases}$$

then ψ_η is a test with

$$E_{\theta_0}\psi_\eta(T) = E_{\theta_0}E_\eta[\varphi^*(X) \mid T] = E_{\theta_0,\eta}\varphi^*(X) \le \alpha.$$

Since the test ψ_0 is most powerful within the class of tests based on T, it follows that

$$E_{\theta_1,\eta}\varphi^*(X) = E_{\theta_1}\psi_\eta(T) \le E_{\theta_1}\psi_0(T) = E_{\theta_1,\eta}\psi_0[T(X)].$$

Hence $\psi_0[T(x)]$ is a UMP level α test which depends only on T.

(ii) If $\mathcal{B}$ denotes the class of Borel subsets of the real line and $(\mathcal{X},\mathcal{A})$ is a measurable space with random variable X, satisfying $\mathcal{X} \in \mathcal{B}$, $\mathcal{A} \subset \mathcal{B}$ and $(-\infty,u] \in \mathcal{A}$ with u as in Example 8, then define $\mathcal{P} = \{P : P$ a probability measure on $(\mathcal{X},\mathcal{A})\}$ and

$$\Delta = \{(P_-,P_+) : P_- \in \mathcal{P}, P_+ \in \mathcal{P}, P_-((-\infty,u]) = P_+((u,\infty)) = 1\}.$$

Because for any $p \in [0,1]$ and $(P_-.P_+) \in \Delta$ the probability measure $P \in \mathcal{P}$ defined by

(27) $$P\{X \in A\} = P(A) = P_-(A)\cdot p + P_+(A)\cdot(1-p)$$

for any $A \in \mathcal{A}$, satisfies

$$p = P\{X \le u\},\ P_-(A) = P\{X \in A \mid X \le u\} \quad (\text{if } p > 0) \text{ and}$$

$$P_+(A) = P\{X \in A \mid X > u\} \quad (\text{if } p < 1),$$

it follows that the mapping from $[0,1] \times \Delta$ to $\mathcal{P}$ as defined by (27), is a surjection. Hence

$$\mathcal{P} = \{P_\theta : \theta = (p,P_-,P_+) \in [0,1] \times \Delta\}$$

yields a parametrization of $\mathcal{P}$. Now consider the product space $(\mathcal{X}^n,\mathcal{A}^n)$ with random variable $(X_1,X_2,\ldots,X_n)$, where $\mathcal{X}^n = \mathcal{X} \times \mathcal{X} \times \ldots \times \mathcal{X}$ and $\mathcal{A}^n$ is the smallest σ-field containing the sets $A_1 \times A_2 \times \ldots \times A_n$ $(A_1,\ldots,A_n \in \mathcal{A})$, and the class of product probability measures

$$\mathcal{P}^n = \{P^n : P^n = P \times \ldots \times P,\ P \in \mathcal{P}\} = \{P_\theta^n : \theta \in [0,1] \times \Delta\}.$$

Define the statistic $M : \mathcal{X}^n \to \{0,1,\ldots,n\}$ by

$$M(x) = M(x_1,x_2,\ldots,x_n) = \sum_{i=1}^{n} I_{(-\infty,u]}(x_i).$$

Now, $M(X)$ has a binomial distribution with parameters n and p and the conditional distribution $P_\theta^n\{X \in A \mid M(X) = m\}$ $(A \in \mathcal{A}^n)$ depends only on $(P_-,P_+) \in \Delta$. We have thus checked that $M(X)$ is sufficient for $p \in [0,1]$ in the presence of a nuisance parameter $(P_-,P_+) \in \Delta$. Hence by (i), the test $\psi_0[M(x)]$ with

$$\psi_0(m) = \begin{cases} 1 & > \\ \gamma & \text{when } \binom{n}{m} p_1^m (1-p_1)^{n-m} = c^*\binom{n}{m} p_0^m (1-p_0)^{n-m}, \\ 0 & < \end{cases}$$

or equivalently

$$\psi_0(m) = \begin{cases} 1 & < \\ \gamma & \text{when } m = C \\ 0 & > \end{cases}$$

is UMP for testing $H^* : p = p_0$ against $K^* : p = p_1 < p_0$ at level of significance α. As in Example 8 it follows from monotonicity and independence of the particular alternative chosen that $\psi_0[M(x)]$ is a UMP level α test for testing $H : p \ge p_0$ against $K : p < p_0$.

(FRASER (1956))

Section 9

Problem 32.

Let $\xi_1 < \eta_1$ be a particular alternative. Consider the testing problem of the simple hypothesis $H^* : \xi = \eta = (m+n)^{-1}(m\xi_1 + n\eta_1)$ against the simple alternative $K^* : \xi = \xi_1$, $\eta = \eta_1$ with level of significance α. By the fundamental lemma of Neyman and Pearson (Theorem 1) the MP test rejects when

$$(2\pi)^{-\frac{1}{2}N} \exp\{-\tfrac{1}{2} \sum_{i=1}^{m} (X_i - \xi_1)^2 - \tfrac{1}{2} \sum_{j=1}^{n} (Y_j - \eta_1)^2\}$$

$$> c_\alpha (2\pi)^{-\frac{1}{2}N} \exp\{-\tfrac{1}{2} \sum_{i=1}^{m} (X_i - (m\xi_1 + n\eta_1)/N)^2 - \tfrac{1}{2} \sum_{j=1}^{n} (Y_j - (m\xi_1 + n\eta_1)/N)^2\}$$

where $N = m+n$, or equivalently when

$$\overline{Y} - \overline{X} > c'_\alpha,$$

where $\overline{X} = m^{-1} \Sigma_{i=1}^{m} X_i$, $\overline{Y} = n^{-1} \Sigma_{j=1}^{n} Y_j$ and c'_α satisfies $P_{H^*}\{\overline{Y} - \overline{X} > c'_\alpha\} = \alpha$. Hence $c'_\alpha = (mnN^{-1})^{\frac{1}{2}} \Phi^{-1}(1-\alpha)$, where Φ^{-1} denotes the inverse of the standard normal distribution function. Since $P_{(\xi,\eta)}\{\overline{Y} - \overline{X} > c'_\alpha\} \le \alpha$ for all (ξ,η) with $\xi \ge \eta$, the test is MP for testing $H : \eta \le \xi$ against K^*. The test does not depend on the particular alternative $\xi_1 < \eta_1$ chosen. Therefore it is UMP for testing H against K.

Remarks: 1. The Kullback-Leibler "distance" from a probability measure P to a probability measure Q is defined as

$$\begin{cases} E_P \log(dP/dQ) & \text{if P is absolutely continuous w.r.t. Q} \\ \infty & \text{otherwise.} \end{cases}$$

Define by $I(\theta_1,\theta_2)$ the Kullback-Leibler "distance" from a normal $N(\theta_1,1)$ distribution to a normal $N(\theta_2,1)$ distribution. Then $I(\theta_1,\theta_2) = \frac{1}{2}(\theta_1 - \theta_2)^2$. Furthermore the Kullback-Leibler "distance" from the probability measure induced by the vector $(X_1,\ldots,X_m,Y_1,\ldots,Y_n)$ under (ξ_1,η_1) to the probability measure induced by the same vector under (ξ,η) is given by $mI(\xi_1,\xi) + nI(\eta_1,\eta)$. This function attains its minimum at $\xi = \eta = (m+n)^{-1}(m\xi_1+n\eta_1)$ if (ξ,η) runs through the set $\{(\xi,\eta);\eta \le \xi\}$. So the point $((m+n)^{-1}(m\xi_1+n\eta_1),(m+n)^{-1}(m\xi_1+n\eta_1))$ is that point of the null hypothesis nearest to the alternative measured in Kullback-Leibler "distance".

2. The same test arises when we reject for large values of X given X + Y (cf. Section 4.5).

Problem 33.

(i) We can restrict attention to the sufficient statistics

$$U = \sum_{i=1}^{m} (X_i - \xi)^2 \quad \text{and} \quad V = \sum_{j=1}^{n} (Y_j - \eta)^2.$$

Their joint density equals

$$f(u,v) = C_n \cdot C_m \sigma^{-m} \tau^{-n} u^{(m/2)-1} v^{(n/2)-1} \exp\{-u(2\sigma^2)^{-1} - v(2\tau^2)^{-1}\}.$$

where C_n and C_m depend only on n and m, respectively. Consider the hypothesis

$$(28) \qquad H : \tau^2 \le \sigma^2 \quad \text{against} \quad K : \tau^2 > \sigma^2.$$

Let α be any significance level ϵ (0,1), and let (σ_1^2, τ_1^2) with $\sigma_1^2 < \tau_1^2$ be any particular alternative. The least favorable distribution λ on H should be concentrated on the line $\sigma^2 = \tau^2$. Choosing λ degenerated at the point (τ_2^2, τ_2^2), with τ_2^2 to be specified later, the testing Problem (28) reduces to

$$(29) \qquad \begin{cases} H_\lambda : f(u,v) = C\tau_2^{-(m+n)} u^{(m/2)-1} v^{(n/2)-1} \exp\{-(u+v)(2\tau_2^2)^{-1}\} \\ K_\theta : f(u,v) = C\sigma_1^{-m} \tau_1^{-n} u^{(m/2)-1} v^{(n/2)-1} \exp\{-u(2\sigma_1^2)^{-1} - v(2\tau_1^2)^{-1}\}. \end{cases}$$

The most powerful level α test for (29) is given by

$$\varphi(u,v) = \begin{cases} 1 \\ 0 \end{cases} \text{when } \exp\{-u(2\sigma_1^2)^{-1} - v(2\tau_1^2)^{-1} + (u+v)(2\tau_2^2)^{-1}\} \gtrless K,$$

which is equivalent to

$$\varphi(u,v) = \begin{cases} 1 \\ 0 \end{cases} \text{when } -(\sigma_1^{-2} - \tau_2^{-2})u + (\tau_2^{-2} - \tau_1^{-2})v \gtrless K^*.$$

The constant K^* satisfies $E\varphi(U,V) = \alpha$, where U/τ_2^2 and V/τ_2^2 have chi-square distributions with m and n degrees of freedom respectively. Choose $\tau_2^2 \in [\sigma_1^2, \tau_1^2]$ such that $K^* = 0$. This is always possible since

$$P\{-(\sigma_1^{-2} - \tau_2^{-2})U + (\tau_2^{-2} - \tau_1^{-2})V > 0\}$$

$$= \begin{cases} 0 & \text{if } \tau_2^2 \ge \tau_1^2 \\ 1 & \text{if } \tau_2^2 \le \sigma_1^2 \\ P\{F_{n,m} > mn^{-1}\tau_0^2\sigma_0^{-2}(\tau_1^2 - \sigma_0^2)/(\tau_0^2 - \tau_1^2)\} & \text{if } \sigma_1^2 < \tau_2^2 < \tau_1^2 \,, \end{cases}$$

where $F_{n,m}$ has an F-distribution with n and m degrees of freedom. Denoting the upper α-percentage point of this distribution by F_α, we find

$$\tau_2^2 = \frac{\sigma_1^2(1 + nF_\alpha m^{-1})}{1+\sigma_1^2 nF_\alpha\tau_1^{-2}m^{-1}} \,.$$

Hence the rejection region for testing H_λ against K_0 is $\{(u,v) : v/u \ge C\}$, where $mn^{-1}C = F_\alpha$. Since

$$P_{(\sigma^2,\tau^2)}\{V/U \ge C\} = P\{F_{n,m} \ge \sigma^2\tau^{-2}mn^{-1}C\},$$

$P_{(\sigma^2,\tau^2)}\{V/U \ge C\}$ attains its maximum α over H when $\sigma^2 = \tau^2$. It follows that the test defined above is also most powerful for testing H against K_0. Since this test does not depend on the particular alternative chosen, it is UMP.

(ii) We can restrict attention to the sufficient statistics

$$U = \sum_{i=1}^{m} (X_i - \bar{X})^2, \quad W = \bar{X}, \quad V = \sum_{j=1}^{n} (Y_j - \bar{Y})^2 \quad \text{and } Z = \bar{Y}.$$

Their joint density equals

$$f(u,v,w,z) = C_{n-1}C_{m-1}\frac{u^{(m-3)/2}}{\sigma^{m-1}}e^{-u/(2\sigma^2)}\frac{\exp\{-m(w-\xi)^2/(2\sigma^2)\}}{(2\pi\sigma^2m^{-1})^{\frac{1}{2}}} \cdot$$
$$\cdot \frac{v^{(n-3)/2}}{\tau^{n-1}}e^{-v/(2\tau^2)}\frac{\exp\{-n(z-\eta)^2/(2\tau^2)\}}{(2\pi\tau^2n^{-1})^{\frac{1}{2}}} \,.$$

Consider again the testing problem (28). Let α be any significance level and let $\sigma_1^2,\tau_1^2,\xi_1,\eta_1$ be any particular alternative ($\sigma_1^2 < \tau_1^2$). The distribution λ over H should be such that

$$\text{(30)} \qquad \int_H g(\sigma^2,\tau^2,\xi,\eta)d\lambda(\sigma^2,\tau^2,\xi,\eta)$$

comes as close as possible to $g(\sigma_1^2,\tau_1^2,\xi_1.\eta_1)$, with

$$g(\sigma^2,\tau^2,\xi,\eta) = \frac{\exp\{-u/(2\sigma^2) - v/(2\tau^2)\}}{\sigma^{m-1}\tau^{n-1}} \cdot \frac{\exp\{-m(w-\xi)^2/(2\sigma^2)\}}{(2\pi\sigma^2 m^{-1})^{\frac{1}{2}}} \cdot$$
$$\cdot \frac{\exp\{-n(z-\eta)^2/(2\tau^2)\}}{(2\pi\tau^2 n^{-1})^{\frac{1}{2}}} .$$

We take λ equal to the product of λ_1,λ_2 and λ_3, where λ_1 is a measure over $\{\sigma^2 \geq \tau^2\}$, λ_2 a measure over $\xi \in \mathbb{R}$ and λ_3 a measure over $\eta \in \mathbb{R}$. It is natural to concentrate λ_1 on $\{\sigma^2 = \tau^2\}$; take λ_1 degenerated in (τ_2^2,τ_2^2), where τ_2^2 will be specified later. For the same reasons as for the testing of H_1 or H_2 against K in Section 9 we take λ_2 degenerated at ξ_1, and λ_3 a normal $N(\eta_1,(\tau_1^2 - \tau_2^2)/n)$-distribution. (Notice that τ_2^2 must be $\leq \tau_1^2$). As in Section 9

$$\int \frac{\exp\{-n(z-\eta)^2/(2\tau_1^2)\}}{(2\pi\tau^2 n^{-1})^{\frac{1}{2}}}\, d\lambda_3(\eta)$$

is the density of the sum of two independent normal $N(0,\tau_2^2/n)$ and $N(\eta_1,(\tau_1^2 - \tau_2^2)/n)$ variables. Therefore, (30) becomes

$$\frac{\exp\{-(u+v)/(2\tau_2^2)\}}{\tau_2^{m+n-2}} \cdot \frac{\exp\{-m(w-\xi_1)^2/(2\tau_2^2)\}}{(2\pi\tau_2^2 m^{-1})^{\frac{1}{2}}} \cdot$$
$$\cdot \frac{\exp\{-n(z-\eta_1)^2/(2\tau_1^2)\}}{(2\pi\tau_1^2 n^{-1})^{\frac{1}{2}}} .$$

The testing problem (28) then reduces to

$$(31)\quad \begin{cases} H_\lambda : f(u,v,w,z) = \\ \dfrac{C_{n-1}C_{m-1}u^{(m-3)/2}v^{(n-3)/2} \cdot \exp\left\{-\dfrac{u+v}{2\tau_2^2} - \dfrac{m(w-\xi_1)^2}{2\tau_2^2} - \dfrac{n(z-\eta_1)^2}{2\tau_1^2}\right\}}{\tau_2^{m-1}\tau_2^{n-1}(2\pi\tau_2^2 n^{-1})^{\frac{1}{2}}(2\pi\tau_1^2 n^{-1})^{\frac{1}{2}}} \\ K_0 : f(u,v,w,z) = \\ \dfrac{C_{n-1}C_{m-1}u^{(m-3)/2}v^{(n-3)/2} \cdot \exp\left\{-\dfrac{u}{2\sigma_1^2} - \dfrac{v}{2\tau_1^2} - \dfrac{m(w-\xi_1)^2}{2\sigma_1^2} - \dfrac{n(z-\eta_1)^2}{2\tau_1^2}\right\}}{\sigma_1^{m-1}\tau_1^{n-1}(2\pi\sigma_1^2 m^{-1})^{\frac{1}{2}}(2\pi\tau_1^2 n^{-1})^{\frac{1}{2}}} . \end{cases}$$

The most powerful level α test for (31) is given by

$$\varphi(u,v,w,z) = \begin{cases} 1 & \\ & \text{when } -(\sigma_1^{-2}-\tau_2^{-2})u - m(\sigma_1^{-2}-\tau_2^{-2})(w-\xi_1)^2 + (\tau_2^{-2}-\tau_1^{-2})v \ \begin{matrix} > \\ \\ < \end{matrix}\ K^*. \\ 0 & \end{cases}$$

The constant K^* is determined such that

$$P\{-(\sigma_1^{-2} - \tau_2^{-2})[U + m(W - \xi_1)^2] + (\tau_2^{-2} - \tau_1^{-2})V > K^*\} = \alpha,$$

where U, W and V are independent and U/τ_2^2, $m(W - \xi_1)^2/\tau_2^2$ and V/τ_2^2 have a chi-square distribution with (m-1), 1 and (n-1) degrees of freedom, respectively. In the same way as in (i) we can take $\tau_2^2 \in (\sigma_1^2, \tau_1^2)$ such that $K^* = 0$. Hence the rejection region for testing H_λ against K_0 is

$$\{V/(U + m(W - \xi_1)^2) \geq C\},$$

where $\frac{m}{n-1}C$ is the upper α percentage point of the F(n-1,m) distribution. To prove that this test is also most powerful for testing H against K_0 we have to show that

$$P = P_{(\eta,\xi,\sigma^2,\tau^2)}\{V/(U + m(W - \xi_1)^2) \geq C\}$$

attains its maximum over H on $\sigma^2 = \tau^2 = \tau_2^2$, $\xi = \xi_1$. Now for all $\sigma^2 \geq \tau^2$ and for all ξ,

$$P = P_{(\xi,\sigma^2,\tau^2)}\left\{\frac{V/\tau^2}{U/\sigma^2 + m(W - \xi_1)^2/\sigma^2} \geq \frac{\sigma^2}{\tau^2}C\right\}$$

$$= P_{(\xi,\sigma^2,\tau^2)}\left\{\frac{V_0}{U_0 + (W_0 - \xi_1 m^{\frac{1}{2}}\sigma^{-1})^2} \geq \frac{\sigma^2}{\tau^2}C\right\}.$$

where V_0 and U_0 have a chi-square distribution with (n-1) and (m-1) degrees of freedom, and W_0 has a normal $N(\xi m^{\frac{1}{2}}\sigma^{-1}, 1)$-distribution. Since $\sigma^2 \geq \tau^2$,

$$P \leq P_{(\xi,\sigma^2)}\{V_0/(U_0 + (W_0 - \xi_1 m^{\frac{1}{2}}\sigma^{-1})^2) \geq C\}$$

$$= P_{(\xi,\sigma^2)}\{(W_0-\xi_1 m^{\frac{1}{2}}\sigma^{-1})^2 \leq V_0/C - U_0\}$$

$$= \int_0^\infty\int_0^\infty P_{(\xi,\sigma^2)}\{(W_0 - \xi_1 m^{\frac{1}{2}}\sigma^{-1})^2 \leq v_0/C - u_0\}dP_{V_0}(v_0)\,dP_{U_0}(u_0).$$

Since the integrand is maximal if $EW_0 = \xi_1 m^{\frac{1}{2}}\sigma^{-1}$, that is if $\xi = \xi_1$,

$$P \leq P_{(\xi_1,\sigma^2)}\{(W_0 - \xi_1 m^{\frac{1}{2}}\sigma^{-1})^2 \leq V_0/C - U_0\}$$

$$= P\{(W_0^*)^2 \leq V_0/C - U_0\},$$

where W_0^* has a standard normal distribution. So

$$P \le P\{V_0/((W_0^*)^2 + U_0) \ge C\}$$

$$= P_{(\eta,\xi_1,\tau_2^2,\tau_2^2)}\{V/(U + m(W - \xi_1)^2) \ge C\}.$$

It follows that the test defined above is most powerful for testing H against K_0. Since this test depends on the alternative, and since the test which is most powerful at a particular alternative is unique by Theorem 1 (iii) and Theorem 7 (ii) there exists no UMP test for testing H against K.

Section 10

Problem 34.

Following the hint this problem is easily solved.

(STEIN (1946))

Problem 35.

(i) Since $EN = \sum_{n=1}^{\infty} nP\{N = n\} = \sum_{n=1}^{\infty} P\{N \ge n\}$, the preceding problem implies that

$$EN \le C \sum_{n=1}^{\infty} \delta^n$$

for some $C > 0$ and $\delta \in (0,1)$, and therefore that $EN < \infty$.

(ii) Similarly we get

$$Ee^{tN} \le \sum_{n=1}^{\infty} e^{tn}P\{N \ge n\} \le C \sum_{n=1}^{\infty} (\delta e^t)^n < \infty$$

for $t \in (0, -\log\delta)$, so that $EN^k < \infty$ for all $k = 1,2,3,\ldots$.

(iii) Suppose $P_i\{Z = 0\} = 1$ for $i = 0$ or 1. Then

$$P_i\{\log (p_1(X)/p_0(X)) = 0\} = 1, \text{ that is } P_i\{p_1(X) = p_0(X)\} = 1,$$

so $p_0(x) = p_1(x)$ a.e.P_i, and hence $P_0 = P_1$.

(STEIN (1946))

Section 11

Problem 36.

(i) In the situation of Examples 9 and 10, (35) p. 99 becomes

$$z_i = \log \left[p_1^{x_i}(1-p_1)^{1-x_i} p_0^{-x_i}(1-p_0)^{x_i-1} \right] = (2x_i - 1)\log(q_0 p_0^{-1}).$$

Hence the test continues as long as

$$-\alpha = (\log A_0)/\log(q_0 p_0^{-1}) < 2\sum_{i=1}^{n} x_i - n < (\log A_1)/\log(q_0 p_0^{-1}) = b.$$

Since a and b are positive integers, we have

$$p_{1n}/p_{0n} = A_1 \text{ on } R_n \quad \text{and } p_{1n}/p_{0n} = A_0 \text{ on } S_n,$$

where R_n and S_n are defined on p. 98. Therefore, in (34) equalities hold, which entails that the approximations in (38) and (40), p. 103, are exact.

(ii) If $p \neq \frac{1}{2}$, then equation (41), p. 103, has a unique nonzero solution h, thus satisfying

$$p(q_0 p_0^{-1})^h + q(p_0 q_0^{-1})^h = 1,$$

which is equivalent to

$$q(p_0 q_0^{-1})^{2h} - (p_0 q_0^{-1})^h + p = 0,$$

from which one concludes that

$$(p_0 q_0^{-1})^h = \frac{1 \pm \sqrt{1-4pq}}{2q} = \begin{cases} 1 \\ p/q \end{cases}.$$

However, by assumption $p_0 < p_1 = q_0$, and $h \neq 0$, implying that $(p_0 q_0^{-1})^h = pq^{-1}$. Substituting this in (38), p. 103, one obtains

$$\beta(p) = \frac{1 - A_0^h}{A_1^h - A_o^h} = \frac{1 - (q_0 p_0^{-1})^{-ah}}{(q_0 p_0^{-1})^{bh} - (q_0 p_0^{-1})^{-ah}}$$

(32)

$$= \frac{1 - (pq^{-1})^a}{(p^{-1}q)^b - (pq^{-1})^a} = \frac{q^a p^b - p^{a+b}}{q^{a+b} - p^{a+b}} \quad \text{for } p \neq \tfrac{1}{2}.$$

$\beta(p)$ is non-decreasing by Lemma 4, and the right-hand side of (32) is continuous, whence

$$\beta(\tfrac{1}{2}) = \lim_{p\to\frac{1}{2}} \beta(p) = \lim_{p\to\frac{1}{2}} \frac{1-(pq^{-1})^a}{(pq^{-1})^{-b}-(pq^{-1})^a} = \frac{a}{-b-a} = \frac{a}{a+b} .$$

(iii) Let w_n be the capital of the gambler who starts with capital a, after playing n times; define $w_0 = a$. The games stop if either $w_n = 0$ or $w_n = a+b$. The relation to the sequential procedure is established by the mapping $w_n \to 2\,\Sigma_{i=1}^n x_i - n + a$.

Problem 37.

First observe that for any i the random variables $I_{\{N\geq i\}}$ and $Z_i^+ = \max(0,Z_i)$ are independent, since $I_{\{N\geq i\}}$ is a function of $(Z_1,Z_2,\ldots,Z_{i-1})$, while Z_i^+ depends only on Z_i. Hence

$$E(Z_1^+ + \ldots + Z_N^+) = E\left(\sum_{n=1}^{\infty} I_{\{N=n\}} \sum_{i=1}^{n} Z_i^+\right) = E\left(\sum_{n=1}^{\infty} \sum_{i=1}^{n} I_{\{N=n\}} Z_i^+\right)$$

$$= E\left(\sum_{i=1}^{\infty} \sum_{n=i}^{\infty} I_{\{N=n\}} Z_i^+\right) = E\left(\sum_{i=1}^{\infty} I_{\{N\geq i\}} Z_i^+\right) = \sum_{i=1}^{\infty} E(I_{\{N\geq i\}} Z_i^+)$$

$$= \sum_{i=1}^{\infty} (EI_{\{N\geq i\}} EZ_i^+) = \sum_{i=1}^{\infty} P(N\geq i) EZ_1^+ = (EN)EZ_1^+ < \infty.$$

In the same way it follows that

$$E(Z_1^- + \ldots + Z_N^-) = (EN)(EZ_1^-) < \infty.$$

Hence

$$E(Z_1 + \ldots + Z_N) = E(Z_1^+ + \ldots + Z_N^+) - E(Z_1^- + \ldots + Z_N^-)$$

$$= (EN)(EZ_1^+ - EZ_1^-) = (EN)(EZ_1),$$

which had to be proved.

(WOLFOWITZ (1947))

Problem 38.

(i) Since $\psi(h) = Ee^{hZ} < \infty$ for all $h \in (-\infty,\infty)$, it follows from part (ii) of Theorem 9, Chapter 2, that

$$\psi''(h) = \int \frac{\partial^2}{\partial h^2} e^{hz} dP_Z(z) =$$

$$= \int z^2 e^{hz} dP_Z(z) = EZ^2 e^{hZ},$$

which is positive, since $P\{Z = 0\} < 1$. For $h > 0$, it holds that

$$\psi(h) = \int e^{hz} dP_Z(z) \geq \int_{(\log(1+\delta),\infty)} e^{hz} dP_Z(z)$$

$$\geq (1+\delta)^h P\{Z > \log(1+\delta)\}.$$

Therefore, $\lim_{h\to\infty} \psi(h) = \infty$. Similarly, $\lim_{h\to-\infty} \psi(h) = \infty$. This means that ψ has a minimum at h_0 and that $\psi'(h_0) = EZe^{h_0 Z} = 0$; furthermore $h_0 \neq 0$ since $EZ \neq 0$. Because $\psi(0) = 1$, there exists a unique $h_1 \neq 0$ for which $\psi(h_1) = 1$.

(ii) Putting $Z = \log\{p_{\theta_1}(X)/p_{\theta_0}(X)\}$ we obtain from part (i) that the following conditions are sufficient for the existence of a non-zero solution of

$$E_\theta \left[\frac{p_{\theta_1}(X)}{p_{\theta_0}(X)}\right]^h = 1$$

a) $E_\theta \log\{p_{\theta_1}(X)/p_{\theta_0}(X)\}$ exists and does not vanish;

b) $E_\theta\{p_{\theta_1}(X)/p_{\theta_0}(X)\}^h$ exists for all $h \in (-\infty,\infty)$;

c) $P_\theta\{(p_{\theta_1}(X)/p_{\theta_0}(X)) < 1-\delta\} P_\theta\{(p_{\theta_1}(X)/p_{\theta_0}(X)) > 1+\delta\} > 0$ for some $\delta > 0$.

Problem 39.

(i) Let $\bar{X} = n^{-1} \Sigma_{i=1}^n X_i$. The most powerful test for H against K rejects when $\bar{x} > x_R(n)$, where $x_R(n)$ satisfies

$$P\{\bar{X} > x_R(n)\} = \alpha \quad \text{when } \theta = 0,$$

that is $x_R(n) = n^{-\frac{1}{2}} u_{1-\alpha}$, where $\Phi(u_{1-\alpha}) = 1-\alpha$, and Φ the distribution function of the standard normal distribution. The power against $\theta > 0$ as a function of n is then

$$f(n) = P_\theta\{\bar{X} > x_R(n)\} = P_\theta\{\bar{X} > n^{-\frac{1}{2}} u_{1-\alpha}\} = 1 - \Phi(u_{1-\alpha} - n^{\frac{1}{2}}\theta).$$

Extend f(n) to all non-negative real values by defining

$$f(x) = 1 - \Phi(u_{1-\alpha} - x^{\frac{1}{2}}\theta), \qquad x \geq 0,$$

then we have

$$f'(x) = \varphi(u_{1-\alpha} - x^{\frac{1}{2}}\theta) \cdot \tfrac{1}{2} x^{-\frac{1}{2}} \theta$$

and

$$f''(x) = \tfrac{1}{4}\theta x^{-\frac{3}{2}}\varphi(u_{1-\alpha} - \theta x^{\frac{1}{2}})\{-\theta^2 x + \theta u_\alpha x^{\frac{1}{2}} - 1\}.$$

The equation $f''(x) = 0$ has solutions

$$x_{1,2}^{\frac{1}{2}} = \{u_{1-\alpha} \pm (u_{1-\alpha}^2 - 4)\}/(2\theta).$$

This means that $f''(x) > 0$ for $x \in (x_1, x_2)$ if $u_{1-\alpha}^2 > 4$. Therefore, $f(n)$ is not necessarily concave. An example is given in part (ii), where indeed $u_{1-\alpha}^2 > 4$.

(ii) For $\alpha = .005$ and $\theta = \frac{1}{2}$, we have

$$f(9) = 1 - \Phi(u_{.995} - 3/2) = 1 - \Phi(1.075) = .1412;$$

$$f(2) = 1 - \Phi(u_{.995} - \tfrac{1}{2}\sqrt{2}) = 1 - \Phi(1.868) = .0308;$$

$$f(16) = 1 - \Phi(u_{.995} - 2) = 1 - \Phi(.575) = .2826.$$

Taking 9 observations we have power .1412; taking 2 or 16 observations with probability $\frac{1}{2}$ each, we have power

$$(.0308)/2 + (.2826)/2 = .1567 > .1412.$$

(iii) a) For $\alpha_1 = .001$ and $n_1 = 2$ we have

$$f(2) = 1 - \Phi(u_{.999} - \tfrac{1}{2}\sqrt{2}) = 1 - \Phi(2.393) = .0084;$$

for $\alpha_2 = .009$ and $n_2 = 16$

$$f(16) = 1 - \Phi(u_{.991} - 2) = 1 - \Phi(.366) = .3557.$$

This gives power $(.0084)/2 + (.3557)/2 = .1821 > .1567 > .1412$.

b) Taking $\alpha_1 = 0$ and $n_1 = 0$ yields $f(0) = 1 - \Phi(u_1 - 0) = 0$, and $\alpha_2 = .01$ and $n_2 = 18$ gives $f(18) = 1 - \Phi(u_{.99} - \frac{1}{2}\sqrt{18}) = 1 - \Phi(.209) = .4168$; this gives power $\frac{1}{2}\cdot 0 + (.4168)/2 = .2084 > .1821 > .1567 > .1412$.

REFERENCES

BIRNBAUM, Z.W. & D.G. CHAPMAN (1950). On optimum selections from multinormal populations. *Ann. Math. Statist.* 21, 443-447.

BLACKWELL, D. (1951). Comparison of experiments. *Proc. Second Berkeley Symp. Math. Stat. Probab.*, Berkeley, Univ. Calif. Press, 93-102.

BLACKWELL, D. (1953). Equivalent comparisons of experiments. *Ann Math. Statist.* 24, 265-272.

EPSTEIN, B. & M. SOBEL (1953). Life testing. *J. Am. Stat. Assoc.* 48, 486-502.

FRASER, D.A.S. (1956). Sufficient statistics with nuisance parameters. *Ann. Math. Statist.* 27, 838-842.

KARLIN, S. (1955). Decision theory for Pólya type distributions. Case of two actions, I. *Proc. Third Berkeley Symp. Math. Stat. Probab.*, Berkeley Univ. Calif. Press, 1, 115-129.

KARLIN, S. (1957). Pólya type distributions, II. *Ann. Math. Statist.* 28, 281-308.

NEYMAN, J. & E.S. PEARSON (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc., Ser. A.* 231, 289-337.

NEYMAN, J. & E.S. PEARSON (1936). Contributions to the theory of testing statistical hypotheses. I. Unbiased critical regions of type A and type A_1. *Stat. Res. Mem.* 1, 1-37.

STEIN, C.M. (1946). A note on cumulative sums. *Ann. Math. Statist.* 17, 489-499.

STEIN, C.M. (1951). A property of some tests of composite hypotheses. *Ann. Math. Statist.* 22, 475-476.

THOMPSON, W.R. (1936). On confidence ranges for the median and other expectation distributions for populations of unknown distribution form. *Ann. Math. Statist.* 7, 122-128.

WOLFOWITZ, J. (1947). The efficiency of sequential estimates and Wald's equation for sequential processes. *Ann. Math. Statist.* 18, 215-230.

CHAPTER 4

Section 1

Problem 1.

Let φ_0 be a level α test which is UMP unbiased and suppose there exists a level α test φ_1 which is more powerful against alternatives in $\Omega' \neq \emptyset \subset \Omega_K$, and at least as powerful against all alternatives in Ω_K, i.e.

$$\beta_{\varphi_1}(\theta) \geq \beta_{\varphi_0}(\theta) \qquad \text{for } \theta \in \Omega_K$$
$$\beta_{\varphi_1}(\theta) > \beta_{\varphi_0}(\theta) \qquad \text{for } \theta \in \Omega' \subset \Omega_K.$$

Then, because φ_0 is unbiased,

$$\beta_{\varphi_1}(\theta) \geq \beta_{\varphi_0}(\theta) \geq \alpha \text{ for } \theta \in \Omega_K$$

and, of course,

$$\beta_{\varphi_1}(\theta) \leq \alpha \qquad \text{for } \theta \in \Omega_H.$$

Hence φ_1 is also unbiased. But φ_1 is more powerful than φ_0 against Ω'. Since this is in contradiction with the fact that φ_0 is UMP unbiased, it follows that such a test φ_1 cannot exist.

Problem 2.

(i) The critical level (see p. 62) is defined as:

$$\begin{aligned}\hat{\alpha}(x) &= \inf\{\alpha : \alpha \text{ is a significance level at which the hypothesis would be rejected for the given observation } x\} \\ &= \inf\{\alpha : x \in S_\alpha\}.\end{aligned}$$

(ii) First of all we show that

(1) $\{x : \hat{\alpha}(x) \le x\} = \{x : x \in S_\alpha\}$ for all $\alpha \in (0,1)$.

If $x \in S_\alpha$ then obviously $\hat{\alpha}(x) \le \alpha$.
If $\hat{\alpha}(x) \le \alpha$ then by condition (b) $x \in S_{\alpha'}$ for all $\alpha' > \alpha$; hence $x \in \bigcap_{\alpha'>\alpha} S_{\alpha'} = S_\alpha$.
Now it follows immediately that $\hat{\alpha}(X)$ is uniformly distributed over (0,1) since by (1) and condition (a) we get:

$$P_{\theta_0}\{\hat{\alpha}(X) \le \alpha\} = P_{\theta_0}\{X \in S_\alpha\} = \alpha \quad \text{for all } \alpha \in (0,1).$$

(iii) Since the tests S_α are unbiased, (1) implies that under any alternative θ

$$P_\theta\{\hat{\alpha}(X) \le \alpha\} = P_\theta\{X \in S_\alpha\} \ge \alpha = P_{\theta_0}\{\hat{\alpha}(X) \le \alpha\}.$$

Section 2

Problem 3.

Let X have the binomial distribution b(p,n).
For testing $H : p = p_0$ against $K : p \ne p_0$ at significance level α consider a test of the form

$$\varphi(x) = \begin{cases} 1 & \text{when } x < C_1 \text{ or } > C_2 \\ \gamma_i & \text{when } x = C_i,\ i = 1,2 \\ 0 & \text{when } C_1 < x < C_2 \end{cases}$$

First consider the case $n = 10$, $p_0 = .2$ and $\alpha = .1$. Then φ is UMP unbiased for $C_1 = 0$, $C_2 = 4$, $\gamma_1 = .5590$ and $\gamma_2 = .0815$ and φ is an equal tails test for $C_1 = 0$, $C_2 = 4$, $\gamma_1 = .4657$ and $\gamma_2 = .195$. The power functions are plotted in Figure 1 (the dotted line is the power function of the equal tails test). Figure 2 is an enlargment of part of Figure 1, which shows that the equal tails test is (slightly) biased.
Secondly let $n = 10$, $p_0 = .4$ and $\alpha = .05$. Then φ is UMP unbiased for $C_1 = 1$, $C_2 = 7$, $\gamma_1 = .5034$ and $\gamma_2 = .2677$ and φ is an equal tails test for $C_1 = 1$, $C_2 = 7$, $\gamma_1 = .4702$ and $\gamma_2 = .2992$. The power functions of these tests are plotted in Figure 3 (again the dotted line is the power

Figure 1

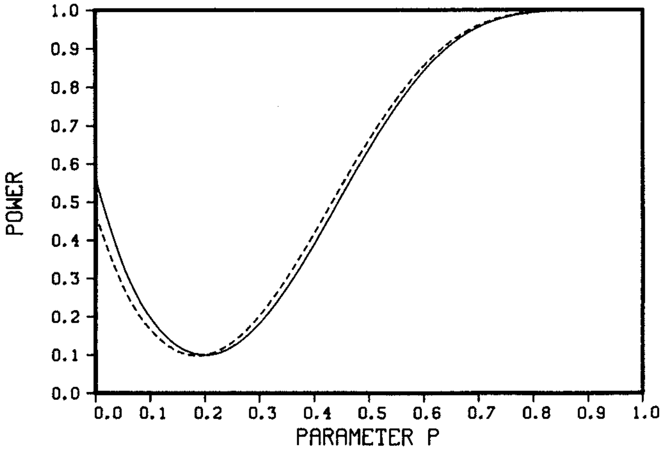


Figure 2

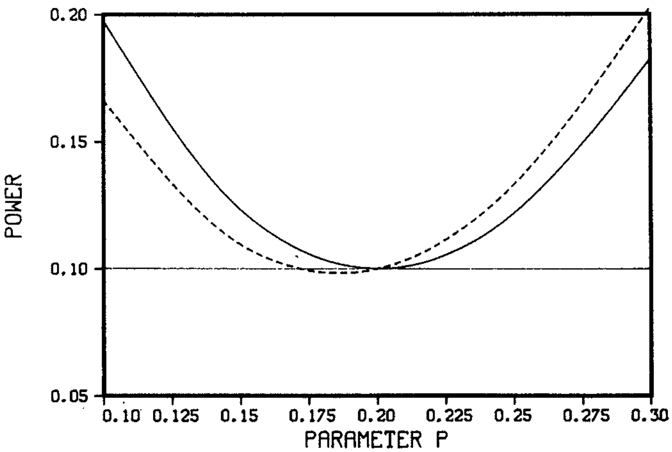
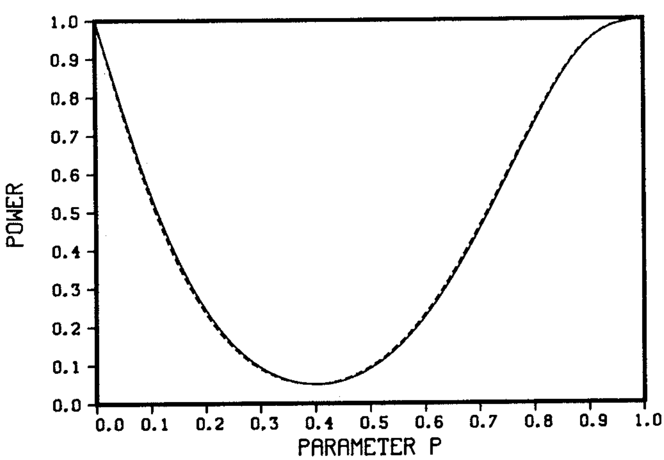


Figure 3



function of the equal tails test).

Problem 4.

Let X have the Poisson distribution $P(\tau)$, hence the density of X with respect to counting measure on $\mathbb{N}$ is $(e^{\tau}x!)^{-1}e^{x\log\tau}$. Since $T(X) = X$ and $E_{\tau_0}(X) = \tau_0$ condition (6) can be rewritten as follows:

$$E_{\tau_0}[T(X)\varphi(X)] = E_{\tau_0}[T(X)]\alpha$$

$$\Leftrightarrow \quad E_{\tau_0}[X\varphi(X)] = \tau_0\alpha$$

$$\Leftrightarrow \quad \tau_0 - E_{\tau_0}[X\varphi(X)] = \tau_0(1-\alpha)$$

$$\Leftrightarrow \quad E_{\tau_0}[X(1-\varphi(X))] = \tau_0(1-\alpha)$$

$$\Leftrightarrow \quad \sum_{x=C_1+1}^{C_2-1} xe^{-\tau_0}\frac{\tau_0^x}{x!} + \sum_{i=1}^{2} C_i(1-\gamma_i)e^{-\tau_0}\frac{\tau_0^{C_i}}{C_i!} = \tau_0(1-\alpha)$$

$\Leftrightarrow$ (iff $C_1 \geq 1$)

$$\sum_{x=C_1+1}^{C_2-1} e^{-\tau_0}\frac{\tau_0^{x-1}}{(x-1)!} + \sum_{i=1}^{2} (1-\gamma_i)e^{-\tau_0}\frac{\tau_0^{C_i-1}}{(C_i-1)!} = 1-\alpha.$$

Problem 5.

T_n/θ has a χ^2-distribution with n degrees of freedom. Hence

$$P\{T_n \leq t\} = P\{T_n/\theta \leq t/\theta\} = \int_0^{t/\theta} \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} dy$$

$$= \int_0^t \frac{\theta^{-\frac{n}{2}}}{2^{n/2}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2\theta}} dx .$$

For varying θ these distributions form a one parameter exponential family. As in Example 2 p. 129 the UMP unbiased test for $H : \theta = 1$ against $K : \theta \neq 1$ has acceptance region $C_1 < T_n < C_2$ where C_1 and C_2 are determined by

$$\int_{C_1}^{C_2} f_n(y)dy = \int_{C_1}^{C_2} f_{n+2}(y)dy = 1-\alpha$$

with f_n the density of a χ^2-distribution with n degrees of freedom. The power of the UMP unbiased test is strictly decreasing for $0 < \theta < 1$ and strictly increasing for $\theta > 1$ (see p. 128). Hence it follows from the table that to ensure a power $\geq .9$ against both $\theta \geq 2$ and $\theta \leq .5$ we need at least $n = 45$ for the UMP unbiased test.

If the test is not required to be unbiased, we can restrict attention to tests in the class C, defined in Problem 4.8 (i). Let $\varphi \in C$ with critical points C_1^* and C_2^*. Then $C_1^* \leq C_1$ or $C_2^* \geq C_2$. Hence from the solution of 4.8 (ii) it follows that for $n \leq 44$ $\beta_\varphi(\frac{1}{2}) \leq \beta(\frac{1}{2}) = .895$ or $\beta_\varphi(2) \leq \beta(2) = .898$. Hence if the test is not required to be unbiased n has to be at least 45.

Remarks

1. The fact that we do not gain one or more observations if we delete the condition of unbiasedness follows from the fact that both $\beta(\frac{1}{2})$ and $\beta(2)$ are less that .9 for $n = 44$ and greater that .9 for $n = 45$.

2. Replacing .9 by another number (e.g. .6 or .2) one or more observations can be gained by using for example the maximin test (see Chapter 8).

Example:

for $C_1^* = 11.650$ and $C_2^* = 38.979$ $\beta_\varphi(\frac{1}{2}) = .615$ and $\beta_\varphi(2) = .615$ for $n = 22$

for $C_1^* = 2.026$ and $C_2^* = 18.937$ $\beta_\varphi(\frac{1}{2}) = .226$ and $\beta_\varphi(2) = .226$ for $n = 7$

Table (UMP unbiased test)

n	C_1	C_2	$\beta(\frac{1}{2})$	$\beta(2)$
1	.003	7.817	.063	.080
2	.085	9.530	.081	.113
3	.296	11.191	.102	.148
4	.607	12.802	.124	.182
5	.989	14.369	.148	.215
6	1.425	15.897	.173	.248
7	1.903	17.392	.198	.280
8	2.414	18.860	.224	.311
9	2.953	20.305	.251	.341
10	3.516	21.729	.278	.370
11	4.099	23.135	.305	.399
12	4.700	24.525	.332	.426
13	5.317	25.900	.359	.453
14	5.948	27.263	.385	.478
15	6.591	28.614	.412	.503
16	7.245	29.955	.438	.527
17	7.910	31.285	.463	.550
18	8.584	32.607	.488	.572
19	9.267	33.921	.513	.593
20	9.958	35.227	.537	.613
21	10.656	36.525	.560	.633
22	11.361	37.818	.582	.651
23	12.073	39.103	.604	.669
24	12.791	40.383	.625	.686
25	13.514	41.658	.646	.702
26	14.243	42.927	.665	.718
27	14.977	44.192	.684	.733
28	15.716	45.451	.702	.747
29	16.459	46.707	.719	.760
30	17.206	47.958	.735	.773
31	17.958	49.205	.751	.785
32	18.713	50.448	.766	.797
33	19.472	51.688	.780	.808
34	20.235	52.924	.794	.818
35	21.001	54.157	.807	.828
36	21.771	55.386	.819	.838
37	22.543	56.613	.830	.847
38	23.319	57.836	.841	.856
39	24.097	59.057	.852	.864
40	24.879	60.275	.861	.871
41	25.663	61.490	.871	.879
42	26.449	62.703	.879	.886
43	27.238	63.913	.887	.892
44	28.029	65.121	.895	.898
45	28.823	66.327	.902	.904
46	29.619	67.530	.909	.910
47	30.417	68.731	.915	.915

Problem 6.

The assertion that a UMP unbiased test for testing $H : \theta_1 = a, \theta_2 = b$ against the alternatives $\theta_1 \neq a$ or $\theta_2 \neq b$ does not exist, is not true in general. In the first part of this solution we give two counterexamples and in the second part we state sufficient conditions under which the assertion is true.

PART 1.

Counterexample 1.

If μ^T or ν^U are concentrated in (one or) two points a UMP unbiased test exists.

PROOF. Without loss of generality let $\mu\{0\} = \mu\{1\} = \frac{1}{2}$, ν arbitrary and $T(x) = x$. Define $p = p(\theta_1) = e^{\theta_1}/(1+e^{\theta_1})$, $q = 1-p$, $p^0 = p(a) = e^a/(1+e^a)$, $q^0 = 1-p^0$.

Unbiasedness of $\varphi(X,Y)$ is equivalent to

$$\begin{cases} pE_{\theta_2}\varphi(1,Y) + qE_{\theta_2}\varphi(0,Y) \geq \alpha & \text{for all } \theta_2, p \\ p^0E_b\varphi(1,Y) + q^0E_b\varphi(0,Y) \leq \alpha \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} E_{\theta_2}\varphi(1,Y) \geq \alpha, \quad E_{\theta_2}\varphi(0,Y) \geq \alpha \\ E_b\varphi(0,Y) = \alpha, \quad E_b\varphi(1,Y) = \alpha \end{cases} \tag{1}$$

Let $\varphi_2(x,y) = \varphi_2(y)$ be a UMP unbiased test of $\theta_2 = b$ against $\theta_2 \neq b$ in the one-paremeter exponential family $K(\theta_2)e^{\theta_2 U(y)}d\nu(\nu)$. Then it will be proved that φ_2 is a UMP unbiased test of $H_0 : \theta_1 = a, \theta_2 = b$ against $H_1 : \theta_1 \neq a$ or $\theta_2 \neq b$.

φ_2 is unbiased (cf. (1)). Let $\varphi^*(X,Y)$ be also an unbiased test of H_0 against H_1. Then by (1) $\varphi^*(0,Y)$ is an unbiased test of $\theta_2 = b$ against $\theta_2 \neq b$ and the same holds for $\varphi^*(1,Y)$.

Therefore

$$E_{\theta_2}\varphi^*(0,Y) \leq E_{\theta_2}\varphi_2(Y) \quad \text{for all } \theta_2$$

and

$$E_{\theta_2}\varphi^*(1,Y) \leq E_{\theta_2}\varphi_2(Y) \quad \text{for all } \theta_2$$

implying

$$E_{\theta_1,\theta_2}\varphi^*(X,Y) = pE_{\theta_2}\varphi^*(1,Y) + qE_{\theta_2}\varphi^*(0,Y)$$
$$\le pE_{\theta_2}\varphi_2(Y) + qE_{\theta_2}\varphi_2(Y)$$
$$= E_{\theta_2}\varphi_2(Y) = E_{\theta_1,\theta_2}\varphi_2(X,Y)$$

for all θ_1,θ_2. So in this case a UMP unbiased test exists.

Counterexample 2.

Let $\mu = \nu$, $T(x) = x$, $U(y) = y$, $\Theta = \{\theta : \int e^{\theta x}d\mu(x) < \infty\} = \mathbb{R}$.
Suppose that μ^T and ν^U are not concentrated in (one or) two points. Furthermore let $C_1 < C_2$ with $\mu((C_1,C_2)) = 0$, $\mu(C_i) > 0$ for $i = 1,2$, $\mu((-\infty,C_1)) > 0$ and $\mu((C_2,+\infty)) > 0$.
We now prove that under these conditions a UMP unbiased test exists.

PROOF. Let $p_i(\theta) = C(\theta)e^{\theta C_i}\mu(C_i)$, $i = 1,2$ with $C(\theta) = 1/\int e^{\theta x}d\mu(x)$. Then

$$\lim_{\theta\to+\infty} p_i(\theta) = 0, \quad \lim_{\theta\to-\infty} p_i(\theta) = 0, \quad p_i(\cdot) \text{ is continuous on } \mathbb{R}.$$

Hence there exists a real number a such that $p_1(a) + p_2(a) = \max_\theta \{p_1(\theta) + p_2(\theta)\}$.
Consider the testing problem $H : \theta_1 = a, \theta_2 = a$ against $K : (\theta_1,\theta_2) \neq (a,a)$.
Let

$$\alpha \ge P_{a,a}\{(X,Y) \notin \{C_1,C_2\} \times \{C_1,C_2\}\} \text{ and let}$$

and let

$$\varphi(x,y) = \begin{cases} \dfrac{\alpha - P_{a,a}\{(X,Y) \notin \{C_1,C_2\}\times\{C_1,C_2\}\}}{P_{a,a}\{(X,Y) \in \{C_1,C_2\}\times\{C_1,C_2\}\}} & \text{if } (x,y) \in \{C_1,C_2\}\times\{C_1,C_2\} \\ 1 & \text{otherwise} \end{cases}$$

Then φ is a UMP unbiased test of H against K, as will be proved.
It follows by direct calculations that $E_{a,a}\varphi(X,Y) = \alpha$ and $E_{\theta_1,\theta_2}\{1-\varphi(X,Y)\} \le 1-\alpha$ and hence $E_{\theta_1,\theta_2}\varphi(X,Y) \ge \alpha$. Therefore φ is unbiased.
Suppose φ^* is also unbiased and suppose $E_{\theta_1,\theta_2}\varphi^*(X,Y) \ge E_{\theta_1,\theta_2}\varphi(X,Y)$ for

all θ_1,θ_2. It can be shown (see part 2) that a.e. on $(x,y) \notin \{C_1,C_2\} \times \{C_1,C_2\}$ $\varphi^* = \varphi$. We therefore have:

$$\sum_{i=1}^{2}\sum_{j=1}^{2} p_i(\theta_1)p_j(\theta_2)\varphi^*(C_i,C_j) \geq \sum_{i=1}^{2}\sum_{j=1}^{2} p_i(\theta_1)p_j(\theta_2)\varphi(C_i,C_j) \tag{2}$$

with equality if $\theta_1 = \theta_2 = a$.
Divide both sides of (2) by $\{p_1(\theta_1)+p_2(\theta_1)\}\{p_1(\theta_2)+p_2(\theta_2)\}$ and write $r_i(\theta_j) = p_i(\theta_j) / (p_1(\theta_j)+p_2(\theta_j))$, $i = 1,2$, $j = 1,2$.
Then

$$\sum_{i=1}^{2}\sum_{j=1}^{2} r_i(\theta_1)r_j(\theta_2)\varphi^*(C_i,C_j) \geq \sum_{i=1}^{2}\sum_{j=1}^{2} r_i(\theta_1)r_j(\theta_2)\varphi(C_i,C_j) \tag{3}$$

with equality if $\theta_1 = \theta_2 = a$.
Note that

$$r_i(\theta_j) = \frac{e^{\theta_j C_i}\mu(C_i)}{e^{\theta_j C_1}\mu(C_1) + e^{\theta_j C_2}\mu(C_2)}$$

and hence

$$\lim_{\theta_j \to +\infty} r_1(\theta_j) = \lim_{\theta_j \to -\infty} r_2(\theta_j) = 0$$

and

$$\lim_{\theta_j \to +\infty} r_2(\theta_j) = \lim_{\theta_j \to -\infty} r_1(\theta_j) = 1.$$

Let $\theta_1 \to +\infty$ in (3). Then, for all θ_2,

$$r_1(\theta_2)\varphi^*(C_2,C_1) + r_2(\theta_2)\varphi^*(C_2,C_2)$$
$$\geq r_1(\theta_2)\varphi(C_2,C_1) + r_2(\theta_2)\varphi(C_2,C_2)$$

implying

$$\varphi^*(C_2,C_1) \geq \varphi(C_2,C_1) \quad \text{and} \quad \varphi^*(C_2,C_1) \geq \varphi(C_2,C_2).$$

Similarly we obtain

$$\varphi^*(C_1,C_1) \geq \varphi(C_1,C_1) \quad \text{and} \quad \varphi^*(C_1,C_2) \geq \varphi(C_1,C_2).$$

If $\theta_1 = \theta_2 = a$, equality holds in (3) (and $r_i(a) > 0$) and hence

$$\varphi^*(C_i,C_j) = \varphi(C_i,C_j), \quad i = 1,2, \quad j = 1,2.$$

Hence φ is a UMP unbiased test.

PART 2.

In this part we formulate sufficient conditions under which the assertion holds. First some preliminary considerations are made after which the solution is split up according to two different situations.

Preliminaries.

In part 2 we require that

(4) μ^T and ν^U are not concentrated in (one or) two points

(5) $(a,b) \in \text{int}\,(\Theta_1 \times \Theta_2)$ (for simplicity)

Consider the problems of testing the hypothesis $H : \theta_1 = a,\ \theta_2 = b$ against the alternatives

$$K_1 : \theta_1 \neq a \text{ or } \theta_2 \neq b$$

$$K_2 : \theta_1 \neq a,\ \theta_2 = b$$

$$K_3 : \theta_1 = a,\ \theta_2 \neq b.$$

For $\theta_2 = b$ is fixed, we have a one-parameter exponential family. Then by Section 2, there exists a UMP unbiased test φ_2, only depending on X, for testing H against K_2. Since φ_2 only depends on X, $E_{a,b}[\varphi_2(X)] = \alpha$ and $E_{\theta_1,b}[\varphi_2(X)] \geq \alpha$, $\theta_1 \neq a$; it follows that $E_{\theta_1,\theta_2}[\varphi_2(X)] \geq \alpha$, $\theta_1 \neq a$ or $\theta_2 \neq b$. So φ_2 is unbiased for testing H against K_1. Analogously there exists a UMP unbiased test φ_3 for testing H against K_3 which is unbiased for testing H against K_1.
Suppose that φ_1 is a UMP unbiased test of H against K_1 then $E_{\theta_1,\theta_2}[\varphi_1(X,Y)] \geq E_{\theta_1,\theta_2}[\varphi_i(X,Y)]$ for all $\theta_1 \neq a$ or $\theta_2 \neq b$ because φ_i is unbiased for H against K_i, $i = 2,3$. This implies that φ_1 is also a UMP unbiased test of H against K_i, $i = 2,3$. In view of Theorem 5 (iv) in Chapter 3 there exist constants k_1 and k_2 such that a.e. $\mu \times \nu$ (product measure)

$$\varphi_1(x,y) = \begin{cases} 1 & \text{if } C(a)(k_1 + k_2T(x))e^{aT(x)} \begin{matrix} < \\ > \end{matrix} C(\theta_1)e^{\theta_1 T(x)} \\ 0 & \end{cases}$$

or, equivalently (see the theory on p. 127)

$$\varphi_1(x,y) = \begin{cases} 1 & \text{if } T(x) \notin [C_1,C_2] \\ 0 & \phantom{\text{if }} T(x) \in (C_1,C_2) \end{cases}$$

for some $-\infty < C_1 \leq C_2 < \infty$; a.e. $\mu \times \nu$. Similarly,

$$\varphi_1(x,y) = \begin{cases} 1 & \text{if } U(y) \notin [D_1,D_2] \\ 0 & \phantom{\text{if }} U(y) \in (D_1,D_2) \end{cases}$$

for some $-\infty < D_1 \leq D_2 < \infty$; a.e. $\mu \times \nu$.

Situation 1: $\mu\{x : T(x) \notin [C_1,C_2]\} = 0$.

Since $\mu\{x : T(x) \notin [C_1,C_2]\} = 0 \Rightarrow \mu\{x : T(x) \in (C_1,C_2)\} > 0$ $\Rightarrow \nu\{y : U(y) \notin [D_1,D_2]\} = 0$, the distribution is concentrated on the rectangle $[C_1,C_2] \times [D_1,D_2]$ and $\varphi_1 = 0$ a.e. $\mu \times \nu$ for the vertices. Unbiasedness of φ_1 is equivalent to

$$(6) \qquad \begin{aligned} \sum_{i=1}^{2} \sum_{j=1}^{2} p_i r_j \varphi_1(C_i,D_j) &\geq \alpha \\ \sum_{i=1}^{2} \sum_{j=1}^{2} p_i^0 r_j^0 \varphi_1(C_i,D_j) &\leq \alpha \end{aligned}$$

where

$$\begin{aligned} p_i &= p_i(\theta_1) = C(\theta_1)e^{\theta_1 C_i}\mu(C_i), \quad i = 1,2 \\ r_j &= r_j(\theta_2) = K(\theta_2)e^{\theta_2 D_j}\nu(D_j), \quad j = 1,2 \\ p_i^0 &= p_i(a), \ r_i^0 = r_i(b). \end{aligned}$$

Note that $\Theta_1 = \Theta_2 = \mathbb{R}$, since the distribution is concentrated on a bounded set. Note also that, if

$$(7) \qquad \mu(C_i) > 0 \text{ and } \nu(D_i) > 0, \quad i = 1,2$$

$\lim_{\theta\to-\infty} p_1(\theta) = 1$, $\lim_{\theta\to+\infty} p_1(\theta) = 0$, $\lim_{\theta\to-\infty} r_1(\theta) = 1$, $\lim_{\theta\to+\infty} r_1(\theta) = 0$, $\lim_{\theta\to+\infty} p_2(\theta) = 1$ etc.

Let $\theta_1 \to +\infty$ in (6) then, for all θ_2,

$$r_1\varphi(C_2,D_1) + r_2\varphi(C_2,D_2) \geq \alpha$$

and hence

$$\varphi(C_2,D_1) \geq \alpha \text{ and } \varphi(C_2,D_2) \geq \alpha.$$

Similarly $\varphi(C_1,D_1) \geq \alpha$ and $\varphi(C_1,D_2) \geq \alpha$ and hence by the second line of (6) we have $\varphi(C_i,D_j) = \alpha$. But this implies $E_{\theta_1,\theta_2}\varphi \equiv \alpha$: a contradiction (cf. p. 128, and Theorem 6 of Chapter 3).
If (7) is violated we have for instance $\mu(C_1) = 0$, then $p_1 = 0$ and (6) reduces to

$$p_2 r_1\varphi(C_2,D_1) + p_2 r_2\varphi(C_2,D_2) \geq \alpha$$
$$p_2^0 r_1^0\varphi(C_2,D_1) + p_2^0 r_2^0\varphi(C_2,D_2) \leq \alpha.$$

But this is impossible (viz. if $\theta_2 \to -\infty$, $p_2 \to 0$ and the first line is violated). So in this case a UMP unbiased test does not exist.
The other cases where (7) is violated are treated in a similar way.

<u>Situation 2: $\mu\{x : T(x) \notin [C_1,C_2]\} > 0$</u>.

$\mu\{x : T(x) \notin [C_1,C_2]\} > 0 \Rightarrow \nu\{y : U(y) \in (D_1,D_2)\} = 0$
$\Rightarrow \nu\{y : U(y) \notin [D_1,D_2]\} > 0 \Rightarrow \mu\{x : T(x) \in (C_1,C_2)\} = 0$. Now $\varphi = 1$ a.e. $\mu \times \nu$ except for $\{C_1,C_2\} \times \{D_1,D_2\}$. In this case a UMP unbiased test is possible (see counterexample 2). So extra conditions are required. For instance μ or ν absolutely continuous w.r.t. Lebesgue measure is sufficient. Then $\varphi \equiv 1$ a.e. $\mu \times \nu$, which contradicts $\alpha < 1$.

<u>Remark</u>.

For discrete distributions counterexamples in the sense of counterexample 2 can occur.

<u>Problem 7</u>.

Since, by the translation $\theta_1^* = \theta_1 - a$, $\theta_2^* = \theta_2 - b$, the given testing problem can be reduced to the equivalent problem

$$H : \theta_1^* \leq 0, \quad \theta_2^* \leq 0 \text{ against } K : \theta_1^* > 0 \text{ or } \theta_2^* > 0 \text{ or both,}$$

we may assume a = b = 0 without loss of generality.

First assume that the natural parameter space Ω is open. The power function of any unbiased test $\phi(x,y)$ satisfies, with $\mathbb{R}^- = (-\infty,0]$

$$\beta(\theta_1,\theta_2) \leq \alpha \quad \text{if } (\theta_1,\theta_2) \in \Omega \cap (\mathbb{R}^- \times \mathbb{R}^-)$$
$$\beta(\theta_1,\theta_2) \geq \alpha \quad \text{if } (\theta_1,\theta_2) \in \Omega \cap (\mathbb{R}^- \times \mathbb{R}^-)^c.$$

These inequalities and the continuity of $\beta(\theta_1,\theta_2)$ (see Theorem 9, Chapter 2) imply that $\beta(0,\theta_2) = \alpha$ for all $\theta_2 : \theta_2 \in \mathbb{R}^-$ and $(0,\theta_2) \in \Omega$. Since, with $b_1 = \inf \{\theta_2 : (0,\theta_2) \in \Omega\}$ and $b_2 = \sup \{\theta_2 : (0,\theta_2) \in \Omega\}$, $\int \phi e^{\theta_2 y} d\mu$ is analytic in (b_1,b_2) and $C(0,\theta_2) = \int e^{\theta_2 y} d\mu$ is analytic and unequal to zero in (b_1,b_2); it follows that

$$\beta(0,\theta_2) = \alpha \text{ on } (b_1,b_2)$$

(see e.g. RUDIN (1970), Theorem 8.5).

Now an induction argument is employed to prove that

$$\frac{\partial^j \beta(0,\theta_2)}{\partial \theta_1^j} = 0 \text{ for all } \theta_2 \in (b_1,b_2) \text{ and all } j \in \mathbb{N}.$$

Fix $\theta_2 \in (0,b_2)$. Since $\beta(0,\theta_2) = \alpha$ and $\beta(\theta_1,\theta_2) \geq \alpha$ for all θ_1, $\beta(\theta_1,\theta_2)$ has a minimum at $\theta_1 = 0$. Hence $\frac{\partial \beta(0,\theta_2)}{\partial \theta_1} = 0$ for all $\theta_2 \in (0,b_2)$. Again by analiticity we get

$$\frac{\partial \beta(0,\theta_2)}{\partial \theta_1} = 0 \text{ for } \theta_2 \in (b_1,b_2).$$

Assume $\frac{\partial^j \beta(0,\theta_2)}{\partial \theta_1^j} = 0$ for all $j = 1,\dots,n-1$.

For any fixed $\theta_2 \in (b_1,b_2)$ and all θ_1 we have

$$\beta(\theta_1,\theta_2) = \alpha + \frac{\theta_1^n}{n!} \; \frac{\partial^n \beta(0,\theta_2)}{\partial \theta_1^n} + \theta_1^n \eta(\theta_1), \tag{8}$$

with $\eta(\theta_1) \to 0$ as $\theta_1 \to 0$.

Let n be odd. Since $\beta(\theta_1,\theta_2) \geq \alpha$ for all $\theta_2 \in (0,b_2)$, we have

$$\beta(\theta_1,\theta_2) - \alpha = \theta_1^n\left(\frac{1}{n!}\,\frac{\partial^n\beta(0,\theta_2)}{\theta_1^n} + \eta(\theta_1)\right) \geq 0$$

The preceding inequality implies

$$\frac{1}{n!}\,\frac{\partial^n\beta(0,\theta_2)}{\partial\theta_1^n} = \begin{cases} \lim\limits_{\theta_1\uparrow 0}\left(\frac{1}{n!}\,\frac{\partial^n\beta(0,\theta_2)}{\partial\theta_1^n} + \eta(\theta_1)\right) \leq 0 \\ \\ \lim\limits_{\theta_1\downarrow 0}\left(\frac{1}{n!}\,\frac{\partial^n\beta(0,\theta_2)}{\partial\theta_1^n} + \eta(\theta_1)\right) \geq 0. \end{cases}$$

Hence, for all $\theta_2 \in (0,b_2)$,

(9) $$\frac{\partial^n\beta(0,\theta_2)}{\partial\theta_1^n} = 0.$$

By analiticity (9) holds for all $\theta_2 \in (b_1,b_2)$.
Let n be even. For all $\theta_2 \in (b_1,0)$ we have $\beta(\theta_1,\theta_2) \leq \alpha$ if $\theta_1 \leq 0$ and $\beta(\theta_1,\theta_2) \geq \alpha$ if $\theta_1 > 0$. Because n is even we obtain in view of (8), for all $\theta_2 \in (b_1,0)$

(10) $$\frac{\partial^n\beta(0,\theta_2)}{\partial\theta_1^n} = 0$$

(by a similar argument as in the case n odd).
By analiticity (10) holds for all $\theta_2 \in (b_1,b_2)$. Hence

$$\beta(\theta_1,\theta_2) = \sum_{k=0}^{\infty} \frac{\theta_1^k}{k!}\,\frac{\partial^k\beta(0,\theta_2)}{\partial\theta_1^k} = \beta(0,\theta_2) = \alpha$$

for all $(\theta_1,\theta_2) \in \Omega$ or equivalently

$$E_{\theta_1,\theta_2}\{\phi(X,Y) - \alpha\} = 0 \quad \text{for all } (\theta_1,\theta_2) \in \Omega.$$

Hence $\phi(x,y) = \alpha$ a.e. μ by Theorem 1 of Chapter 4 (since Ω is open and hence it contains a 2-dimensional rectangle).

In a second part we drop the assumption that Ω is open, however we suppose $(0,0) \in \text{int}\,\Omega$.
For all $(\theta_1,\theta_2) \in \text{int}\,\Omega$ we have $\beta(\theta_1,\theta_2) = \alpha$, by the first part of the proof. Now let $(\theta_1^0,\theta_2^0) \in \Omega \setminus \text{int}\,\Omega$. Let $\theta_{in} = \theta_i^0(1-n^{-1})$, $n = 1,2,\ldots$, then $(\theta_{1n},\theta_{2n}) \in \text{int}\,\Omega$ as a convex combination of $(0,0) \in \text{int}\,\Omega$ and

$(\theta_1^0,\theta_2^0) \in \Omega$. Moreover, since exp $(\cdot)$ is a positive convex function,

$$\exp\,(\theta_{1n}x + \theta_{2n}y) \le 1 + \exp\,(\theta_1^0 x + \theta_2^0 y)$$

for all $x,y \in \mathbb{R}$. The function $1 + \exp\,(\theta_1^0 x + \theta_2^0 y)$ is integrable because $(0,0) \in \Omega$ and $(\theta_1^0,\theta_2^0) \in \Omega$. Hence for each bounded, measurable function ϕ we have by the dominated convergence theorem

$$\lim_{n\to\infty} \int \phi(x,y)\exp\,(\theta_{1n}x + \theta_{2n}y)d\mu = \int \phi(x,y)\exp\,(\theta_1^0 x + \theta_2^0 y)d\mu,$$

implying $\beta(\theta_1^0,\theta_2^0) = \lim_{n\to\infty} \beta(\theta_{1n},\theta_{2n}) = \alpha$.

(LEHMANN (1952))

Problem 8.

Throughout we assume $\theta_0 \in \text{int}\,\Omega$ (otherwise (ii) is incorrect).
(i) Let ϕ_0 be any level α test with $\beta'_{\phi_0}(\theta_0) = \rho$. Without loss of generality assume that ϕ_0 has size α.
Let $\theta_1 \neq \theta_0$ and consider the problem of maximizing $E_{\theta_1}[\psi(X)]$, subject to $E_{\theta_0}[\psi(X)] = \alpha$ and $\beta'_\psi(\theta_0) = \rho$. Since, by Theorem 9 of Chapter 2 and the argument leading to (6) on p. 127,

$$\beta'_\psi(\theta_0) = E_{\theta_0}[T(X)\psi(X)] - E_{\theta_0}[T]E_{\theta_0}[\psi(X)]$$

on equivalent formulation of the maximization problem is given by:

$$\text{(11)}\qquad \begin{aligned} &\text{maximize} \quad E_{\theta_1}[\psi(X)] \\ &\text{subject to } E_{\theta_0}[\psi(X)] = \alpha \text{ and } E_{\theta_0}[T(X)\psi(X)] = \rho + \alpha E_{\theta_0}[T]. \end{aligned}$$

Define $M = \{(E_{\theta_0}[\psi(X)],\ E_{\theta_0}[T(X)\psi(X)]) : \psi \text{ is a critical function}\}$, which is a convex set.
Now we shall prove that either it is possible to construct a test $\phi \in C$ which has the same power as ϕ_0 (we say for short that there exists an "equivalent" test $\phi \in C$) or $(\alpha, \rho + \alpha E_{\theta_0}T)$ is an inner point of M.

First suppose that for all critical functions ψ, with $E_{\theta_0}[\psi(X)] = \alpha$,

$$E_{\theta_0}[T(X)\psi(X)] \le \rho + \alpha E_{\theta_0}[T];$$

then ϕ_0 maximizes $E_{\theta_0}[T(X)\psi(X)]$ among all critical functions of size α. Then, by Theorem 5 (iv) of Chapter 3, there exists a constant k such that a.e.

$$\phi_0(x) = \begin{cases} 1 & \text{if } T(x) > k \\ 0 & \text{if } T(x) < k. \end{cases}$$

Hence an equivalent test ϕ exists in C. Similarly, if

$$E_{\theta_0}[T(X)\psi(X)] \geq \rho + \alpha E_{\theta_0}[T]$$

for all critical functions ψ with $E_{\theta_0}[\psi(X)] = \alpha$, we have that ϕ_0 minimizes $E_{\theta_0}[T(X)\psi(X)]$ and again by Theorem 5 (iv) of Chapter 3, the existence of an equivalent test $\phi \in C$ follows.

The final possibility is the existence of size α critical functions ψ_1 and ψ_2 such that $E_{\theta_0}[\psi_1(X)T(X)] > \rho + \alpha E_{\theta_0}[T]$ and $E_{\theta_0}[\psi_2(X)T(X)] < \rho + \alpha E_{\theta_0}[T]$. Then, using the fact that M constains all points $(u, uE_{\theta_0}[T])$ with $0 \leq u \leq 1$, it is easily seen that $(\alpha, \rho + \alpha E_{\theta_0}[T])$ is an inner point of M.

Hence by Theorem 5 (ii,iv) of Chapter 3, there exist constants k_1 and k_2 and a test ϕ, such that:

$$\phi(x) = \begin{cases} 1 \\ 0 \end{cases} p_{\theta_1}(x) \begin{matrix} > \\ < \end{matrix} k_1 p_{\theta_0}(x) + k_2 T(x) p_{\theta_0}(x)$$

is a solution of the maximization Problem (11).

An equivalent definition for ϕ is

$$\phi(x) = \begin{cases} 1 \\ 0 \end{cases} a_1 + a_2 T(x) \begin{matrix} < \\ > \end{matrix} e^{bT(x)},$$

which implies that the rejection region is the complement of an interval. Indeed, a one-sided rejection region cannot occur, because this would imply that ϕ maximizes or minimizes $E_{\theta_0}[T(X)\psi(X)]$ subject to $E_{\theta_0}[\psi(X)] = \alpha$ by Theorem 5 (ii) of Chapter 3, which contradicts $(\alpha, \rho + \alpha E_{\theta_0}[T]) \in \text{int } M$.

Since for some C_1, C_2

$$p_{\theta_1}(x) > k_1 p_{\theta_0}(x) + k_2 T(x) p_{\theta_0}(x) \Leftrightarrow T(x) \notin [C_1, C_2],$$

$\tilde{k}_1$ and $\tilde{k}_2$ can be chosen such that

$$p_{\theta_2}(x) > \tilde{k}_1 p_{\theta_0}(x) + \tilde{k}_2 T(x) p_{\theta_0}(x)$$
$$\Leftrightarrow T(x) \notin [C_1, C_2]$$
$$\Leftrightarrow p_{\theta_1}(x) > k_1 p_{\theta_0}(x) + k_2 T(x) p_{\theta_0}(x)$$

implying by Theorem 5 (ii) of Chapter 3 that ϕ also is a solution of (1) with θ_1 replaced by θ_2. Therefore $E_\theta[\phi(X)] \geq E_\theta[\phi_0(X)]$ for all θ.

(ii) T is distributed according to the exponential family $dP_\theta(t) = C(\theta)e^{t\theta}d\nu(t)$.

Let $\phi, \tilde{\phi} \in \mathcal{C}$, i.e. ϕ and $\tilde{\phi}$ satisfy (3) and (5) of Section 2, with parameters $C_1, C_2, \gamma_1, \gamma_2$ and $\tilde{C}_1, \tilde{C}_2, \tilde{\gamma}_1, \tilde{\gamma}_2$ respectively. Without restriction we may suppose that either "$C_1 < \tilde{C}_1$" or "$C_1 = C_1$ and $\gamma_1 \leq \tilde{\gamma}_1$".

Define $\psi(t) = \phi(t) - \tilde{\phi}(t)$, $J = \{t : \psi(t) < 0\}$ and $I = \{t : \psi(t) > 0\}$. Then it is immediate that $t \leq C_2$ for all $t \in J$. Also we have that $t \geq C_2$ for all $t \in I$. This follows easily from the remark that $\tilde{\phi}(t) = 1$ if $t < \tilde{C}_1$, $\phi(t) = \gamma_1 \leq \tilde{\gamma}_1 = \tilde{\phi}(t)$ if $C_1 = \tilde{C}_1 = t$ and $\phi(t) = 0$ if $C_1 < t < C_2$.

Let θ be such that $\theta > \theta_0$ and define $s = p_\theta(C_2)/p_{\theta_0}(C_2)$. Then, since $p_\theta(t)/p_{\theta_0}(t)$ is increasing with respect to t, we have $p_\theta(t)/p_{\theta_0}(t) \leq s$ for $t \in J$ and $p_\theta(t)/p_{\theta_0}(t) \geq s$ for $t \in I$. Hence

$$\int_J (-\psi) dP_\theta \leq s \int_J (-\psi) dP_{\theta_0} = s \int_I \psi dP_{\theta_0} \leq \int_I \psi dP_\theta,$$

where the equality follows from the fact that

$$0 = \int \psi dP_{\theta_0} = \int_J \psi dP_{\theta_0} + \int_I \psi dP_{\theta_0}$$

(ϕ and $\tilde{\phi}$ are size α tests).

First suppose that $\int_J (-\psi) dP_\theta = \int_I \psi dP_\theta$, then

$$\int_J (-\psi)(s p_{\theta_0} - p_\theta) d\nu = \int_I \psi (s p_{\theta_0} - p_\theta) d\nu = 0.$$

This implies that $-\psi(s p_{\theta_0} - p_\theta) = 0$ a.e. on J and $\psi(s p_{\theta_0} - p_\theta) = 0$ a.e. on I. Since $\psi \neq 0$ on J and I and T is distributed according to the exponential family $dP_\theta(t) = C(\theta)e^{t\theta}d\nu(t)$ it follows that $e^{(\theta-\theta_0)t}$ is constant a.e. on $I \cup J$. Hence $\nu(I \cup J) = 0$ which implies $\phi = \tilde{\phi}$ a.e.

Secondly suppose that $\int_J (-\psi) dP_\theta < \int_I \psi dP_\theta$ or equivalently $0 < \int \psi dP_\theta$. Then $\beta_\phi(\theta) > \beta_{\tilde{\phi}}(\theta)$. Similarly we prove that $\beta_\phi(\theta') < \beta_{\tilde{\phi}}(\theta')$ for all

$\theta' < \theta_0$, unless $\phi = \tilde{\phi}$ a.e.

(iii) Let C^* denote the class of all two- or one-sided tests based on T(x): i.e. the class of all tests of the form:

$$\phi(x) = \begin{cases} 1 & T(x) < C_1 \text{ or } > C_2 \\ \gamma_i \quad \text{if} & T(x) = C_i, \quad i = 1,2 \\ 0 & C_1 < T(x) < C_2 \end{cases}$$

where $-\infty \le C_1 \le C_2 \le \infty$ and $0 \le \gamma_1, \gamma_2 \le 1$.
To prove (iii) we add the obvious condition $L_0(\theta_0) < L_1(\theta_0)$.
For any two test (i.e. decision procedures) ϕ and ϕ'

$$(12) \qquad R(\theta,\phi') - R(\theta,\phi) = [L_1(\theta) - L_0(\theta)] \cdot [\beta_{\phi'}(\theta) - \beta_\phi(\theta)],$$

which follows by the proof of Theorem 3 in Chapter 3. Let ϕ be any test and let $\alpha = E_{\theta_0}\phi$. By (i) there exists a test $\phi' \in C^*$ with $E_{\theta_0}\phi' = \alpha$ such that $\beta_{\phi'}(\theta) - \beta_\phi(\theta) \ge 0$ for all θ. Hence, if $L_1(\theta) \le L_0(\theta)$ for all $\theta \ne \theta_0$ we have $R(\theta,\phi') \le R(\theta,\phi)$ for all θ. Hence C^* is essentially complete.
To prove that C^* is minimal, let $\phi_0 \in C^*$ and define $C' =$
$C^* \setminus \{\phi \in C^* : \phi = \phi_0 \text{ a.e.}\}$.
Let $\alpha_0 = E_{\theta_0}\phi_0$. Part (ii) ensures that for any $\phi \in C'$ with $E_{\theta_0}\phi' = \alpha_0$ parameters θ_1, θ_2, both $\ne \theta_0$, exist, such that

$$(13) \qquad \beta_\phi(\theta_1) > \beta_{\phi_0}(\theta_1) \quad \text{and} \quad \beta_\phi(\theta_2) < \beta_{\phi_0}(\theta_2).$$

If $L_1(\theta) < L_0(\theta)$ for all $\theta \ne \theta_0$, (12) and (13) imply that no test $\phi \in C'$, with $E_{\theta_0}\phi = \alpha_0$, exists such that $R(\theta,\phi) \le R(\theta,\phi_0)$ for all θ. For any $\phi \in C'$, with $E_{\theta_0}\phi > \alpha_0$, we have (by the additional assumption) $R(\theta_0,\phi) > R(\theta_0,\phi_0)$. For any $\phi \in C'$, with $E_{\theta_0}\phi < \alpha_0$, we have, by continuity of the power function, $R(\theta,\phi) > R(\theta,\phi_0)$ for $\theta - \theta_0$ sufficiently small, but unequal to zero.

(iv) In this part we extend (iii) to the problem of testing $H' : \theta \in [\theta_1,\theta_2]$ against the alternative $K' : \theta \notin [\theta_1,\theta_2]$. As in (iii) the problem is considered as a two-decision problem with decisions d_0 and d_1 corresponding to acceptance and rejection of H' and with loss function $L(\theta,d_i) = L_i(\theta)$, $i = 0,1$.
For this two-decision problem, the following results hold:

(14) if $L_1(\theta) - L_0(\theta) \geq 0$ for all $\theta \in [\theta_1,\theta_2]$ and
$L_1(\theta) - L_0(\theta) \leq 0$ for all $\theta \notin [\theta_1,\theta_2]$,
then the class C^* is essentially complete.

(15) if $L_1(\theta) - L_0(\theta) > 0$ for all $\theta \in [\theta_1,\theta_2]$ and
$L_1(\theta) - L_0(\theta) < 0$ for all $\theta \notin [\theta_1,\theta_2]$,
then the class C^* is minimal essentially complete.

Proof of (14). Take an arbitrary test ϕ_0 and define $\alpha = E_{\theta_1}[\phi_0(X)]$ and $\beta = E_{\theta_2}[\phi_0(X)]$. Given ϕ_0 and a test $\phi \in C^*$ the difference between the risk functions satisfies

$$R(\theta,\phi) - R(\theta,\phi_0) = [L_1(\theta) - L_0(\theta)][\beta_\phi(\theta) - \beta_{\phi_0}(\theta)]$$

by the proof of Theorem 3 of Chapter 3. Hence, under condition (14), the inequality $R(\theta,\phi) \leq R(\theta,\phi_0)$ holds for all θ and some $\phi \in C^*$ (i.e. C^* is essentially complete) if we can prove that either there exists an equivalent test $\phi \in C^*$ or there exists a test $\phi \in C^*$ such that

$$E_\theta[\phi(X)] \geq E_\theta[\phi_0(X)] \quad \text{for all } \theta \notin [\theta_1,\theta_2]$$
$$E_\theta[\phi(X)] \leq E_\theta[\phi_0(X)] \quad \text{for all } \theta \in]\theta_1,\theta_2[$$
$$E_{\theta_1}[\phi(X)] = \alpha, \quad E_{\theta_2}[\phi(X)] = \beta.$$

First suppose that for all critical functions ϕ, satisfying $E_{\theta_1}[\phi(X)] = 1-\alpha$, $E_{\theta_2}[\phi(X)] \leq 1-\beta$. Then $1-\phi_0$ is a level $(1-\alpha)$ test, which is most powerful for the testing problem $H : \theta = \theta_1$ against $K : \theta = \theta_2$. Hence from the Neyman-Pearson fundamental lemma $1-\phi_0$ (and hence ϕ_0) is a.e. one-sided. This implies that there exists $\phi \in C^*$, which is equivalent to ϕ_0. Similarly if for all critical functions ϕ, satisfying $E_{\theta_1}[\phi(X)] = 1-\alpha$, $E_{\theta_2}[\phi(X)] > 1-\beta$.
Since the two preceding arguments can be repeated with the roles of α and β interchanged, it remains to consider as a final case the situation where $(1-\alpha,1-\beta)$ is an inner point of $M = \{(E_{\theta_1}\psi, E_{\theta_2}\psi) : \psi \text{ a critical function}\}$. Let θ be any value between θ_1 and θ_2 and consider the maximization problem:

(16)
$$\text{maximize} \quad E_\theta[\phi(X)]$$
$$\text{subject to } E_{\theta_1}[\phi(X)] = 1-\alpha \text{ and } E_{\theta_2}[\phi(X)] = 1-\beta.$$

Then, by Theorem 5 (ii,iv) of Chapter 3, there exist constants k_1 and k_2 and a test ϕ^*, such that

$$\phi^*(x) = \begin{cases} 1 & p_\theta(x) > k_1 p_{\theta_1}(x) + k_2 p_{\theta_2}(x) \\ 0 & p_\theta(x) < k_1 p_{\theta_1}(x) + k_2 p_{\theta_2}(x) \end{cases}$$

is the solution of the maximization problem (16).
Just as in (i) it is seen that ϕ^* is also a solution of (16) with θ replaced by $\tilde{\theta} \in (\theta_1,\theta_2)$. Hence $E_\theta[(1-\phi^*)(X)] \leq E_\theta[\phi_0(X)]$ for all $\theta \in (\theta_1,\theta_2)$.
Now let θ' be a value of the parameter less then θ_1. Since for some C_1,C_2

$$\begin{aligned} &p_\theta(x) < k_1 p_{\theta_1}(x) + k_2 p_{\theta_2}(x) \\ &\Leftrightarrow T(X) \notin [C_1,C_2] \Leftrightarrow p_{\theta'}(x) > \tilde{k}_1 p_{\theta_1}(x) + \tilde{k}_2 p_{\theta_2}(x) \end{aligned}$$

for some $\tilde{k}_1,\tilde{k}_2$; $1-\phi^*$ maximizes $E_{\theta'}[\phi(X)]$ subject to $E_{\theta_1}[\phi(X)] = \alpha$ and $E_{\theta_2}[\phi(X)] = \beta$ by Theorem 5 (ii) of Chapter 3. Therefore $E_\theta[(1-\phi^*)(X)] \geq E_\theta[\phi_0(X)]$ for all $\theta < \theta_1$.
Repeating once more the preceding arguments for a value of the parameter θ greater then θ_2, we end up with the conclusion that the test $1-\phi^*$ has the desired properties.

Proof of (15). Let $\phi_0 \in C^*$ and define $C' = C \setminus \{\phi \in C : \phi = \phi_0 \text{ a.e.}\}$.
Suppose C' is essentially complete. Then a test $\phi \in C'$ exists such that $R(\phi,\theta) - R(\phi_0,\theta) \leq 0$ for all θ. Hence, by the assumptions in (15), it follows that $\beta_\phi(\theta) - \beta_{\phi_0}(\theta) \leq 0$ for all $\theta \in [\theta_1,\theta_2]$ and $\beta_\phi(\theta) - \beta_{\phi_0}(\theta) \geq 0$ for all $\theta \notin [\theta_1,\theta_2]$. Now, the continuity of $\beta_\phi - \beta_{\phi'}$ (by Theorem 9, Chapter 2) implies that:

(17) $$\beta_\phi(\theta_1) = \beta_{\phi_0}(\theta_1) = \alpha \text{ say}$$

and

(18) $$\beta_\phi(\theta_2) = \beta_{\phi_0}(\theta_2) = \beta \text{ say.}$$

From (17) it follows that ϕ and ϕ_0 belong to the class C as defined in (i) (with θ_0 replaced by θ_1). Further, from the proof of (ii), it follows that $\beta_\phi(\theta_2) \neq \beta_{\phi_0}(\theta_2)$ which contradicts (18). Hence the reduced class C' is no more essentially complete, or equivalently C^* is minimal.

(LEHMANN (1947); FERGUSON (1967) pp. 217-223)

Section 3

Problem 9.

Let $X_1,\ldots,X_n$ be a sample from

(i) the normal distribution $N(a\sigma,\sigma^2)$, with a fixed and $0 < \sigma < \infty$

(ii) the uniform distribution $R(\theta-\frac{1}{2},\theta+\frac{1}{2})$, $-\infty < \theta < \infty$

(iii) the uniform distribution $R(\theta_1,\theta_2)$, $-\infty < \theta_1 < \theta_2 < \infty$

then the joint density of $X_1,\ldots,X_n$ is given by

(i)
$$p_\sigma(x_1,\ldots,x_n) = \frac{1}{(2\pi)^{n/2}}\,\frac{1}{\sigma^n}\prod_{i=1}^{n}\exp\left\{-\frac{1}{2}\left(\frac{x_i - a\sigma}{\sigma}\right)^2\right\}$$
$$= \frac{1}{(2\pi)^{n/2}}\,\frac{1}{\sigma^n}\exp\left\{-\frac{1}{2}\left(\frac{\Sigma x_i^2}{\sigma^2} - 2\frac{a}{\sigma}\Sigma x_i + na^2\right)\right\} = g_\sigma\left(\Sigma x_i,\Sigma x_i^2\right)$$

(ii)
$$p_\theta(x_1,\ldots,x_n) = \begin{cases} 1 & \text{if } \theta-\frac{1}{2} < x_1,\ldots,x_n < \theta+\frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} 1 & \text{if } \theta-\frac{1}{2} < \min(x_1,\ldots,x_n) \le \max(x_1,\ldots,x_n) < \theta+\frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$
$$= g_\theta(\min(x_1,\ldots,x_n),\ \max(x_1,\ldots,x_n))$$

(iii)
$$p_{\theta_1,\theta_2}(x_1,\ldots,x_n) = \begin{cases} (\theta_2-\theta_1)^{-n} & \text{if } \theta_1 < \min(x_1,\ldots,x_n) \le \max(x_1,\ldots,x_n) < \theta_2 \\ 0 & \text{otherwise} \end{cases}$$
$$= g_{\theta_1,\theta_2}(\min(x_1,\ldots,x_n),\max(x_1,\ldots,x_n)).$$

Hence

(i) $(\Sigma X_i,\ \Sigma X_i^2)$ is sufficient for $N(a\sigma,\sigma^2)$

(ii) $(\min(X_1,\ldots,X_n),\max(X_1,\ldots,X_n))$ is sufficient for $R(\theta-\frac{1}{2},\theta+\frac{1}{2})$

(iii) $(\min(X_1,\ldots,X_n),\max(X_1,\ldots,X_n))$ is sufficient for $R(\theta_1,\theta_2)$

by the factorization theorem.

Let P^T be the family of distributions of $T = (T_1,T_2) = (\Sigma X_i,\Sigma X_i^2)$, where $X_1,\ldots,X_n$ is a sample from $N(a\sigma,\sigma^2)$.
Since

$$P\{\Sigma X_i / \sqrt{\Sigma X_i^2} \le x\} = P\{\Sigma X_i / \sigma)(\sqrt{\Sigma(X_i / \sigma)^2} \le x\}$$

$$= P\{\Sigma Z_i / \sqrt{\Sigma Z_i^2} \le x\},$$

where $Z_1,\ldots,Z_n$ are independent $N(a,1)$ distributed random variables, the distribution of $\Sigma X_i / \sqrt{\Sigma X_i^2}$ does not depend on σ.
Define

$$f(t_1,t_2) = \begin{cases} t_1 / \sqrt{t_2} - E[\Sigma X_i / \sqrt{\Sigma X_i^2}] & \text{if } (t_1 / \sqrt{t_2})^2 \text{ and } t_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then $E_\sigma[f(\Sigma X_i, \Sigma X_i^2)] = 0$ for all $\sigma > 0$ and $f \not\equiv 0$, hence the family P^T is not complete. Since the function f is bounded, it also follows that P^T is not boundedly complete.

Let P^T be the family of distributions of $T = (T_1,T_2) =$ $(\min(X_1,\ldots,X_n), \max(X_1,\ldots,X_n))$ where $X_1,\ldots,X_n$ is a sample from the uniform distribution $R(\theta-\frac{1}{2},\theta+\frac{1}{2})$.
It is easily seen that $E_\theta[\min(X_1,\ldots,X_n)] = \frac{1}{n+1} + (\theta-\frac{1}{2})$ and $E_\theta[\max(X_1,\ldots,X_n)] = \frac{1}{n+1} + (\theta-\frac{1}{2})$.
Hence we have for all θ that $E_\theta\left[\max(X_1,\ldots,X_n) - \min(X_1,\ldots,X_n) - \frac{n-1}{n+1}\right] = 0$.
Since

$$f(t_1,t_2) = \begin{cases} t_2-t_1 - \frac{n-1}{n+1} & \text{if } 0 < t_1 < t_2 < t_1+1 \\ 0 & \text{otherwise} \end{cases}$$

is a bounded function and $E_\theta f(T_1,T_2) = 0$ for all $-\infty < \theta < \infty$, it follows that the family P^T is neither complete nor boundedly complete.

Finally, let P_T denote the family of distributions of $T = (T_1,T_2) =$ $(\min(X_1,\ldots,X_n), \max(X_1,\ldots,X_n))$, where $X_1,\ldots,X_n$ is a sample from the uniform distribution $R(\theta_1,\theta_2)$. From the distribution theory of order statistics it follows that the density of T is given by

$$p(t_1,t_2) = \begin{cases} \dfrac{n(n-1)(t_2-t_1)^{n-2}}{(\theta_2-\theta_1)^n} & \text{if } \theta_1 < t_1 < t_2 < \theta_2 \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that for all (θ_1,θ_2), with $\theta_1 < \theta_2$,

$$\int_{\theta_1}^{\theta_2}\int_{\theta_1}^{t_2} f(t_1,t_2)(t_2-t_1)^{n-2}dt_1dt_2 = 0$$

Let $f(t_1,t_2) = f^+(t_1,t_2) - f^-(t_1,t_2)$ where f^+ and f^- denote the positive and negative parts of f respectively. Then for all Borel sets A of X

$$\nu^+(A) = \iint_A f^+(t_1,t_2)(t_2-t_1)^{n-2}dt_1dt_2$$

and

$$\nu^-(A) = \iint_A f^-(t_1,t_2)(t_2-t_1)^{n-2}dt_1dt_2$$

are two measures over the Borel sets on $X = \{(x_1,x_2) : x_1 < x_2\} \subset \mathbb{R}^2$, which agree for all triangles $\Delta_{\theta_1,\theta_2} = \{(t_1,t_2) : \theta_1 \le t_1 \le t_2; \theta_1 \le t_2 \le \theta_2\}$ Since these triangles generate the Borel sets of X, $\nu^+(A) = \nu^-(A)$ for all Borel sets A of X. This implies $f^+(t_1,t_2) = f^-(t_1,t_2)$ on X except possibly on a set of Lebesgue measure zero and hence $f(t_1,t_2) = 0$ a.e. P^T.

Problem 10.

Let $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ be independent samples from $N(\xi,\sigma^2)$ and $N(\xi,\tau^2)$. The statistic $T = (\Sigma X_i, \Sigma Y_j, \Sigma X_i^2, \Sigma Y_j^2)$ is sufficient but not complete (Example 5). Here it will be shown that T is also not boundedly complete.

Let

$$f(t_1,t_2,t_3,t_4) = \begin{cases} 1 & \text{if } \frac{t_2}{n} - \frac{t_1}{m} \ge 0 \\ -1 & \text{if } \frac{t_2}{n} - \frac{t_1}{m} < 0. \end{cases}$$

Then

$$E_{(\xi,\sigma^2,\tau^2)}[f(T)] = E_{(\xi,\sigma^2,\tau^2)}[f(\Sigma X_i,\Sigma Y_j,\Sigma X_i^2,\Sigma Y_j^2)] =$$
$$P_{(\xi,\sigma^2,\tau^2)}\{\bar{Y}-\bar{X} \ge 0\} - P_{(\xi,\sigma^2,\tau^2)}\{\bar{Y}-\bar{X} < 0\} = \tfrac{1}{2} - \tfrac{1}{2} = 0.$$

Problem 11.

First suppose that f is an arbitrary bounded (measurable) function and that for all θ

$$0 = E_\theta[f(X)] = f(-1)\theta + \sum_{x=0}^{\infty} f(x)(1-\theta)^2\theta^x$$

$$= f(0) + \sum_{x=1}^{\infty} (f(x-2) - 2f(x-1) + f(x))\theta^x.$$

The above power series in θ converges absolutely for every interior point θ of the unit circle (f is bounded), so all the coefficients of the power series are equal to zero.

Hence $f(0) = 0$ and $f(x) - 2f(x-1) + f(x-2) = 0$ for $x = 1,2,\ldots$.

By induction it follows that $f(x) = -xf(-1)$ for $x = 0,1,2,\ldots$.

But, since f is bounded, we conclude that $f(x) \equiv 0$ a.e. $\mathcal{P}$. To show that $\mathcal{P}$ is not complete, consider the function $\tilde{f}(x) = x$.

For $0 < \theta < 1$

$$E_\theta[\tilde{f}(X)] = -\theta + \theta(1-\theta)^2 \sum_{x=0}^{\infty} x\theta^{x-1}$$

$$= -\theta + \theta(1-\theta)^2 \sum_{x=0}^{\infty} \frac{d}{d\theta}\theta^x$$

$$= -\theta + \theta(1-\theta)^2 \frac{d}{d\theta}\frac{1}{1-\theta} = 0$$

Hence $\mathcal{P}$ is not complete.

(GIRSHICK, MOSTELLER and SAVAGE (1946))

Problem 12.

Given $\mathcal{P} = \{P\}$ a family of distributions with the property that for any $P,Q \in \mathcal{P}$ there exists a $0 < p < 1$ such that $pP + (1-p)Q \in \mathcal{P}$; it can easily be shown by induction over m that for any $P,Q \in \mathcal{P}$ there exists $\alpha_1,\alpha_2,\ldots$ satisfying $1 > \alpha_1 > \alpha_2 > \ldots > 0$ such that

$$\alpha_m P + (1-\alpha_m)Q \in \mathcal{P} \quad \text{for } m = 1,2,\ldots .$$

Then for any n, $P,Q \in \mathcal{P}$ and h satisfying

$$(19) \qquad \int h(x_1,\ldots,x_n)d\tilde{P}(x_1)\ldots d\tilde{P}(x_n) = 0 \quad \text{for all } \tilde{P} \in \mathcal{P}$$

it holds for $m = 1,2,\ldots$ that

$$(20) \qquad 0 = \int h(x_1,\ldots,x_n) \prod_{k=1}^{n} d[\alpha_m P(x_k) + (1-\alpha_m)Q(x_k)]$$

$$= \int h(x_1,\dots,x_n) \prod_{k=1}^{n} [\alpha_m dP(x_k) + (1-\alpha_m) dQ(x_k)]$$

$$= \sum_{k=0}^{n} \binom{n}{k} \alpha_m^k (1-\alpha_m)^{n-k} J_k^{(n)}(P,Q,h)$$

$$= (1-\alpha_m)^n \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\alpha_m}{1-\alpha_m}\right)^k J_k^{(n)}(P,Q,h)$$

by Fubini's theorem and symmetry of h; where for $0 \le k \le n$

$$J_k^{(n)}(P,Q,h) = \int h(x_1,\dots,x_n) dP(x_1)\dots dP(x_k) dQ(x_{k+1})\dots dQ(x_n).$$

Note that by (19), (20) and Fubini's theorem $J_k^{(n)}(P,Q,h)$ is well defined. Since in (20) we end up with a polynomial with an infinite number of zeros $\left(\frac{\alpha_m}{1-\alpha_m}, m = 1,2,\dots\right)$ all the coefficients must by zero (see also Example 3). Hence $J_k^{(n)}(P,Q,h) = 0$ for $k = 0,1,\dots,n$.

Now, using the preceding result, we will prove for $n = 1,2,\dots$

(21) for any h satisfying (1)
$\int h(x_1,\dots,x_n) dP_1(x_1)\dots dP_n(x_n) = 0$ for each n-tuple $(P_1,\dots,P_n) \in \mathcal{P}$.

The proof is by induction over n. For $n = 1$ it is trivial. Suppose that (21) is true for some n. Consider any h satisfying (19) (with n replaced by n+1) and any $P_1,\dots,P_{n+1} \in \mathcal{P}$.
For all $P \in \mathcal{P}$

$$J_1^{(n+1)}(P_{n+1},P,h) =$$
$$\int h(x_1,\dots,x_{n+1}) dP_{n+1}(x_{n+1}) dP(x_1)\dots dP(x_n) = 0.$$

From Fubini's theorem and symmetry of h it follows that

$$\int h^*(x_1,\dots,x_n) dP(x_1)\dots dP(x_n) = 0$$

where

$$h^*(x_1,\dots,x_n) = \int h(x_1,\dots,x_{n+1}) dP_{n+1}(x_{n+1}).$$

Hence by the induction argument

$$\int h(x_1,\ldots,x_{n+1})dP_1(x_1)\ldots dP_n(x_n)dP_{n+1}(x_{n+1}) = 0.$$

(FRASER (1953); WALSH (1949))

Problem 13.

The distributions in $\mathcal{P}$ are continuous. By the remarks on p. 48 and Example 7 of Chapter 2, it follows that T is sufficient for $\mathcal{P}$.
Suppose $E_Q[h(T)] = 0$ for all $Q \in \mathcal{P}^T$. Remark that $E_Q[h(T)] = 0$ for all $Q \in \mathcal{P}^T$ implies $E_P[h(T(X))] = 0$ for all $P \in \mathcal{P}$ by Lemma 2 of Chapter 2.
Since $h(T(X_1,\ldots,X_n))$ is a symmetric function and since (by independence) for all $P \in \mathcal{P}$

$$E_P[h(T(X))] = \int h(T(x_1,\ldots,x_n))dP(x_1)\ldots\, dP(x_n) = 0,$$

Problem 12 implies

$$\int h(T(x_1,\ldots,x_n))dP_1(x_1)\ldots dP_n(x_n) = 0 \tag{22}$$

for each n-tuple $(P_1,\ldots,P_n) \in \mathcal{P}^n$.
Let $B_1,\ldots,B_n$ be n finite intervals and let $P_1,\ldots,P_n$ be the uniform distributions over these intervals. They can be written as

$$dP_i(x) = c_i I_{B_i}(x)d\lambda(x), \qquad i = 1,\ldots,n$$

where λ is the Lebesgue measure.
Hence

$$\int_{B_1\times\ldots\times B_n} h(T(x_1,\ldots,x_n))d\lambda(x_1)\ldots d\lambda(x_n) = 0. \tag{23}$$

Let h^+ and h^- denote the positive and negative parts of h respectively.
Then

$$\nu^+(A) = \int_A h^+(T(x_1,\ldots,x_n))d\lambda(x_1)\ldots d\lambda(x_n)$$

$$\nu^-(A) = \int_A h^-(T(x_1,\ldots,x_n))d\lambda(x_1)\ldots d\lambda(x_n)$$

are two measures over the Borel sets $A \in \mathcal{B}^n$.

Since by (23) it follows that $\nu^+(B_1 \times \ldots \times B_n) - \nu^-(B_1 \times \ldots \times B_n) = 0$ for each n-dimensional bounded rectangle $B_1 \times \ldots \times B_n \in \mathcal{B}^n$, the measures ν^+ and ν^- agree on bounded rectangles and hence for all $A \in \mathcal{B}^n$. This implies $h^+(T(x_1,\ldots,x_n)) = h^-(T(x_1,\ldots,x_n))$ a.e. λ^n. Therefore $h(T(x_1,\ldots,x_n)) = 0$ a.e. λ^n.
Let $B = \{t : h(t) \neq 0\}$. Then

$$Q(B) = P^n\{T^{-1}(B)\} = P^n\{\{(x_1,\ldots,x_n) : h(T(x_1,\ldots,x_n)) \neq 0\}\} = 0,$$

because P^n is absolutely contiuous with respect to λ^n.
Hence $h(t) = 0$ a.e. P^T.

(FRASER (1953))

Section 4

Problem 14.

With $C_1 = v$ and $C_2 = w$, the determining equations for v, w, γ_1 and γ_2 are:

$$F_t(v-) + 1 - F_t(w) + \gamma_1[F_t(v) - F_t(v-)] + \gamma_2[F_t(w) - F_t(w-)] = \alpha \tag{24}$$

and

$$G_t(v-) + 1 - G_t(w) + \gamma_1[G_t(v) - G_t(v-)] + \gamma_2[G_t(w) - G_t(w-)] = \alpha \tag{25}$$

where

$$F_t(u) = \int_{-\infty}^{u} C_t(\theta_1)e^{\theta_1 y} d\nu_t(y) \tag{26}$$

$$G_t(u) = \int_{-\infty}^{u} C_t(\theta_2)e^{\theta_2 y} d\nu_t(y)$$

denote the conditional cumulative distribution function of U given t when $\theta = \theta_1$ and $\theta = \theta_2$ respectively. Note that in view of Lemma 8, Chapter 2 these distributions are independent of $\theta_1,\ldots,\theta_k$.
For each $y \in [0,\alpha]$ let

$$v(y,t) = \inf\{u : F_t(u) \geq y\}$$

and

$$w(y,t) = \inf \{u : F_t(u) \geq 1-\alpha+y\},$$

where $\inf \phi = \infty$.

Define $\gamma_1(y,t)$ and $\gamma_2(y,t)$ so that for $v = v(y,t)$ and $w = w(y,t)$

$$F_t(v-) + \gamma_1[F_t(v) - F_t(v-)] = y \tag{27}$$

and

$$1-F_t(w) + \gamma_2[F_t(w) - F_t(w-)] = \alpha-y \tag{28}$$

and such that $\gamma_1 = 0$ and $\gamma_2 = 0$, respectively, if $F_t(v) = F_t(v-)$ and $F_t(w) = F_t(w-)$, respectively.

Let $H(y,t)$ denote the left hand-side of (25) with $v = v(y,t)$ etc.

By Theorem 1 (ii) of Chapter 3 and (28) it follows that the test

$$\phi_t(u) = \begin{cases} 1 & > \\ \gamma_2(0,t) \text{ when } C_t(\theta_2)e^{\theta_2 u} & = k(0,t)C_t(\theta_1)e^{\theta_1 u} \\ 0 & < \end{cases}$$

or, equivalently,

$$\phi_t(u) = \begin{cases} 1 & > \\ \gamma_2(0,t) \text{ when } u & = w(0,t) \\ 0 & < \end{cases}$$

is most powerful at level α for testing F_t against G_t. Hence by Corollary 1 on p. 67 it holds that $H(0,t) > \alpha$.

Now let $\alpha^* = H(\alpha,t)$. As in the previous case it follows by Theorem 1 (ii) of Chapter 3 that the test

$$\phi_t^*(u) = \begin{cases} 1 & < \\ \gamma_1(\alpha,t) \text{ when } u & = v(\alpha,t) \\ 0 & > \end{cases}$$

is most powerful at level α^* for testing G_t against F_t. Hence (27) and Corollary 1 on p. 67 imply that $H(\alpha,t) = \alpha^* < \alpha$.

Define $H_1(y,t)$ and $H_2(y,t)$ by

$$H_1(y,t) = G_t(v-) + \gamma_1[G_t(v) - G_t(v-)]$$

and

$$H_2(y,t) = 1 - G_t(w) + \gamma_2[G_t(w) - G_t(w-)],$$

where $v = v(y,t)$ etc.

We shall prove that, for fixed t, $H_1(y,t)$ is a continuous function of y. Since $G_t(x-)$ as a function of x is left continuous and since $v(y,t)$ as a function of y is left continuous and nondecreasing, $G_t(v(y,t)-)$ is left continuous. The left-continuity of $F_t(v(y,t)-)$ can be shown in the same way and hence $\gamma_1(y,t)[F_t(v(y,t)) - F_t(v(y,t)-)]$ is left continuous, since by (27) it is the difference of two left continuous functions. Finally, since

$$G_t(v(y,t)) - G_t(v(y,t)-) = \tag{29}$$

$$= \frac{C_t(\theta_2)}{C_t(\theta_1)} e^{(\theta_2-\theta_1)v(y,t)}[F_t(v(y,t)) - F_t(v(y,t)-)]$$

the proof of the left-continuity of $H_1(y,t)$ is complete. As to the proof of the right-continuity, we remark that for $v(y,t)$ we always have one of the following three situations:

(a) There exists a $\delta > 0$ such that $v(y,t) = v(y+\delta,t)$

Then $F_t(v(y,t)-) \leq y < y+\delta \leq F_t(v(y,t))$. Hence from (29) we get

$$G_t(v(y,t)-) < G_t(v(y,t)). \tag{30}$$

Further remark that (cf. (27)) for $h < \delta$

$$[\gamma_1(y+h,t) - \gamma_1(y,t)][F_t(v(y,t)) - F_t(v(y,t)-)] = h;$$

and since $F_t(v(y,t)-) < F_t(v(y,t))$ it follows that

$$\lim_{h\downarrow 0} \gamma_1(y+h,t) = \gamma_1(y,t). \tag{31}$$

Finally, since for $h < \delta$, the difference $H_1(y+h,t) - H_1(y,t)$ reduces to

$$[\gamma_1(y+h,t) - \gamma_1(y,t)][G_t(v(y,t)) - G_t(v(y,t)-)]$$

the right-continuity (i.e. $\lim_{h\downarrow 0} H_1(y+h,t) = H_1(y,t)$)follows from (30) and (31).

(b) $v(y,t) \le v(y+,t) < v(y+\delta,t)$ for all $\delta > 0$. In this case

$$\limsup_{h\downarrow 0} [G_t(v(y+h,t)) - G_t(v(y+h,t)-)] = 0,$$

because otherwise G_t has infinitely many jumps $\ge \varepsilon > 0$, which is impossible. Therefore

$$\begin{aligned}
&\lim_{h\downarrow 0} [H_1(y+h,t) - H_1(y,t)] \\
&= \lim_{h\downarrow 0} G_t(v(y+h,t)) - G_t(v(y,t)-) - \\
&\qquad\qquad \gamma_1(y,t)[G_t(v(y,t)) - G_t(v(y,t)-)] \\
&= [1 - \gamma_1(y,t)][G_t(v(y,t) - G_t(v(y,t)-)].
\end{aligned}$$

If $G_t(v(y,t)) > G_t(v(y,t)-)$, then by (29) also $F_t(v(y,t)) > F_t(v(y,t)-)$; moreover $F_t(v(y,t)) = y$ and hence by (27) it follows that $\gamma_1(y,t) = 1$.

(c) There exists a $\delta > 0$ such that $v(y,t) < v(y+,t) = v(y+\delta,t)$. For $h < \delta$ we have $F_t(v(y+h,t)) = F_t(v(y+,t)) > F_t(v(y,t)) = F_t(v(y+h,t)-) = y$ and hence, by (27), $\lim_{h\downarrow 0} \gamma_1(y+h,t) = 0$. Therefore (cf. (b))

$$\lim_{h\downarrow 0} H_1(y+h,t) - H_1(y,t) = [1-\gamma_1(y,t)][G_t(v(y,t)) - G_t(v(y,t)-)] = 0.$$

Similarly one can show that $H_2(y,t)$ is both left- and right-continuous. To make the analogy with the previous case ($H_1(y,t)$) clear, it has to be noted that

$$1 - H_2(y,t) = G_t(w-) + (1-\gamma_2)[G_t(w) - G_t(w-)]$$

and that (28) is equivalent with

$$F_t(w-) + (1-\gamma_2)[F_t(w) - F_t(w-)] = 1 - \alpha + y$$

where $w = w(y,t)$ etc.

Consequently $H(y,t) = H_1(y,t) + H_2(y,t)$ is a continuous function of y for fixed t.

It follows that $H(y,t) \ge c$ if and only if for each n there exists a rational number r such that $y - n^{-1} < r < y + n^{-1}$ and $H(r,t) > c - n^{-1}$.

Therefore if the rationals are denoted by $r_1, r_2, \ldots$

$$\{(y,t) : H(y,t) \geq c\} = \bigcap_n \bigcup_i \{(y,t) : -n^{-1} < r_i - y < n^{-1}, H(r_i,t) > c - n^{-1}\}. \tag{32}$$

Since $v(y,t) \leq u \Leftrightarrow y \leq F_t(u)$ and $w(y,t) \leq u \Leftrightarrow 1-\alpha+y \leq F_t(u)$ and since $F_t(u) = E_{\theta_1}\{I_{\{U \leq u, T \in \mathbb{R}\}} | T = t\}$ is a measurable function of t for each fixed u, $v(y,t)$ and $w(y,t)$ are measurable functions of t for each fixed y. Moreover γ_1 and γ_2 are measurable functions of t for each fixed y by (27) and (28). Hence $H(r_i,t)$ is measurable in t for each i. Now formula (32) shows that $H(y,t)$ is jointly measurable in y and t. Define

$$y(t) = \inf \{y : H(y,t) < \alpha\}.$$

Note that $H(0,t) > \alpha > H(\alpha,t)$ and that $H(y,t)$ is a continuous function of y for fixed t. Then (24) is satisfied with $v = v(y(t),t)$ etc. By the definitions of v and w and by (27) and (28), (25) is satisfied, since $H(y(t),t) = \alpha$. The measurability of $C_1(t) = v(y(t),t)$, $C_2(t) = w(y(t),t)$, $\gamma_1(t) = \gamma_1(y(t),t)$ and $\gamma_2(t) = \gamma_2(y(t),t)$ now follows from the following relations, which hold for all real c,

$$\{t : y(t) < c\} = \bigcup_{r<c} \{t : H(r,t) < \alpha\}$$

where r denotes a rational number,

$$\{t : v(y(t),t) \leq c\} = \{t : y(t) \leq F_t(c)\} = \{t : y(t) - F_t(c) \leq 0\},$$

$$\{t : w(y(t),t) \leq c\} = \{t : 1-\alpha+y(t) \leq F_t(c)\} = \{t : y(t) - F_t(c) \leq \alpha-1\}$$

and from (27) and (28), the defining equations of $\gamma_1(t)$ and $\gamma_2(t)$. Hence the function ϕ_3 defined in Section 4.4 by (16) and (17) in the sense as indicated above is jointly measurable in u and t.

Problem 15.

The solution of this problem is essentially the same as that outlined in the preceding problem. Therefore the comments upon the different steps in the solution as well as the references to the text are reduced to a minimum.

With $C_1 = v$ and $C_2 = w$, the determining equations for v, w, γ_1 and γ_2 are:

(33) $$F_t(v-) + 1 - F_t(w) + \gamma_1[F_t(v) - F_t(v-)] + \gamma_2[F_t(w) - F_t(w-)] = \alpha$$

and

(34) $$\int_{u<v} u dF_t(u) + \int_{u>w} u dF_t(u) + \gamma_1 v[F_t(v) - F_t(v-)] + \gamma_2 w[F_t(w) - F_t(w-)]$$
$$= \alpha \int u dF_t(u).$$

Since $F_t(u)$ belongs to an exponential family $\int |u| dF_t(u) < \infty$ (see Theorem 9, Chapter 2).
For each $y \in [0,\alpha]$ let

$$v(y,t) = \inf \{u : F_t(u) \geq y\}$$

and

$$w(y,t) = \inf \{u : F_t(u) \geq 1-\alpha+y\},$$

where $\inf \phi = \infty$.
Define $\gamma_1(y,t)$ and $\gamma_2(y,t)$ so that for $v = v(y,t)$ and $w = w(y,t)$

(35) $$F_t(v-) + \gamma_1[F_t(v) - F_t(v-)] = y,$$

(36) $$1 - F_t(w) + \gamma_2[F_t(w) - F_t(w-)] = \alpha-y$$

and such that $\gamma_1 = 0$ if $F_t(v) = F_t(v-)$ and $\gamma_2 = 0$ if $F_t(w) = F_t(w-)$.
Let $H(y,t)$ denote the left-hand side of (34) with $v = v(y,t)$ etc.
Since $v(0,t) = -\infty$ and $F(w(\alpha,t)) = 1$, it follows from Problem 3.18 that

$$H(0,t) = \int_{u>w(0,t)} u dF_t(u) + \gamma_2(0,t)w(0,t)[F_t(w(0,t)) - F_t(w(0,t)-)]$$
$$> \alpha \int u dF_t(u)$$

and that $H(\alpha,t) < \alpha \int u dF_t(u)$.
Define $H_1(y,t)$ and $H_2(y,t)$ by

$$H_1(y,t) = \int_{u<v(y,t)} u dF_t(u) + \gamma_1(y,t)v(y,t)[F_t(v(y,t)) - F_t(v(y,t)-)],$$

and

$$H_2(y,t) = \int_{u>w(y,t)} u dF_t(u) + \gamma_2(y,t)w(y,t)[F_t(w(y,t)) - F_t(w(y,t)-)].$$

We shall prove that, for t fixed, $H_1(y,t)$ is a continuous function of y. Since both $\int_{u<x} u dF_t(u)$ as a function of x and v(y,t) as a function of y are left-continuous and v(y,t) is non-decreasing, $\int_{u<v(y,t)} u dF_t(u)$ is left-continuous. Furthermore the left-continuity of $F_t(x-)$ implies the left-continuity of $F_t(v(y,t)-)$, and in view of (35) the same is true for $\gamma_1(y,t)[F_t(v(y,t)) - F_t(v(y,t)-)]$, and therefore also for $H_1(y,t)$. To prove that $H_1(y,t)$ is right-continuous, we consider three cases:

(a) There exists a $\delta > 0$ such that $v(y,t) = v(y+,t)$.

Then $F_t(v(y,t)-) \le y < y+h \le F_t(v(y,t))$ and

$$\lim_{h\downarrow 0} H_1(y+h,t) - H_1(y,t) =$$

$$\lim_{h\downarrow 0} [\gamma_1(y+h,t) - \gamma_1(y,t)]v(y,t)[F_t(v(y,t)) - F_t(v(y,t)-)] = 0$$

since in this case $\lim_{h\downarrow 0} \gamma_1(y+h,t) = \gamma_1(y,t)$.

(b) $v(y,t) \le v(y+h,t) < v(y+\delta,t)$ for all $\delta > 0$. In this case $\limsup_{h\downarrow 0} [F_t(v(y+h,t)) - F_t(v(y+h,t)-)] = 0$, because otherwise F_t has infinitely many jumps $\ge \varepsilon > 0$, which is impossible. Therefore

$$\lim_{h\downarrow 0} H_1(y+h,t) - H_1(y,t) =$$

$$\lim_{h\downarrow 0} \int_{[v(y,t),v(y+h,t))} u dF_t(u)$$

$$- \gamma_1(y,t)v(y,t)[F_t(v(y,t)) - F_t(v(y,t)-)].$$

Since $F_t(v(y+,t)-) - F_t(v(y,t)) = 0$

$$\lim_{h\downarrow 0} H_1(y+h,t) - H_1(y,t) =$$

$$= (1-\gamma_1(y,t))v(y,t)[F_t(v(y,t)) - F_t(v(y,t)-)].$$

Because $\gamma_1(y,t) = 1$ if $F_t(v(y,t)) > F_t(v(y,t)-)$, it follows that also in this case $H_1(y,t)$ is right-continuous.

(c) There exists a $\delta > 0$ such that $v(y,t) < v(y+,t) = v(y+\delta,t)$.

Then $\lim_{h\downarrow 0} \gamma_1(y+h,t) = 0$, and we can argue as in (b) to obtain the right-continuity of $H_1(y,t)$.

Since $v(y,t)$ always satisfies one of the three cases (a), (b) or (c), the continuity of $H_1(y,t)$ follows.
Next it has to be noted that

$$\int_{-\infty}^{+\infty} u dF_t(u) - H_2(y,t)$$

$$= \int_{u<w(y,t)} u dF_t(u) + (1-\gamma_2(y,t))w(y,t)[F_t(w(y,t)) - F_t(w(y,t)-)]$$

and that (36) is equivalent to

$$F_t(w(y,t)-) + (1-\gamma_2(y,t))[F_t(w(y,t)) - F_t(w(y,t)-)] = 1-\alpha+y.$$

Comparison of these equations with (35) and the definition of H_1, shows that the same arguments as in the proof of the continuity of $H_1(y,t)$ apply, yielding the continuity of $H_2(y,t)$ as a function of y for fixed t. Just as in Problem 14 it now follows that $H(y,t)$ is jointly measurable in y and t. Define $y(t) = \inf\{y : H(y,t) < \alpha \int u dF_t(u)\}$, and let $v(t) = v(y(t),t)$ etc. Then (33) and (34) are satisfied for all t. Furthermore as in Problem 14 $y(t)$, $v(t)$ and $w(t)$ are measurable, and in view of (35) and (36) $\gamma_1(t)$ and $\gamma_2(t)$ are measurable. Finally the argument of the hint shows that $\int_{(-\infty,z)} u dF_t(u)$ is measurable in z and t. Hence the function ϕ_4 defined in Section 4.4 by (16), (18) and (19) in the sense as indicated above is jointly measurable in u and t.

Problem 16.

Let $\phi_i(u,t)$, $i = 1,\dots,4$ be the UMP unbiased tests of the hypotheses $H_1,\dots,H_4$ of Theorem 3. Suppose $\phi_i^*(u,t)$ is another UMP unbiased test of H_i, $i = 1,\dots,4$. Then

$$E_{\theta,\vartheta}[\phi_i(U,T) - \phi_i^*(U,T)] = 0 \quad \text{for all } (\theta,\vartheta) \in K_i.$$

The parameter space Ω is convex and it has dimension k+1. Moreover, there are points in Ω with $\theta <$ as well as points in Ω with $\theta > \theta_0,\theta_1$ and θ_2 respectively, implying that the family

$$P_i = \{P_{\theta,\vartheta}^{U,T} : (\theta,\vartheta) \in K_i\}$$

is complete. Hence $\phi_i^*(u,t) = \phi_i(u,t)$ a.e. P_i or equivalently $\phi_i^*(u,t) = \phi_i(u,t)$ a.e. ν.

Section 5

Problem 17.

As indicated in the hint, the power of the UMP unbiased test under consideration is given by

$$\beta = \sum_{t=0}^{\infty} \beta(t) \frac{(\lambda+\mu)^t}{t!} e^{-(\lambda+\mu)}$$

where $\beta(t)$ is the power of the conditional test given t against the alternative in question.
Remark that, since $0 \le \beta(t) \le 1$, $t = 0,1,2,\ldots$, we have for each $T \in \mathbb{N}$

$$\left|\beta - \sum_{t=0}^{T} \beta(t) \frac{(\lambda+\mu)^t}{t!} e^{-(\lambda+\mu)}\right|$$

$$= \left|\sum_{t=T+1}^{\infty} \beta(t) \frac{(\lambda+\mu)^t}{t!} e^{-(\lambda+\mu)}\right|$$

$$\le \sum_{t=T+1}^{\infty} \frac{(\lambda+\mu)^t}{t!} e^{-(\lambda+\mu)} \ .$$

Hence the truncation error in β is bounded by a Poisson tail probability. For each (λ,μ) T is chosen large enough to ensure the truncation error to be less that 10^{-3}.
The results are summarized in the following table ($\alpha = .1$)

(λ,μ)	β
(.1 , .2)	.110
(1 , 2)	.217
(10 , 20)	.710
(.1 , .4)	.135

These results illustrate the discussion on p. 142, concerning the ratio $\rho = \frac{\mu}{\lambda}$ as a measure of the extent to which the two Poisson populations differ (the first three (λ,μ) have the same ratio).

Problem 18.

Consider two (independent) sequences of binomial trials $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_m$ with probabilities of success p_1 and p_2 respectively and let

$$\rho = \frac{p_2\, q_1}{p_1\, q_2} \quad (q_i = 1 - p_i,\ i = 1,2), \quad X = \sum_{i=1}^{n} X_i \quad \text{and} \quad Y = \sum_{i=1}^{m} Y_i .$$

(i) Let $\phi(X_1,\ldots,X_n,Y_1,\ldots,Y_m)$ be a level α test for testing $H : \rho = \rho_0$ against $K : \rho = \rho_1$. For each p_1,p_2 satisfying $\rho = \rho_0$

$$\begin{aligned}\alpha &\geq E_{p_1,p_2}\phi(X_1,\ldots,X_n,Y_1,\ldots,Y_m)\\ &= \phi(0,\ldots,0)P\{X+Y=0\} + E_{p_1,p_2}[\phi(X_1,\ldots,X_n,Y_1,\ldots,Y_m)I_{\{X+Y>0\}}]\\ &= \phi(0,\ldots,0)q_1^n q_2^n + f(p_1,p_2)\end{aligned}$$

with $0 \leq f(p_1,p_2) \leq 1 - q_1^n q_2^n$ (we write P_{p_1,p_2} as P). Since $\frac{p_2\, q_1}{p_1\, q_2} = \rho_0$, under the hypothesis, $p_1 \to 0$ inplies $p_2 \to 0$, and hence $\phi(0,\ldots,0) \leq \alpha$. Analogously for each p_1,p_2 satisfying $\rho_1 = \frac{p_2\, q_1}{p_1\, q_2}$ we have

$$\begin{aligned}\beta(p_1,p_2) &= E_{p_1,p_2}\phi(X_1,\ldots,X_n,Y_1,\ldots,Y_m)\\ &= \phi(0,\ldots,0)q_1^n q_2^n + f(p_1,p_2).\end{aligned}$$

Now let $p_1 \to 0$ then $\beta(p_1,p_2) \to \phi(0,\ldots,0) \leq \alpha$ and hence no test ϕ exists which has power $\geq \beta > \alpha$ against all alternatives p_1,p_2 with $\rho = \rho_1$.

(ii) Let $M_1,M_2,M_3,\ldots$ denote the consecutive indices M for which $X_M \neq Y_M$ and let $\varepsilon_1,\varepsilon_2,\varepsilon_3,\ldots \in \{0,1\}$. Then

$$\begin{aligned}&P\{X_{M_1} = \varepsilon_1,\ldots,X_{M_k} = \varepsilon_k\} =\\ &\sum_{m_1=1}^{\infty}\ \sum_{m_2=m_1+1}^{\infty} \cdots \sum_{m_k=m_{k-1}+1}^{\infty} P\{X_{M_1}=\varepsilon_1,\ldots,X_{M_k}=\varepsilon_k, M_1=m_1,\ldots,M_k=m_k\} =\\ &\sum_{m_1=1}^{\infty}\ \sum_{m_2=m_1+1}^{\infty} \cdots \sum_{m_k=m_{k-1}+1}^{\infty} P\{X_{m_1}=\varepsilon_1,\ldots,X_{M_k}=\varepsilon_k, M_1=m_1,\ldots,M_k=m_k\} =\\ &\sum_{m_1=1}^{\infty}\ \sum_{m_2=m_1+1}^{\infty} \cdots \sum_{m_k=m_{k-1}+1}^{\infty} (p_1p_2+q_1q_2)^{m_1-1}P\{X_{m_1}=\varepsilon_1, Y_{m_1}=1-\varepsilon_1\}\cdot\\ &\cdot(p_1p_2+q_1q_2)^{m_2-m_1-1}P\{X_{m_2}=\varepsilon_2, Y_{m_2}=1-\varepsilon_2\}\ldots\\ &\cdot(p_1p_2+q_1q_2)^{m_k-m_{k-1}-1}P\{X_{m_k}=\varepsilon_k, Y_{m_k}=1-\varepsilon_k\} =\end{aligned}$$

$$\frac{\prod_{i=1}^{k} P\{X_1=\varepsilon_i, Y_1 \neq X_1\}}{(1-(p_1p_2+q_1q_2))^k} = \prod_{i=1}^{k} P\{X_1=\varepsilon_i \mid X_1 \neq Y_1\}$$

where (X,Y) has the same distribution as (X_i,Y_i), $i = 1,2,\ldots$. By summation over all ε_ℓ, $\ell = 1,\ldots,i-1,i+1,\ldots,k$ it follows that $X_{M_1}, X_{M_2},\ldots$ are independent with distribution

$$P\{X_{M_i} = 1\} = 1 - P\{X_{M_i} = 0\} = \frac{p_1q_2}{p_1q_2+p_2q_1} = \frac{1}{1+\rho}\ .$$

Now we restrict attention to $X_{M_1}, X_{M_2},\ldots$. Experimentation is continued until N (N being a given integer) such pairs are available. From the preceding it follows that $X_{M_1} + \ldots + X_{M_N}$ is $B(N,\frac{1}{1+\rho})$ distributed. Testing $H : \rho = \rho_0$ against $K : \rho = \rho_1$ is therefore equivalent to the testing problem $H : p = p_0$ against $K : p = p_1$ for a binomial distribution. For this situation tests of arbitrary high power can be obtained by choosing N large enough.

(iii) is immediate by Example 9, Chapter 3, since $X_{M_1}, X_{M_2},\ldots$ are independent random variables (see (ii)).

(WALD (1947))

Section 6

Problem 19.

(i) By stationarity we have

$$\pi_1 = P\{X_1 = 1\} = P\{X_i = 1\}$$

$$= \sum_{k=0}^{1} P\{X_{i-1} = k\}P\{X_i = 1 \mid X_{i-1} = k\} = (1-\pi_1)p_0 + \pi_1 p_1.$$

Hence

$$\pi_1 = p_0 / (p_0 + q_1)$$

$$\pi_0 = 1 - \pi_1 = q_1 / (p_0 + q_1).$$

(ii) Let $P(x_1,\ldots,x_N)$ denote the probability of any particular sequence of outcomes $(x_1,\ldots,x_N)$, then we have to prove that

(37) $$P(x_1,\ldots,x_N) = (p_0 + q_1)^{-1} p_0^v p_1^{n-v} q_1^u q_0^{m-u}$$

where m and n denote the number of zeros and ones and u and v the number of runs of zeros and ones in the sequence $(x_1,\ldots,x_N)$. The proof is by induction over N. For N = 1 it follows immediately by (i). Now suppose that (37) holds for N, and consider the sequence $(x_1,x_2,\ldots,x_{N+1})$. Then the following four cases can occur.

(a) $x_N = 1$, $x_{N+1} = 1$

$$P(x_1,\ldots,x_{N-1},1,1) = P\{X_{N+1} = 1 \mid x_1,\ldots,x_{N-1},1\}P(x_1,\ldots,x_{N-1},1)$$
$$= p_1 \cdot (p_0 + q_1)^{-1} p_0^v p_1^{(n-1)-v} q_1^u q_0^{m-u}.$$

The cases

(b) $x_N = 1$, $x_{N+1} = 0$

(c) $x_N = 0$, $x_{N+1} = 1$

(d) $x_N = 0$, $x_{N+1} = 0$

are treated similarly.

(DAVID (1947))

Problem 20.

(i) The most powerful similar test.

Let $p_0^* < p_1^*$ be a fixed alternative. Consider the testing problem $H : p = p_0$ against $K^* : p_0^* < p_1^*$. Under H, independence of the sequence $X_1,\ldots,X_N$, we have that

$$P_{p_0,p_0}\{X_1 = x_1,\ldots,X_N = x_N\} = p_0^{N-m}(1-p_0)^m = p_0^N e^{m \log \frac{1-p_0}{p_0}}$$

where m is the value of the random variable M, denoting the number of zeros. Hence M is a sufficient statistic for

$$\{P_{(p_0,p_0)}^{X_1,\ldots,X_N} : 0 < p_0 < 1\}.$$

Moreover M is complete by Example 3 p. 131. Sufficiency and completeness of M under H, together with an application of Theorem 2 p. 134 make the following equivalences immediate:

a test φ is similar

$\Leftrightarrow$

$E_{p_0,p_0}\varphi(X_1,\ldots,X_N) = \alpha$ for all $0 < p_0 < 1$

$\Leftrightarrow$

$E_{p_0,p_0}E\{\varphi(X_1,\ldots,X_N) \mid M\} = \alpha$ for all $0 < p_0 < 1$

$\Leftrightarrow$

$E\{\varphi(X_1,\ldots,X_N) \mid M\} = \alpha$ a.e.

So conditionally on M a simple hypothesis has to be tested against a simple alternative. Therefore, a test

$$\varphi^*(u,v,m) = \begin{cases} 1 & > \\ \gamma^* \text{ if } P_{(p_0^*,p_1^*)}\{X_1=x_1,\ldots,X_N=x_N|M=m\} & = kP_H\{X_1=x_1,\ldots,X_N=x_N|M=m\} \\ 0 & < \end{cases}$$

or, equivalently,

$$\varphi^*(u,v,m) = \begin{cases} 1 & > \\ \gamma^* \text{ if } (p_0^*+q_1^*)^{-1}p_0^{*v}p_1^{*n-v}q_1^{*u}q_o^{*m-u} & = \tilde{k} \\ 0 & < \end{cases}$$

or, equivalently, with $\Delta(u,v) = (p_0^*/p_1^*)^v(q_1^*/q_0^*)^u$

$$\varphi^*(u,v,m) = \begin{cases} 1 & > \\ \gamma^* \text{ if } \Delta(u,v) & = k^* \\ 0 & < \end{cases}$$

where $\gamma^* = \gamma^*(p_0^*,p_1^*,m) \in [0,1)$ and $k^* = k^*(p_0^*,p_1^*,m)$ satisfy $P_H\{\Delta(U,V) > k^* \mid M=m\} + \gamma^*P_H\{\Delta(U,V) = k^* \mid M=m\} = \alpha$, is a most powerful similar test. Also note that k^* can be chosen as one of the possible values of $\Delta(u,v)$ and the for all possible pairs u and v $|u-v| \le 1$. Therefore from now on $k^* = \Delta(u,v)$ for some u,v with $|u-v| \le 1$.

The run test.

As to the run test we first remark that the conditional distribution of R, given M, is independent of p_0 (an explicit expression of this distribution is derived in (iii) of the problem). Now define the integer $C(m) \ge 0$ by

$$P_H\{R<C(m) \mid M=m\} < \alpha \le P_H\{R<C(m) \mid M=m\}$$

and $\gamma(m) \in [0,1)$ by

$$P_H\{R<C(m) \mid M=m\} + \gamma(m)P_H\{R=C(m) \mid M=m\} = \alpha; \tag{38}$$

then the run test φ is defined as follows: $(r=u+v)$

$$\varphi(u,v,m) = \begin{cases} 1 & < \\ \gamma(m) & \text{if} \quad r = C(m) \\ 0 & > \end{cases}$$

Comparison of the run test and the most powerful similar test

As a consequence of $p_0^* < p_1^*$ (and $q_1^* < q_0^*$) the function Δ has the following properties:

$\Delta(i,i)$ is stricly decreasing in i

$\Delta(i-1,i) > \Delta(i,i) > \Delta(i+1,i)$

$\Delta(i,i-1) > \Delta(i,i) > \Delta(i,i+1)$

Moreover,

$$R < 2r \Leftrightarrow \Delta(U,V) > \Delta(r,r).$$

This can be shown, using the properties of Δ, as follows. Suppose $R < 2r$. If $U = V$, then $U < r$ and hence $\Delta(U,V) = \Delta(U,U) > \Delta(r,r)$. If $U = V-1$, then $V \le r$ and $\Delta(U,V) = \Delta(V-1,V) > \Delta(V,V) \ge \Delta(r,r)$. If $V = U-1$, then $U \le r$ and $\Delta(U,V) = \Delta(U,U-1) > \Delta(U,U) \ge \Delta(r,r)$. This proves the '$\Rightarrow$' part.
Now suppose $\Delta(U,V) > \Delta(r,r)$. If $U = V$, then $\Delta(U,U) > \Delta(r,r)$ and hence $U \le r-1$ implying $R \le 2r-2 < 2r$. If $U = V+1$, then $\Delta(r,r) < \Delta(U,V) = \Delta(V+1,V) < \Delta(V,V)$ and hence $V \le r-1$ implying $R \le 2r-1 < 2r$. If $U = V-1$, then $\Delta(r,r) < \Delta(U,V) = \Delta(U,U+1) < \Delta(U,U)$ and hence $U \le r-1$ implying $R \le 2r-1 < 2r$. This proves the '$\Leftarrow$' part.
Further we have that $R = 2r$ iff $U = V = r$ iff $\Delta(U,U) = \Delta(r,r)$. We now can compare the tests φ^* and φ.
As noted before k^* equals $\Delta(r,r)$, $\Delta(r,r-1)$ or $\Delta(r-1,r)$ for some r.
If $k^* = \Delta(r,r)$ then φ and φ^* coincide.
If $k^* = \Delta(r,r-1)$ then $C(m) = 2r-1$. This can be seen as follows.
Suppose $C(m) \ge 2r$, then $P_H\{R < C(m) \mid M=m\} \ge P_H\{R < 2r \mid M=m\}$ $= P_H\{\Delta(U,V) > \Delta(r,r) \mid M=m\} \ge P_H\{\Delta(U,V) \ge k^* \mid M=m\} > \alpha$, because $P\{\Delta(U,V) = k^*\} > 0$ and $\gamma^* < 1$. This contradicts (38).
Suppose $C(m) \le 2r-2$. Then, since $P_H\{R = C(m)\} > 0$ and $\gamma < 1$,
$P_H\{R < C(m) \mid M=m\} + \gamma(m)P_H\{R = C(m) \mid M=m\} < P_H\{R \le C(m) \mid M=m\}$
$\le P_H\{R \le 2r-2 \mid M=m\} = P_H\{\Delta(U,V) \ge \Delta(r-1,r-1) \mid M=m\}$
$\le P_H\{\Delta(U,V) > k^* \mid M=m\} \le \alpha$, again in contradiction with (38).
Therefore, if the run test rejects H with probability one, i.e. if

$R < C(m) = 2r-1$, then $R \le 2r-2$ and hance $\Delta(U,V) \ge \Delta(r-1,r-1) > \Delta(r,r-1) = k^*$. Hence the most powerful similar test φ^* also rejects H with probability one. The ordering of $\Delta(r,r-1)$ and $\Delta(r-1,r)$ depends on the alternative under consideration. Hence only the supplementary rule for bringing the conditional probability of rejection (given m) up to α depends on the specific alternative.

The case $k^* = \Delta(r-1,r)$ is analogous to the previous one.

(ii) Let $p_0^* < p_1^*$ be a fixed alternative. The power of the run test φ is given by

$$E_{p_0^*,p_1^*}\varphi = E_{p_0^*,p_1^*}\beta(p_0^*,p_1^* \mid M),$$

where

$$\begin{aligned}&\beta(p_0^*,p_1^* \mid M=m) = \\ &= (1-\gamma(m))P_{p_0^*,p_1^*}\{R < C(m) \mid M=m\} + \gamma(m)P_{p_0^*,p_1^*}\{R \le C(m) \mid M=m\}.\end{aligned}$$

Now fix $m = m_0$ and define $\alpha_1 = P_H\{R < C(m_0) \mid M = m_0\}$. Let $\varphi_1 = \varphi_1(u,v,m)$ and $\varphi_1^* = \varphi_1^*(u,v,m)$ be the run test and the most powerful similar test of H against K^* : $p_0^* < p_1^*$ with level α_1. Then $\varphi_1(u,v,m) = 1$ iff $R < C(m_0)$. By (i) it follows that $R < C(m_0)$ implies $\varphi_1^*(u,v,m_0) = 1$ and hence, since $P_H\{R < C(m_0) \mid M = m\} = \alpha_1$, $\varphi_1^*(u,v,m_0) = 1$ iff $R < C(m_0)$; i.e. $\varphi_1(u,v,m_0) = \varphi_1^*(u,v,m_0)$. Therefore

$$P_{p_0^*,p_1^*}\{R < C(m_0) \mid M = m_0\} = E_{p_0^*,p_1^*}\{\varphi_1^* \mid M = m_0\} \ge \alpha_1.$$

Since m_0 was chosen arbitrarely, we have for all m

$$P_{p_0^*,p_1^*}\{R < C(m) \mid M = m\} \ge P_H\{R < C(m) \mid M = m\}.$$

Similarly one shows

$$\begin{aligned}P_{p_0^*,p_1^*}\{R \le C(m) \mid M = m\} &= P_{p_0^*,p_1^*}\{R < C(m)+1 \mid M = m\} \\ &\ge P_H\{R \le C(m) \mid M = m\}\end{aligned}$$

and hence

$$\begin{aligned}&\beta(p_0^*,p_1^* \mid M = m) \\ &\ge (1-\gamma(m))P_H\{R < C(m) \mid M=m\} + \gamma(m)P_H\{R \le C(m) \mid M=m\} = 0\end{aligned}$$

implying

$$E_{p_0^*, p_1^*}\varphi \geq \alpha$$

as was to be shown.

(iii) See the hint for $P_H\{R = 2r+1 \mid M = m\}$; in the case $R = 2r$ a similar argument yields the result.

Problem 21.

(i) The joint distribution of $Y_1,\ldots,Y_N$ is given by:

$$P\{Y_1 = y_1,\ldots,Y_N = y_N\} = \prod_{i=1}^{N} P\{Y_i = y_i\}$$

$$= \prod_{i=1}^{N} \binom{n_i}{y_i} p_i^{y_i}(1-p_i)^{n_i-y_i} = C(\alpha,\beta)\exp\left\{\sum_{i=1}^{N} y_i \log \frac{p_i}{1-p_i}\right\}d\mu(y)$$

$$= C(\alpha,\beta)\exp\{\alpha T_1(y) + \beta T_2(y)\}d\mu(y)$$

with $y = (y_1,\ldots,y_N)$, μ a suitably defined σ-finite measure, $T_1(y) = \Sigma_{i=1}^N y_i$ and $T_2(y) = \Sigma_{i=1}^N x_i y_i$.

Hence for tests concerning both α and β Theorem 3 can be applied.

(ii) By Theorem 3 the UMP unbiased test for testing $H : \beta = 0$ against the alternative $\beta > 0$ is given by

$$\phi(u,t) = \begin{cases} 1 & \text{when } u > C(t) \\ \gamma(t) & \text{when } u = C(t) \\ 0 & \text{when } u < C(t) \end{cases}$$

where the function $C(t)$ and $\gamma(t)$ are determined by

$$E_{\beta=0}[\phi(U,T) \mid t] = \alpha \quad \text{for all } t.$$

Remark that, if n denotes the number of successes in N trials,

$$U(y) = \sum_{i=1}^{N} x_i y_i = \Delta \sum_{i=1}^{N} i y_i = \Delta \sum_{j=1}^{n} s_j$$

since $y_i = 1$ when the i'th trial is a succes and $y_i = 0$ otherwise. Hence $\Sigma_{i=1}^N iy_i$ just adds the ranks of the trials at which a succes occurs.

(iii) The joint distribution of $Y_1,\ldots,Y_M$, $Z_1,\ldots,Z_N$ can be written in the form (10) of Section 4.4 with

$$\theta = \gamma-\alpha \qquad ; U(y,z) = \sum_{i=1}^{N} z_i$$

$$\vartheta_1 = \delta-\beta \qquad ; T_1(y,z) = \sum_{i=1}^{N} v_i z_i$$

$$\vartheta_2 = \alpha \qquad ; T_2(y,z) = \sum_{i=1}^{N} z_i + \sum_{i=1}^{M} y_i$$

$$\vartheta_3 = \beta \qquad ; T_3(y,z) = \sum_{i=1}^{N} v_i z_i + \sum_{i=1}^{M} u_i y_i .$$

Hence Theorem 3 can be used for the construction of tests concerning $\gamma-\alpha$. A similar argument holds for tests concerning $\delta-\beta$.

(HALDANE and SMITH (1948); KRUSKAL (1957))

REFERENCES

DAVID, F.N. (1947). A power function for tests of randomness in a sequence. *Biometrika* 34, 335-339.

FERGUSON, T.S. (1967). *Mathematical Statistics, a Decision Theoretic Approach*. Acad. Press, New York.

FRASER, D.A.S. (1953). Completeness of order statistics. *Canad. J. Math.* 6, 42-45.

GIRSHICK, M.A., F. MOSTELLER & L.J. SAVAGE (1946). Unbiased estimates for certain binomial sampling problems with application. *Ann. Math. Stat.* 17, 13-23.

HALDANE, J.B.S. & C.A.B. SMITH (1948). A simple exact test for birth-order effect. *Ann. Eugenics.* 14, 117-124.

KRUSKAL, W.H. (1957). Historical notes on the Wilcoxon unpaired two-sample test. *J. Am. Stat. Assoc.* 52, 356-360.

LEHMANN, E.L. (1947). On families of admissible tests. *Ann. Math. Stat.* 18, 97-104.

LEHMANN, E.L. (1952). Testing multiparameter hypotheses. *Ann. Math. Stat.* 23, 541-552.

RUDIN, W.R. (1970). *Real and Complex Analysis*. International Student Edition. Mc. Graw- Hill.

WALD, A. (1947). *Sequential Analysis*. Wiley, New York.

WALSH, J.E. (1949). Some significance tests for the median which are valid under very general conditions. *Ann. Math. Stat.* 20, 64-81.

CHAPTER 5

Section 2

Problem 1.

Student's t-test for $H : \xi \le 0$, $K : \xi > 0$ is given by $\varphi(X) = 1 \Leftrightarrow t(X) \ge C_0$, with $t(X) = \sqrt{n}\,\bar{X}/S$.

Let $\theta = (\xi,\sigma)$ and note that $(U,V) = (\sqrt{n}(\bar{X}-\xi)/\sigma,\ S/\sigma)$ has a joint distribution which does not depend on θ. Then

$$\beta_\varphi(\theta) = E_\theta\varphi(X) = P_\theta\{t(X) \ge C_0\} = P_\theta\{\sqrt{n}\,\bar{X}/S \ge C_o\}$$

$$= P_\theta\{\sqrt{n}(\bar{X}-\xi)/S \ge C_0 - \sqrt{n}\,\xi/S\}$$

$$= P_\theta\{\sqrt{n}(\bar{X}-\xi)/\sigma \ge C_0 S/\sigma - \sqrt{n}\,\xi/\sigma\} = P\{U \ge C_0 V - \sqrt{n}\,\xi/\sigma\}$$

$$= P\{C_0 V - U \le \sqrt{n}\,\xi/\sigma\},$$

which is increasing in ξ/σ.

Problem 2.

Let $\beta(\xi_1,\sigma)$ be the power of any level α test of H and let $\beta(\sigma)$ denote the most powerful test for testing H against $\xi = \xi_1$ when σ is known. (The existence is guaranteed by the Neyman-Pearson lemma.) Then

$$\beta(\xi_1,\sigma) \le \beta(\sigma)$$

and hence

$$\inf_\sigma \beta(\xi_1,\sigma) \le \inf_\sigma \beta(\sigma) = \alpha.$$

The last equality holds for the same reason as given in Section 2 of Chapter 5, p. 167, for when $\sigma \to \infty$ then $\beta(\sigma) \to \alpha$ by the continuity of β and the fact that $\beta(\sigma) > \alpha$ for all σ.

Problem 3.

(i) The density f_δ of the joint distribution of Z and V is

(1) $$f_\delta(z,v) = \frac{1}{\sqrt{2\pi}\; 2^{\frac{1}{2}f}\Gamma(\frac{1}{2}f)} \exp\left[-\tfrac{1}{2}(z-\delta)^2\right] v^{\frac{1}{2}f-1} \exp(-\tfrac{1}{2}v).$$

Consider the invertible transformation $t = z/\sqrt{\frac{\sigma}{f}}$, $y = v$, with Jacobian,

(2) $$\left|\frac{\partial(z,v)}{\partial(t,y)}\right| = \sqrt{\frac{y}{f}}\ .$$

The joint distribution of T and Y then has density

(3) $$g_\delta(t,y) = f_\delta\left(t\sqrt{\frac{y}{f}}, y\right)\sqrt{\frac{y}{f}} =$$
$$= \frac{1}{2^{\frac{1}{2}(f+1)}\Gamma(\frac{1}{2}f)\sqrt{\pi f}}\, y^{\frac{1}{2}(f-1)} \exp(-\tfrac{1}{2}y) \exp\left[-\tfrac{1}{2}\left(t\sqrt{\frac{y}{f}} - \delta\right)^2\right].$$

The integration $p_\delta(t) = \int_0^\infty g_\delta(t,y)dy$ gives the first result. Substitution of $v = \sqrt{(f+t^2)/f}\cdot\sqrt{y}$ leads by some elementary calculus to the second result. Finally the substitution $w = t\sqrt{\frac{y}{f}}$ yields the slightly different form

(4) $$p_\delta(t) = \frac{1}{2^{\frac{1}{2}(f-1)}\Gamma(\frac{1}{2}f)\sqrt{\pi f}} \exp\left(-\frac{1}{2}\frac{\delta f}{f+t^2}\right)\left(\frac{\sqrt{f}}{t}\right)^{f-1} \cdot$$
$$\cdot \int_0^\infty w^f \exp\left[-\frac{1}{2}\cdot\frac{f+t^2}{t^2}\left(w - \frac{\delta t^2}{f+t^2}\right)^2\right]dw.$$

(ii) Consider an orthogonal transformation $Z_i = \Sigma_{j=1}^n a_{ij}X_j$, which satisfies $a_{1i} = \frac{1}{\sqrt{n}}$, $i = 1,\ldots,n$. Then $Z_1 = \overline{X}\sqrt{n}$ and

(5) $$\sum_{i=2}^n Z_i^2 = \sum_{i=1}^n Z_i^2 - Z_1^2 = \sum_{i=1}^n X_i^2 - n\overline{X}^2 = \sum_{i=1}^n (X_i - \overline{X})^2$$

by the orthogonality of the transformation.
In Problem 6 it will be shown that $Z_1,\ldots,Z_n$ are independently normally distributed with common variance σ^2 and means $\xi_i = \xi\Sigma_{i=1}^n a_{ij}$. We have $\xi_1 = \xi\sqrt{n}$ and $\xi_i = 0$ for $i = 2,3,\ldots,n$ (by orthogonality). It follows that $\frac{1}{\sigma}Z_1$ and $\frac{1}{\sigma^2}\Sigma_{i=2}^n Z_i^2$ are independently distributed as $N(\frac{\xi}{\sigma}\sqrt{n}, 1)$ and χ^2 with (n-1) degrees of freedom respectively. Hence

$$\frac{\overline{X}\sqrt{n}}{\left\{\frac{\Sigma_{i=1}^n (X_i - \overline{X})^2}{n-1}\right\}^{\frac{1}{2}}} = \frac{\frac{1}{\sigma}Z_1}{\{\frac{1}{\sigma^2}\Sigma_{i=2}^n Z_i^2\}^{\frac{1}{2}}}$$

has a non-central t-distribution with (n-1) degrees of freedom and non-centrality parameter $\delta = \frac{\xi}{\sigma}\sqrt{n}$.

Problem 4.

With the aid of the following formulae we obtain the required powers:

I. One-sided test.

a. σ unknown,

$$\beta(\xi/\sigma) = P\{T_{n-1}(\sqrt{n}\,\xi/\sigma) \geq t_{n-1;\alpha}\},$$

where $T_{n-1}(\sqrt{n}\,\xi/\sigma)$ has a non-central t-distribution with (n-1) degrees of freedom, and non-centrality parameter $\sqrt{n}\,\xi/\sigma$, and $t_{n-1;\alpha}$ is determined by $\beta(0) = \alpha$.

b. σ known,

$$\beta^*(\xi/\sigma) = P\{U \geq u_\alpha - \sqrt{n}\,\xi/\sigma\},$$

where U has a standard normal distribution and u_α is determined by $\beta^*(0) = \alpha$.

II. Two-sided test.

a. σ unknown,

$$\beta(\xi/\sigma) = P\{T_{n-1}(\sqrt{n}\,\xi/\sigma) \geq t_{n-1;\alpha/2}\} + P\{T_{n-1}(\sqrt{n}\,\xi/\sigma) \leq -t_{n-1;\alpha/2}\}.$$

with $T_{n-1}(\sqrt{n}\,\xi/\sigma)$ and $t_{n-1,\alpha/2}$ as under I.a.

b. σ known,

$$\beta^*(\xi/\sigma) = P\{U \geq u_{\alpha/2} - \sqrt{n}\,\xi/\sigma\} + P\{U \leq -u_{\alpha/2} - \sqrt{n}\,\xi/\sigma\}.$$

with U and $u_{\alpha/2}$ as under I.b.

The following tables give the required powers. We notice that $\beta^* \geq \beta$, and that the difference between β^* and β becomes smaller as n becomes larger.

One sided test

ξ/σ \ n	5		10		15	
	β^*	β	β^*	β	β^*	β
.7	.4683	.3660	.7152	.6548	.8568	.8243
.8	.5573	.4364	.8119	.7544	.9270	.9030
.9	.6434	.5084	.8852	.8360	.9672	.9521
1.0	.7228	.5797	.9354	.8975	.9871	.9789
1.1	.7924	.6482	.9666	.9402	.9955	.9918
1.2	.8505	.7119	.9842	.9675	.9987	.9971

Two sided test

ξ/σ \ n	5		10		15	
	β^*	β	β^*	β	β^*	β
.7	.3467	.2278	.6001	.5064	.7737	.7129
.8	.4322	.2810	.7156	.6162	.8725	.8213
.9	.5210	.3393	.8122	.7172	.9365	.8995
1.0	.6088	.4014	.8854	.8031	.9721	.9491
1.1	.6914	.4656	.9356	.8708	.9893	.9764
1.2	.7653	.5302	.9667	.9203	.9964	.9906

Problem 5.

$Z_1,\ldots,Z_n$ are independently normally distributed with common variance σ^2 and means $E(Z_i) = \zeta_i$ $(i = 1,\ldots,s)$, $E(Z_i) = 0$ $(1 = s+1,\ldots,n)$.
Consider the problems of testing $H : \zeta_1 \le \zeta_1^0$ against $K : \zeta_1 > \zeta_1^0$ and $H' : \zeta_1 = \zeta_1^0$ agains $K' : \zeta_1 \ne \zeta_1^0$.
As is seen by transforming the variable Z_1 into $Z_1 - \zeta_1^0$, there is no loss of generality in assuming that $\zeta_1^0 = 0$.
The joint density of $Z_1,\ldots,Z_n$ is

$$(6) \qquad (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} z_i^2 - 2\sum_{i=1}^{s} z_i\zeta_i + \sum_{i=1}^{s} \zeta_i^2\right)\right\} .$$

We make the identification of (6) with (1) of p. 160 of the book through the correspondence

$$\theta = \zeta_1/\sigma^2; \quad \vartheta = (\zeta_2/\sigma^2,\ldots,\zeta_s/\sigma^2,\ 1/\sigma^2)' :$$

$$U(z) = z_1; \quad T(z) = (z_2,\ldots,z_s,\ \sum_{i=1}^{n} z_i^2)' = (T_1,\ldots,T_s)' .$$

Theorem 3 of Chapter 4 then shows that UMP unbiased tests exist for the hypotheses $\theta \le 0$ and $\theta = 0$, which are equivalent to $\zeta_1 \le 0$ and $\zeta_1 = 0$ respectively. Now

$$V = \frac{U}{(T_s - \Sigma_{i=1}^{s-1} T_i^2 - U^2)^{\frac{1}{2}}} = \frac{Z_1}{(\Sigma_{i=s+1}^{n} Z_i^2)^{\frac{1}{2}}}$$

is independent of T when $\zeta_1 = 0$ (θ fixed $\Rightarrow \sigma^2$ fixed $\Rightarrow$ distribution of V does not depend on $\vartheta \Rightarrow$ V is independent of T by Corollary 2).
It follows from Theorem 1 that the UMP unbiased rejection region for $H : \zeta_1 \le 0$ is $v \ge C_0'$ or equivalently

$$t(z) \ge C_0, \text{ where } t(z) = \frac{z_1}{\left(\Sigma_{i=s+1}^{n} \frac{z_i^2}{n-s}\right)^{\frac{1}{2}}} .$$

In order to apply the theorem to $H' : \zeta_1 = 0$, let

$$W = \frac{U}{(T_s - \Sigma_{i=1}^{s-1} T_i)^{\frac{1}{2}}} = \frac{Z_1}{(\Sigma_{i=s+1}^{n} Z_i^2 + Z_1^2)^{\frac{1}{2}}} .$$

Now W is also independent of T when $\zeta_1 = 0$, and is moreover linear in U. The distribution of W is symmetric about 0 when $\zeta_1 = 0$. It follows from Theorem 1 that the UMP unbiased rejection region for $H' : \zeta_1 = 0$ is $|w| \ge C'$. Since

$$t(z) = \frac{\sqrt{n-s}\, w}{\sqrt{1-w^2}} ,$$

the absolute value of t(z) is an increasing function of $|w|$, and the rejection region is thus equivalent to

$$|t(z)| = \frac{|z_1|}{\left(\Sigma_{i=s+1}^{n} \frac{z_i^2}{n-s}\right)^{\frac{1}{2}}} \ge C.$$

From the definition of t(z) it is seen that t(Z) is the ratio of two independent random variables Z_1/σ and $(\Sigma_{i=s+1}^{n} Z_i^2 / (n-s)\sigma^2)^{\frac{1}{2}}$. The denominator is distributed as the square root of a χ^2 variable with n-s degrees of freedom divided by n-s; the distribution of the numerator, when $\zeta_1 = 0$, is the standard normal distribution. It follows that the distribution of t(Z) under $\zeta_1 = 0$ is Student's t distribution with (n-s) degrees of freedom

Problem 6.

The result of this problem may be obtained as a straightforward application of a standard result from multivariate analysis. See, for instance, Theorem 2.4.1 on pp. 19-20 in ANDERSON (1958).

Problem 7.

$X_1,\ldots,X_n$ is a sample from $N(\xi,\sigma^2)$. Consider the same orthogonal transformation as in Problem 3 (ii), i.e. we have $Z_1 = \overline{X}\sqrt{n} \sim N(\sqrt{n}\,\xi,\sigma^2)$ and $Z_2,\ldots,Z_n$ are normally distributed random variables with common variance ^{2} and mean zero. Moreover, $Z_1,\ldots,Z_n$ are independent.

By Problem 5, it follows that the UMP unbiased tests for testing $\xi \le 0$ and $\xi = 0$ are given by the rejection regions

$$\frac{Z_1}{\left\{\Sigma_{i=2}^n \; Z_i^2/(n-1)\right\}^{\frac{1}{2}}} > C_0 \quad \text{and} \quad \frac{|Z_1|}{\left\{\Sigma_{i=2}^n \; Z_i^2/(n-1)\right\}^{\frac{1}{2}}} > C$$

respectively.

Using $Z_1 = \overline{X}\sqrt{n}$ and (5) in the solution to Problem 3, it follows that these rejection regions may also be written as

$$\frac{\sqrt{n}\,\overline{X}}{\left\{\Sigma_{i=1}^n \; (X_i-\overline{X})^2/(n-1)\right\}^{\frac{1}{2}}} > C_0 \quad \text{and} \quad \frac{|\sqrt{n}\,\overline{X}|}{\left\{\Sigma_{i=1}^n \; (X_i-\overline{X})^2/(n-1)\right\}^{\frac{1}{2}}} > C.$$

These results correspond to formulae (16) and (17) of Section 2.

Problem 8.

(i) The random variables $Y_1,Y_2,\ldots$ are independently distributed as $N(0,\sigma^2)$.

PROOF. For each fixed $n \ge 1$, $(Y_1,Y_2,\ldots,Y_n)'$ is a linear function of $(X_1,\ldots,X_{n+1})'$ and hence has a multivariate normal distribution which is completely specified by its expectation and covariance structure. We have

$$EY_n = \{n(n+1)\}^{-\frac{1}{2}}[nEX_{n+1} - (EX_1 + \ldots + EX_n)] = 0, \quad n = 1,2,\ldots,$$

and, for $m > n$,

$$\mathrm{cov}(Y_m,Y_n) =$$

$$= \{n(n+1)m(m+1)\}^{-\frac{1}{2}} \,\mathrm{cov}\,(mX_{m+1} - \sum_{i=1}^{m} X_i,\, nX_{n+1} + \sum_{j=1}^{n} X_j) =$$

$$= \{n(n+1)m(m+1)\}^{-\frac{1}{2}}\Big[mn\,\mathrm{cov}\,(X_{m+1},X_{n+1}) - m\,\mathrm{cov}\Big(X_{m+1}, \sum_{j=1}^{n} X_j\Big)$$

$$- n\,\mathrm{cov}\Big(X_{n+1}, \sum_{i=1}^{m} X_i\Big) + \mathrm{cov}\Big(\sum_{i=1}^{m} X_i, \sum_{j=1}^{n} X_j\Big)\Big] =$$

$$= \{n(n+1)m(m+1)\}^{-\frac{1}{2}}[0 - 0 - n\sigma^2 + n\sigma^2] = 0.$$

By symmetry of course $\mathrm{cov}(Y_m,Y_n) = 0$ for $m < n$.
Furthermore we find for $\sigma^2(Y_n)$:

$$\sigma^2(Y_n) = \mathrm{cov}(Y_n,Y_n) =$$

$$= \{n(n+1)\}^{-1}\,\mathrm{cov}\,(nX_{n+1} - \sum_{i=1}^{n} X_i,\, nX_{n+1} - \sum_{j=1}^{n} X_j) =$$

$$= \{n(n+1)\}^{-1}[n^2\sigma^2 - 0 - 0 + n\sigma^2] = \sigma^2.$$

This proves the proposition.

(ii) $H : \sigma = \sigma_0$, $K : \sigma = \sigma_1$.
Assume $\sigma_1 > \sigma_0$. The case $\sigma_1 < \sigma_0$ can be treated analogously.
The joint density of $Y_1,\ldots,Y_n$ is equal to

$$p_{in}(y) = (\sqrt{2\pi}\,\sigma_i)^{-n} \exp\Big\{-\frac{1}{2\sigma_i^2}\sum_{j=1}^{n} y_j^2\Big\} \qquad i = 0,1,$$

under H and K respectively. Therefore

$$\frac{p_{1n}(Y)}{p_{0n}(Y)} = \Big(\frac{\sigma_0}{\sigma_1}\Big)^n \exp\Big\{\tfrac{1}{2}\Big(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\Big)\sum_{j=1}^{n} Y_j^2\Big\}.$$

So we have

$$A_0 < \frac{p_{1n}(Y)}{p_{0n}(Y)} < A_1$$

$$\Leftrightarrow\; A_0\Big(\frac{\sigma_1}{\sigma_0}\Big)^n < \exp\Big\{\tfrac{1}{2}\Big(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\Big)\sum_{j=1}^{n} Y_j^2\Big\} < A_1\Big(\frac{\sigma_1}{\sigma_0}\Big)^n$$

$$\Leftrightarrow\; 2\Big(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\Big)^{-1}\log\Big(A_0\Big(\frac{\sigma_1}{\sigma_0}\Big)^n\Big) < \sum_{j=1}^{n} Y_j^2 < 2\Big(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\Big)^{-1}\log\Big(A_1\Big(\frac{\sigma_1}{\sigma_0}\Big)^n\Big).$$

H is rejected or accepted at the first violation of this equality, i.e. H is accepted when $\Sigma_{j=1}^{n} Y_j^2$ is too small and rejected when $\Sigma_{j=1}^{n} Y_j^2$ is too large. For the determination of A_0 and A_1 for given confidence level and power the reader is referred to Section 10 of Chapter 3.

(GIRSHICK (1946))

Section 3

Problem 9.

We test $H : \tau \le \sigma$ against $K : \tau > \sigma$, at level $\alpha = .05$, based on two independent samples $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ from $N(\xi,\sigma^2)$ and $N(\eta,\tau^2)$ respectively. The hypothesis H is rejected when

$$\frac{\Sigma\,(Y_i - \overline{Y})^2}{\Sigma\,(X_i - \overline{X})^2} \ge C_n,$$

where C_n is determined by

$$(7) \qquad \int_{C_n}^{\infty} F_{n-1,n-1}(y)dy = .05.$$

(cf. (20) and (22) in Section 3, p. 169).
For an alternative with $\tau\sigma^{-1} = \Delta$, the power is given by

$$P_\Delta\left\{\frac{\Sigma\,(Y_i - \overline{Y})^2}{\Sigma\,(X_i - \overline{X})^2} \ge C_n\right\} = P_\Delta\left\{\frac{\Sigma\,(Y_i - \overline{Y})^2\tau^{-2}}{\Sigma\,(X_i - \overline{X})^2\sigma^{-2}} \ge C_n\sigma^2\tau^{-2}\right\} =$$

$$= P\{V_n \ge C_n\Delta^{-2}\} = P\{V_n \le C_n^{-1}\Delta^2\} = 1 - P\{V_n \ge C_n^{-1}\Delta^2\},$$

where V_n has an F-distribution with n-1 and n-1 degrees of freedom.
Note that the power function is increasing in Δ.
To determine the sample size n necessary to obtain power $\ge .9$ against the alternatives with $\tau\sigma^{-1} > \Delta$, the number n has to satisfy (7) and

$$\int_{C_n^{-1}\Delta^2}^{\infty} F_{n-1,n-1}(y)dy \le 0.1.$$

By trial and error we find for $\Delta = 1.5$ that $n \ge 56$, for $\Delta = 2$ that $n \ge 20$ and for $\Delta = 3$ that $n \ge 9$.

Problem 10.

The test $\varphi(W)$, with

$$W = \frac{S_Y^2}{\Delta_0 S_X^2 + S_Y^2},$$

which is given by

$$\varphi(w) = \begin{cases} 0 & \text{when } C_1 \le w \le C_2, \\ 1 & \text{when } w < C_1 \text{ or } w > C_2, \end{cases}$$

and where C_1 and C_2 satisfy

$$E_{H'}[\varphi(W)] = \alpha \quad \text{and} \quad E_{H'}[W\varphi(W)] = \alpha E_{H'}[W],$$

is UMP unbiased for $H' : \frac{\tau^2}{\sigma^2} = \Delta_0$.
Under H' and if $m = n$, the distribution of W is the beta-distribution with density $B_{(\frac{1}{2}(n-1),\frac{1}{2}(n-1))}(w)$. So, in that case, the distribution of W is symmetric about $\frac{1}{2}$.
Take $C_2 = 1 - C_1$ and let C_1 satisfy $\int_0^{C_1} B_{(\frac{1}{2}(n-1),\frac{1}{2}(n-1))}(w)dw = \frac{\alpha}{2}$, then

$$E_{H'}[\varphi(W)] = \int_0^{C_1} B_{(\frac{1}{2}(n-1),\frac{1}{2}(n-1))}(w)dw + \int_{1-C_1}^{1} B_{(\frac{1}{2}(n-1),\frac{1}{2}(n-1))}(w)dw =$$

$$= 2\int_0^{C_1} B_{(\frac{1}{2}(n-1),\frac{1}{2}(n-1))}(w)dw = \alpha,$$

and

$$E_{H'}[W\varphi(W)] = E_{H'}[(W-\tfrac{1}{2})\varphi(W)] + \tfrac{1}{2}E_{H'}[\varphi(W)] = \tfrac{1}{2}\alpha = E_{H'}(W)\cdot\alpha,$$

since φ is symmetric about $\frac{1}{2}$ and the distribution of W is symmetric about $\frac{1}{2}$, so that $E_{H'}[(W-\frac{1}{2})\varphi(W)] = 0$.
Finally we have

$$\left\{C_1 \le \frac{S_Y^2}{\Delta_0 S_X^2 + S_Y^2} \le 1 - C_1\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{S_Y^2 \ge C_1\Delta_0 S_X^2 + C_1 S_Y^2 \text{ and } S_Y^2 \le (1-C_1)\Delta_0 S_X^2 + (1-C_1)S_Y^2\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{(1-C_1)S_Y^2 \le C_1\Delta_0 S_X^2 \text{ and } C_1 S_Y^2 \le (1-C_1)\Delta_0 S_X^2\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{\frac{\Delta_0 S_X^2}{S_Y^2} \le \frac{1-C_1}{C_1} \text{ and } \frac{S_Y^2}{\Delta_0 S_X^2} \le \frac{1-C_1}{C_1}\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{\max\left(\frac{\Delta_0 S_X^2}{S_Y^2}, \frac{S_Y^2}{\Delta_0 S_X^2}\right) \le \frac{1-C_1}{C_1}\right\}.$$

Problem 11.

Define $R = (X_1,\ldots,X_m,Y_1,\ldots Y_n)'$. Let A be the orthogonal $(m+n)\times(m+n)$ matrix which defines the orthogonal transformation $Z = AR$ such that

(8) $$Z_1 = \sum_{j=1}^{m} a_{1j}X_j + \sum_{j=1}^{n} a_{1,j+m}Y_j = \left\{\frac{1}{m}+\frac{1}{n}\right\}^{-\frac{1}{2}}(\overline{Y}-\overline{X}),$$

(9) $$Z_2 = \sum_{j=1}^{m} a_{2j}X_j + \sum_{j=1}^{n} a_{2,j+m}Y_j = \{m+n\}^{-\frac{1}{2}}\left(\sum_{i=1}^{m} X_i + \sum_{j=1}^{n} Y_j\right).$$

This defines the first two rows of A. The orthogonality conditions are easily checked. The matrix A may now be completed using the Gram-Schmidt method. The covariance matrix of R is equal to $D(R) = \sigma^2 I_{m+n}$ so that

$$D(Z) = D(AR) = AD(R)A' = \sigma^2 AA' = \sigma^2 I_{m+n}.$$

It follows that the components of Z are independent and have common variance σ^2. For the expectation of Z_1 and Z_2 we find

$$EZ_1 = \left\{\frac{1}{m}+\frac{1}{n}\right\}^{-\frac{1}{2}}(\eta-\xi), \quad EZ_2 = \{m+n\}^{-\frac{1}{2}}(m\xi-n\eta).$$

From (8) and (9) we obtain

(10) $$a_{1j} = -\frac{1}{m}\left\{\frac{1}{m}+\frac{1}{n}\right\}^{-\frac{1}{2}},\ j=1,\ldots,m; \quad a_{1j} = \frac{1}{n}\left\{\frac{1}{m}+\frac{1}{n}\right\}^{-\frac{1}{2}},$$
$$j = m+1,\ldots,m+n;$$

(11) $$a_{2j} = \{m+n\}^{-\frac{1}{2}},\quad j=1,\ldots,m+n.$$

Because A is orthogonal we must have for $i = 3,\ldots,m+n$,

(12) $$\sum_{j=1}^{m+n} a_{ij}a_{1j} = 0,\quad \sum_{j=1}^{m+n} a_{ij}\,a_{2j} = 0,$$

therefore, from (10) and (12), (11) and (12),

(13) $$-\frac{1}{m}\sum_{j=1}^{m} a_{ij} + \frac{1}{n}\sum_{j=m+1}^{m+n} a_{ij} = 0,$$

$$(14) \qquad \sum_{j=1}^{m} a_{ij} = - \sum_{j=m+1}^{m+n} a_{ij}.$$

So, from (13) and (14) it follows for $i = 3,4,\ldots,m+n$, that

$$(15) \qquad \sum_{j=1}^{m} a_{ij} = 0, \quad \sum_{j=m+1}^{m+n} a_{ij} = 0.$$

We find for EZ_i, $i = 3,4,\ldots,m+n$, using (15),

$$EZ_i = E\left(\sum_{j=1}^{m} a_{ij}X_j + \sum_{j=1}^{m} a_{i,j+m}Y_j\right) =$$

$$= \xi \sum_{j=1}^{m} a_{ij} + \eta \sum_{j=m+1}^{m+n} a_{ij} = 0.$$

So $Z_1,\ldots,Z_{m+n}$ satisfy the conditions of Problem 5. To test $H : \eta-\xi = 0$, the critical region is defined by

$$\frac{|Z_1|}{\sqrt{\sum_{i=3}^{m+n} Z_i^2/(m+n+2)}} > C.$$

Using $Z'Z' = R'A'AR = R'R$, $S_X^2 = \Sigma_{j=1}^m (X_j - \overline{X})^2$, $S_Y^2 = \Sigma_{j=1}^n (Y_j - \overline{Y})^2$, we may rewrite this as

$$\frac{|\overline{Y} - \overline{X}|/\sqrt{\frac{1}{m} + \frac{1}{n}}}{\sqrt{(S_X^2 + S_Y^2)/(m+n-2)}} > C.$$

The test statistic has, under H, a Student distribution with m+n-2 degrees of freedom, so C is determined by $P(|t_{m+n-2}| > C) = \alpha$.

Problem 12.

We restrict attention to the ordered variables $X^{(1)} < \ldots < X^{(n)}$ since these are sufficient for a and b, and transform to new variables $Z_1 = nX^{(1)}$, $Z_i = (n-i+1)(X^{(i)} - X^{(i-1)})$ for $i = 2,\ldots,n$ as in Problem 13 of Chapter 2. The joint density of $Z_1,Z_2,\ldots,Z_n$ is then

$$p(z_1,z_2,\ldots,z_n) = a^{-n} \exp\{-a^{-1}(z_1 - nb) - a^{-1} \sum_{i=2}^{n} z_i\},$$

$$z_1 \geq nb, \; z_2,\ldots,z_n \geq 0.$$

(i) Under the hypothesis $H : a = 1$, Z_1 is easily seen to be sufficient for b. We shall prove that Z_1 is also complete. Consider

$$E_b(f(Z_1)) = \int_{nb}^{\infty} f(z)e^{-(z-nb)}dz = e^{nb}\int_{nb}^{\infty} f(z)e^{-z}dz.$$

Now suppose that $E_b(f(Z_1)) = 0$ for all $b \in \mathbb{R}$. This means that $\int_{nb}^{\infty} f(z)e^{-z}dz$ is equal to zero for each $b \in \mathbb{R}$. This implies that $\int_I f(z)e^{-z}dz = 0$ for each interval I, from which it follows that $f(z) = 0$, except possibly on a set of Lebesgue-measure zero. This proves completeness.

Considering only unbiased tests, and thus only similar tests, by Theorem 2, Chapter 4, all considered tests φ then have Neyman-structure w.r.t. Z_1, i.e. for all such tests we have

$$E_{a=1}[\varphi(Z_1,\dots,Z_n) \mid Z_1 = z_1] = \alpha.$$

Following the reasoning of Section 3 of Chapter 4, we only have to solve the "optimum problem" on each surface $Z_1 = z_1$ separately.
The conditional distribution of $Z_2,\dots,Z_n$ given $Z_1 = z_1$ has density

(16) $$h(z_2,\dots,z_n \mid z_1) = c\exp\{-\frac{1}{a}\sum_{i=2}^{n} z_i\}.$$

from which it follows in particular that Z_1 and $Z_2,\dots,Z_n$ are independent. Moreover, (16) constitutes an exponential family for which the UMP test for the hypothesis $H : a = 1$ has critical function (cf. Section 2 of Chapter 4)

(17) $$\varphi(z_1,\dots,z_n) = \begin{cases} 1 & \text{when } \Sigma_{i=2}^{n} z_i < C_1(z_1) \text{ or } > C_2(z_2) \\ 0 & \text{when } C_1(z_1) < \Sigma_{i=2}^{n} z_i < C_2(z_1) \end{cases}.$$

Because of the above mentioned independence of Z_1, and $Z_2,\dots,Z_n$, the constants $C_1(z_1)$ and $C_2(z_1)$ are independent of z_1 and are determined by

(18) $$E_{a=1}\varphi(z_1,Z_2,\dots,Z_n) = \alpha,$$

(19) $$E_{a=1}\left[\left\{\sum_{i=2}^{n} Z_i\right\}\varphi(z_1,Z_2,\dots,Z_n)\right] = E_{a=1}\left[\sum_{i=2}^{n} Z_i\right]\alpha .$$

The test $\varphi(z_1,\dots,z_n)$ considered as an unconditional test is then UMP unbiased for $H : a = 1$.

The acceptance region derived from (17) may be rewritten as follows

$$\{C_1 < \sum_{i=2}^{n} Z_i < C_2\} \Leftrightarrow \{2C_1 < \sum_{i=2}^{n} 2Z_i < 2C_2\} \Leftrightarrow$$

$$\{2C_1 < 2 \sum_{i=2}^{n} (X_i - \min(X_1,\ldots,X_n)) < 2C_2\}.$$

By Problem 13 (ii) of Chapter 2, for a = 1, the statistics $2Z_2,\ldots,2Z_n$ are independently distributed as χ^2 with 2 degrees of freedom and hence $\Sigma_{i=2}^{n} 2Z_i$ has a χ^2 distribution with 2n-2 degrees of freedom. Hence $k_1 = 2C_1$ and $k_2 = 2C_2$ must satisfy

$$\int_{k_1}^{k_2} \chi^2_{2n-2}(y)dy = 1-\alpha \qquad \text{(i.e. 18)}$$

$$\int_{k_1}^{k_2} \tfrac{1}{2}y\chi^2_{2n-2}(y)dy = (n-1)(1-\alpha) \qquad \text{(i.e. 19)}.$$

The last equality is equivalent to (cf. Example 2 Section 2 of Chapter 4)

$$\int_{k_1}^{k_2} \chi^2_{2n}(y)dy = 1-\alpha.$$

(ii) When b = 0, the statistic $\Sigma_{i=1}^{n} Z_i$ is sufficient for a and also complete, which can be shown using the fact that for b = 0 the statistic $\frac{z}{a}\Sigma_{i=1}^{n} Z_i$ has a χ^2-distribution with 2n degrees of freedom.
We can therefore follow the same reasoning as under (i).
It is easy to show that the density of $Z_1,\ldots,Z_n$ given $\Sigma_{i=1}^{n} Z_i = c$ is a constant on the following section of the hyperplane $\Sigma_{i=1}^{n} z_i = c$:

$$\{z_1 \geq nb,\ z_2,\ldots,z_n \geq 0,\ \sum_{i=1}^{n} z_i = c\} = S_b \qquad \text{(say)}.$$

Hence $(Z_1,\ldots,Z_n \mid \Sigma_{i=1}^{n} Z_i = c)$ has a homogeneous distribution over S_b and a homogeneous distribution over S_0 when b = 0, i.e. when H holds.
Application of Neyman & Pearsons fundamental lemma to this conditional situation shows that a test with critical function:

$$\varphi(z_1,\ldots,z_n \mid \sum_{i=1}^{n} z_i = c) = \begin{cases} 1 & \text{when } z_1 > k(c) \text{ or } z_1 < 0 \\ 0 & \text{when } 0 < z_1 < k(c) \end{cases}$$

is MP for H : b = 0 against K : b = b_1,
where k(c) is determined from

$$\text{(20)} \qquad E_{b=0}\varphi(Z_1,\ldots,Z_n \mid \sum_{i=1}^{n} Z_i = c) = \alpha.$$

Because the same test is obtained for any $b_1 \neq 0$, the test is UMP for $H : b = 0$ against $K : b \neq 0$.

Now, (20) implies

$$P_{b=0}\{0 < Z_1 < k(c) \mid \sum_{i=1}^{n} Z_i = c\} = 1-\alpha.$$

This may be rewritten as

$$P_{b=0}\{0 < Z_1 / \sum_{i=1}^{n} Z_i < k(c) / c \mid \sum_{i=1}^{n} Z_i = c\} = 1-\alpha$$

or, by the independence of $\Sigma_{i=1}^{n} Z_i$ and $Z_1 / \Sigma_{i=1}^{n} Z_i$, as

$$P_{b=0}\{0 < Z_1 / \sum_{i=1}^{n} Z_i < C\} = 1-\alpha.$$

This shows that $k(c) / c = C$ does not depend on c. Hence the test with acceptance region

(21) $$0 < z_1 / \sum_{i=1}^{n} z_i < C$$

now considered as an unconditional test is UMP unbiased for $H : b = 0$ against $K : b \neq 0$.

A trivial calculation shows that (21) may be written equivalently as

$$0 < \frac{n \min(x_1,\ldots,x_n)}{\sum_{i=1}^{n} \{x_i - \min(x_1,\ldots,x_n)\}} < C' \quad (\text{i.e. } 0 < z_i / \sum_{i=2}^{n} z_i < C').$$

Furthermore, from Problem 13 (ii) of Chapter 2 we know that, when $b = 0$, $2Z_1$ and $\Sigma_{i=2}^{n} 2Z_i$ are independently distributed as χ_2^2 and χ_{2n-2}^2 respectively. Hence

$$(n-1)Z_1 / \sum_{i=2}^{n} Z_i \sim F_{2,2n-2},$$

an F-distribution with 2 and 2n-2 degrees of freedom. It follows then easily that $U = Z_1 / \Sigma_{i=2}^{n} Z_i$ has probability density

$$(n-1)(1+u)^{-n}, \qquad u \geq 0.$$

(PAULSON (1941), LEHMANN (1947))

Problem 13.

Notice that the joint density of $(X^{(1)},\dots,X^{(r)})$ is equal to

$$h(x_1,\dots,x_r) = C\cdot\exp\left\{-\frac{1}{a}\sum_{i=1}^{r}(x_i-b)-(n-r)\left(\frac{x_r-b}{a}\right)\right\},$$
$$x_1 < x_2 < \dots < x_r.$$

as follows from Problem 13 (i) of Chapter 2. Transformation to new variables $Z_1 = nX^{(1)}$, $Z_i = (n-i+1)(X^{(i)}-X^{(i-1)})$ for $i=2,\dots,r$ $(2 \le r \le n)$ as again in Problem 13 (ii) of Chapter 2 gives the joint density of $Z_1,\dots,Z_r$

$$p(z_1,\dots,z_r) = C\cdot\exp\left\{-\frac{1}{a}(z_1-nb)-\frac{1}{a}\sum_{i=2}^{r}z_i\right\}$$
$$z_1 \ge nb,\ z_2,\dots,z_r \ge 0.$$

Thus, the joint density of $Z_1,\dots,Z_r$ has the same structure as the joint density of $Z_1,\dots,Z_n$ in the preceding problem. (This result could also have been obtained by integrating out the variables $z_{r+1},\dots,z_n$ of the density of $Z_1,\dots,Z_n$ as one can see easily). Notice that this time $X^{(1)},\dots,X^{(r)}$ are (of course) sufficient for a and b, as no further information is available, and that $Z_1,\dots,Z_n$ is a one to one transformation of the $X^{(1)},\dots,X^{(r)}$.

It follows that the results of the preceding problem hold when we replace n by r throughout.

This means that the UMP unbiased test for $H : a = 1$ has acceptance region

$$(22)\qquad C_1 < \sum_{i=2}^{r} z_i < C_2$$

where the constants $k_1 = 2C_1$ and $k_2 = 2C_2$ are determined by

$$\int_{k_1}^{k_2}\chi^2_{2r-2}(y)\,dy = 1-\alpha,$$
$$\int_{k_1}^{k_2}\chi^2_{2r}(y)\,dy = 1-\alpha.$$

Expression of the critical region in terms of the original variables $X_1,\dots,X_n$, as was done in the preceding problem, is in this case of course impossible.

The critical region can however be expressed in terms of $X^{(1)},\dots,X^{(r)}$. Using the identity

$$\sum_{i=2}^{r} Z_i = (n-r)X^{(r)} - nX^{(1)} + \sum_{i=1}^{r} X^{(i)},$$

(22) may be rewritten as

$$C_1 < (n-r)x^{(r)} - nx^{(1)} + \sum_{i=1}^{r} x^{(i)} < C_2.$$

The UMP unbiased test for H : b = 0 has acceptance region

$$0 < z_1 / \sum_{i=2}^{r} z_i < C$$

where $U = Z_1 / \Sigma_{i=2}^{r} Z_i$ has, for b = 0, probability density

$$(n-1)(1+u)^{-r}, \quad u \geq 0.$$

Section 4

Problem 14.

Let $\delta : \mathbb{R}^n \to \mathbb{R}$ be any measurable function and L any constant such that $(\delta(X) - L/2, \delta(X) + L/2)$ is a confidence interval for ξ. Furthermore, let N be any integer and $\xi_1,\ldots,\xi_{2N}$ such that $|\xi_i - \xi_j| > L$ whenever $i \neq j$. Then the sets

$$S_i = \{(x_1,\ldots,x_n) \mid |\delta(x_1,\ldots,x_n) - \xi_i| \leq L/2\},\ i = 1,\ldots,2N$$

are mutually exclusive. Let $Y^{(i)} = (Y_1^{(i)},\ldots,Y_n^{(i)})$ be a sample of fixed size n from $N((\xi_i - \xi_1)/\sigma;1)$, and $f^{(i)}(y)$ the density function of $Y^{(i)}$, $i = 1,\ldots,2N$. Since $f^{(i)}(y) \to f^{(1)}(y)$ as $\sigma \to \infty$ for all i, it follows from the Lemmas 4 and 2 of the Appendix that there exists a $\sigma_N > 0$ such that for $\sigma > \sigma_N$,

$$|P_{\xi_i,\sigma}\{X \in S_i\} - P_{\xi_1,\sigma}\{X \in S_i\}| =$$

$$= |P\{(\sigma Y_1^{(i)} + \xi_1,\ldots,\sigma Y_n^{(i)} + \xi_1) \in S_i\} +$$

$$- P\{(\sigma Y_1^{(1)} + \xi_1,\ldots,\sigma Y_n^{(1)} + \xi_1) \in S_i\}| =$$

$$= |P\{Y^{(i)} \in T_i\} - P\{Y^{(1)} \in T_i\}| \leq (2N)^{-1},$$

where

$$T_i = \{(y_1,\ldots,y_n) \mid (\sigma y_1 + \xi_1,\ldots,\sigma y_n + \xi_n) \in S_i\}.$$

Since

$$\min_i P_{\xi_1,\sigma}\{X \in S_i\} \le (2N)^{-1},$$

it follows for $\sigma > \sigma_N$ that

$$\min_i P_{\xi_i,\sigma}\{X \in S_i\} \le N^{-1},$$

and hence that

$$\inf_{\xi,\sigma} P_{\xi,\sigma}\{|\delta(X) - \xi| \le L/2\} \le N^{-1}.$$

Because N is arbitrary, the confidence coefficient associated with the intervals $(\delta(x) - L/2, \delta(X) + L/2)$ is zero, and the same must be true a fortiori of any set of confidence intervals of length $\le L$.

Problem 15.

(i) mS^2/σ^2 is χ^2-distributed with m degrees of freedom and, conditional on $S = s$, Y is $N(0,\sigma^2/s^2)$ distributed. Therefore, conditional on $S = s$, $sY/\sigma = SY/\sigma$ is $N(0,1)$ distributed. Since this distribution does not depend on s, SY/σ is $N(0,1)$ distributed independently of S^2. Therefore $Y = (SY/\sigma) / \sqrt{S^2/\sigma^2}$ is t-distributed with m degrees of freedom.

(ii) The random variables $\overline{X}_0, S, X_{n_0+1}, X_{n_0+2}, \ldots$ are independent; $\overline{X}_0$ is $N(\xi,\sigma^2/n_0)$ distributed, X_i, $i \ge n_0+1$, is $N(\xi,\sigma^2)$ distributed, and $(n_0-1)S^2/\sigma^2$ is χ^2-distributed with n_0-1 degrees of freedom. Conditional on $S = s$, a, b, and n are fixed and $\overline{X}_0, X_{n_0+1}, X_{n_0+2}, \ldots$ are still independent with the same normal distributions. Since

$$\sum_{i=1}^{n} a_i(X_i - \xi) = n_0 a(\overline{X}_0 - \xi) + \sum_{i=n_0+1}^{n} b(X_i - \xi),$$

we have that conditional on $S = s$, $\Sigma_{i=1}^n a_i(X - \xi)$ is $N(0, n_0 a^2\sigma^2 + (n-n_0)b^2\sigma^2) = N(0, \Sigma_{i=1}^n a_i^2\sigma^2)$ distributed. Therefore, still conditional on $S = s$,

$$Y = \sum_{i=1}^{n} a_i(X_i - \xi) / \sqrt{(S^2 \Sigma_{i=1}^n a_i^2)}$$

is $N(0,\sigma^2/s^2)$ distributed. It follows by (i) that Y is (unconditionally) t-distributed with n_0-1 degrees of freedom.

(iii) There exist numbers a and b such that $n_0 a + (n-n_0)b = 1$ and, for given c : $n_0 a^2 + (n-n_0)b^2 = c/S^2$ if $n > n_0$ and if the following equation in a, obtained after substitution of $b = (1-n_0a)/(n-n_0)$ in $n_0a^2 + (n-n_0)b^2 = c/S^2$, has a solution

$$n_0a^2 + (n-n_0)\frac{(1-n_0a)^2}{(n-n_0)^2} = c/S^2.$$

This equation is equivalent to

$$nn_0a^2 - 2n_0a + 1 - (n-n_0)c/S^2 = 0$$

and this equation has a solution if

$$4n_0^2 - 4nn_0(1-(n-n_0)c/S^2) \geq 0 \Leftrightarrow n \geq \frac{S^2}{c}.$$

Hence for $n = \max\{n_0+1, \left[\frac{S^2}{c}\right]+1\} > n_0, \frac{S^2}{c}$, the problem is solved if we take $a_1 = \dots = a_{n_0} = a$, $a_{n_0+1} = \dots = a_n = b$ where a and b are the solutions of the equations above.

(iv) Apply (ii), where $a_i = \frac{1}{n}$ $(i = 1,\dots,n)$.

(STEIN (1945), CHAPMAN (1959))

Problem 16.

(i) Clearly, the interval $\Sigma_{i=1}^n a_iX_i \pm L/2$ has length L. Moreover

$$P\left\{\sum_{i=1}^n a_iX_i - L/2 \leq \xi \leq \sum_{i=1}^n a_iX_i + L/2\right\} =$$

$$= P\left\{-L/2 \leq \sum_{i=1}^n a_iX_i - \xi \leq L/2\right\} =$$

$$= P\left\{-L/2\sqrt{c} \leq \sum_{i=1}^n a_i(X_i-\xi)/\sqrt{c} \leq L/2\sqrt{c}\right\} =$$

$$= \int_{-L/2\sqrt{c}}^{L/2\sqrt{c}} t_{n_0-1}(y)dy = \gamma,$$

using $\Sigma_{i=1}^n a_i = 1$, and the fact that $\Sigma_{i=1}^n a_i(X_i-\xi)/\sqrt{c} \sim t(n_0-1)$, which follows from (iii) of Problem 15.

(ii) $$P\{\overline{X} - L/2 \leq \xi \leq \overline{X} + L/2\} = P\{|\overline{X}-\xi| \leq L/2\} =$$

$$= P\left\{\frac{\sqrt{n}\,|\overline{X}-\xi|}{S} \leq \frac{\sqrt{n}\,L}{2S}\right\} \geq P\left\{\frac{\sqrt{n}\,|\overline{X}-\xi|}{S} \leq \frac{L}{2\sqrt{c}}\right\},$$

since by definition (Problem 15, (iv)), $n \geq \left[\frac{S^2}{c}\right] + 1 \Rightarrow \frac{\sqrt{n}}{S} \geq \sqrt{c}$.

Furthermore,

$$P\left\{\frac{\sqrt{n}\,|\overline{X}-\xi|}{S} \leq \frac{L}{2\sqrt{c}}\right\} = \int_{-L/2\sqrt{c}}^{L/2\sqrt{c}} t_{n_0-1}(y)dy = \gamma,$$

since $\sqrt{n}(\overline{X}-\xi)/S \sim t(n_0-1)$ by (iv) of Problem 15.

We now show that the expected number of observations under Π_2 is slightly lower than under Π_1. Define

$$n_{\Pi_1} = \max\{n_o + 1, \left[\frac{S^2}{c}\right] + 1\},$$

$$n_{\Pi_2} = \max\{n_0, \left[\frac{S^2}{c}\right] + 1\}.$$

Since $[S^2/c] + 1$ is an integer,

$$n_{\Pi_1} = n_{\Pi_2} = [S^2/c] + 1 \quad \text{if } [S^2/c] + 1 \geq n_0 + 1;$$

$$n_{\Pi_1} - 1 = n_{\Pi_2} = n_0 \quad \text{if } [S^2/c] + 1 \leq n_0$$

So

$$\begin{aligned} E(n_{\Pi_1}) - E(n_{\Pi_2}) &= E(n_{\Pi_1} - n_{\Pi_2}) \\ &= P\{[S^2/c] + 1 \leq n_0\} \\ &= P\{S^2/c < n_0\} \\ &= P\{U < n_0 c/\sigma^2\} \end{aligned}$$

where U is χ^2-distributed with n_0-1 degrees of freedom.

(STEIN (1945), CHAPMAN (1950))

Problem 17.

(i) Define the critical function

$$\varphi(X_1,\ldots,X_n) = \begin{cases} 1 & \text{if } (\Sigma_{i=1}^n a_i X_i - \xi_0)/\sqrt{c} \geq C, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have to prove that

(1) $\quad \beta_c(\xi) = E_\xi \varphi(X_1,\ldots,X_n) \leq \alpha$ for $\xi \leq \xi_0$

(2) $\quad \beta_c(\xi)$ is a strictly increasing function depending only on ξ.

By Problem 15 (iii) $(\Sigma_{i=1}^{n} a_iX_i - \xi)/\sqrt{c} = \Sigma_{i=1}^{n} a_i(X_i - \xi)/\sqrt{c}$ (since $\Sigma_{i=1}^{n} a_i = 1$) has a t-distribution with n_0-1 degrees of freedom. Hence

$$\begin{aligned}\beta_c(\xi) &= E_\xi\varphi(X_1,\ldots,X_n) \\ &= P_\xi\left\{\frac{\Sigma a_iX_i - \xi_0}{\sqrt{c}} > C\right\} \\ &= P_\xi\left\{\frac{\Sigma a_iX_i - \xi}{\sqrt{c}} > C - \frac{\xi - \xi_0}{\sqrt{c}}\right\} \\ &= \int_{C-(\xi-\xi_0)/\sqrt{c}}^{\infty} t_{n_0-1}(y)\,dy,\end{aligned}$$

which is strictly increasing in ξ.

Also, by the definition of C, it follows immediately from the preceding equality that $\beta_c(\xi) \le \alpha$ for all $\xi \le \xi_0$.

(ii) From the equality

$$\beta_c(\xi_1) = \int_{C-(\xi_1-\xi_0)/\sqrt{c}}^{\infty} t_{n_0-1}(y)\,dy = \beta$$

we get

$$1 - \beta = F_{n_0-1}\left(C - \frac{\xi_1 - \xi_0}{\sqrt{c}}\right),$$

where $F_\nu(t)$ denotes the probability integral of a t-distribution with ν degrees of freedom. Hence

$$c = \left(\frac{\xi_1 - \xi_0}{F_{n_0-1}^{-1}(1-\beta) - C}\right)^2$$

(note that since $\alpha < \beta < 1$, we have $C < F_{n_0-1}^{-1}(1-\beta) < \infty$).

(iii) Since $n = \max\{n_0, \left[\frac{S^2}{c}\right] + 1\}$ it follows that $n > \frac{S^2}{c}$ or equivalently

$$\frac{\sqrt{n}}{S} > \frac{1}{\sqrt{c}}.$$

This implies

$$\frac{\sqrt{n}(\xi - \xi_0)}{S} > \frac{\xi - \xi_0}{\sqrt{c}} \quad \text{for all } \xi > \xi_0 \tag{23}$$

$$\frac{\sqrt{n}(\xi - \xi_0)}{S} \le \frac{\xi - \xi_0}{\sqrt{c}} \quad \text{for all } \xi \le \xi_0. \tag{24}$$

First we show that the test determined by $\sqrt{n}(\bar{X}-\xi_0)/S > C$ is a level α test. Indeed we have for all $\xi \le \xi_0$

$$P_\xi\left\{\frac{\sqrt{n}(\bar{X}-\xi_0)}{S} > C\right\} = P_\xi\left\{\frac{\sqrt{n}(\bar{X}-\xi)}{S} > C - \frac{\sqrt{n}(\xi-\xi_0)}{S}\right\}$$

$$\le P_\xi\left\{\frac{\Sigma\, a_iX_i - \xi}{\sqrt{c}} > C - \frac{\xi-\xi_0}{\sqrt{c}}\right\} = P_\xi\left\{\frac{\Sigma\, a_iX_i - \xi}{\sqrt{c}}\right\} \le \alpha$$

where the first inequality follows by (24) and the fact that $\sqrt{n}(\bar{X}-\xi)/S$ and $(\Sigma\, a_iX_i - \xi)/\sqrt{c}$ have a t-distribution with n_0-1 degrees of freedom, by Problem 15 (iii) and 15 (iv). The second inequality follows by part (i). Analogously it follows from (23) that for all $\xi > \xi_0$

$$P_\xi\left\{\frac{\sqrt{n}(\bar{X}-\xi_0)}{\sqrt{c}} > C\right\} = P_\xi\left\{\frac{\sqrt{n}(\bar{X}-\xi)}{S} > C - \frac{\sqrt{n}(\xi-\xi_0)}{S}\right\}$$

$$> P_\xi\left\{\frac{\Sigma\, a_iX_i - \xi}{\sqrt{c}} > C - \frac{\xi-\xi_0}{\sqrt{c}}\right\} = P_\xi\left\{\frac{\Sigma\, a_iX_i - \xi_0}{\sqrt{c}} > C\right\}.$$

Hence the test with rejection region $\sqrt{n}(\bar{X}-\xi_0)/S > C$ based on Π_2 and the same c as in (i) is a level α test of H which is uniformly more powerful than the test given in (i).

(iv) <u>part 1</u> (extension of (i)).

For the procedure Π_1 with any given c, let C now be defined by $\int_C^\infty t_{n_0-1}(y)dy = \frac{\alpha}{2}$. Then the rejection region $|(\Sigma\, a_iX_i - \xi_0)/\sqrt{c}| > C$ defines a level α test for H : $\xi = \xi_0$ against K : $\xi \ne \xi_0$ with power function $\beta_c(\xi)$ strictly increasing in $|\xi - \xi_0|$ and depending only on ξ.

The power of the test is

$$\beta_c(\xi) = P_\xi\left\{\left|\frac{\Sigma\, a_iX_i - \xi_0}{\sqrt{c}}\right| > C\right\}.$$

Hence

$$1 - \beta_c(\xi) = P_\xi\left\{-C \le \frac{\Sigma\, a_iX_i - \xi_0}{\sqrt{c}} \le C\right\}$$

$$= P_\xi\left\{-C - \frac{\xi-\xi_0}{\sqrt{c}} \le \frac{\Sigma\, a_iX_i - \xi}{\sqrt{c}} \le C - \frac{\xi-\xi_0}{\sqrt{c}}\right\}.$$

Since $(\Sigma_{i=1}^n\, a_iX_i - \xi)/\sqrt{c}$ has a t-distribution with n_0-1 degrees of freedom it follows by elementary consideration that $1-\beta_c(\xi)$ decreases strictly as ξ tends away form ξ_0 in either direction.

part 2 (extension of (ii)).

We must show that, given any alternative ξ_1 and any β, $\alpha < \beta < 1$, the number c can be chosen so that $\beta_c(\xi_1) = \beta$.

From the equality

$$\beta_c(\xi_1) = \int_{\infty}^{-c-\frac{\xi_1-\xi_0}{\sqrt{c}}} t_{n_0-1}(y)dy + \int_{c-\frac{\xi_1-\xi_0}{\sqrt{c}}}^{\infty} t_{n_0-1}(y)dy$$

it follows that c is the solution of the equation

$$1-\beta = F_{n_0-1}\left(c - \frac{\xi_1-\xi_0}{\sqrt{c}}\right) - F_{n_0-1}\left(-c - \frac{\xi_1-\xi_0}{\sqrt{c}}\right).$$

Note that in this case we obtain an integral equation for c, but this equation always has a unique solution for $\alpha < \beta < 1$.

part 3 (extension of (iii)).

We must show that the test with rejection region $|\sqrt{n}(\bar{X}-\xi_0)/S| > c$ based on Π_2 and the same c as in part 1 is a level α test for $H : \xi = \xi_0$ against $K : \xi \neq \xi_0$ which is more powerful than the test given in part 1.

Let $\tilde{\beta}_c(\xi)$ denote the power function of the test with rejection region $|\sqrt{n}(\bar{X}-\xi_0)/S| > c$.

Case (1): $\xi > \xi_0$.

We have

$$1-\tilde{\beta}_c(\xi) = P_\xi\left\{-c - \frac{\sqrt{n}(\xi-\xi_0)}{S} < T < c - \frac{\sqrt{n}(\xi-\xi_0)}{S}\right\},$$

$$1-\beta_c(\xi) = P_\xi\left\{-c - \frac{\xi-\xi_0}{\sqrt{c}} < T < c - \frac{\xi-\xi_0}{\sqrt{c}}\right\}$$

where T has a t-distribution with n_0-1 degrees of freedom. Since $\sqrt{n}(\xi-\xi_0)/S > (\xi-\xi_0)/\sqrt{c}$, it follows from the preceding equations that $1-\tilde{\beta}_c(\xi) < 1-\beta_c(\xi)$.

Case (2): $\xi < \xi_0$.

Since now $\sqrt{n}(\xi-\xi_0)/S < (\xi-\xi_0)/\sqrt{c}$, we get the same relation between $\tilde{\beta}_c(\xi)$ and $\beta_c(\xi)$ as in case (1).

(STEIN (1945), CHAPMANN (1950))

Section 5

Problem 18.

Define

$$f_{a,b}(x) = \begin{cases} a^{-1}e^{-(x-b)/a} & \text{if } x \geq b, \\ 0 & \text{if } x < b. \end{cases}$$

If X has density $f_{a,b}(x)$, then X/a' has density $f_{a/a',b/a'}(x)$. Let φ be the UMP unbiased test of Problem 12 for the hypothesis $H : a = 1$. Then $\varphi'(x) = \varphi(x/a_0)$ is a UMP unbiased test for $H' : a = a_0$. We have

$$E_{(a,b)}\varphi'(X) = E_{(a,b)}\varphi(X/a_0) =$$
$$= E_{(a/a_0,b/b_0)}\varphi(X) = \begin{cases} \alpha & \text{if } a = a_0 \\ \geq \alpha & \text{if } a \neq a_0. \end{cases}$$

Hence, φ' is unbiased.
Since each test ψ for $H' : a = a_0$ can be transformed by $\psi'(x) = \psi(x \cdot a_0)$ to a test for $H : a = 1$ and conversely, φ' is UMP unbiased.

The acceptance region of φ' is

$$c_1 \leq 2\Sigma\left[\frac{x_i}{a_0} - \min\left(\frac{x_1}{a_0},\ldots,\frac{x_n}{a_0}\right)\right] \leq c_2$$

and hence the most accurate unbiased confidence interval for a is given by

$$\frac{2}{c_2}\,\Sigma\left[x_i - \min(x_1,\ldots,x_n)\right] \leq a \leq \frac{2}{c_1}\,\Sigma\left[x_i - \min(x_1,\ldots,x_n)\right].$$

Problem 19.

(i) Let X and Y be independently distributed according to the binomial distributions $b(p_1,m)$ and $b(p_2,n)$ respectively. The conditional distribution of Y given $Y + X = t$ is given by (21) on p. 143 of the book

$$p\{Y = y \,/\, X + Y = t\} = C_t(\rho)\binom{m}{t-y}\binom{n}{y}\rho^y, \quad y = 0,1,\ldots,t$$

where $\rho = (p_2q_1)/(p_1q_2)$ and $C_t(\rho)$ is a norming constant. The UMP unbiased test of the hypothesis $\rho = \rho_0$ is the same for each t on the line segment $X + Y = t$ as the UMP unbiased conditional test of the hypothesis $\rho = \rho_0$. The conditional distribution of Y given $X + Y = t$ constitutes a one-parameter exponential family, so Lemma 1 can be applied. The conditions of of this lemma are satisfied because the above test is strictly unbiased as is shown on p. 128, and the conditional distribution can be made continuous by addition of a uniform variable (cf. p. 180, p. 81 of the book). Therefore most accurate unbiased confidence regions exist and are indeed intervals.

(ii) In the same way, using the conditional distribution of X given $X+X'$ and $X+Y$ as stated on p. 145, it follows that most accurate unbiased confidence intervals exist in a 2×2 table for the parameter Δ of Section 6 of Chapter 4.

Section 6

Problem 20.

(i) Consider the following transformation:

$$Z_1 = \left(\frac{a^2}{n} + b\right)^{-\frac{1}{2}} \sum_{i=1}^n \left(bv_i + \frac{a}{n}\right) y_i = \sum_{i=1}^n a_{1i} y_i$$

$$Z_2 = (a + nb^2)^{-\frac{1}{2}} \sum_{i=1}^n (av_i - b) y_i = \sum_{i=1}^n a_{2i} y_i .$$

Then, since $\Sigma_{i=1}^n v_i = 0$ and $\Sigma_{i=1}^n v_i^2 = 1$,

$$\sum_{i=1}^n a_{1i}^2 = \left(\frac{a^2}{n} + b^2\right)^{-1} \left(b^2 \sum_{i=1}^n v_i^2 + \frac{2ab}{n} \sum_{i=1}^n v_i + n\cdot\frac{a^2}{n^2}\right) = 1$$

and analogously, $\Sigma_{i=1}^n a_{2i}^2 = 1$ and $\Sigma_{i=1}^n a_{1i}a_{2i} = 0$.
Following the Gram-Schmidt procedure we can now construct an orthogonal transformation

$$Z_j = \sum_{i=1}^n a_{ji} Y_i, \qquad j = 1,2,\ldots,n$$

with Z_1 and Z_2 as above.
For $j = 3,4,\ldots,n$ we have

$$\sum_{i=1}^n a_{1i}a_{ji} = \left(\frac{a^2}{n} + b^2\right)^{-\frac{1}{2}} \sum_{i=1}^n \left(bv_i + \frac{a}{n}\right) a_{ji} = 0$$

$$\sum_{i=1}^n a_{2i}a_{ji} = (a^2 + nb^2)^{-\frac{1}{2}} \sum_{i=1}^n (av_i - b) a_{ji} = 0$$

or

(25) $$b \sum_{i=1}^n v_i a_{ji} = -\frac{a}{n} \sum_{i=1}^n a_{ji}$$

(26) $$a \sum_{i=1}^n v_i a_{ji} = b \sum_{i=1}^n a_{ji}$$

so that for $j = 3,4,\ldots,n$

(27) $$\sum_{i=1}^{n} v_i a_{ji} = 0 \quad \text{and} \quad \sum_{i=1}^{n} a_{ji} = 0.$$

For ($a = 0$ and $b \neq 0$) or ($a \neq 0$ and $b = 0$) this follows immediately from (25) and (26). For $a \neq 0$ and $b \neq 0$ it follows from (25) and (26) because

$$\left(\frac{b^2}{a} + \frac{a^2}{n}\right) \sum_{i=1}^{n} a_{ji} = 0, \quad \text{or} \quad \frac{nb^2 + a^2}{na} \sum_{i=1}^{n} a_{ij} = 0.$$

By orthogonality of the transformation $\Sigma_{i=1}^{n} Z_i^2 = \Sigma_{i=1}^{n} Y_i^2$ and hence $\Sigma_{i=3}^{n} Z_i^2 = \Sigma_{i=1}^{n} Y_i^2 - (Z_1^2 + Z_2^2)$.
Now

$$\begin{aligned} Z_1^2 + Z_2^2 &= \left(\frac{a^2}{n} + b^2\right)^{-1}\left[b \sum_{i=1}^{n} v_i Y_i + a\overline{Y}\right]^2 + \\ &\quad + (a^2 + nb^2)^{-1}\left[a \sum_{i=1}^{n} v_i Y_i - nb\overline{Y}\right]^2 = \\ &= n\overline{Y}^2 + \left(\sum_{i=1}^{n} v_i Y_i\right)^2 \end{aligned}$$

so

(28) $$\sum_{i=3}^{n} Z_i^2 = \sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2 - \left(\sum_{i=1}^{n} v_i Y_i\right)^2 = \sum_{i=1}^{n} (Y_i - \overline{Y})^2 - \left(\sum_{i=1}^{n} v_i Y_i\right)^2.$$

By Problem 6, $Z_1,\ldots,Z_n$ are independently normally distributed with common variance σ^2 and means

$$\begin{aligned} EZ_1 &= \left(\frac{a^2}{n} + b^2\right)^{-\frac{1}{2}}\left[b \sum_{i=1}^{n} v_i EY_i + aE\overline{Y}\right] = \\ &= \left(\frac{a^2}{n} + b^2\right)^{-\frac{1}{2}}\left[b \sum_{i=1}^{n} v_i(\gamma + \delta v_i) + \frac{a}{n} \sum_{i=1}^{n} (\gamma + \delta v_i)\right] = \\ &= \left(\frac{a^2}{n} + b^2\right)^{-\frac{1}{2}}(b\cdot\delta + a\cdot\gamma) = \left(\frac{a^2}{n} + b^2\right)^{-\frac{1}{2}}\rho \end{aligned}$$

$$EZ_2 = (a^2 + nb^2)^{-\frac{1}{2}}(a\delta - nb\gamma) \quad \text{(analogously)}$$

and for $j = 3,4,\ldots,n$

$$EZ_j = \sum_{i=1}^{n} a_{ji} EY_j = \sum_{i=1}^{n} a_{ji}(\gamma + \delta v_i) = 0, \text{ by (27)}.$$

Putting $s = 2$ in Problem 5 the UMP unbiased test for testing $H : (a^2/n + b^2)^{-\frac{1}{2}}\rho = (a^2/n + b^2)^{-\frac{1}{2}}\rho_0$ (or equivalently $H : \rho = \rho_0$) is given by the acceptance region

$$\frac{|Z_1 - (a^2/n + b^2)^{-\frac{1}{2}}\rho_0|}{\sqrt{\Sigma_{i=3}^n Z_i^2/(n-2)}} \le C$$

Using (28) this can be written as

$$\frac{|b\ \Sigma_{i=1}^n\ v_i Y_j + a\bar{Y} - \rho_0|\ /\ \sqrt{a^2/n + b^2}}{\sqrt{\left[\Sigma_{i=1}^n\ (Y_i - \bar{Y})^2 - (\Sigma_{i=1}^n\ v_i Y_i)^2\right]/(n-2)}} \le C$$

where C is determined by $\int_{-c}^{c} t_{n-2}(y)dy = 1-\alpha$ since under H the test statistic has a t-distribution with n-1 degrees of freedom (cf. Problem 5).

(ii) Putting $S^2 = \Sigma_{i=1}^n\ (Y_i - \bar{Y})^2 - (\Sigma_{i=1}^n\ v_i Y_i)^2$ we have for all $t > 0$

$$|\bar{\rho}(a,b) - \rho(a,b)| \le t$$

$$\Leftrightarrow \left|b \sum_{i=1}^n v_i Y_i + a\bar{Y} + C\sqrt{a^2/n + b^2}\ \sqrt{S^2/(n-2)} - \rho(a,b)\right| \le t$$

$$\Leftrightarrow \left|\frac{b\ \Sigma_{i=1}^n\ v_i Y_i + a\bar{Y} - \rho(a,b)}{\sigma\sqrt{a^2/n + b^2}} + \frac{C}{\sqrt{n-2}}\sqrt{S^2/\sigma^2}\right| \le \frac{t/\sigma}{\sqrt{a^2/n + b^2}}\ .$$

Since $(b\ \Sigma_{i=1}^n\ v_i Y_i + a\bar{Y} - \rho(a,b))/\sigma\sqrt{a^2/n + b^2}$ is standard normally distributed and independent of S^2/σ^2 which has the χ^2-distribution with n-2 degrees of freedom, it follows that for (a_1,b_1) and (a_2,b_2) satisfying $\sqrt{a_1^2/n + b_1^2} < \sqrt{a_2^2/n + b_1^2}$,

$$P_{a_1,b_1}\{|\bar{\rho}(a_1,b_1) - \rho(a_1,b_1)| \le t\} > P_{a_2,b_2}\{|\bar{\rho}(a_2,b_2) - \rho(a_2,b_2)| \le t\}.$$

Since f_1 is an increasing function it follows from Problem 11 of Chapter 3 that

$$E_{a_1,b_1}[f_1(|\bar{\rho}(a_1,b_1) - \rho(a_1,b_1)|)] \le E_{a_2,b_2}[f_1(|\bar{\rho}(a_2,b_2) - \rho(a_2,b_2)|)].$$

The similar result for f_2 and $\underline{\rho}$ can be shown in the same way, giving the required conclusion.

Section 7

Problem 21.

The set of order statistics

$$T(Z) = (Z_1^{(1)},\ldots,Z_1^{(N_1)};Z_2^{(1)},\ldots,Z_2^{(N_2)};\ldots;Z_c^{(1)},\ldots,Z_c^{(N_c)})$$

is a complete sufficient statistic for $\mathcal{F}$ (Chapter 4, Example 6). A necessary and sufficient condition for (48) on p. 184 is therefore

(29) $$E[\varphi(Z) \mid T(z)] = \alpha \quad \text{a.e.}$$

The set S(z) consists of the $N_1!\ldots N_c!$ points obtained from z through permutation of the coordinates z_{ij} $(j+1,2,\ldots,N_i)$ within the ith subgroup $(i=1,2,\ldots,c)$ in all $N_1!\ldots N_c!$ possible ways, so that $S(z) = \{z' : T(z') = T(z)\}$.

It follows from Section 4 of Chapter 2 that the conditional distribution of Z given T(z) assigns probability $1/(N_1!\ldots N_c!)$ to each of the $N_1!\ldots N_c!$ points of S(z). Thus (29) is equivalent to (49) on p. 184, as was to be proved.

Section 8

Problem 22.

For c = 1 and m = n = 4 the rejection region given by (54) on p. 188 becomes

$$\sum_{j=5}^{8} z_j > C[T(z)].$$

The test statistic takes on only $\binom{8}{4} = 70$ distinct values over all permutations of $Z^{(1)},\ldots,Z^{(8)}$ so for $\alpha = .10$ the test rejects H for the 7 largest ones. The rejection region turns out to consist of the points z for which $\Sigma_{j=5}^{8} z_j \geq 12.09$, corresponding to $\{Z_5,Z_6,Z_7,Z_8\} = \{Z^{(i)} : i \in I\}$ with $I = \{3,4,7,8\},\{4,5,6,8\},\{3,5,7,8\},\{3,6,7,8\},\{4,5,7,8\},\{4,6,7,8\}$ or $\{5,6,7,8\}$.

Problem 23.

A point $(x_1,\ldots,x_m,y_1,\ldots,y_n)$ is in the acceptance region $A(\Delta)$ for $H(\Delta)$ if and only if

$$\left|\frac{1}{n}\sum_{j=m+1}^{m+n} z_j - \frac{1}{m}\sum_{j=1}^{m} z_j\right| = \left|\frac{1}{n}\sum_{j=1}^{n}(y_i - \Delta) - \frac{1}{m}\sum_{j=1}^{m} x_j\right| = |\bar{y} - \bar{x} - \Delta|$$

is exceeded by at least $\binom{m+n}{n} - k = \alpha\binom{m+n}{n}$ of the quantities

$$\left|\frac{1}{n}\sum_{j=m+1}^{m+n} z_{i_j} - \frac{1}{m}\sum_{j=1}^{m} z_{i_j}\right|$$

where $i_1 < \ldots < i_m; i_{m+1} < \ldots < i_{m+n}$ is a permutation of the integers

$1,2,\ldots,m+n$.

Denote for any finite set A the number of elements of A by N(A) and let $J = \{j : i_j \in \{m+1,\ldots,m+n\}\}$.

Then

$$\frac{1}{n}\sum_{j=m+1}^{m+n} z_{i_j} - \frac{1}{m}\sum_{j=1}^{m} z_{i_j} =$$

$$= \frac{1}{n}\left[\sum_{j=1}^{n} y'_j - N(J \cap \{m+1,\ldots,m+n\})\Delta\right] - \frac{1}{m}\left[\sum_{j=1}^{m} x'_j - N(J \cap \{1,2,\ldots,m\})\Delta\right] =$$

$$= \bar{y}' - \bar{x}' - \gamma\,\Delta$$

where $(x'_1,\ldots,x'_m,y'_1,\ldots,y'_n)$ is the permutation of $(x_1,\ldots,x_m,y_1,\ldots,y_n)$ that corresponds to $(i_1,\ldots,i_{m+n})$ and

$$\gamma = \frac{1}{n}\,N(J \cap \{m+1,\ldots,m+n\}) - \frac{1}{m}\,N(J \cap \{1,2,\ldots,m\})$$

so that $|\gamma| \leq 1$.

The $(1-\alpha)$ confidence region for Δ is given by

$$S(x_1,\ldots,x_m,y_1,\ldots,y_n) = \{\Delta : (x_1,\ldots,x_m,y_1,\ldots,y_n) \in A(\Delta)\}.$$

That this is indeed an interval can be seen as follows:

Let a, b, γ, Δ_1 and Δ_2 be real values with $|\gamma| \leq 1$, $\Delta_1 < \Delta_2$ and $(a-\Delta_i)^2 \leq (b-\gamma\Delta_i)^2$, $i = 1,2$. Then, because $(b-\gamma\Delta)^2 - (a-\Delta)^2 = (\gamma^2-1)\Delta^2 + 2(a-b\gamma)\Delta + b^2 - a^2$ and $\gamma^2 - 1 \leq 0$, it follows that $(a-\Delta)^2 \leq (b-\gamma\Delta)^2$ for any Δ with $\Delta_1 \leq \Delta \leq \Delta_2$.

Now consider any $\Delta_1 \in S(x_1,\ldots,x_m,y_1,\ldots,y_n)$. Then

$$(\bar{y}-\bar{x}-\Delta)^2 \leq (\bar{y}'-\bar{x}'-\gamma\Delta)^2$$

holds for $\Delta = \Delta_1$ for at least $\binom{m+n}{n} - k$ permutations $(x'_1,\ldots,x'_m,y'_1,\ldots,y'_n)$. The inequality also holds for $\Delta = \bar{y}-\bar{x}$ for all permutations and hence it holds for all Δ between $\bar{y}-\bar{x}$ and Δ_1 for at least $\binom{m+n}{n} - k$ permutations $(x'_1,\ldots,x'_m,y'_1,\ldots,y'_n)$. So any Δ between $\bar{y}-\bar{x}$ and Δ_1 is an element of $S(x_1,\ldots,x_m,y_1,\ldots,y_n)$ which yields that $S(x_1,\ldots,x_m,y_1,\ldots,y_n)$ is an interval.

Section 9

Problem 24.

From Problem 7 (iii) of Chapter 2 it follows that under the hypothesis the set T of order statistics of $(Z_1^2,\dots,Z_n^2)$ is sufficient and from Example 6 of Chapter 4 that T is also complete. (T is equivalent to $(\Sigma_{i=1}^n Z_i^2,\dots,\Sigma_{i=1}^n Z_i^{2n})$, which is suggested in the hint).
Let $Z = (Z_1,\dots,Z_n)$, φ any unbiased level α test and h_ζ the density of an n-variate normal distribution with mean $(\zeta,\dots,\zeta)$ and covariance matrix $\sigma^2 I$.
As in Example 7 of Chapter 2 under the alternative the conditional expectation of $\varphi(Z)$ given $T = t$ is

$$E[\varphi(Z) \mid T = t] = \frac{\Sigma\, \varphi(z')h_\zeta(z')}{\Sigma\, h_\zeta(z')}$$

where the summation is over the $2^n n!$ point $z' \in S(t) = \{(z_1',\dots,z_n') : (z_1'^2,\dots,z_n'^2) \text{ is a permutation of } t\}$. Furthermore, as in the proof of Theorem 3 under the Hypothesis

$$E[\varphi(Z) \mid T = t] = \frac{1}{2^n n!} \sum_{z' \in S(t)} \varphi(z').$$

Therefore, the problem is quite similar to the one described in Section 8. To carry out a most powerful unbiased test the $2^n n!$ points of each set S(t) are ordered according to the values of the density h_ζ.
Since $h_\zeta(z')$, for fixed $\Sigma_{i=1}^n z_i'^2$, is an increasing function of $\Sigma_{i=1}^n z_i'$, and since $\Sigma_{i=1}^n z_i$ is constant over the n! permutations of $(z_1',\dots,z_n')$, the test reduces to the following:
Order the 2^n values of $\Sigma_{i=1}^n \pm Z_i$ (almost surely there will be no ties) and reject the hypothesis if $\Sigma_{i=1}^n Z_i$ is one of the k largest values; reject with probability γ if $\Sigma_{i=1}^n Z_i$ is the (k+1)'th largest value. Here k and γ are defined by $k+\gamma = \alpha 2^n$. As in Section 8, the rejection region has the form of the t-test (15) of Chapter 5 in which the constant cutoff point C_0 has been replaced by a random one.

Problem 25.

(i) Let $X_1,\dots,X_n,Y_1,\dots,Y_n$ be independently normally distributed with common variance σ^2 and means $EX_i = \xi_i$ and $EY_i = \xi_i + \Delta$, $i = 1,2,\dots,n$.
Consider the following transformation:

$$X_i' = (Y_i - X_i)/\sqrt{2}, \quad Y_i' = (Y_i + X_i)/\sqrt{2}, \quad i = 1,2,\dots,n.$$

Since this transformation is orthogonal, $X_1',\dots,X_n',Y_1',\dots,Y_n'$ are independently normally distributed with common variance σ^2 and means $EX_i' = \Delta/\sqrt{2}$ and $EY_i' = (2\xi_i + \Delta)/\sqrt{2}$, $i = 1,2,\dots,n$. (cf. Problem 6). Therefore the joint density of $X_1',\dots,X_n',Y_1',\dots,Y_n'$ is given by

$$f(x',y') = \frac{1}{(2\pi\sigma^2)^n} \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^n (x_i' - \Delta/\sqrt{2})^2 + \sum_{i=1}^n (y_i' - (2\xi_i + \Delta)/\sqrt{2})^2\right\}\right] =$$

$$(30) \qquad = C(\sigma^2,\Delta,\xi)\exp\left[\frac{\Delta}{\sigma^2\sqrt{2}}\sum_{i=1}^n x_i' - \frac{1}{2\sigma^2}\sum_{i=1}^n \{(x_i')^2 + (y_i')^2\} + \frac{1}{\sigma^2\sqrt{2}}\sum_{i=1}^n (2\xi_i + \Delta)y_i'\right].$$

This density is of the form (1) on p. 160 with

$$U = \sum_{i=1}^n X_i', \quad T_0 = \sum_{i=1}^n \{(X_i')^2 + (Y_i')^2\} \quad \text{and } T_i = Y_i', \quad i = 1,2,\dots,n$$

and

$$\theta = \frac{\Delta}{\sqrt{2}\,\sigma^2}, \quad \vartheta_0 = -\frac{1}{2\sigma^2} \quad \text{and } \vartheta_i = \frac{2\xi_i + \Delta}{\sigma^2\sqrt{2}}, \quad i = 1,2,\dots,n.$$

When $\Delta = 0$ (or equivalently $\theta = 0$) the statistic

$$V = \frac{U}{\sqrt{T_0 - \Sigma_{i=1}^n T_i^2 - \frac{1}{n}U^2}} = \frac{\Sigma_{i=1}^n X_i'}{\sqrt{\Sigma_{i=1}^n (X_i' - \bar{X}')^2}}$$

is independent of T_0 (see Example 1) and of $T_1,\dots,T_n$. Furthermore V is an increasing function of U for each $T_0,T_1,\dots,T_n$, so by Theorem 1 the UMP unbiased test for the hypothesis $H : \Delta = 0$ is given by the rejection region $V > C$.

In terms of the differences $W_i = Y_i - X_i$, $i = 1,2,\dots,n$ the rejection region of the UMP unbiased test for testing $H : \Delta = 0$ agianst $K : \Delta > 0$ can also be written as

$$\frac{\sqrt{n}\,\bar{W}}{\sqrt{\Sigma_{i=1}^n (W_i - \bar{W})^2/(n-1)}} > C_0$$

Under H this statistic has a t-distribution with n-1 degrees of freedom, so C_0 is determined by $\int_C^\infty t_{n-1}(y)dy = \alpha$.

(ii) In order to obtain the most accurate unbiased confidence intervals for Δ we first consider the hypothesis $H' : \Delta = \Delta_0$ against $K' : \Delta \neq \Delta_0$. With $X_i'' = (Y_i - \Delta_0 - X_i)/\sqrt{2}$ and $Y_i'' = (Y_i - \Delta_0 + X_i)/\sqrt{2}$, $i = 1,2,\dots,n$ the joint density of $X_1'',\dots,X_n'',Y_1'',\dots,Y_n''$ is given by (30) with Δ replaced by $\Delta - \Delta_0$. When $\Delta = \Delta_0$ (or equivalently $\theta' = \frac{\Delta-\Delta_0}{\sigma^2\sqrt{2}} = 0$) the statistic

$$W = \left(\sum_{i=1}^{n} X_i''\right) / \sqrt{\Sigma_{i=1}^{n} (X_i'')^2}$$

is independent of the sufficient statistic $T_0' = \Sigma_{i=1}^{n} \{(X_i'')^2 + (Y_i'')^2\}$ and $T_i' = Y_i''$, $i=1,2,\dots,n$ and is linear in $U' = \Sigma_{i=1}^{n} X_i''$. The distribution of W is symmetric about 0 when $\Delta = \Delta_0$, so Theorem 1 implies that the UMP unbiased test for testing $H' : \Delta = \Delta_0$ against $K' : \Delta \neq 0$ is given by rejection region $|W| > C$.

Since

$$|W'| = \frac{\sqrt{n-1}\,|W|}{\sqrt{n-W^2}}$$

is an increasing function of $|W|$ this rejection region is equivalent to $|W'| > C_0$.

In terms of the differences $W_i = Y_i - X_i$, $i = 1,2,\dots,n$ the rejection region of the UMP unbiased test for testing $H' : \Delta = \Delta_0$ against $K' : \Delta \neq \Delta_0$ can also be written as

$$\frac{\sqrt{n}\,|\overline{W} - \Delta_0|}{\sqrt{\Sigma_{i=1}^{n} (W_i - \overline{W})^2/(n-1)}} > C_p .$$

Under H' this statistic has a t-distribution with n-1 degrees of freedom, so C_p is determined by $\int_{c_1}^{\infty} t_{n-1}(y) = \alpha/2$.

The corresponding most accurate unbiased confidence intervals for Δ are given by

$$\overline{W} - C_1 \frac{1}{\sqrt{n(n-1)}} \sqrt{\Sigma_{i=1}^{n} (W_i - \overline{W})^2} \leq \Delta \leq$$

$$\leq \overline{W} + C_1 \frac{1}{\sqrt{n(n-1)}} \sqrt{\Sigma_{i=1}^{n} (W_i - \overline{W})^2} .$$

Problem 26.

Let $U_1,\dots,U_{2n}$ and $V_1,\dots,V_n$ and $V_{n+1},\dots,V_{2n}$ be independently distributed as $N(\mu,\sigma_1^2)$, $N(\xi,\sigma^2)$ and $N(\eta,\sigma^2)$ respectively. Consider the hypothesis $H : \eta \leq \xi$ or, equivalently, $H : \Delta \leq 0$ where $\Delta = \eta - \xi$.

In the case of complete randomization we have

$$X_i = U_i + V_i \qquad i = 1,2,\ldots,n$$

$$Y_i = U_{n+i} + V_{n+i} \qquad i = 1,2,\ldots,n$$

with X_i and Y_i independently distributed as $N(\mu+\xi,\sigma_1^2+\sigma^2)$ and $N(\mu+\xi,\sigma_1^2+\sigma^2)$ respectively. The UPM unbiased test given by (27) on p. 172 has rejection region

$$T_1 = \frac{(\overline{Y}-\overline{X})/\sqrt{2/n}}{\sqrt{[\Sigma\ (X_i-\overline{X})^2 + \ (Y_i-\overline{Y})^2]/2(n-1)}} > C_1.$$

The test statistic T_1 has a noncentral t-distribution with 2(n-1) degrees of freedom and noncentrality parameter $\Delta / \sqrt{(2/n)(\sigma_1^2+\sigma^2)}$.

In the case of matched pairs we have

$$X_i = U_i + V_i \qquad i = 1,2,\ldots,n$$

$$Y_i = U_i + V_{n+i} \qquad i = 1,2,\ldots,n.$$

Define $W_i = Y_i - X_i = V_{n+1} - V_i$. Then $W_1,\ldots,W_n$ are independently distributed as $(\Delta, 2\sigma^2)$.

The UMP unbiased test given by (59) on p. 192 has rejection region

$$T_2 = \frac{\sqrt{n}\,\overline{W}}{\sqrt{\Sigma_{i=1}^n\ (W_i-\overline{W})^2/(n-1)}} > C_2$$

The test statistic T_2 has a noncentral t-distribution with (n-1) degrees of freedom and noncentrality parameter $\Delta/\sqrt{2\sigma^2/n}$.

In table I the power of the two methods is given for a number of values of n with $\Delta = 4$ and $\alpha = .05$ when $\sigma_1 = 1$ and $\sigma = 2$. The procedure used is:

(a) Compute C_i:

In the case of complete randomization C_1 follows from $P\{t_{2(n-1)} > C_1\} = .05$.

In the case of matched pairs C_2 follows from $P\{t_{n-1} > C_2\} = .05$.

(b) Compute the power β for the alternative with $\Delta = 4$:

In the case of complete randomization the noncentrality parameter becomes $4\sqrt{n/10}$ so the power is given by $\beta_1 = P\{t_{2(n-1)}(4\sqrt{n/10}) > C_1\}$.

In the case of matched pairs the noncentrality parameter becomes $\sqrt{2n}$ so the power is given by $\beta_2 = P\{t_{n-1}(\sqrt{2n}) > C_2\}$.

n	β_1	β_2
2	.33477	.24669
3	.56548	.49087
4	.72157	.69359
5	.82471	.82595
6	.89147	.90464
7	.93380	.94911
8	.96016	.97340
9	.97630	.98634
10	.98605	.99308
12	.99530	.99829
14	.99846	.99959
16	.99951	.99991
18	.99985	.99998
20	.99995	1.00000
30	1.00000	1.00000

Table I

The method of complete randomization only yields greater power for $n \leq 4$. For larger n, the two methods are very close.
In table II the power of the two methods is given for the same values of n, Δ and α when $\sigma_1 = 2$ and $\sigma_2 = 1$. In the case of complete randomization the results remain the same as in table I, because the distribution of T_1 doesn't change. In the case of matched pairs, only the noncentrality parameter changes and becomes $\sqrt{8n}$.

n	β_1	β_2
2	.33477	.46851
3	.56548	.90794
4	.72157	.99292
5	.82471	.99961
6	.89147	.99998
7	.93380	1.00000
8	.96016	1.00000
9	.97630	1.00000
10	.98605	1.00000
12	.99530	1.00000
14	.99846	1.00000
16	.99951	1.00000
18	.99985	1.00000
20	.99995	1.00000
30	1.00000	1.00000

Table II

In the second example the method of matched pairs performs much better than the method of complete randomization and its power tends to 1 more rapidly. This is due to the fact that the method of matched pairs eliminates the effect of the units which have a greater variance in the second example. Values of C_1 and C_2 for the two examples are given in the next problem.

Problem 27.

If a random variable Z has a χ^2-distribution with n degrees of freedom, then

$$E\sqrt{Z} = \int_0^\infty z^{\frac{1}{2}} \frac{z^{n/2-1} e^{-z/2}}{2^{n/2}\Gamma(n/2)} dz = \frac{\sqrt{2}}{\Gamma(n/2)} \int 4^{n/2-\frac{1}{2}} e^{-u} du =$$

$$= \sqrt{2}\,\frac{\Gamma((n+1)/2)}{\Gamma(n/2)} .$$

In the case of complete randomization in Problem 26 the most accurate unbiased confidence intervals for $\Delta = \eta - \xi$ are given in Example 6 on p. 178

$$\bar{Y} - \bar{X} - C_1 S \le \Delta \le \bar{Y} - \bar{X} + C_1 S$$

where

$$S^2 = \frac{2}{n}\,\frac{\Sigma_{i=1}^n (X_i - \bar{X})^2 + \Sigma_{i=1}^n (Y_i - \bar{Y})^2}{2(n-1)}$$

and C_1 is determined by

$$\int_{C_1}^\infty t_{2(n-1)}(y)dy = \alpha/2.$$

The expected length of this confidence interval is given by $2C_1E(S)$ where S^2 is distributed as χ^2 with $2(n-1)$ degrees of freedom multiplied by $(\sigma_1^2+\sigma^2)/n(n-1)$.

In view of (31) we have

$$ES = \sqrt{\frac{\sigma_1^2+\sigma^2}{n(n-1)}} \cdot \sqrt{2}\,\frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-1)}$$

so $2C_1E(S) = 2\sqrt{2}\,C_1 \sqrt{(\sigma_1^2+\sigma^2)/n(n-1)}\;\Gamma(n-\frac{1}{2})/\Gamma(n-1)$.

In the case of matched pairs it follows from Problem 26 and Example 4 on p. 175 that the most accurate unbiased confidence intervals for $\Delta = \eta - \xi$ are given by

$$\overline{W} - \frac{C_2}{\sqrt{n}}\sqrt{\frac{\Sigma_{i=1}^{n}(W_i-\overline{W})^2}{n-1}} \le \Delta \le \overline{W} + \frac{C_2}{\sqrt{n}}\sqrt{\frac{\Sigma_{i=1}^{n}(W_i-\overline{W})^2}{n-1}}$$

where C_2 is determined by $\int_{c_2}^{\infty} t_{n-1}(y)dy = \alpha/2$. The expected length of this confidence interval is given by

$$\frac{2C_2}{\sqrt{n(n-1)}}\ E\sqrt{\frac{\Sigma_{i=1}^{n}(W_i-\overline{W})^2}{\sigma^2}} = 2\sqrt{2}\ \frac{\sigma\cdot C_2}{\sqrt{n(n-1)}}\ \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

in view of (31), because $\Sigma_{i=1}^{n}(W_i-\overline{W})^2/\sigma^2$ has a χ^2-distribution with n-1 degrees of freedom.

In table III the expected lengths of both confidence intervals are given in the same situations as in Problem 26 (when $\sigma_1 = 1$ and $\sigma = 2$ and when $\sigma_1 = 2$ and $\sigma = 1$) with $\alpha = .05$ and with various values of n.
Note that in the case of complete randomization the results for the two examples are the same as in Problem 26. In the case of matched pairs, the expected length is halved when going from the first example to the second. The qualitative conclusions are the same as in the power comparisons of Problem 26.

n	complete randomization		matched pairs		
				expected length	
	C_1	expected length	C_2	$\sigma_1=1,\sigma_2=2$	$\sigma_1=2,\sigma_2=1$
2	4.3027	17.0528	12.7062	28.6748	14.3374
3	2.7764	9.5297	4.3027	8.8060	4.4030
4	2.4469	7.4234	3.1824	5.8641	2.9320
5	2.3060	6.3222	2.7764	4.6686	2.3343
6	2.2281	5.6112	2.5706	3.9943	1.9971
7	2.1788	5.1011	2.4469	3.5491	1.7745
8	2.1448	4.7111	2.3646	3.2271	1.6136
9	2.1199	4.3999	2.3060	2.9803	1.4902
10	2.1009	4.1439	2.2622	2.7832	1.3916
12	2.0739	3.7436	2.2010	2.4844	1.2422
14	2.0555	3.4412	2.1604	2.2656	1.1328
16	2.0423	3.2023	2.1314	2.0962	1.0481
18	2.0322	3.0073	2.1098	1.9601	.9801
20	2.0244	2.8442	2.0930	1.8476	.9238
30	2.0017	2.3014	2.0452	1.4808	.7404
40	1.9908	1.9845	2.0227	1.2711	.6355
50	1.9845	1.7704	2.0096	1.1310	.5655
60	1.9803	1.6135	2.0010	1.0289	.5145
70	1.9773	1.4920	1.9949	.9503	.4752
80	1.9751	1.3944	1.9905	.8873	.4437
90	1.9734	1.3137	1.9870	.8354	.4177
100	1.9720	1.2456	1.9842	.7917	.3958

Table III.

Problem 28.

Let $\gamma = (N_1! \dots N_c!)^{-1}$.
By induction it is seen that

(32) $$0 \le \alpha - \psi_m(z) \le (1-\gamma)^m[\alpha - \psi_0(z)] \quad \text{a.e.} \qquad (m=0,1,\dots).$$

In the case $m = 0$ this reduces to $0 \le \alpha - \psi_0(z) \le \alpha - \psi_0(z)$ a.e. which is true since ϕ_0 satisfies (65) on p. 193.
Suppose that (32) holds for $m = m_0$. Using $\psi_m(z) = \psi_m(z')$ for all $z' \in S(z)$ and $m \in \mathbb{N}$ we have

$$\psi_{m_0+1}(z) = \gamma \sum_{z' \in S(z)} \phi_{m_0+1}(z') =$$

$$= \gamma \sum_{z' \in S(z)} \{\phi_{m_0}(z') + [\alpha - \psi_{m_0}(z')] \cdot I_A(z')\} \le$$

$$\le \psi_{m_0}(z) + [\alpha - \psi_{m_0}(z)] = \alpha \qquad \text{a.e.}$$

by the induction hypothesis, where $A = \{z : \phi_{m_0}(z) < \alpha, \psi_{m_0}(z) < \alpha\}$. So by induction the first inequality of (32) is true for $m = 0,1,\dots$. To prove the second inequality of (32) it should be noticed that, because if $\psi_{m_0}(z) = \alpha$ then $\psi_{m_0+1}(z) = \alpha$ and hence $0 = \alpha - \psi_{m_0+1}(z) \le (1-\gamma)^{m_0+1}[\alpha - \psi_0(z)]$, we may restrict our attention to z with $\phi_{m_0}(z) < \alpha$. If $\psi_{m_0}(z) < \alpha$ then there exists $z_0 \in S(z)$ with $\psi_{m_0}(z_0) < \alpha$, and

$$\psi_{m_0+1}(z) = \psi_{m_0}(z) + \gamma \sum_{z' \in S(z)} [\alpha - \psi_{m_0}(z')] \cdot I_A(z') \ge$$

$$\ge \psi_{m_0}(z) + \gamma[\alpha - \psi_{m_0}(z_0)] =$$

$$\psi_{m_0}(z) + \gamma[\alpha - \psi_{m_0}(z)] = (1-\gamma)\psi_{m_0}(z) + \gamma\alpha \qquad \text{a.e.}$$

The inequality holds a.e. by the first part of the induction hypothesis and the second equality holds because $z_0 \in S(z)$. Hence $\alpha - \psi_{m_0+1}(z) \le (1-\gamma)[\alpha - \psi_{m_0}(z)] \le (1-\gamma)^{m_0+1}[\alpha - \psi_0(z)]$ a.e. by the induction hypothesis.
By induction also the second inequality of (8) is now proved.
Also by induction it can be proved that beginning with a critical function ϕ_0 the construction provides measurable functions ϕ_m and ψ_m.
From (32) and the construction of ϕ_m it follows that the functions ϕ_m are nondecreasing between 0 and 1 a.e. Therefore $\phi = \lim_{m\to\infty} \phi_m$ and $\psi = \lim_{m\to\infty} \psi_m$

are a.e. bounded measurable functions.
$\lim_{m\to\infty} (1-\gamma)^m = 0$ so from (32) it follows that $\psi(z) = \alpha$ a.e. Hence

$$\alpha = \psi(z) = \lim_{m\to\infty} \psi_m(z) = \lim_{m\to\infty} \gamma \sum_{z'\in S(z)} \phi_m(z') =$$

$$= \gamma \sum_{z'\in S(z)} \lim_{m\to\infty} \phi_m(z') = \gamma \sum_{z'\in S(z)} \phi(z') \quad \text{a.e.}$$

as was to be proved.

Problem 29.

The formulation of the problem is incomplete because only the hypothesis H is described. "Unbiasedness of a test ϕ of H" is undefined without specification of the alternative.
So let H be the class of densities

$$\{p_{\sigma,\zeta}(z) : \sigma > c_o > 0, \ \zeta \in R\}$$

where $p_{\sigma,\zeta}$ is given by (63), c_0 is a given constant and R is a given bounded region containing a rectangle. In this solution we shall consider the problem of testing H against a class of alternative densities K, where K is such that for each $\sigma > c_0 > 0$ and $\zeta \in R$ there exists a sequence of densities $p^{(j)}_{\sigma,\zeta} \in K$ with $p^{(j)}_{\sigma,\zeta} \to p_{\sigma,\zeta}$ a.e. as $j \to \infty$.
Theorem 4 is no longer applicable because σ and ζ are restricted and so Lemma 3 cannot be applied.
Let ϕ be any unbiased level α test of H against K. Unbiasedness of ϕ implies

$$\int \phi(z)p_{\sigma,\zeta}(z)dz \le \alpha$$

for all $p_{\sigma,\zeta} \in H$, and

$$\int \phi(z)p(z)dz \ge \alpha$$

for all $p \in K$.
Let $p_{\sigma,\zeta} \in H$. Then there exists a sequence $p^{(j)}_{\sigma,\zeta} \in K$ such that $p^{(j)}_{\sigma,\zeta} \to p_{\sigma,\zeta}$ a.e. as $j \to \infty$. Since $(\phi(z)-1)p^{(j)}_{\sigma,\zeta}(z) \le 0$ Fatou's lemma yields

$$\limsup_{j\to\infty} \int (\phi(z)-1)p^{(j)}_{\sigma,\zeta}(z)dz \le \int \limsup_{j\to\infty} (\phi(z)-1)p^{(j)}_{\sigma,\zeta}(z)dz.$$

The left hand side of this inequality is greater than or equal to $\alpha-1$

while the right hand side equals $\int (\phi(z)-1)p_{\sigma,\zeta}(z)dz$ which is smaller than or equal to $\alpha-1$.

Hence $\int \phi(z)p_{\sigma,\zeta}(z)dz = \alpha$ for all $p_{\sigma,\zeta} \in H$. Now let

$$\psi(z) = \frac{1}{N_1!\ldots N_c!} \sum_{z'\in S(z)} \phi(z'),$$

then

$$\alpha = \int \phi(z)p_{\sigma,\zeta}(z)dz = \int \psi(z)p_{\sigma,\zeta}(z)dz =$$

$$= \int \psi(z)\frac{1}{(\sqrt{2\pi}\,\sigma)^N} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{c}\sum_{j=1}^{N_i}(z_{ij}-\zeta_{ij})^2\right]dz$$

for all $\sigma > c_0 > 0$ and $\zeta \in R$. Since $\alpha > c_0 > 0$ and R contains a rectangle, application of Theorem 1 of Chapter 4 yields the completeness of this last family. Therefore $\psi(z) = \alpha$ a.e. as was to be proved.

Problem 30.

Let $G = \{g_1,\ldots,g_r\}$ be a group of orthogonal transformations of $\mathbb{R}^N$, let ϕ be an critical function and

$$\psi(z) = \frac{1}{r}\sum_{z'\in S(z)} \phi(z') = \frac{1}{r}\sum_{k=1}^{r} \phi(g_k(z)).$$

If (78) on p. 207 does not hold, there exists $\eta > 0$ such that $\psi(z) > \alpha+\eta$ on a set A of positive measure. By Lemma 3 there exists $\sigma > 0$ and $\zeta = (\zeta_1,\ldots,\zeta_N)$ such that $P\{Z \in A\} \geq 1-\eta$ when $Z_1,\ldots,Z_N$ are independently normally distributed with common variance σ^2 and means $EZ_i = \zeta_i$, $i = 1,2,\ldots,N$.

We have for $k = 1,2,\ldots,r$

$$p_{\sigma,\zeta}(g_k(z) = \frac{1}{r}\sum_{l=1}^{r}\frac{1}{(\sqrt{2\pi}\,\sigma)^N}\exp\left(-\frac{1}{2\sigma^2}\,|g_k(z) - g_l(\zeta)|^2\right) =$$

$$= \frac{1}{r}\sum_{l=1}^{r}\frac{1}{(\sqrt{2\pi}\,\sigma)^N}\exp\left(-\frac{1}{2\sigma^2}\,|g_k(r) - g_k g_k^{-1} g_l(\zeta)|^2\right) =$$

$$= \frac{1}{r}\sum_{l=1}^{r}\frac{1}{(\sqrt{2\pi}\,\sigma)^N}\exp\left(-\frac{1}{2\sigma^2}\,|g_k(z - g_k^{-1} g_l(\zeta))|^2\right) =$$

$$= \frac{1}{r}\sum_{l=1}^{r}\frac{1}{(\sqrt{2\pi}\,\sigma)^N}\exp\left(-\frac{1}{2\sigma^2}\,|z - g_k^{-1} g_l(\zeta)|^2\right) =$$

$$= \frac{1}{r}\sum_{m=1}^{r}\frac{1}{(\sqrt{2\pi}\,\sigma)^N}\exp\left(-\frac{1}{2\sigma^2}\,|z - g_m(\zeta)|^2\right) = p_{\sigma,\zeta}(z).$$

The third and fourth equalities hold because g_k is a linear and orthogonal transformation; the fifth equality holds because every g_m, $m = 1,2,\dots,r$ can be written as $g_k^{-1} g_1$ for some $1 \in \{1,2,\dots,r\}$. (Take $g_1 = g_k g_m$). Hence

$$\int_{\mathbb{R}^N} \psi(z) p_{\sigma,\zeta}(z)dz = \frac{1}{r} \sum_{k=1}^{r} \int_{\mathbb{R}^N} \phi(g_k(z)) p_{\sigma,\zeta}(z)dz =$$

$$= \frac{1}{r} \sum_{k=1}^{r} \int_{\mathbb{R}^N} \phi(g_k(z)) p_{\sigma,\zeta}(g_k(z))dz.$$

Applying the transformation $\tilde{z} = g_k(z)$ this leads to

$$\int_{\mathbb{R}^N} \psi(z) p_{\sigma,\zeta}(z)dz = \int_{\mathbb{R}^N} \phi(z) p_{\sigma,\zeta}(z)dz.$$

Further, since $\psi(g_m(z)) = \frac{1}{r} \Sigma_{k=1}^{r} \phi(g_k g_m(z)) = \frac{1}{r} \Sigma_{1=1}^{r} \phi(g_1(z)) = \psi(z)$, $m = 1,2,\dots,r$ it follows that

$$\int_{\mathbb{R}^N} \psi(z) p_{\sigma,\zeta}(z)dz =$$

$$= \frac{1}{r} \sum_{1=1}^{r} \int_{\mathbb{R}^N} \psi(z) \frac{1}{(\sqrt{2\pi}\,\sigma)^N} \exp\left(-\frac{1}{2\sigma^2} |z - g_1(\zeta)|^2\right)dz =$$

$$= \frac{1}{r} \sum_{1=1}^{r} \int_{\mathbb{R}^N} \psi(g_1(\tilde{z})) \frac{1}{(\sqrt{2\pi}\,\sigma)^N} \exp\left(-\frac{1}{2\sigma^2} |g_1(\tilde{z}) - g_1(\zeta)|^2\right)d\tilde{z} =$$

$$= \frac{1}{r} \sum_{1=1}^{r} \int_{\mathbb{R}^N} \psi(\tilde{z}) \frac{1}{(\sqrt{2\pi}\,\sigma)^N} \exp\left(-\frac{1}{2\sigma^2} |g_1(\tilde{z} - \zeta)|^2\right)d\tilde{z} =$$

$$= \frac{1}{r} \sum_{1=1}^{r} \int_{\mathbb{R}^N} \psi(\tilde{z}) \frac{1}{(\sqrt{2\pi}\,\sigma)^N} \exp\left(-\frac{1}{2\sigma^2} |\tilde{z} - \zeta|^2\right)d\tilde{z} \geq$$

$$\geq (\alpha + \eta)(1 - \eta) > \alpha$$

and hence $\int_{\mathbb{R}^N} \phi(z) p_{\sigma,\zeta}(z)dz > \alpha$.
This proves that (77) on p. 207 implies (78) on p. 207.

Problem 31.

Let ϕ_0 be any level α test. By the preceding problem the average value of ϕ_0 over each set $S(z)$ is $\leq \alpha$. On the sets for which this inequality is strict, one can increase ϕ_0 to obtain a critical function ϕ satisfying (79) on p. 207, and such that $\phi_0(z) \leq \phi(z)$ for all z. Since against all alternatives the power of ϕ is at least that of ϕ_0, this establishes the result. An explicit construction of ϕ similar to the one in Problem 28 is possible.

Problem 32.

(i) Since the marginal distribution of X is normal with mean ξ and variance σ^2 we have

$$p^{Y|x}(y) = \frac{p^{X,Y}(x,y)}{p^X(x)} = (2\pi\sigma\tau\sqrt{1-\rho^2})^{-1} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{1}{\sigma^2}(x-\xi)^2 + \right.\right.$$

$$\left.\left. -\frac{2\rho}{\sigma\tau}(x-\xi)(y-\eta) + \frac{1}{\tau^2}(y-\eta)^2\right\}\right] \Big/ \left[(\sqrt{2\pi}\,\sigma)^{-1}\exp\left\{-\frac{1}{2\sigma^2}(x-\xi)^2\right\}\right] =$$

$$= (2\pi\tau\sqrt{1-\rho^2})^{-1}\exp\left[-\frac{1}{2\tau^2(1-\rho^2)}(y-\eta-\frac{\rho\tau}{\sigma}(x-\xi))^2\right]$$

so Y given x has the normal distribution with mean $\eta + \rho\cdot\frac{\tau}{\sigma}(x-\xi)$ and variance $\tau^2(1-\rho^2)$.

(ii) If $v_i = (x_i - \bar{x})/\sqrt{\Sigma\,(x_j-\bar{x})^2}$ so that $\Sigma_{i=1}^n v_i = 0$ and $\Sigma_{i=1}^n v_i^2 = 1$ the statistic R can be written as

$$R = \frac{\Sigma\,(x_i-\bar{x})(Y_i-\bar{Y})}{\sqrt{\Sigma\,(x_i-\bar{x})^2\,\Sigma\,(Y_i-\bar{Y})^2}} = \frac{\Sigma\, v_iY_i}{\sqrt{\Sigma\, Y_i^2 - n\bar{Y}^2}}$$

so that

$$T = \frac{\sqrt{n-2}\,R}{\sqrt{1-R^2}} = \frac{\Sigma\, v_iY_i}{\sqrt{[\Sigma\, Y_i^2 - n\bar{Y}^2 - (\Sigma\, v_iY_i)^2]/(n-2)}}$$

where all summations are taken over $i = 1,2,\dots,n$.
The distribution of this statistic is seen to be independent of η and τ^2 so one can assume $\eta = 0$ and $\tau^2 = 1$.
Consider the following transformation:

$$Z_1 = \sum_{i=1}^n \frac{1}{\sqrt{n}} Y_i = \sum_{i=1}^n a_{1i}Y_i$$

$$Z_2 = \sum_{i=1}^n v_iY_i = \sum_{i=1}^n a_{2i}Y_i.$$

Then $\Sigma_{i=1}^n a_{1i}^2 = \Sigma_{i=1}^n a_{2i}^2 = 1$ and $\Sigma_{i=1}^n a_{1i}a_{2i} = 0$ so we can construct an orthogonal transformation from $Y_1,\dots,Y_n$ to $Z_1,\dots,Z_n$ with Z_1 and Z_2 as above.
Since $Y_1,\dots,Y_n$ given $x_1,\dots,x_n$ are independently standard normally distributed, by Problem 6 we have that $Z_1,\dots,Z_n$ given $x_1,\dots,x_n$ are independently standard normally distributed.

Because $\Sigma_{i=1}^n Y_i^2 - n\bar{Y}^2 - \left(\Sigma_{i=1}^n v_i Y_i\right)^2 = \Sigma_{1=3}^n Z_i^2$ we have

$$T = \frac{Z_2}{\sqrt{\Sigma_{i=3}^n Z_i^2/(n-2)}} .$$

Given $x_1,\dots,x_n$ the numerator is standard normally distributed and the denominator has a χ^2-distribution with n-2 degrees of freedom divided by n-2. Furthermore both are independent, so given $x_1,\dots,x_n$ T has a t-distribution with n-2 degrees of freedom. Since the conditional distribution of T doesn't depend on $x_1,\dots,x_n$ the result also holds for the unconditional distribution of T.

(iii) For $-1 < r < 1$

$$P\{R \le r\} = P\left\{\frac{\sqrt{n-2}\,R}{\sqrt{1-R^2}} \le \frac{\sqrt{n-2}\,r}{\sqrt{1-r^2}}\right\} =$$

$$= \int_{-\infty}^{\frac{\sqrt{n-2}\,r}{\sqrt{1-r}}} (\pi(n-2))^{-\frac{1}{2}} \frac{\Gamma(\frac{1}{2}(n-1))}{\Gamma(\frac{1}{2}(n-2))} \left(1 + \frac{Y^2}{n-2}\right)^{-\frac{1}{2}(n-1)} dy$$

So $p(r) = \frac{d}{dr} p\{R \le r\} = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{1}{2}(n-1))}{\Gamma(\frac{1}{2}(n-1))} (1-r^2)^{\frac{1}{2}n-2}$.

<u>Problem 33.</u>

(i) Let $(X_1,Y_1),\dots,(X_n,Y_n)$ be a sample from the bivariate normal distribution (70) on p. 197 with $|\rho| \neq 1$.
Consider the hypothesis $H : \frac{\tau}{\sigma} = \Delta$. Making a transformation $U_i = \Delta X_i + Y_i$, $V_i = X_i - \frac{1}{\Delta} Y_i$, $i = 1,2,\dots,n$ the variables $(U_1,V_1),\dots,(U_n,V_n)$ are independently identically bivariate normally distributed with $\text{cov}(U_1,V_1) = \Delta\sigma^2 - \frac{1}{\Delta}\tau^2$. Hence the hypothesis $H : \frac{\tau}{\sigma} = \Delta$ is equivalent to
$H : \text{cov}(U_1,V_1) = 0$.
By (73) on p. 198 the UMP unbiased test for testing H against the alternative $K : \text{cov}(U_1,V_1) \neq 0$ has acceptance region

$$\frac{|R|}{\sqrt{(1-R^2)(n-2)}} \le C$$

or equivalently $|R| \le C_0$, where

$$R = \frac{\Sigma_{i=1}^n (U_i - \bar{U})(V_i - V)}{\sqrt{\Sigma_{i=1}^n (U_i - \bar{U})^2 \, \Sigma_{i=1}^n (V_i - \bar{V})^2}} = \frac{\Delta^2 S_1^2 - S_2^2}{\sqrt{(\Delta^2 S_1^2 + S_2^2)^2 - 4\Delta^2 S_{12}}}$$

with

$$S_1^2 = \sum_{i=1}^n (X_i - \overline{X})^2, \quad S_2^2 = \sum_{i=1}^n (Y_i - \overline{Y})^2, \quad S_{12} = \sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})$$

Under H the probability density of R is given by (80) on p. 208 (cf. Problem 22).

(ii) Assume $\tau = \sigma$. Making a transformation $U_i' = Y_i + X_i$, $V_i = Y_i - X_i$, $i = 1,2,\dots,n$ the variables $(U_1',V_1'),\dots,(U_n',V_n')$ are independent and bivariate normally distributed with means $EU_i' = \eta + \xi$, $EV_i' = \eta - \xi$ and covariance matrices $2\sigma^2(1+\rho)I$.
Because the transformation $\gamma = \eta - \xi$, $\delta = \eta + \xi$ is 1-1 onto and $(U_1',\dots,U_n')$ and $(V_1',\dots,V_n')$ are independent $(V_1',\dots,V_n')$ is sufficient for γ.
Hence, by Problem 7, the UMP unbiased test for testing $H : \xi = \eta$ (or equivalently $H : \gamma = 0$) is given by the acceptance region

$$|T| \le C$$

where

$$T = \frac{\overline{V}'}{\sqrt{\Sigma_{i=1}^n (V_i' - \overline{V}')^2}} = \frac{\overline{Y} - \overline{X}}{\sqrt{S_1^2 + S_2^2 - 2S_{12}}}.$$

From Problem 3 (ii) it follows that under H the statistic $T\sqrt{n(n-1)}$ has a t-distribution with n-1 degrees of freedom.

Note that the test statistic T differs slightly from the test statistic given in the problem.

(HSU (1940), MORGAN (1939), PITMAN (1939))

Problem 34.

(i) Make an orthogonal transformation $X' = AX$, $A = (a_{ij})$, $i,j = 1,2,\dots,n$ such that $a_{nj} = \frac{1}{\sqrt{n}}$, $j = 1,\dots,n$ and apply the same transformation to $Y : Y' = AY$. Now we have the orthogonal transformation

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

By orthogonality it follows that the covariance matrix of $\binom{X'}{Y'}$ is the same as the of $\binom{X}{Y}$. Therefore the pairs of variables $(X_1,Y_1),\ldots,(X_n,Y_n)$ are independently bivariate normally distributed with the same covariance matrix as that of (X,Y) and with means

$$EX_i = \sum_{j=1}^{n} a_{ij}\xi = \sqrt{n}\,\xi \sum_{j=1}^{n} a_{ij}a_{nj} = 0,\ i=1,2,\ldots,n-1$$

and similarly $EY_i = 0,\ i=1,2,\ldots,n-1$.
By orthogonality $\Sigma_{i=1}^{n} X_i^2 = \Sigma_{i=1}^{n} (X_i')^2$ and hence

$$\sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 = \sum_{i=1}^{n} (X_i')^2 - (X_n')^2 = \sum_{i=1}^{n-1} (X_i')^2$$

and similarly

$$\sum_{i=1}^{n} (Y_i - \overline{Y})^2 = \sum_{i=1}^{n-1} (Y_i')^2 \quad \text{and} \quad \sum_{i=1}^{n} (x_i - \overline{X})(Y_i - \overline{Y}) = \sum_{i=1}^{n-1} X_i'Y_i' .$$

So we have

$$S_1^2 = \sum_{i=1}^{n-1} (X_i')^2,\ S_2^2 = \sum_{i=1}^{n-1} (Y_i')^2 \text{ and } S_{12} = \sum_{i=1}^{n-1} X_i'Y_i' .$$

Since $(X_n',Y_n') = \sqrt{n}(\overline{X},\overline{Y})$ the result follows.

(ii) Following the hint we see that

$$P\{S_{12} \le s_{12},\ S_2^2 \le s_2^2 \mid x_1,\ldots,x_m\} =$$

$$= P\{s_1 Z_1 \le s_{12},\ (S_2^2 - Z_1^2) + Z_1^2 \le s_2^2 \mid x_1,\ldots,x_m\} =$$

$$= \int_{-\infty}^{s_{12}s_2^{-1}} \left(\int_0^{s_2^2 - x^2} f(x)\chi_{m-1}^2(y)\,dy\right) dx$$

where f and χ_{m-1}^2 are the density functions of Z_1 (standard normal) and $S_2^2 - Z_1^2$ (chi-square with m-1 degrees of freedom) respectively.
Therefore the joint density of S_{12} and S_2^2 given $x_1,\ldots,x_m$ is given by

$$p^{S_{12},S_2^2|x}(s_{12},s_2^2) = s_1^{-1} f(s_{12}s_1^{-1})\chi_{m-1}^2(s_2^2 - s_{12}^2 s_1^{-2}) = p^{S_{12},S_2^2|s_1}.$$

The last equality holds because the conditional distribution depends on the x's only through s_1^2. The joint density of S_1^2, S_{12} and S_2^2 is therefore found by multiplying the above conditional density by the marginal density of S_1^2, which is χ_m^2:

$$p^{S_1^2,S_{12},S_2^2}(s_1^2,s_{12},s_2^2) = \chi_m^2(s_1^2)s_1^{-1}d(s_{12}s_1^{-1})\chi_{m-1}^2(s_2^2 - s_{12}^2 s_1^{-2}) =$$

$$= \begin{cases} \{4\pi\Gamma(m-1)\}^{-1}(s_1^2 s_2^2 - s_{12}^2)^{\frac{1}{2}(m-3)} \exp\{-\frac{1}{2}(s_1^2 + s_2^2)\} & \text{for } s_{12}^2 \le s_1^2 s_2^2 \\ 0 & \text{elsewhere.} \end{cases}$$

(iii) If $(X',Y') = (X_1',Y_1';\ldots;X_m',Y_m')$ is a sample from the bivariate normal distribution (70) on p. 197 with $\xi = \eta = 0$ then $T = (\Sigma_{i=1}^m (X_i')^2, \Sigma_{i=1}^m X_i'Y_i', \Sigma_{i=1}^m (Y_i')^2)$ is sufficient for $\theta = (\sigma,\rho,\tau)$ and the probability density of (X',Y') can be written as $g_\theta(t)$.
Now for any Borel set $\mathcal{B}$ we have by Lemma 2 of Chapter 2

$$P_\theta^T(\mathcal{B}) = P\{(X',Y') \in T^{-1}(\mathcal{B})\} = \int_{T^{-1}(\mathcal{B})} g_\theta(T(x,y))dxdy =$$

$$= \int_{T^{-1}(\mathcal{B})} g_{\theta_0}(T(x,y)\cdot g_\theta(T(x,y))/g_{\theta_0}(T(x,y))dxdy =$$

$$= \int_{T^{-1}(\mathcal{B})} g_\theta(T(x,y))/g_{\theta_0}(T(x,y))dP_{\theta_0}^{(X',Y')}(x,y) =$$

$$= \int_{\mathcal{B}} g_\theta(t)/g_{\theta_0}(t)dP_{\theta_0}^T(t).$$

From part (ii) we know that $P_{(1,0,1)}^T$ is absolutely continuous with respect to Lebesgue measure so that for all θ

$$p_\theta^T(t) = p_{(1,0,1)}^T(t)g_\theta(t)/g_{(1,0,1)}(t)$$

where p_θ^T is the density of P_θ^T with respect to Lebesgue measure.
Using part (i) with $m = n-1$ and part (ii) it follows that

$$p_\theta^T(t) = p_\theta^T(s_1^2,s_{12},s_2^2) =$$
$$= \frac{(s_1^2 s_2^2 - s_{12}^2)^{\frac{1}{2}(n-4)}}{4\pi\Gamma(n-2)(\sigma\tau\sqrt{1-\rho^2})^{n-1}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{s_1^2}{\sigma^2} - \frac{2\rho s_{12}}{\sigma\tau} + \frac{s_2^2}{\tau^2}\right)\right\}$$

if $s_{12}^2 \le s_1^2 s_2^2$ and $p_\theta^T(t) = 0$ elsewhere.

Problem 35.

Let $(X_1,Y_1),\ldots,(X_n,Y_n)$ be a sample from the bivariate normal distribution given by (70) on p. 197. As in the preceding problems let

$S_1^2 = \Sigma_{i=1}^n (X_i - \overline{X})^2$, $S_{12} = \Sigma_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})$ and $S_2^2 = \Sigma_{i=1}^n (Y_i - \overline{Y})^2$. The joint density of S_1^2, S_{12} and S_2^2 is given in Problem 34 (iii).
The sample correlation coefficient is defined by $R = S_{12}/\sqrt{S_1^2 S_2^2}$ so the joint density of (S_1^2, S_2^2, R) is given for $r^2 \le 1$ by

$$\frac{(s_1^2 s_2^2 - r^2 s_1^2 s_2^2)^{\frac{1}{2}(n-4)}}{4\pi\Gamma(n-2)(\sigma\tau\sqrt{1-\rho^2})^{n-1}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{s_1^2}{\sigma^2} - \frac{2\rho\sqrt{s_1^2 s_2^2}\, r}{\sigma\tau} + \frac{s_2^2}{\tau^2}\right\}\right] \cdot \sqrt{s_1^2 s_2^2}$$

because the Jacobian of this transformation equals $\sqrt{s_1^2 s_2^2}$. This density equals (for $r^2 \le 1$)

$$\frac{(1-r^2)^{\frac{1}{2}(n-4)} (s_1^2 s_2^2)^{\frac{1}{2}(n-3)}}{4\pi\Gamma(n-2)(\sigma\tau\sqrt{1-\rho^2})^{n-1}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{s_1^2}{\sigma^2} + \frac{s_2^2}{\tau^2}\right\}\right] \cdot \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{\rho r\sqrt{s_1^2 s_2^2}}{\sigma\tau(1-\rho^2)}\right)^k .$$

Integrating with respect to s_1^2 and s_2^2 gives the marginal density of R:

$$\frac{(1-r^2)^{\frac{1}{2}(n-4)}}{4\pi\Gamma(n-2)(\sigma\tau\sqrt{1-\rho^2})^{n-1}} \sum_{k=0}^{\infty} \left(\frac{\rho r}{\sigma\tau(1-\rho^2)}\right)^k \frac{1}{k!} \times$$

$$\times \int_0^\infty\int_0^\infty \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{s_1^2}{\sigma^2} + \frac{s_2^2}{\tau^2}\right\}\right](s_1^2)^{\frac{1}{2}(n-3+k)} (s_2^2)^{\frac{1}{2}(n-3+k)} ds_1 ds_2 =$$

$$= \frac{(1-r^2)^{\frac{1}{2}(n-4)}}{4\pi\Gamma(n-2)} (\sqrt{1-\rho^2})^{n-1} 2^{n-1} \times$$

$$\times \sum_{k=0}^{\infty} \frac{(2\rho r)^k}{k!} \int_0^\infty\int_0^\infty \exp\{-u_1 - u_2\}(u_1 u_2)^{\frac{1}{2}(n-3+k)} du_1 du_2 =$$

$$= \frac{(1-r^2)^{\frac{1}{2}(n-4)}}{\pi(n-3)!}(1-\rho^2)^{\frac{1}{2}(n-1)} 2^{n-3} \sum_{k=0}^{\infty} \frac{(2\rho r)^k}{k!} \Gamma^2(\tfrac{1}{2}(n+k-1))$$

for $r^2 \le 1$, which was to be proved.
To see the equivalence of (82) and (83) on p. 210 we expand

$$\frac{1}{(1-\rho r t)^{n-1}} = \sum_{k=0}^{\infty} \frac{(n-1)\ldots(n+k-2)}{k!}(\rho r t)^k = \sum_{k=0}^{\infty} \frac{\Gamma(n+k-1)}{\Gamma(n-1)} \frac{(\rho r t)^k}{k!}$$

so that

$$\int_0^1 \frac{t^{n-2}}{(1-\rho r t)^{n-1}} \frac{1}{\sqrt{1-t}}\, dt = \int_0^1 \tfrac{1}{2} \frac{u^{\frac{1}{2}(n-3)}}{(1-\rho r\sqrt{u})^{n-1}} \frac{1}{\sqrt{1-u}}\, du =$$

$$= \int_0^1 \sum_{k=0}^{\infty} \tfrac{1}{2} \frac{(\rho r)^k}{k!} \frac{\Gamma(n+k-1)}{\Gamma(n-1)} u^{\frac{1}{2}(n+k-3)} \frac{1}{\sqrt{1-u}}\, du =$$

$$= \sum_{k=0}^{\infty} \tfrac{1}{2} \frac{\Gamma(n+k-1)}{\Gamma(n-1)} \frac{(\rho r)^k}{k!} B(\tfrac{1}{2}(n+k-1),\tfrac{1}{2}) =$$

$$= \sum_{k=0}^{\infty} \tfrac{1}{2} \frac{\Gamma(n+k-1)}{\Gamma(n-1)} \frac{(\rho r)^k}{k!} \frac{\Gamma(\frac{1}{2}(n+k-1)\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}(n+k))} =$$

$$= \sum_{k=0}^{\infty} \tfrac{1}{2} \frac{(2\pi)^{-\frac{1}{2}} 2^{n+k-\frac{3}{2}} \Gamma(\frac{1}{2}(n+k-1))\Gamma(\frac{1}{2}(n+k))}{\Gamma(n-1)} \frac{(\rho r)^k}{k!} \frac{\Gamma(\frac{1}{2}(n+k-1)\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}(n+k)} =$$

$$= \sum_{k=0}^{\infty} 2^{n+k-3} \frac{\Gamma^2(\frac{1}{2}(n+k-1))}{\Gamma(n-1)} \frac{(\rho r)^k}{k!}.$$

The fifth equation holds because $\Gamma(2x) = (2\pi)^{-\frac{1}{2}} 2^{2x-\frac{1}{2}} \Gamma(x)\Gamma(x+\frac{1}{2})$ and the last equation because $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Substituting this expression in (83) on p. 210 yields the required result.

Making a transformation $t = \dfrac{1-v}{1-\rho r v}$ we have

$$\int_0^1 \frac{t^{n-2}}{(1-\rho r t)^{n-1}} \frac{1}{\sqrt{1-t^2}}\, dt =$$

$$= \int_1^0 \frac{(1-v)^{n-2}}{(1-\rho r v)^{n-2}} \frac{(1-\rho r v)^{n-1}}{(1-\rho r)^{n-1}} \frac{(1-\rho r v)}{\sqrt{(2-\rho r v-v)(v-\rho r v)}} \frac{\rho r-1}{(1-\rho r v)^2}\, dv =$$

$$= \int_0^1 \frac{1}{(1-\rho r)^{n-3/2}} \frac{(1-v)^{n-2}}{\sqrt{2v}} \frac{1}{\sqrt{1-\frac{1}{2}v(1+\rho r)}}\, dv.$$

Using the expansion $\dfrac{1}{\sqrt{1-\frac{1}{2}v(1+\rho r)}} = \sum_{k=0}^{\infty} \dfrac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})} \dfrac{v^k}{k!} \left(\dfrac{1+\rho r}{2}\right)^k$

this equals

$$\frac{1}{(1-\rho r)^{n-3/2}} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})} \frac{1}{\sqrt{2}\, k!} \left(\frac{1+\rho r}{2}\right)^k \int_0^1 (1-v)^{n-2} v^{k-\frac{1}{2}} dv =$$

$$= \frac{\Gamma(n-1)}{\Gamma(n-\frac{1}{2})} \frac{1}{(1-\rho r)^{n-3/2}} \sqrt{\pi/2} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})} \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n+k-\frac{1}{2})} \frac{1}{k!} \left(\frac{1+\rho r}{2}\right)^k.$$

Substituting this expression in (83) on p. 210 yields the required result.

Problem 36.

If X and Y have a bivariate normal distribution given by (70) on p. 197 with $\rho > 0$, then Y given x has a normal distribution with mean $\eta + \rho\tau\sigma^{-1}(x-\xi)$ and variance $\tau^2(1-\rho^2)$ (cf. Problem 32).

Let Y_x denote the random variable Y given x. Then it follows that the distribution of $Y_{x'}$ is the same as that of $Y_x + \rho\tau\sigma^{-1}(x' - x)$. Since $\rho > 0$ we have for $x < x'$ $Y_{x'}$ stochastically larger than Y_x and hence X and Y are positively dependent.

REFERENCES

ANDERSON, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

CHAPMAN, D.G. (1950). Some two-sample tests. *Ann. Math. Stat.* 21, 601-606.

GIRSHICK, M.A. (1946). Contributions to the theory of sequential analysis I. *Ann. Math. Stat.* 17, 123-143.

HSU, C.T. (1940). On samples from a normal bivariate population. *Ann. Math. Stat.* 11, 410-426.

LEHMANN, E.L. (1947). On optimum tests of composite hypotheses with one constraint. *Ann. Math. Stat.* 18, 473-494.

MORGAN, W.A. (1939). A test for the significance of the differece between the two variances in a sample from a normal bivariate population. *Biometrika* 31, 13-19.

PAULSON, E. (1941). On certain likelihood ratio tests associated with the exponential distribution. *Ann. Math. Stat.* 12, 301-306.

PITMAN, E.J.G. (1939). A note on normal correlation. *Biometrika* 31, 9-12.

STEIN, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Stat.* 16, 243-258.

CHAPTER 6

Section 1

Problem 1.

Let G be a group of measurable transformations (bijections) of $(\mathcal{X},\mathcal{A})$. Suppose that $T : (\mathcal{X},\mathcal{A}) \to (\mathcal{T},\mathcal{B})$ is such that:

(i) T is a measurable transformation from $\mathcal{X}$ onto $\mathcal{T}$;

(ii) for all $g \in G$, for all $x_1,x_2 \in \mathcal{X}$

$$T(x_1) = T(x_2) \Rightarrow T(gx_1) = T(gx_2).$$

For all $g \in G$ define $g^* : \mathcal{T} \to \mathcal{T}$, by $g^*T(x) = T(gx)$.

We will first show that $G^* = \{g^* \mid g \in G\}$ is a group of measurable transformations of $(\mathcal{T},\mathcal{B})$. By assumption (ii) and since $T(\mathcal{X}) = \mathcal{T}$, g^* is a well defined function. Moreover g^* is a bijection. To see this, we notice that:

(a) for all $x,y \in \mathcal{X}$

$$\begin{aligned} g^*T(x) = g^*T(y) &\Rightarrow T(gx) = T(gy) \\ &\Rightarrow T(g^{-1}gx) = T(g^{-1}gy), \quad \text{(by assumption (ii))} \\ &\Rightarrow T(x) = T(y), \end{aligned}$$

(hence g^* is injective)

(b) for all $z \in \mathcal{T}$

$$z = T(y) = T(gx) = g^*T(x), \quad \text{for some } x,y \in \mathcal{X},$$

(hence g^* is surjective).

Let G_1 be the group of all bijections from $\mathcal{T}$ to $\mathcal{T}$. Define φ by

$$\begin{aligned} \varphi : G &\to G_1 \\ g &\mapsto g^*. \end{aligned}$$

Since $G^* = \varphi(G)$ and since φ is a homeomorphism, G^* is a group. (That φ is a homeomorphism follows from the fact that for all $g_1, g_2 \in G$, and for all $T(x) \in \mathcal{T}$

$$\varphi(g_1 g_2)T(x) = (g_1 g_2)^* T(x) = T(g_1 g_2 x) =$$
$$= g_1^* T(g_2 x) = g_1^* g_2^* T(x) = \varphi(g_1)\varphi(g_2)T(x).)$$

Now we show that G^* leaves $\mathcal{P}^T$ invariant, whenever G leaves $\{P_\theta, \theta \in \Omega\}$ invariant (i.e. for all $g \in G$: $\bar{g}\Omega = \Omega$). It is tacitly assumed that for all $\theta_1, \theta_2 \in \Omega$

$$\theta_1 \neq \theta_2 \Rightarrow P_{\theta_1} \neq P_{\theta_2}$$

(otherwise $\bar{g}\theta$ is not uniquely defined). However this does not imply that $\theta_1 \neq \theta_2 \Rightarrow P^T_{\theta_1} \neq P^T_{\theta_2}$ (e.g. if T is a constant function, then all P^T_θ are the same). Let Ω^* be a subset of Ω such that for all θ in Ω there exist precisely one θ^* in Ω^* with the property

"if X has distribution P_θ then T has distribution $P^T_{\theta^*}$"

(the existence of such a set Ω^* follows from the axiom of choice). Now we have

$$\{P^T_\theta, \theta \in \Omega\} = \{P^T_{\theta^*}, \theta^* \in \Omega^*\},$$
$$\text{for all } \theta_1^*, \theta_2^* \in \Omega^* : \theta_1^* \neq \theta_2^* \Rightarrow P^T_{\theta_1^*} \neq P^T_{\theta_2^*}.$$

It remains to show that for all $g^* \in G^*$ we can define a function $\overline{g^*} : \Omega^* \to \Omega^*$ such that

(a) if T has distribution $P^T_{\theta^*}$, then g^*T has distribution $P^T_{\overline{g^*}\theta^*}$,

(b) $\overline{g^*}\Omega^* = \Omega^*$.

Let T have distribution $P^T_{\theta^*}$. Then for all $B \in \mathcal{B}$

$$P\{g^*T \in B\} = P_{\theta^*}\{g^*T(X) \in B\} = P_{\theta^*}\{gX \in T^{-1}(B)\} =$$
$$= P_{\bar{g}\theta^*}\{X \in T^{-1}(B)\}.$$

Hence, if we define $\overline{g^*}$ by

$$\overline{g^*}\theta^* = \theta^{**},$$

where θ^{**} is the unique element in Ω^* with the property

"if X has distribution $P_{\bar{g}\theta^*}$ then T has distribution $P^T_{\theta^{**}}$";

then the first requirement (a) is satisfied; and moreover: $\overline{g^*}\Omega^* \subset \Omega^*$. Conversely, if $\theta^{**} \in \Omega^*$ then $\theta^{**} = \overline{g^*}\theta^*$, with $\theta^* = \overline{(g^{-1})^*}\theta^{**}$ as is easily seen. This proves that $\overline{g^*}\Omega^* = \Omega^*$.

Section 2

Problem 2.

(i) Let $T(x) = (\operatorname{sgn} x_n, x_1/x_n, \ldots, x_{n-1}/x_n)$. T is evidently invariant under G. If $T(x) = T(y)$ then $\operatorname{sgn} x_n = \operatorname{sgn} y_n$ and hence $x_n/y_n > 0$. Since $T(x) = T(y)$ also implies that $x_i/x_n = y_i/y_n$ for $i = 1,2,\ldots,n-1$, we have $x = cy$, with $c = x_n/y_n$, and hence T is a maximal invariant.

(ii) Let $x,x' \in \mathcal{X}$. Let f be a function which satisfies the conditions stated in the hint, and which in addition maps I_j one to one onto I_j'. Then $f \in G$ and $f(x) = x'$. So for each pair of distinct points $\mathcal{X}$, there exists a $f \in G$ which maps the one onto the other. Hence G is transitive over $\mathcal{X}$.

Problem 3.

(i) Let D be a normal subgroup of G. Since s is maximal invariant, the equality $s(x_1) = s(x_2)$ implies $x_2 = dx_1$ for some $d \in D$. Since D is a normal subgroup there exists for all $e \in E$ an element $d' \in D$ such that $ex_2 = edx_1 = d'ex_1$. By the invariance of s w.r.t. D it follows that $s(ex_2) = s(d'ex_1) = s(ex_1)$.

(ii) We first show that the subgroup D_0 of translations $x' = x+b$ is normal. Let $d_0 \in D_0$ be given by $x' = x + b_o$, and let $g \in G$ be given by $x' = ax + b$. The translation $d_0' \in D_0$ given by $x' = x + ab_0$ satisfies $gd_0 = d_0'g$.

To show that the subgroup D_1 of transformations $x' = ax$ is not normal, let $d_1 \in D_1$ be given by $x' = 2x$ and let $g \in G$ be given by $x' = 2x+1$. If D_1 is normal then there exists a transformation $d_1' \in D_1$ given by $x' = a'x$ such that $gd_1 = d_1'g$, i.e.

$$4x + 1 = a'(2x + 1), \quad \text{for all } x.$$

Since it is impossible to choose an a' satisfying this condition, D_1 can not be normal.

Section 3

Problem 4.

Define $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$

$$(x,y) \mapsto (z,v) = (y-x,y).$$

Since the Jacobian of this transformation equals one, the joint probability density of $Z = Y - X$ and Y is equal to $f(v-z,v)$. Hence Z has probability density

$$h(z) = \int_{-\infty}^{+\infty} f(v-z,v)dv.$$

Since $\int_{-\infty}^{+\infty} h(z)dz = 1$, $h(z)$ is finite a.e.

Problem 5.

(i) The invariance of the testing problem is established easily. It follows from the argument on p. 217 that a maximal invariant is given by $(z_1,\dots,z_{n-2})$, with

$$z_i = (x_i - x_n)/(x_{n-1} - x_n), \quad i = 1,2,\dots,n-2.$$

By Theorem 1 we may restrict attention to tests depending only on $(Z_1,\dots,Z_{n-2})$, with

$$Z_i = (X_i - X_n)/(X_{n-1} - X_n), \quad i = 1,2,\dots,n-2.$$

Defining W and Y by $W = X_{n-1} - X_n$, $Y = X_n$ and expressing $X_1,\dots,X_n$ as functions of $Z_1,\dots,Z_n,W,Y$ we find that the density of $Z_1,\dots,Z_{n-2},W,Y$ satisfies

$$g(z_1,\dots,z_{n-2},w,y) =$$
$$= |w|^{n-2}\cdot \frac{1}{\theta^n}\cdot f\left(\frac{wz_1+y-\xi}{\theta},\dots,\frac{wz_{n-2}+y-\xi}{\theta},\frac{y+w-\xi}{\theta},\frac{y-\xi}{\theta}\right).$$

The marginal density of $Z_1,\dots,Z_{n-2}$ is thus given by

$$g(z_1,\dots,z_{n-2}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(z_1,\dots,z_{n-2},w,y)\,dwdy =$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(\cdot)\,dwdy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(\cdot)\,dwdy.$$

Applying the transformations $\left(\frac{w}{\theta}, \frac{y-\xi}{\theta}\right) \to (w,y)$ and $\left(\frac{w}{\theta}, \frac{y-\xi}{\theta}\right) \to (-w,-y)$, and using the fact that f is even we find

$$g(z_1,\ldots,z_{n-2}) = 2\int_{-\infty}^{+\infty}\int_{0}^{+\infty} w^{n-2} f(wz_1+y,\ldots,wz_{n-2}+y,w+y,y)\,dwdy.$$

Since this density is independent of ξ and θ, the testing problem is reduced to the testing of the simple hypothesis $f = f_0$ against the simple alternative $f = f_1$. By the Neyman-Pearson lemma (Chapter 3, Theorem 1), the most powerful rejection region is given by

$$\int_{-\infty}^{+\infty}\int_{0}^{\infty} w^{n-2} f_1(wz_1+y,\ldots,wz_{n-2}+y,w+y,y)\,dwdy$$
$$> c\int_{-\infty}^{+\infty}\int_{0}^{\infty} w^{n-2} f_0(wz_1+y,\ldots,wz_{n-2}+y,w+y,y)\,dwdy.$$

The result now follows from the substitution $z_i = (x_i - x_n)/(x_{n-1} - x_n)$, $i = 1,\ldots,n-2$; and from the transformations $w = (x_{n-1} - x_n)u$, $y = vx_n + u$.

(ii) Let W denote the $n \times k$-matrix whose (i,j)-th element is given by w_{ij}. Write $x,\tilde{x},\beta,\gamma$ for $(x_1,\ldots,x_n)'$, $(\tilde{x}_1,\ldots,\tilde{x}_n)'$, $(\beta_1,\ldots,\beta_k)'$, $(\gamma_1,\ldots,\gamma_k)'$ respectively. G is the group of transformations $x \mapsto \tilde{x}$, given by $\tilde{x} = x + W\gamma$, $\gamma \in \mathbb{R}^k$.

(A). We first assume that W has full rank k.

Let W_1 denote the $(n-k) \times k$-matrix, consisting of the (n-k) first rows of W. Let W_2 denote the $k \times k$-matrix consisting of the k last rows of W. Without loss of generality, we may assume that W_2 is nonsingular. Given a vector $x = (x_1,\ldots,x_n)'$ we will write $x_{(1)},x_{(2)}$ for $(x_1,\ldots,x_{n-k})'$ and $(x_{n-k+1},\ldots,x_n)'$ respectively. We will use the shorthand notation

$$x = \begin{pmatrix} x_{(1)} \\ x_{(2)} \end{pmatrix}.$$

We first show that a maximal invariant with respect to G is given by

$$t(x) = x_{(1)} - Ax_{(2)},$$

where A is a suitable $(n-k) \times k$-matrix.

Let $\tilde{x} = x + W\gamma$. Then we may write

$$\tilde{x} = \begin{pmatrix} \tilde{x}_{(1)} \\ \tilde{x}_{(2)} \end{pmatrix} = \begin{pmatrix} x_{(1)} + W_1\gamma \\ x_{(2)} + W_2\gamma \end{pmatrix}.$$

t will be invariant if and only if $W_1\gamma - AW_2\gamma = 0$, for all $\gamma \in \mathbb{R}^k$. This condition is equivalent to

$$A = W_1W_2^{-1}.$$

Now suppose that $t(x) = t(x^*)$, i.e. $x_{(1)} - Ax_{(2)} = x^*_{(1)} - Ax^*_{(2)}$. Put $\gamma = W_2^{-1}(x^*_{(2)} - x_{(2)})$. Then we have

$$x^*_{(2)} = x_{(2)} + W_2\gamma;$$
$$x_{(1)} - x^*_{(1)} = A(x_{(2)} - x^*_{(2)}) = -AW_2\gamma = -W_1\gamma,$$

so

$$x^*_{(1)} = x_{(1)} + W_1\gamma.$$

It follows that t is maximal invariant. By Theorem 1 we can restrict attention to tests depending only on

$$T = X_{(1)} - W_1W_2^{-1}X_{(2)}$$
$$= (T_1,\dots,T_{n-k})'.$$

Then density of T will now be derived. Put $Y = X_{(2)}$. Then we have

$$\begin{pmatrix} X_{(1)} \\ X_{(2)} \end{pmatrix} = \begin{pmatrix} T + W_1W_2^{-1}Y \\ Y \end{pmatrix}.$$

The Jacobian of this transformation equals one. Since the density of $(X_1,\dots,X_n)$ is given by $f(z_1,\dots,z_n)$, with $z = (z_1,\dots,z_n)' = x - W\beta$, the density of $(T_1,\dots,T_{n-k},Y_1,\dots,Y_k)$ is given by

$$h(t_1,\dots,t_{n-k},y_1,\dots,y_k) = f(z^*_1,\dots,z^*_n),$$

with

$$\begin{pmatrix} z^*_{(1)} \\ z^*_{(2)} \end{pmatrix} = \begin{pmatrix} t + W_1W_2^{-1}y - W_1\beta \\ y - W_2\beta \end{pmatrix} = \begin{pmatrix} t + W_1[W_2^{-1}y - \beta] \\ W_2[W_2^{-1}y - \beta] \end{pmatrix}.$$

The marginal density of $(T_1,\dots,T_{n-k})$ is given by

$$g(t_1,\ldots,t_{n-k}) = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(z_1^*,\ldots,z_n^*)dy_1\ldots dy_k,$$

with z^* as above. Transforming to the new variables $u=(u_1,\ldots,u_k)'$ given by

$$u = W_2^{-1}y - \beta,$$

we get

$$g(t_1,\ldots,t_{n-k}) = |\det(W_2)| \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(z_1'',\ldots,z_n'')du_1\ldots du_k,$$

where

$$\begin{pmatrix} z''_{(1)} \\ z''_{(2)} \end{pmatrix} = \begin{pmatrix} t + W_1 u \\ W_2 u \end{pmatrix} .$$

The density g does not depend on $(\beta_1,\ldots,\beta_k)$. Thus the testing problem is reduced to the testing of the simple hypothesis $H : g = g_0$ against the single alternative $K : g = g_1$. By the Neyman-Pearson lemma (Chapter 3, Theorem 1), the most powerful rejection region is given by

$$(1) \qquad \frac{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_1(z_1'',\ldots,z_n'')du_1\ldots du_k}{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_0(z_1'',\ldots,z_n'')du_1\ldots du_k} > C,$$

with z'' given as above. In terms of the original variables, the rejection region is also given by (1), but now

$$\begin{pmatrix} z''_{(1)} \\ z''_{(2)} \end{pmatrix} = \begin{pmatrix} x_{(1)} - W_1 W_2^{-1} x_{(2)} + W_1 u \\ W_2 u \end{pmatrix} .$$

Transforming to new variables $\delta=(\delta_1,\ldots,\delta_k)'$ given by $\delta = W_2^{-1}x_{(2)} - u$, we get (for $i=0,1$)

$$(2) \qquad \begin{aligned} &\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_i(z_1'',\ldots,z_n'')du_1\ldots du_k = \\ &= c\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_i(z_1''',\ldots,z_n''')d\delta_1\ldots d\delta_k \end{aligned}$$

where c is some constant and

$$(3)\qquad \begin{pmatrix} z'''_{(1)} \\ z'''_{(2)} \end{pmatrix} = \begin{pmatrix} x_{(1)} - W_1\delta \\ x_{(2)} - W_2\delta \end{pmatrix} = (x - W\delta).$$

The desired result now follows by substitution of (2) and (3) in (1) and replacement of δ by β troughout.

(B). If W has rank $\ell < k$, then via a linear transformation $(\beta_1,\ldots,\beta_k) \mapsto (\beta_1^*,\ldots,\beta_k^*)$, the density of $(x_1,\ldots,x_n)$ can be written as

$$f(x_1 - \sum_{j=1}^{\ell} w_{1j}^*\beta_j^*,\ldots,x_n - \sum_{j=1}^{\ell} w_{nj}^*\beta_j^*),$$

where $W^* = (w_{ij}^*)$ is an $n\times\ell$-matrix which is explicitly known (as a function of W). It follows from (A) that in this case the most powerful invariant rejection region is given by

$$\frac{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_1(x_1 - \Sigma\, w_{1j}^*\beta_j^*,\ldots,x_n - \Sigma\, w_{nj}^*\beta_j^*)d\beta_1^*\ldots d\beta_\ell^*}{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_0(x_1 - \Sigma\, w_{1j}^*\beta_j^*,\ldots,x_n - \Sigma\, w_{nj}^*\beta_j^*)d\beta_1^*\ldots d\beta_\ell^*} > C.$$

(It should be noticed that this rejection region is most powerful invariant w.r.t. the group of transformations $G^* : x \to x + W^*\gamma^*,\ \gamma^* \in \mathbb{R}^\ell$.)

Problem 6.

For convenience of notation, we will write Δ for the ratio τ/σ. Let Δ_0 be positive. We are dealing with the testing problem: $H : \Delta \le \Delta_0$ vs $K : \Delta > \Delta_0$.

(i) Let φ be a test which is invariant with respect to G. Denote the ordered variables by $X^{(1)} < \ldots < X^{(m)}$, and $Y^{(1)} < \ldots < Y^{(n)}$. Since

$$E[\varphi(X_1,\ldots,X_m,Y_1,\ldots,Y_n) \mid x^{(1)},\ldots,x^{(m)},y^{(1)},\ldots,y^{(n)}] =$$
$$= (m!n!)^{-1}\, \Sigma\, \varphi(x_{i_1},\ldots,x_{i_m},y_{j_1},\ldots,y_{j_n}),$$

where the summation extends over the m!n! permutations of $(x_1,\ldots,x_m)$ and $(y_1,\ldots,y_n)$; $E[\varphi(\cdot) \mid x^{(1)},\ldots,x^{(m)},y^{(1)},\ldots,y^{(n)}]$ is also an invariant test which has the same power function as φ. The transformation

$$U_1 = mX^{(1)};\quad U_i = (m-i+1)[X^{(i)} - X^{(i-1)}],\quad i=2,\ldots,m,$$
$$V_1 = mY^{(1)};\quad V_j = (n-j+1)[Y^{(j)} - Y^{(j-1)}],\quad j=2,\ldots,n,$$

is one to one. Hence attention can be restricted to invariant test based on $U_1,\dots,U_m,V_1,\dots,V_n$. Now it will be shown that attention can further be restricted to invariant test based on the sufficient statistics

$$T_1 = V_1, \quad T_2 = V_1, \quad T_3 = \sum_{i=2}^{m} U_i, \quad T_4 = \sum_{j=2}^{n} V_j.$$

Since $U_1,\dots,U_m$ and $V_1,\dots,V_n$ are independent random variables, following an exponential distribution with parameters σ^{-1} and τ^{-1} respectively (Chapter 2, Problem 13), the joint density of $(U_2,\dots,U_m,V_2,\dots,V_n)$ and the joint density of (T_3,T_4) are given respectively by

$$f(u_2,\dots,u_m,v_2,\dots,v_n) =$$

$$= \sigma^{-(m-1)}\tau^{-(n-1)} \exp\left(-\sum_{i=2}^{m} u_i/\sigma - \sum_{j=2}^{n} v_j/\tau\right); \qquad u_i,v_j \geq 0;$$

$$g(t_3,t_4) = \sigma^{-(m-1)}\tau^{-(n-1)} \frac{t_3^{m-2}}{(m-2)!}\,\frac{t_4^{n-2}}{(n-2)!} \exp(-t_3/\sigma - t_4/\tau);$$

$$t_3,t_4 \geq 0.$$

Suppose that φ is an invariant test based on $U_1,\dots,U_m,V_1,\dots,V_n$. The conditional expectation of φ given $(T_1,T_2,T_3,T_4) = (t_1,t_2,t_3,t_4)$ satisfies

$$\varphi^*(t_1,t_2,t_3,t_4) =$$

$$= E[\varphi(U_1,\dots,U_m,V_1,\dots,V_n) \mid T_1 = t_1,\dots,T_4 = t_4] =$$

$$= E[\varphi(t_1,U_2,\dots,U_m,t_2,V_2,\dots,V_n) \mid T_3 = t_3, T_4 = t_4] =$$

$$= \int_0^\infty \dots \int_0^\infty \varphi(t_1,u_2,\dots,u_{m-1},t_3 - \sum_{i=2}^{m-1} u_i,t_2,v_2,\dots,v_{n-1},t_4 - \sum_{j=2}^{n-1} v_j)$$

$$\cdot\, [f(u_2,\dots,u_{m-1},t_3 - \sum_{i=2}^{m-1} u_i,v_2,\dots,v_{n-1},t_4 - \sum_{j=2}^{n-1} v_j)]/g(t_3,t_4)$$

$$\cdot\, du_2\dots du_{m-1}dv_2\dots dv_{n-1} =$$

$$= (n-2)!(m-2)! \int\dots\int [\varphi(t_1,u_2,\dots,t_3 - \sum_{i=2}^{m-1} u_i,t_2,v_2,\dots,t_4 - \sum_{j=2}^{n-1} v_j)$$

$$\cdot\, t_3^{-(m-2)} t_4^{-(n-2)}]du_2\dots du_{m-1}dv_2\dots dv_{n-1}.$$

(the last integration is on the region $\Sigma_{i=2}^{m-1} u_i \leq t_3$, $\Sigma_{j=2}^{n-1} v_j \leq t_4$). The group G induces the group G_1 of transformations given by

$$T_1' = aT_1 + b_1, \quad T_2' = aT_2 + c_1, \quad T_3' = aT_3, \quad T_4' = aT_4.$$

It is easily seen that invariance of φ w.r.t. G implies invariance of φ^* w.r.t. G_1. Moreover φ and φ^* have the same power function. This proves that attention can be restricted to tests which depend only on (T_1,T_2,T_3,T_4) and which are invariant w.r.t. G_1.

A maximal invariant w.r.t. G_1 is given by $Z = T_4/T_3$ (Example 6). By Theorem 1 attention can be restricted to tests depending on Z. Since $2T_3/\sigma$ and $2T_4/\tau$ are independent random variables, following a χ^2-distribution with 2(m-1) and 2(n-1) degrees of freedom (respectively), $Z \cdot \sigma(m-1)/[\tau(n-1)]$ has an F-distribution with 2(n-1) and 2(m-1) degrees of freedom. Hence the density of Z equals

$$C(\Delta)\frac{z^{n-2}}{(z+\Delta)^{m+n-2}}, \quad z > 0, \quad \Delta = \frac{\tau}{\sigma}.$$

For varying Δ these densities constitute a family with monotone likelihood ratio. Hence among all invariant tests of H there exists a UMP invariant test given by the rejection region

$$Z = \frac{T_4}{T_3} = \frac{\Sigma\ [Y_j - \min\ (Y_1,\ldots,Y_n)]}{\Sigma\ [X_i - \min\ (X_1,\ldots,X_m)]} > C.$$

(ii) In order to construct a UMP unbiased test we first note that attention can be restricted to tests based on the sufficient statistics T_1,T_2,T_3,T_4 (Chapter 5, Problem 22). Under $\theta = (\xi,\eta,\sigma,\tau)$ the joint density of (T_1,T_2,T_3,T_4) is given by

$$f_\theta(t) = \sigma^{-m}\tau^{-n}\exp\left(-\frac{t_1-m\xi}{\sigma}-\frac{t_2-n\eta}{\tau}-\frac{t_3}{\sigma}-\frac{t_4}{\tau}\right)\frac{t_3^{m-2}}{(m-2)!}\,\frac{t_4^{n-2}}{(n-2)!},$$

where $t = (t_1,t_2,t_3,t_4)$ satisfies $t_1 \ge m\xi$, $t_2 \ge n\eta$, $t_3 \ge 0$, $t_4 \ge 0$. Now it will be shown that the power function of every test is continuous (then Lemma 1 of Chapter 5 will be applied). Let $\{\theta_j\} = \{(\xi_j,\eta_j,\sigma_j,\tau_j)\}$ be a sequence of parameter values such that $\lim_{j\to\infty} \theta_j = \theta_0 = (\xi_0,\eta_0,\sigma_0,\tau_0)$. Let $\delta = \min(\sigma_0,\tau_0)/2$. For all j large enough we have: $|\xi_j - \xi_0| < \delta$, $|\eta_j - \eta_0| < \delta$, $|\sigma_j - \sigma_0| < \delta$, $|\tau_j - \tau_0| < \delta$. Hence $t_1 \ge m\xi_j$ implies

$$-\frac{t_1-m\xi_j}{\sigma_j} \le -\frac{t_1-m\xi_j}{\sigma_0+\delta} \le -\frac{t_1-m(\xi_0+\delta)}{\sigma_0+\delta} = -\frac{t_1-m(\xi_0-\delta)}{\sigma_0+\delta}+\frac{2m\delta}{\sigma_0+\delta}.$$

Moreover, if $t_1 \ge m\xi_j$ then $t_1 \ge m(\xi_0-\delta)$. Therefore, if φ is some critical function, then

(4) $$0 \le \varphi(t) f_{\theta_j}(t) \le f_{\theta_j}(t) \le \left(\frac{\sigma_0+\delta}{\sigma_0-\delta}\right)^m \left(\frac{\tau_0+\delta}{\tau_0-\delta}\right)^n \exp\left(\frac{2m\delta}{\sigma_0+\delta} + \frac{2n\eta}{\tau_0+\delta}\right) f_{\tilde\theta}(t),$$

where $\tilde\theta = (\xi_0-\delta, \eta_0-\delta, \sigma_0+\delta, \tau_0+\delta)$. Since the right hand side of (4) is independent of j and integrable w.r.t. Lebesgue measure, application of the dominated convergence theorem yields

$$\lim_{j\to\infty} E_{\theta_j}\varphi(T) = \lim_{j\to\infty} \int \varphi(t) f_{\theta_j}(t)dt = \int \varphi(t) f_{\theta_0}(t)dt = E_{\theta_0}\varphi(T).$$

By Lemma 1 of Chapter 4 the problem of finding a UMP unbiased level α test is reduced to the problem of finding a level α test which is UMP among all similar tests.

We make the following one to one transformation: $S_1 = T_1$, $S_2 = T_2$, $S_3 = T_4/T_3$, $S_4 = T_3 + T_4/\Delta_0$. Under $\theta = (\xi,\eta,\sigma,\sigma\Delta)$ the joint density of (S_1,S_2,S_3,S_4) is given by

$$g_\theta(s) = g_\theta(s_1,s_2,s_3,s_4) =$$
$$= \Delta_0^{m+n-2} \frac{\sigma^{-m}}{(m-2)!} \frac{(\sigma\Delta)^{-n}}{(n-2)!} \exp\left\{-\frac{s_1-m\xi}{\sigma} - \frac{s_2-n\eta}{\sigma\Delta}\right\}$$
$$\cdot\ s_4^{m+n-3} s_3^{n-2} (s_3+\Delta_0)^{-m-n+2} \exp\left\{-\frac{s_4\Delta_0(s_3+\Delta)}{\sigma\Delta(s_3+\Delta_0)}\right\},$$

where $s_1 \ge m\xi$, $s_2 \ge n\eta$, $s_3 \ge 0$, $s_4 \ge 0$. Similarity of a test φ means that for all $\theta = (\xi,\eta,\sigma,\sigma\Delta_0)$: $\int \varphi(s) g_\theta(s)ds = \alpha$, or equivalently

$$\int_{m\xi}^{\infty} \sigma^{-1} \exp\left(-\frac{s_1-m\xi}{\sigma}\right) h(s_1;\eta,\sigma)ds_1 = \alpha$$

where

$$h(s_1;\eta,\sigma) = \int_{n\eta}^{\infty} (\sigma\Delta_0)^{-1} \exp\left(-\frac{s_2-n\eta}{\sigma\Delta_0}\right) k(s_1,s_2;\sigma)ds_2,$$

$$k(s_1,s_2;\sigma) = \int_0^{\infty} \sigma^{-(m+n-2)} \frac{s_4^{m+n-3}}{(m+n-3)!} \exp\left(-\frac{s_4}{\sigma}\right) \ell(s_1,s_2,s_4)ds_4,$$

$$\ell(s_1,s_2,s_4) = \int_0^{\infty} \varphi(s_1,s_2,s_3,s_4) \frac{\Delta_0^{m-1}(m+n-3)!}{(m-2)!(n-2)!} s_3^{n-2}(s_3+\Delta_0)^{-m-n+2} ds_3.$$

From the completeness of the class of exponential distributions with unknown starting point and fixed parameter σ (Chapter 5, Problem 12(i)), it follows that for all η and σ there exists a Lebesgue null set $A(\eta,\sigma)$ such that

$$s_1 \notin A(\eta,\sigma) \Rightarrow h(s_1;\eta,\sigma) = \alpha.$$

Let $\{(\eta_i,\sigma_i) \mid i \in \mathbb{N}\}$ be a countable dense subset of $\mathbb{R}\times\mathbb{R}_0^+$. Define $A = \bigcup_{i=1}^{\infty} A(\eta_i,\sigma_i)$. Then A is a Lebesgue null set such that $h(s_1,\eta_i,\sigma_i) = \alpha$ for all $s_1 \notin A$ and all $i \in \mathbb{N}$. Since h depends continuously on η and σ (see the proof of the continuity of $E_\theta\varphi(T)$), and since A is a dense subset of $\mathbb{R}\times\mathbb{R}_0^+$, we have

$$h(s_1,\eta,\sigma) = \alpha \text{ for all } s_1 \notin A \text{ and all } (\eta,\sigma) \in \mathbb{R}\times\mathbb{R}_0^+.$$

Applying the completeness argument once again we find that for all $s_1 \notin A$, for all $\sigma \in \mathbb{R}_0^+$ there exists a Lebesgue null set $B(s_1,\sigma)$ such that

$$s_2 \notin B(s_1,\sigma) \Rightarrow k(s_1,s_2;\sigma) = \alpha.$$

Using the continuity of k w.r.t. σ we arrive at the following statement: for all $s_1 \notin A$ there exists a Lebesgue null set $B(s_1)$ such that

$$k(s_1,s_2;\sigma) = \alpha \text{ for all } s_2 \notin B(s_1) \text{ and all } \sigma > 0.$$

Applying now the completeness of gamma distributions yields the following result: for all $s_1 \notin A$, for all $s_2 \notin B(s_2)$ there exists a Lebesgue null set $C(s_1,s_2)$ such that

$$s_4 \notin C(s_1,s_2) \Rightarrow \ell(s_1,s_2,s_4) = \alpha.$$

Define

$$V_1 = A \times \mathbb{R} \times \mathbb{R}^+,$$

$$V_2 = \{(s_1,s_2,s_4) \mid s_1 \notin A,\ s_2 \in B(s_1),\ s_4 \in \mathbb{R}^+\},$$

$$V_3 = \{(s_1,s_2,s_4) \mid s_1 \notin A,\ s_2 \notin B(s_1),\ s_4 \in C(s_1,s_2)\}.$$

Then $\ell(s_1,s_2,s_4) = \alpha$ for all $(s_1,s_2,s_4) \notin V_1 \cup V_2 \cup V_3$. Since V_1,V_2,V_3 are Lebesgue null sets on $\mathbb{R}\times\mathbb{R}\times\mathbb{R}^+$ we have

$$(5) \qquad \int_0^\infty \varphi(s_1,s_2,s_3,s_4)\frac{(m+n-3)!}{(m-2)!(n-2)!}\Delta_0^{m-1}s_3^{n-2}(s_3+\Delta_0)^{-m-n+2}ds_3 = \alpha, \quad \text{a.e.}$$

We have shown that similarity of a test implies (5).

Now let $(\Delta_1,\xi,\eta,\sigma)$ be an arbitrary (fixed) alternative $(\Delta_1 > \Delta_0)$. The power function of a test φ, at this alternative equals

$$(6) \qquad c\int_{m\xi}^{\infty}\int_{n\eta}^{\infty}\int_{0}^{\infty} q(s_1,s_2,s_4,\sigma)\exp\left(-\frac{s_1-m\xi}{\sigma}-\frac{s_2-n\eta}{\sigma\Delta_1}\right)\cdot s_4^{m+n-3}ds_4 ds_2 ds_1,$$

with

$$(7) \qquad q(s_1,s_2,s_4,\sigma) = \int_0^{\infty} \varphi(s_1,s_2,s_3,s_4)\frac{(m+n-3)!}{(m-2)!(n-2)!}\Delta_0^{m-1} s_3^{n-2}(s_3+\Delta_0)^{-m-n+2} \cdot \exp\left(-\frac{s_4}{\sigma}\frac{\Delta_0}{\Delta_1}\frac{s_3+\Delta_1}{s_3+\Delta_0}\right)ds_3.$$

If a test φ_0 satisfies (5) and maximizes q for each fixed s_1,s_2,s_4 (among all tests satisfying (5)) then certainly φ_0 has maximal power at the alternative $(\Delta_1,\xi,\eta,\sigma)$ among all similar tests. A test φ_0 with this property will now be constructed.

The problem of maximizing (7) w.r.t. (5) is identical to the problem of finding (for each fixed s_1,s_2,s_4) a most powerful test of level $\alpha\cdot\exp(-s_4/\sigma)$ for the simple hypothesis $\Delta = \Delta_0$ against the alternative $\Delta = \Delta_1$. By the fundamental lemma of Neyman and Pearson (Chapter 3, Theorem 1) it follows that the test

$$\varphi_0(s_1,s_2,s_3,s_4) = \begin{cases} 1 & \\ \gamma(\Delta_1,\sigma,s_1,s_2,s_4) & \text{if } -\dfrac{s_3+\Delta_1}{s_3+\Delta_0} \gtreqless k(\Delta_1,\sigma,s_1,s_2,s_4) \\ 0 & \end{cases}$$

(where γ and k are such that (2) is satisfied) exists, and maximizes (7) subject to (5). Since $-(s_3+\Delta_1)/(s_3+\Delta_0) = -1-(\Delta_1-\Delta_0)/(s_3+\Delta_0)$ is an increasing function of s_3 when $\Delta_1 > \Delta_0$, the test φ can be written as

$$\varphi_0(s_1,s_2,s_3,s_4) = \begin{cases} 1 & \\ \gamma & \text{if } z = s_3 \gtreqless k' \\ 0 & \end{cases}.$$

Since $\{z = k\}$ is a Lebesgue null set, γ can be taken zero. The number k' is given by (5). Since the test φ_0 does not depend on the particular alternative (i.e. not on Δ_1, nor on σ), it is uniform most powerful among the tests satisfying (5), i.e. it is UMP unbiased. It is identical to the UMP invariant test of (i).

(iii) To extend the results of (i) and (ii) we restrict our attention to

$$U_1 = mX^{(1)}, \quad U_i = (m-i+1)[X^{(i)}-X^{(i-1)}], \quad i=2,\ldots,r \le m,$$

and

$$V_1 = nY^{(1)}, \quad V_j = (n-j+1)[Y^{(j)} - Y^{(j-1)}], \quad j = 2,\dots,s \le n,$$

and proceed in the same way as we did in (i) and (ii). Observing that

$$\sum_{i=2}^{r} U_i = (m-r)[X^{(r)} - X^{(1)}] + \sum_{i=1}^{r} [X^{(i)} - X^{(1)}]$$

we arrive at the following two statements

(i)' for testing $\tau/\sigma \le \Delta_0$ against $\tau/\sigma > \Delta_0$ there exists a UMP invariant test w.r.t. G and its rejection region is given by

$$\frac{(n-s)[Y^{(s)} - Y^{(1)}] + \Sigma_{j=1}^{s} [Y^{(j)} - Y^{(1)}]}{(m-r)[X^{(r)} - X^{(1)}] + \Sigma_{i=1}^{r} [X^{(i)} - X^{(1)}]} > C;$$

(ii)' this test is also UMP unbiased.

Problem 7.

We take as our sample space the set
$\mathcal{X} = \mathbb{R}^{2n} \setminus \{(x_1,\dots,x_n,y_1,\dots,y_n) : x_1 = x_2 = \dots = x_n \text{ or } y_1 = y_2 = \dots = y_n\}$.
Suppose T is maximal invariant under G. By Theorem 1 attention can be restricted to tests depending only on T. By Theorem 3 the distribution of T depends on $\theta = (\xi,\sigma^2,\eta,\tau^2)$ only through a maximal invariant $v(\theta)$ w.r.t. $\overline{G}$. Let $\Gamma = v(\Omega)$. The family of possible distributions of T can be written as $\mathcal{P}^T = \{P_\gamma^T : \gamma \in \Gamma\}$. Let $\Sigma(T)$ be a sufficient statistic for γ. For every invariant test $\psi(T)$ there exists an invariant test ψ_0 which depends on T only through $\Sigma(T)$ and which has the same power function as $\psi(T)$
define

$$\psi_0(s) = E_\gamma[\psi(T) \mid \Sigma(T) = s].$$

Hence we can restrict attention to those tests which depend only on $\Sigma(T)$. A statistic $\Sigma(T(X,Y))$ with $T(X,Y)$ maximal invariant under G, and $\Sigma(T)$ sufficient for $v(\theta)$ (as above) will now be constructed.

Define $S(x,y) = (\bar{x},\Sigma_{i=1}^{n} (x_i - \bar{x})^2,\bar{y},\Sigma_{j=1}^{n} (y_j - \bar{y})^2)$. S is sufficient for $(\xi,\sigma^2,\eta,\tau^2)$. It is easily seen that

$$\text{for all } g \in G\text{: } S(x,y) = S(x',y') \Rightarrow S(g(x,y)) = S(g(x,y))$$

((x,y) stands for $(x_1,\dots,x_n,y_1,\dots,y_n)$). Hence S induces a group G_S on $S(\mathcal{X}) = \mathbb{R} \times \mathbb{R}_0^+ \times \mathbb{R} \times \mathbb{R}_0^+$. G_S consists of the transformations

$$(s_1,s_2,s_3,s_4) \mapsto as_1+b, a^2s_2, as_3+c, a^2s_4, a \neq 0, b \in \mathbb{R}, c \in \mathbb{R},$$
$$(s_1,s_2,s_3,s_4) \mapsto as_3+b, a^2s_4, as_1+c, a^2s_2, a \neq 0, b \in \mathbb{R}, c \in \mathbb{R}.$$

A maximal invariant w.r.t. G_S is given by

$$W(s_1,s_2,s_3,s_4) = \max(s_2/s_4, s_4/s_2).$$

This can be seen as follows: if $W(s_1,s_2,s_3,s_4) = W(s_1',s_2',s_3',s_4')$ then $s_2/s_4 = s_2'/s_4'$ or $s_2/s_4 = s_4'/s_2'$. In the first case $(s_1,s_2,s_3,s_4) = g(s_1',s_2',s_3',s_4')$ where $g \in G_S$ is defined by

$$g(\tilde{s}_1,\tilde{s}_2,\tilde{s}_3,\tilde{s}_4) = (a\tilde{s}_1+b, a^2\tilde{s}_2, a\tilde{s}_3+c, a^2\tilde{s}_4),$$

with $a = (s_2/s_2')^{\frac{1}{2}}$, $b = s_1-as_1'$ and $c = s_3-as_3'$. The second case can be treated similarly.

Condition (C) of HALL, WIJSMAN and GHOSH (1965) (see the remarks in the solution of Problem 11), is easily verified here. Hence $W(S(X,Y))$ can be written as $\Sigma(T(X,Y))$, where Σ and T are as above, and attention can be restricted to tests depending on W.

Let $\Delta = \tau^2/\sigma^2$. The testing problem is equivalent to $H : \Delta = 1$ vs $K : \Delta \neq 1$. Let

$$F = \frac{S_2/\sigma^2}{S_4/\tau^2} = \Delta S_2/S_4.$$

F has the distribution $F_{n-1,n-1}$ (Chapter 5, Section 3). The cummulative distribution function of W equals

$$\begin{aligned} H(w;\Delta) &= P\{\max(\Delta/F, F\Delta) \leq w\}, \quad (w \geq 1) \\ &= P\{\Delta/w \leq F \leq w\Delta\} \\ &= K(w\Delta) - K(\Delta/w), \end{aligned}$$

where $K(x) = \int_0^x k(y)dy$, $k(y) = C_n y^{(n-3)/2}/(1+y)^{n-1}$, $C_n = \Gamma(n-1)/[\Gamma((n-1)/2)]^2$. Hence the density of W equals

$$h(w;\Delta) = C_n\Delta^{(n-1)/2}w^{(n-3)/2}[(\Delta+w)^{-n+1} + (\Delta w+1)^{n-1}].$$

Let $f(w,\Delta) = h(w,\Delta)/h(w,1)$. Since $\frac{df}{dw}(w,\Delta) = ((n-1)/2)\Delta^{(n-1)/2}(\Delta-1)(1+w)^{n-2}[(\Delta+w)^{-n} - (\Delta w+1)^{-n}] > 0$ for all $\Delta > 0$; $f(w,\Delta)$ is for all $\Delta > 0$ a nondecreasing function of w $(w \geq 1)$.

Let $\Delta_1 \neq 1$ be an arbitrary (fixed) alternative. By the fundamental lemma of Neyman and Pearson the test which is most powerful at the alternative Δ_1 is given by $f(w,\Delta_1) \geq k$, or equivalently by

$$w \geq k^*.$$

Since this test does not depend on the particular alternative it is UMP among all tests depending on W, i.e. UMP invariant.

Section 4

Problem 8.

Let p be the probability of an item being defective. We consider the testing problem $H : p \leq p_0$ vs $K : p > p_0$.

(i) Inspection of the item by variables

An item is considered satisfactory if a variable Y exceeds a given constant u. Hence $p = P\{Y \leq u\}$. We assume that $Y_1,\ldots,Y_n$ constitute a sample from a normal distribution. The UMP invariant test rejects when

$$t = \sqrt{n}(\bar{y} - u)/\{\Sigma\ (y_i - \bar{y})^2/(n-1)\}^{\frac{1}{2}} < C,$$

where C satisfies $\int_{-\infty}^{C} g(t)dt = \alpha$. g(t) is the noncentral t-density with n-1 degrees of freedom and noncentrality parameter $-\sqrt{n}\,\Phi^{-1}(p_0)$ (see Section 4). For given n, p_0, α we can calculate C. For an alternative p the power of the test, $\beta(p)$ is then given by $\int_{-\infty}^{C} g_1(t)dt$, with $g_1(t)$ the non-central t-density with degrees of freedom n-1 and noncentrality parameter $-\sqrt{n}\,\Phi^{-1}(p)$. Since the UMP invariant test is based on all the observations $y_1,\ldots,y_n$, the power $\beta(p)$ is nondecreasing w.r.t. n. Hence the sample size required to obtain power β is the smallest n for which $\beta(p) \geq \beta$.

(ii)

(a) Inspection of the item by attributes

Each item is classified directly as satisfactory or defective. The number of defectives D in a sample of size n is distributed as b(n,p). The UMP test rejects H with probability one when D > k, and with probability γ when D = k; γ and k satisfy

$$\sum_{i=k+1}^{n} \binom{n}{i} p_0^i (1-p_0)^{n-i} + \gamma \binom{n}{k} p_0^k (1-p_0)^{n-k} = \alpha.$$

The sample size required to obtain power β is determined analogously as in (i).

(b) Sequential probability ratio test

Let $E_p(N)$ denote the expected sample size under p. By Section 11 of Chapter 3, $E_p(N)$ is approximately equal to

$$\{\beta(p) \log(\beta/\alpha) + [1-\beta(p)] \log[(1-\beta)/(1-\alpha)]\}/E_p(Z),$$

where

$$E_p(Z) = p \log(p_1/p_0) + (1-p) \log[(1-p_1)/(1-p_0)],$$

$$\beta(p_0) = \alpha; \quad \beta(p_1) = \beta.$$

Table 1. Numerical results for $\alpha = .05$ and $\beta = .9$.

		required sample sizes		expected sample size sequential test	
		inspection by variables	inspection by attributes	$E_{p_0}(N)$	$E_{p_1}(N)$
$p_0 = .1;$	$p_1 = .15$	237	433	183.0	194.2
	$p_1 = .20$	69	132	54.4	53.5
	$p_1 = .25$	34	69	27.5	25.7
$p_0 = .05;$	$p_1 = .10$	134	267	111.4	115.0
	$p_1 = .15$	44	93	39.3	33.8
	$p_1 = .20$	24	51	21.2	17.0
	$p_1 = .25$	16	35	13.8	10.6
$p_0 = .01;$	$p_1 = .02$	390	1519	639.3	607.2
	$p_1 = .05$	55	209	80.6	57.5
	$p_1 = .10$	21	65	28.0	16.4
	$p_1 = .15$	13	43	16.1	8.6
	$p_1 = .20$	9	25	11.0	4.0

Problem 9.

(i) Since the time N of the first violation of the inequalities $A_0 < p_{\delta_1}(t_1,\dots,t_n)/p_{\delta_0}(t_1,\dots,t_n) < A_1$ satisfies $p_{\delta_i}\{N<\infty\} = 1$ for $i=0,1$, the same arguments as those on p. 98 can be used to prove (34) of Chapter 3.

(ii) Define $y(x) = (y_1(x),\dots,y_n(x))$ by $y_i(x) = x_i\cdot|x_1|^{-1}$, $i=1,2,\dots,n$ and $t(x) = (t_1(x),\dots,t_n(x))$ by $t_1 = \operatorname{sgn} x_1$,

$$t_k(x) = \sqrt{\frac{k-1}{k}}\ \frac{x_1+x_2+\dots+x_k}{\left[\Sigma_{j=1}^k \left(x_j - \frac{x_1+\dots+x_k}{k}\right)^2\right]^{\frac{1}{2}}}\ ,\quad k=2,\dots,n.$$

Let $Y = y(X)$ and $T = t(X)$. Then

$$\begin{aligned}
&P_\delta\{Y_1 = 1, Y_2 \le y_2,\dots,Y_n \le y_n\} = \\
&= P_\delta\{X_1 > 0, X_2 \le y_2X_1,\dots,X_n \le y_nX_1\} = \\
&= \int P_\delta\{X_1 > 0, X_2 \le y_2X_1,\dots,X_n \le y_nX_1 \mid X_1 = x\} f_\delta(x)\,dx = \\
&= \int_0^\infty P_\theta\{X_2 \le y_2x\}\dots P_\delta\{X_n \le y_nx\}\cdot f_\delta(x)\,dx,
\end{aligned}$$

where f_δ is the density function of X_1 under δ, and similarly

$$\begin{aligned}
&P_\delta\{Y_1 = -1, Y_2 \le y_2,\dots,Y_n \le y_n\} = \\
&= \int_{-\infty}^0 P_\delta\{X_2 \le -y_2x\}\dots P_\delta\{X_n \le -y_nx\} f_\delta(x)\,dx = \\
&= \int_0^\infty P_\delta\{X_2 \le y_2x\}\dots P_\delta\{X_n \le y_nx\} f_\delta(-x)\,dx.
\end{aligned}$$

This means that the density h_δ of Y w.r.t. $\mu^* = \mu\times\lambda^{n-1}$, where μ is counting measure on $\mathbb{Z}$ and λ^{n-1} is Lebesgue-measure on $\mathbb{R}^{n-1}$, is given by

$$h_\delta(1,y_2,\dots,y_n) = \int_0^\infty x^{n-1} f_\delta(x)\cdot\prod_{i=2}^n f_\delta(y_ix)\,dx,$$

and

$$h_\delta(-1,y_2,\dots,y_n) = \int_0^\infty x^{n-1} f_\delta(-x)\prod_{i=2}^n f_\delta(y_ix)\,dx,$$

or

$$h_\delta(y_1,\dots,y_n) = \int_0^\infty x^{n-1}\prod_{i=1}^n f_\delta(y_ix)\,dx =$$

(8) $$= \int_0^\infty x^{n-1}(\sigma\sqrt{2\pi})^{-n}\exp\left[-(2\sigma^2)^{-1}\sum_{i=1}^n (y_ix-\delta\sigma)^2\right]dx$$

for $y_1 = \pm 1$, $-\infty < y_2,\dots,y_n < \infty$. Putting $w = x\sigma^{-1}\sqrt{\Sigma_{i=1}^n y_i^2}$ and $z_n = \{\Sigma_{i=1}^n y_i\}\cdot\{\Sigma_{i=1}^n y_i^2\}^{-\frac{1}{2}}$ we find after some computations that

$$h_\delta(y_1,\dots,y_n) = \left[2\pi \sum_{i=1}^n y_i^2\right]^{-\frac{n}{2}} \exp\left[-(\delta^2/2)(n-z_n^2)\right] \cdot$$

$$\cdot \int_0^\infty w^{n-1} \exp\left[-\tfrac{1}{2}(w-\delta z_n)^2\right]dw.$$

Since

$$z_n = \frac{n\bar{y}}{[S_n(y)^2 + n(\bar{y})^2]^{\frac{1}{2}}} = \frac{n\bar{x}}{[S_n(x)^2 + n(\bar{x})^2]^{\frac{1}{2}}} =$$

$$= \frac{t_n S_n(x)[n/(n-1)]^{\frac{1}{2}}}{\left[S_n(x)^2 + \frac{t_n^2 S_n(x)^2}{n-1}\right]^{\frac{1}{2}}} = \frac{t_n\sqrt{n}}{\left[n-1+t_n^2\right]^{\frac{1}{2}}},$$

where $\bar{y} = \frac{1}{n}\Sigma_{i=1}^n y_i$, $\bar{x} = \frac{1}{n}\Sigma_{i=1}^n x_i$, $S_n(y)^2 = \Sigma_{i=1}^n (y_i - \bar{y})^2$ and $S_n(x)^2 = \Sigma_{i=1}^n (x_i - \bar{x})^2$, it follows that

$$n - z_n^2 = n - \frac{nt_n^2}{n-1+t_n^2} = \frac{n(n-1)}{n-1+t_n^2}.$$

Hence

$$h_\delta(y_1,\dots,y_n) = c(y_1,\dots,y_n)\cdot\exp\left[-\frac{1}{2}\frac{(\delta\sqrt{n})^2(n-1)}{n-1+t_n^2}\right]\cdot$$

$$\cdot \int_0^\infty w^{n-1}\exp\left[-\frac{1}{2}\left(w - \frac{t_n\delta\sqrt{n}}{(n-1+t_n^2)^{\frac{1}{2}}}\right)^2\right]dw,$$

where c is independent of δ. Comparison with (76) of Chapter 5 yields that $h_\delta(y_1,\dots,y_n) = c(y_1,\dots,y_n)\cdot p_\delta(t_n(y_1,\dots,y_n))$, where p is the density function of the non-central t-distribution with n-1 degrees of freedom and noncentrality parameter $\delta\sqrt{n}$. Now by the factorization theorem $T_n = t_n(Y_1,\dots,Y_n)$ is sufficient for δ on the basis of $Y_1,\dots,Y_n$. Since $(T_1,\dots,T_n)$ is a function of $(Y_1,\dots,Y_n)$, it follows that T_n is sufficient for δ on the basis of $T_1,\dots,T_n$ (cf. Problem 10 of Chapter 2), which is equivalent to the statement we have to prove.

(iii) Making the transformation $v = \sigma x^{-1}$ in (8), we obtain that

$$h\,(y_1,\dots,y_n) =$$

$$= \int_0^\infty \sigma^{n-1}v^{1-n}(\sigma\sqrt{2\pi})^{-n}\exp\left[-(2\sigma^2)^{-1}\sum_{i=1}^n (y_i\sigma v^{-1} - \delta\sigma)^2\right]\sigma v^{-2}dv =$$

$$= \int_0^\infty (v\sqrt{2\pi})^{-n} \exp\left[-(2v^2)^{-1} \sum_{i=1}^n (y_i - \delta v)^2\right] v^{-1} dv.$$

writing $v = w|x_1|^{-1}$ this becomes

$$h_\delta(y_1,\ldots,y_n) =$$
$$= \int_0^\infty (w\sqrt{2\pi})^{-n} |x_1|^n \exp\left[-(2w^2)^{-1} \sum_{i=1}^n (x_i - \delta w)^2\right] w^{-1} dw,$$

which implies the required result.

(ARNOLD (1951), RUSHTON (1952))

Problem 10.

(i) We take as our sample space the set

$$X = \mathbb{R}^n \setminus \{(x_1,\ldots,x_n) : x_1 = x_2 = \ldots = x_n\}.$$

We write x for $(x_1,\ldots,x_n)$. As in the solution of Problem 7 we can restrict attention to tests depending on a statistic $\Sigma(V(X))$, where

(a) V is maximal invariant under G,

(b) the distribution of V depends on (ξ,σ^2) only through $v(\xi,\sigma^2)$ where v is maximal invariant under $\overline{G}$,

(c) $\Sigma(V)$ is sufficient for v.

(G is the group of transformations: $x \mapsto cx$ $(c \neq 0)$.) The statistic $\Sigma(V)$ will be constructed as in the solution of Problem 7.

Define

$$S : X \to \mathbb{R} \times \mathbb{R}_0^+$$
$$x \mapsto (s_1(x), s_2(x)) = (\bar{x}, \sum_{i=1}^n (x_i - \bar{x})^2).$$

Clearly S(X) is sufficient for (ξ,σ^2).
Since for all g in G: $S(x) = S(x') \Rightarrow S(gx) = S(gx')$, S induces the group of transformations G_S on $\mathbb{R} \times \mathbb{R}_0^+$, where $G_S = \{g_s : g_s(s_1,s_2) = (cs_1, c^2 s_2);$ $c \neq 0\}$. A maximal invariant under G_S is given by $|s_1|/\sqrt{s_2}$.

Since condition C of HALL, WIJSMANN and GHOSH (1965) (cf. Problem 11) is satisfied it follows that $|s_1|/\sqrt{s_2} = |\bar{x}|/(\Sigma\ (x_i - \bar{x})^2)^{\frac{1}{2}}$ can be written as

$\Sigma(V(X))$. Hence attention can be restricted to tests depending on

$$T = (n(n-1))^{\frac{1}{2}} \cdot |\overline{X}|/(\Sigma\ (X_i - \overline{X})^2)^{\frac{1}{2}}.$$

Write θ for ξ/σ. The testing problem is equivalent to $H : \theta = 0$ vs $K : \theta \neq 0$. The density of T is given by: $p_\delta(t) + p_\delta(-t)$, $t > 0$; with $\delta = \sqrt{n}\,\theta$, and p_δ as on p. 223. Now it will be shown that the ratio $r(t) = [p_\delta(t) + p_\delta(-t)]/[p_0(t) + p_0(-t)]$ is an increasing function of t (on the interval $t \in (0,\infty)$). Since $p_0(-t) = p_0(t)$, $r(t) = \frac{1}{2}[p_\delta(t)/p_0(t) + p_\delta(-t)/p_0(-t)]$. The substitution $v = t(w/(n-1))^{\frac{1}{2}}$ in the integral of $p_\delta(t)$ yields

$$p_\delta(t)/p_0(t) = \exp(-\delta^2) \cdot \int_0^\infty \exp(\delta_1 v) g_{t^2}(v)\,dv$$

with $g_{t^2}(v)$ as on p. 223, and $f(v) = \exp(-v^2/2)\cdot v^{n-1}$ (i.e. f as on p. 223, but with $\delta_0 = 0$). Since $p_\delta(-t) = p_{-\delta}(t)$ we have

$$r(t) = \tfrac{1}{2} \exp(-\delta^2/2) \int_0^\infty [\exp(\delta v) + \exp(-\delta v)] g_{t^2}(v)\,dv.$$

That r(t) is monotone on $t \in (0,\infty)$ follows now from the argument at the bottom of p. 223.

Hence for any particular alternative δ_1, the most powerful test depending on T is given by the rejection region $T > C$. Since this test does not depend on δ_1, it is uniform most powerful among the tests depending on T, and hence uniform most powerful invariant.

(ii) We take as our sample space the set

$$\mathcal{X} = \mathbb{R}^{m+n} \setminus \{(x_1,\ldots,x_m,y_1,\ldots,y_n) : x_1 = x_2 = \ldots = x_m \text{ or } y_1 = y_2 = \ldots = y_n \text{ or } \bar{x} = \bar{y}\}.$$

Write x, y for $(x_1,\ldots,x_m)$, $(y_1,\ldots,y_n)$ respectively.
As in (i) we will construct a statistic $\Sigma(V(X,Y))$ with the properties (a),(b),(c). (Now G consists of the transformations $(x,y) \mapsto (ax+b.ay+b)$, $a \in \mathbb{R}_0$, $b \in \mathbb{R}$.) Define S by

$$S = \mathcal{X} \to \mathbb{R} \times \mathbb{R}_0 \times \mathbb{R}_0^+$$
$$(x,y) \mapsto (s_1(x,y), s_2(x,y), s_3(x,y)) =$$
$$= (\bar{x},\ \bar{x}-\bar{y},\ |\bar{x}-\bar{y}|/[\Sigma_{i=1}^m (x_i - \bar{x})^2 + \Sigma_{j=1}^n (y_j - \bar{y})^2].$$

S induces a group of transformations G_S on $\mathbb{R} \times \mathbb{R}_0 \times \mathbb{R}_0^+$ given by $G_S = \{g_s : g_s(s_1,s_2,s_3) = (as_1+b, as_2, s_3), a \in \mathbb{R}_0, b \in \mathbb{R}\}$. s_3 is a maximal invariant, because for all $s_1', s_2', s_3', s_1'', s_2''$ we have $(s_1'', s_2'', s_3') = g_0(s_1', s_2', s_3')$, with g_0 given by $g_0(s_3, s_2, s_3) = (as_1+b, as_2, s_3)$; $a = s_2'/s_2$, $b = s_1' - as_1$.

Since condition C of HALL, WIJSMANN and GHOSH (1965) (cf. Problem 11) is satisfied $s_3 = |\overline{X}-\overline{Y}|/[\Sigma\ (X_i - \overline{X})^2 + \Sigma\ (Y_j - \overline{Y})^2]^{\frac{1}{2}}$ can be written as $\Sigma(U(X,Y))$. Hence attention can be restricted to tests depending on

$$T = |\overline{X}-\overline{Y}|/\{[\Sigma\ (X_i - \overline{X})^2 + \Sigma\ (Y_j - \overline{Y})^2](\tfrac{1}{m} + \tfrac{1}{n})/(m+n-2)\}^{\frac{1}{2}}.$$

Write θ for $(\eta - \xi)/\sigma$. The testing problem is equivalent to $H : \theta = 0$ vs $K : \theta \neq 0$. The density of T is given by: $p_\delta(t) + p_\delta(-t)$, $t > 0$, with $\delta = \theta/(\frac{1}{m} + \frac{1}{n})^{\frac{1}{2}}$ and p_δ the density of a noncentral t-distribution with m+n-2 degrees of freedom and noncentrality parameter δ. In (i) it has been shown that the density of T has monotone likelihood ratio. Continuing as in (i) yields: the UMP invariant test for testing $\eta = \xi$ is given by the rejection region $T > C$.

Problem 11.

In many problems of this chapter, concerning normal distributions and groups of linear transformations, a sufficiency reduction precedes a reduction through invariance. From a theoretical point of view the reverse order is the correct one.
HALL, WIJSMAN and GHOSH (1965), however, proved that under certain conditions the final result is independent of the order chosen.
As an example, we will check Assumption C from section II.7 of their famous paper, which guarantees that (with their notation) the subfield A_{SI} is sufficient for A_I, which is a reformulation of the assertion in terms of σ-fields. We quote:

Assumption C. X is an n-dimensional Borel set, A the Borel subsets of X, $P = \{P_\theta, \theta \in \Theta\}$ with Θ an arbitrary index set, and with respect to n-dimension Lebesgue measure P_θ has a density

(9) $$p_\theta(x) = g_\theta(s(x))h(x), \qquad x \in X,$$

in which s is a measurable function from X into k-space $(k < n)$ with range S, g_θ and h are positive, real-valued measurable functions on S, X, respectively, and s and h satisfy the conditions below. Let G, A_S, A_I and A_{SI} be as in Section 3 and suppose that there is an open set $A_{SI} \in A_{SI}$ of P-measure 1, such that on A_{SI}

(i) for each $g \in G$ the transformation $x \to gx$ is continuously differentiable, and the Jacobian depends only on s(x),

(ii) for each $g \in G$, $s(x) = s(x')$ implies $s(gx) = s(gx')$,

(iii) s is continuously differentiable, and the matrix D(x), whose ij element is $\partial s_j/\partial x_i$, is of rank k,

(iv) for each $g \in G$, $h(gx)/h(x)$ depends only on s(x).

In our problem

$$X = \left\{\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x_1\ldots x_n\\y_1\ldots y_n\end{pmatrix} : x_i, y_i \in \mathbb{R},\ i = 1,\ldots,n\right\} \quad (n \geq 2),$$

where we suppose (without restricting generality) that the x_i's (resp. y_i's) are not all equal and that x and y are linearly independent (then we can take $A_{SI} = X$). Note that $P_\theta\{X \text{ and } Y \text{ are linearly independent}\} = 1$, for all θ (cf. Problem 24 of Chapter 7).

Let A denote the (induced) borel σ-field.

$P = \{P_\theta : \theta \in \Theta\}$ with

$$\Theta = \{\theta = (\xi,\eta,\sigma,\tau,\rho) : \xi,\eta \in \mathbb{R},\ \sigma,\tau > 0,\ \rho \in (-1,+1)\},$$

where P_θ denotes a probability measure which has, with respect to 2n-dimensional Lebesgue-measure, density

$$p_\theta(x,y) = (2\pi\sigma\tau\sqrt{1-\rho^2})^{-n} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\sigma^2}\Sigma\ (x_i-\xi)^2 + \right.\right.$$
$$\left.\left. - \frac{2\rho}{\sigma\tau}\Sigma\ (x_i-\xi)(y_i-\eta) + \frac{1}{\tau^2}\Sigma\ (y_i-\eta)^2\right]\right\} .$$

Take

$$s(x,y) = (\bar{x},\bar{y},\sqrt{\Sigma\ (x_i-\bar{x})^2},\sqrt{\Sigma\ (y_i-\bar{y})^2},\ \Sigma\ (x_i-\bar{x})(y_i-\bar{y})) =$$
$$= (t_1,t_2,t_3,t_4,t_5).$$

To check (9) remark that

$$p_\theta(x,y) = C(\theta)\cdot\exp\left\{\frac{\rho}{\sigma\tau(1-\rho^2)}(t_5+nt_1t_2) - \frac{1}{2\sigma^2(1-\rho^2)}(t_3^2+nt_1^2) + \right.$$

$$- \frac{1}{2\tau^2(1-\rho^2)}(t_4^2 + nt_2^2) - \frac{1}{1-\rho^2}\left(\frac{\xi}{\sigma^2} - \frac{\eta\rho}{\sigma\tau}\right)nt_1 + \frac{1}{1-\rho^2}\left(\frac{\eta}{\tau^2} - \frac{\xi\rho}{\sigma\tau}\right)nt_2\Big\} =$$

$$= C(\theta)g_\theta[s(x,y)].$$

From this relation it follows that $h(x,y) \equiv 1$, hence condition (iv) is trivial.

Let G_1 be the group of transformations from part (i) and G_2 the group from part (ii). Then

$$G_1 = \{g : g(x,y) = \begin{pmatrix} ax+b \\ cy+d \end{pmatrix}, \ a > 0, \ c > 0\} \text{ and}$$

$$G_2 = \{g : g(x,y) = \begin{pmatrix} ax+b \\ cy+d \end{pmatrix}, \ a \neq 0, \ c \neq 0\}.$$

Since $G_1 \subset G_2$ it suffices to check conditions (i) and (ii) only for G_2.

Condition (i): the g's are continuously differentiable with a constant Jacobian.

Condition (ii): suppose that $s(x,y) = s(x^*,y^*)$, where

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \dots x_n \\ y_1 \dots y_n \end{pmatrix} \text{ and } \begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} x_1^* \dots x_n^* \\ y_1^* \dots y_n^* \end{pmatrix}.$$

Denoting $s(x,y) = (t_1,\dots,t_5)$ and $s(x^*,y^*) = (t_1^*,\dots,t_5^*)$, we have $t_i = t_i^*$, $i = 1,\dots,5$.

Since

$$s[g(x,y)] = s(ax+b, cy+d) = (at_1+b, ct_2+d, |a|t_3, |c|t_4, act_5)$$

and

$$s[g(x^*,y^*)] = (at_1^*+b, ct_2^*+d, |a|t_3^*, |c|t_4^*, act_5^*),$$

condition (ii) is immediate.

Condition (iii):

$$D(\underset{\sim}{x},\underset{\sim}{y}) = \begin{pmatrix} \frac{\partial t_1}{\partial x_1} & \cdots & \frac{\partial t_5}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial t_1}{\partial x_n} & \cdots & \frac{\partial t_5}{\partial x_n} \\ \frac{\partial t_1}{\partial y_1} & \cdots & \frac{\partial t_5}{\partial y_1} \\ \vdots & & \vdots \\ \frac{\partial t_1}{\partial y_n} & \cdots & \frac{\partial t_5}{\partial y_n} \end{pmatrix} = \begin{pmatrix} \frac{1}{n} & 0 & 2(x_1-\bar{x})t_3^{-1} & 0 & y_1-\bar{y} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{n} & 0 & 2(x_n-\bar{x})t_3^{-1} & 0 & y_n-\bar{y} \\ 0 & \frac{1}{n} & 0 & 2(y_1-\bar{y})t_4^{-1} & x_1-\bar{x} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \frac{1}{n} & 0 & 2(y_n-\bar{y})t_4^{-1} & x_n-\bar{x} \end{pmatrix}$$

Since the five columns of $D(\underset{\sim}{x},\underset{\sim}{y})$ are linearly independent (this is the reason that we have restricted the $\mathcal{X}$-space), this matrix is of rank 5. Hence condition (iii) is satisfied.

Conclusion: when in search for a UMP invariant test, we may reduce the observations first to the sufficient statistic $S = s(\underset{\sim}{X},\underset{\sim}{Y}) = (T_1,T_2,T_3,T_4,T_5)$, where $T_1 = \bar{X}$, $T_2 = \bar{Y}$, $T_3 = \{\Sigma\ (X_i - \bar{X})^2\}^{\frac{1}{2}}$, $T_4 = \{\Sigma\ (Y_i - \bar{Y})^2\}^{\frac{1}{2}}$ and $T_5 = \Sigma\ (X_i - \bar{X})(Y_i - \bar{Y})$.

This reduction induces a new sample space

$$\mathcal{T} = \{(t_1,t_2,t_3,t_4,t_5) : t_3 > 0,\ t_4 > 0,\ |t_5| < t_3t_4\}$$

and groups of transformations G_1^T and G_2^T, where

$$G_1^T = \{g : g(t_1,t_2,t_3,t_4,t_5) = (at_1+b,ct_2+d,at_3,ct_4,act_5), a > 0, c > 0\}$$

and

$$G_2^T = \{g : g(t_1,t_2,t_3,t_4,t_5) = (at_1+b,ct_2+d,|a|t_3,|c|t_4,act_5),$$
$$a \neq 0, c \neq 0\}.$$

(i) With respect to G_1^T, $r(t_1,\ldots,t_5) = t_5/t_3t_4$ is a maximal invariant: Clearly r is invariant and if $t_5/t_3t_4 = t_5^*/t_3^*t_4^*$, then $t_3^* = at_3$, $t_4^* = ct_4$ and $t_5^* = act_5$, where $a = t_3^*t_3^{-1}$ and $c = t_4^*t_4^{-1}$. Hence $r(t_1,\ldots,t_5) = r(t_1^*,\ldots,t_5^*)$ implies $(t_1^*,\ldots,t_5^*) = g(t_1,\ldots,t_5)$ for some $g \in G_1^T$.

$R = r(T_1,\ldots,T_5) = T_5/T_3T_4$ equals the sample correlation coefficient.

We will now show that R has a density (w.r.t. Lebesgue measure) $p_\rho(r)$

with monotone likelihood ratio in r. By (85) of Chapter 5,

$$p_\rho(r) = \frac{n-2}{\sqrt{2\pi}} \frac{\Gamma(n-1)}{\Gamma(n-\frac{1}{2})}(1-\rho^2)^{\frac{1}{2}(n-1)}(1-r^2)^{\frac{1}{2}(n-4)}(1-\rho r)^{-n+\frac{3}{2}} \cdot$$
$$\cdot\ F(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1+\rho r}{2}),$$

where

$$F(a,b,c,x) = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)}{\Gamma(a)} \frac{\Gamma(b+j)}{\Gamma(b)} \frac{\Gamma(c)}{\Gamma(c+j)} \frac{x^j}{j!}$$

is a hypergeometric function and $|\rho r| < 1$.
We must show that, for $\rho_1 > \rho_2$, $p_{\rho_1}(r)/p_{\rho_2}(r)$ is a nondecreasing function of r.

PROOF.

$$\frac{p_{\rho_1}(r)}{p_{\rho_2}(r)} = \frac{(1-\rho_1^2)^{\frac{1}{2}(n-1)}}{(1-\rho_2^2)^{\frac{1}{2}(n-1)}} \frac{(1-\rho_1 r)^{-n+\frac{3}{2}}}{(1-\rho_2 r)^{-n+\frac{3}{2}}} \frac{F\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1+\rho_1 r}{2}\right)}{F\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1+\rho_2 r}{2}\right)}$$

First remark that, given $\rho_1 > \rho_2$,

$$\left(\frac{1-\rho_2 r}{1-\rho_1 r}\right)^{n-\frac{3}{2}}$$

is an increasing function of r, since $n-\frac{3}{2} > 0$ and

$$\frac{d}{dr}\left(\frac{1-\rho_2 r}{1-\rho_1 r}\right) = \frac{\rho_1-\rho_2}{(1-\rho_1 r)^2} > 0.$$

Hence to show that $p_\rho(r)$ has monotone likelihood ratio in r, it suffices to prove that, given $\rho_1 > \rho_2$,

$$\frac{F\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1+\rho_1 r}{2}\right)}{F\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1+\rho_2 r}{2}\right)}$$

is a nondecreasing function of r, or, equivalently (cf. the solution of Problem 6 (i), Chapter 3), that

$$(10) \qquad \frac{\partial^2}{\partial\rho\partial r} \log F(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1+\rho r}{2}) \geq 0, \text{ for all } \rho \text{ and } r.$$

To prove this inequality, note that $F(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1+\rho r}{2})$ can be written as

(11) $$F(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1-\rho r}{2}) = \sum_{j=0}^{\infty} c_j(1+\rho r)^j,$$

where

$$c_j = \left[\frac{\Gamma(\frac{1}{2}+j)}{\Gamma(\frac{1}{2})}\right]^2 \cdot \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-\frac{1}{2}+j)} \cdot \frac{1}{2^j j!}, \quad j \geq 0.$$

Remark also that $c_0 = 1$ and that $c_{j+1} < \frac{1}{2}c_j$, $j \geq 0$, since

$$\frac{c_{j+1}}{c_j} = \left[\frac{\Gamma(\frac{3}{2}+j)}{\Gamma(\frac{1}{2}+j)}\right]^2 \cdot \frac{\Gamma(n-\frac{1}{2}+j)}{\Gamma(n+\frac{1}{2}+j)} \cdot \frac{2^j j!}{2^{j+1}(j+1)!} =$$

$$= \frac{(\frac{1}{2}+j)^2}{2(n-\frac{1}{2}+j)(j+1)} = \frac{1}{2}\,\frac{j^2+j+\frac{1}{4}}{j^2+(n+\frac{1}{2})j+(n-\frac{1}{2})} < \frac{1}{2}.$$

From (10) and (11) we get, with $t = 1+\rho r$,

$$\frac{\partial^2 \log F}{\partial\rho\partial r} = \left[\sum_{i=0}^{\infty} c_i t^i\right]^{-2} \cdot \sum_{j,i=0}^{\infty} c_i c_j t^{i+j-2}[j+(t-1)j^2-(t-1)ji].$$

Changing the roles of the indices i and j we get

$$\frac{\partial^2 \log F}{\partial\rho\partial r} = \left[\sum_{j=0}^{\infty} c_j t^j\right]^{-2} \cdot \sum_{j,i=0}^{\infty} c_j c_i t^{j+i-2}[i+(t-1)i^2-(t-1)ij].$$

Hence, since

$$\tfrac{1}{2}[j+(t-1)j^2-(t-1)ji+i+(t-1)i^2-(t-1)ij] =$$

$$= \tfrac{1}{2}[(t-1)(j-i)^2-(i+j)],$$

we get

$$\frac{\partial^2 \log F}{\partial\rho\partial r} = \frac{1}{2}\cdot\left[\sum_{i=0}^{\infty} c_i t^i\right]^{-2} \cdot \sum_{i,j=0}^{\infty} c_i c_j t^{i+j-2}[(j-i)^2(t-1)+(i+j)].$$

Hence it remains to show that the last factor is nonnegative. Remark that it is greater than

(12) $$2\sum_{i=0}^{\infty} c_i t^{i-2} \sum_{j=i+1}^{\infty} c_j t^j[(j-i)^2(t-1)+(i+j)],$$

since terms with $i = j$ were deleted in (12).
The coefficient of t^j in the interior sum of (12), which is given by

$$a_j^{(i)} = c_j[-(j-i)^2+(i+j)] + c_{j-1}(j-1-i)^2,$$

is nonnegative. Indeed, from $c_j < \frac{1}{2}c_{j-1}$, $j \geq 1$, it follows that

(13) $$a_j^{(i)} \geq c_j[-(j-1)^2 + (i+j) + 2(j-1-i)^2] = c_j[(j-i)^2 + 5i - 3j + 2]$$

Define $\tilde{a}_j^{(i)} = (j-i)^2 + 5i - 3j + 2$, then (as can be shown easily) for any fixed i $\{\tilde{a}_j^{(i)}\}_{j=i+1}^{\infty}$ is a nondecreasing sequence with $\tilde{a}_{i+1}^{(i)} = 2i \geq 0$. Hence $\tilde{a}_j^{(i)} \geq 0$ $(j \geq i+1)$. Now (13) implies $a_j^{(i)} \geq 0$ $(j \geq i+1)$, so that (12) is nonnegative and as a consequence $(\partial^2 \log F)/(\partial\rho\partial r) \geq 0$, as was to be proved. Now by Theorem 2 of Chapter 3, there exists a UMP invariant test with rejection region $R > C$ for testing $H : \rho \leq \rho_0$ against $K : \rho > \rho_0$. □

(ii) Clearly $|r(t_1,\ldots,t_5)| = |t_5/t_3t_4|$ is G_2^T-invariant. To show that $|t_5/t_3t_4|$ is maximal G_2^T-invariant, suppose that $|t_5/t_3t_4| = |t_5^*/t_3^*t_4^*|$. If $t_5/t_3t_4 = t_5^*/t_3^*t_4^*$ there exists a transformation $g_1 \in G_1^T \subset G_2^T$ with $g_1(t_1,\ldots,t_5) = (t_1^*,\ldots,t_5^*)$, by part (i).
Therefore suppose that $t_5/t_3t_4 = -t_5^*/t_3^*t_4^*$.
Take $a = t_3^*t_3^{-1}$ and $c = -t_4^*t_4^{-1}$. Then $a > 0$ and $c < 0$, and $t_5 = act_5$, $t_3^* = |a|t_3$ and $t_4^* = |c|t_4$. Hence there exists a transformation $g_2 \in G_2^T$ such that $g_2(t_1,\ldots,t_5) = (t_1^*,\ldots,t_5^*)$.
It now remains to show that the test with rejection region $|R| \geq C$ is UMP within the class of invariant tests for testing $H : \rho = 0$ against $K : \rho \neq 0$. Remark that the density $q_\rho(r)$ of $|R|$ is $p_\rho(r) + p_\rho(-r)$ for $r \in [0,1)$ and 0 elsewhere.
We now show that

(14) for $r \in [0,1)$, the ratio $q_\rho(r)/q_0(r)$ is a nondecreasing function of r.

Since $p_0(r) = p_0(-r)$, the ratio in (14) equals

(15) $$\frac{q_\rho(r)}{q_0(r)} = \frac{p_\rho(r) + p_\rho(-r)}{2p_0(r)} =$$

$$= \tfrac{1}{2}(1-\rho^2)^{\frac{1}{2}(n-1)}F(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2})^{-1}\Big[(1-\rho r)^{-n+\frac{1}{2}}F(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1-\rho r}{2}) +$$

$$+ (1+\rho r)^{-n+\frac{1}{2}}F(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1-\rho r}{2})\Big] =$$

$$= C(\rho^2,n)[G(\rho r) + G(-\rho r)],$$

where $C(\rho^2,n) = \frac{1}{2}(1-\rho^2)^{\frac{1}{2}(n-1)}F(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2})^{-1} > 0$ and

$G(t) = (1-t)^{-n+\frac{1}{2}}F(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1+t}{2}) = (1-t)^{-n+\frac{1}{2}} \Sigma_{j=0}^{\infty} c_j(1+t)^j$, $t \in (-1,1)$.
Observe that

$$G'(t) = (1-t)^{-n-\frac{1}{2}}\left[(n-\tfrac{1}{2}) \sum_{j=0}^{\infty} c_j(1+t)^j + (1-t) \sum_{j=1}^{\infty} jc_j(1+t)^{j-1}\right] =$$

$$= [(1-t)(1+t)]^{-n-\frac{1}{2}}\left[(n-\tfrac{1}{2}) \sum_{j=0}^{\infty} c_j(1+t)^{j+n+\frac{1}{2}} + (1-t)(1+t) \sum_{j=1}^{\infty} jc_j(1+t)^{j+n-\frac{3}{2}}\right] =$$

$$= (1-t^2)^{-n-\frac{1}{2}}\left[(n-\tfrac{1}{2}) \sum_{j=0}^{\infty} c_j(1+t)^{j+n+\frac{1}{2}} + (1-t^2) \sum_{j=1}^{\infty} jc_j(1+t)^{j+n-\frac{3}{2}}\right].$$

Hence

$$\frac{\partial}{\partial r}[G(\rho r) + G(-\rho r)] = \rho[G'(\rho r) - G'(-\rho r)] =$$

$$= \rho(1-\rho^2r^2)^{-n-\frac{1}{2}}\left\{(n-\tfrac{1}{2}) \sum_{j=0}^{\infty} c_j\left[(1+\rho r)^{j+n+\frac{1}{2}} - (1-\rho r)^{j+n+\frac{1}{2}}\right] +\right.$$

$$\left. + (1-\rho^2r^2) \sum_{j=1}^{\infty} jc_j\left[(1+\rho r)^{j+n-\frac{3}{2}} - (1-\rho r)^{j+n-\frac{3}{2}}\right]\right\}.$$

Suppose $\rho \geq 0$. Then $0 \leq \rho r < 1$ $(r \geq 0)$. Hence $(1+\rho r)^{\gamma} \geq (1-\rho r)^{\gamma}$, for any $\gamma > 0$. It follows that $\frac{\partial}{\partial r}[G(\rho r) + G(-\rho r)] \geq 0$ when $\rho \geq 0$.
By symmetry, this inequality also holds when $\rho < 0$.
Hence $G(\rho r) + G(-\rho r)$ is nondecreasing in r. Now (15) implies (14).

Finally, consider any fixed alternative $\rho_0 \neq 0$. Then the most powerful test depending on $|R|$ for testing $H : \rho = 0$ against $K_0 : \rho = \rho_0$ rejects when $|R| \geq C$. Since $|R|$ is maximal invariant and since the test does not depend on the particular alternative chosen, it is UMP invariant for the original testing problem.

(ANDERSON (1958)).

Problem 12.

The requested powers are computed from tables in DAVID (1938).

Table 2. Power of the test for the hypothesis $\rho \leq \rho_0$. ($\alpha = .05$).

n	significance level ($\rho_0 = .3$)	power for $\rho = .5$
50	.50	.51
100	.45	.75
200	.40	.96

Section 5

Problem 13.

We have to check the conditions of Theorem 4 in the case of Problem 6(i) and Example 6.

In both cases the sample space $\mathcal{X}$ equals $\mathbb{R}^{m+n}$, while the group of transformations G is given by $\{(a,b,c) : a > 0;\ b,c \in \mathbb{R}\}$, where $(a,b,c) : \mathcal{X} \to \mathcal{X}$ is defined by

$$(a,b,c)(x_1,\dots,x_m,y_1,\dots,y_n) =$$
$$= (ax_1+b,\dots,ax_m+b,ay_1+c,\dots,ay_n+c).$$

Let $\mathcal{A}$ and $\mathcal{B}$ be the Borel σ-fields on $\mathcal{X}$ and G respectively.

The function $h : \mathcal{X} \times G \to \mathcal{X}$ defined by $h(x_1,\dots,x_m,y_1,\dots,y_n,a,b,c) = (ax_1+b,\dots,ax_m+b,ay_1+c,\dots,ay_n+c)$ is continuous and hence for any set $A \in \mathcal{A}$ the set $R^{-1}(A) =$
$= \{(x_1,\dots,x_m,y_1,\dots,y_n,a,b,c) : (a,b,c)(x_1,\dots,x_m,y_1,\dots,y_n) \in A\} \in \mathcal{A} \times \mathcal{B}$.

Finally define a measure ν over G by

$$\nu(B) = \ell(B^{-1}), \text{ for all } B \in \mathcal{B},$$

where ℓ is the Lebesgue measure and $B^{-1} = \{b : b^{-1} \in B\}$. For any $g = (g_1,g_2,g_3) \in G$ and any $B \in \mathcal{B}$ we have $\nu(Bg) = \ell(g^{-1}B^{-1})$. Writing g^{-1} as (g_1^*,g_2^*,g_3^*), the set $g^{-1}B^{-1}$ is given by

$$g^{-1}B^{-1} = \{(g_1^*,g_2^*,g_3^*)(b_1,b_2,b_3) : (b_1,b_2,b_3) \in B^{-1}\} =$$

$$= \{(g_1^*b_1,g_1^*b_2+g_2^*,g_1^*b_3+g_3^*) : (b_1,b_2,b_3) \in B^{-1}\}.$$

Since $g^{-1}B^{-1}$ arises from B^{-1} by multiplication of each coördinate by the fixed number g_1^*, followed by a translation over the fixed vector $(0,g_2^*,g_3^*)$, we have: $\ell(B^{-1}) = 0 \Rightarrow \ell(g^{-1}B^{-1}) = 0$, and hence also $\nu(B) = 0 \Rightarrow \nu(Bg) = 0$.

Section 6

Problem 14.

Let the testing problem be $\theta \in H$ vs $\theta \in K$. Denote by d_0, d_1 the decisions "$\theta \in H$", "$\theta \in K$" respectively. Define the loss function as follows

$$L(\theta, d_0) = \begin{cases} 0 \text{ if } \theta \in H, \\ 1 \text{ if } \theta \in K; \end{cases}$$

$$L(\theta, d_1) = \begin{cases} 0 \text{ if } \theta \in K, \\ 1 \text{ if } \theta \in H. \end{cases}$$

For all $g \in G$ define g^* by $g^* d_0 = d_0$, $g^* d_1 = d_1$.
Every test ϕ defines a decision procedure in the following way: given $X = x$, decision d_1 is taken with probability $\phi(x)$.

Let C be a class of tests which is closed under a group of transformations G. The C is also a class of procedures which is closed under G in the sense of Problem 5 of Chapter 1.

Let ϕ_0 be an a.e. unique UMP member of C. Since for all $\theta \in K$

$$R(\theta, \phi_0) = 1 - E_\theta[\phi_0(X)]$$

we have that ϕ_0 uniformly minimizes the risk within the class C. Hence by Problem 5 of Chapter 1 we have for all $g \in G$:

$$\phi_0(gx) = \phi_0(x) \text{ excepts for } x \in N_g,$$

where $P_\theta(N_g) = 0$, for all θ. It follows that ϕ_0 is almost invariant.

Problem 15.

For any test ϕ, let ϕg be the test defined by $(\phi g)(x) = \phi(gx)$. It follows from (1) of Chapter 6 that

$$\beta_\phi(\bar{g}\theta) = E_{\bar{g}\theta}[\phi(X)] = E_\theta[\phi(gX)] = \beta_{\phi g}(\theta),$$

for all θ and g.

If in particular $\theta \in S(\alpha)$, and $\theta \in \Omega_H$ then

$$\beta_{\phi g}(\theta) = \beta_\phi(\bar{g}\theta) \le \alpha$$

($\theta \in \Omega_H$ implies that $\bar{g}\theta \in \Omega_H$, by the assumption that the testing problem is invariant under G). Hence if $\phi \in S(\alpha)$ then $\phi g \in S(\alpha)$ for all $g \in G$, and since G is a group we have for all $g \in G$

$$\phi \in S(\alpha) \Leftrightarrow \phi g \in S(\alpha).$$

As a consequence $\sup_{\phi g^{-1} \in S(\alpha)} \beta_\phi(\theta) = \sup_{\phi \in S(\alpha)} \beta_\phi(\theta)$ and hence:

$$\beta^*_\alpha(\bar{g}\theta) = \sup_{\phi \in S(\alpha)} \beta(\bar{g}\theta) = \sup_{\phi \in S(\alpha)} \beta_{\phi g}(\theta) = \sup_{\phi g^{-1} \in S(\alpha)} \beta_\phi(\theta) = \beta^*_\alpha(\theta).$$

Problem 16.

(i) The equation $\int_A f(x)dP_\theta(x) = \int_{gA} f(g^{-1}x)dP_{\bar{g}\theta}(x)$ is a generalization of equation (1), of Chapter 6, as is easily seen, by setting $f \equiv 1$. It is a consequence of Lemma 2 of Chapter 2 take $T = g^{-1} : (X,\mathcal{A}) \to (X,\mathcal{A})$, $g = f$, $\mu = P_{\bar{g}\theta}$. Then by (1) $\mu^* = P_\theta$.

(ii) Suppose P_{θ_1} is absolute continuous with respect to P_{θ_0}. Let $P_{\bar{g}\theta_0}\{X \in A\}$ be equal to zero. Then, by (1), $P_{\theta_0}\{gX \in A\} = 0$, and hence $P_{\theta_1}\{gX \in A\} = 0$. Applying (1) once more we find $P_{\bar{g}\theta_1}\{X \in A\} = 0$. This shows that $P_{\bar{g}\theta_1}$ is absolute continuous with respect to $P_{\bar{g}\theta_0}$. Furthermore, for all $A \in \mathcal{A}$

$$\int_{g^{-1}A} \frac{dP_{\theta_1}}{dP_{\theta_0}}(x)dP_{\theta_0}(x) = P_{\theta_1}\{X \in g^{-1}A\} = P_{\bar{g}\theta_1}\{X \in A\}, \quad \text{(by (1))},$$

$$= \int_A \frac{dP_{\bar{g}\theta_1}}{dP_{\bar{g}\theta_0}}(x)dP_{\bar{g}\theta_0}(x) = \int_{g^{-1}A} \frac{dP_{\bar{g}\theta_1}}{dP_{\bar{g}\theta_0}}(gx)dP_{\bar{g}\theta_0}(x), \quad \text{(by (i))}.$$

This proves the second assertion of (ii) (since $\mathcal{A} = \{g^{-1}A \mid A \in \mathcal{A}\}$).

(iii) When X is distributed as P_{θ_0}, then gX is distributed as $P_{\bar{g}\theta_0}$. Hence (iii) follows from the second result of (ii).

Problem 17.

(i) Let $g \in G$. Let $d\mu/d\mu g^{-1}$ denote a fixed version of the Radon-Nikodym derivative of μ with respect to μg^{-1}. Then for any measurable set A and any $\theta \in \Omega$ we have by Lemma 2 of Chapter 2

$$\int_A p_{\bar{g}\theta}(gx)\frac{d\mu}{d\mu g^{-1}}(gx)d\mu(x) = \int_{gA} p_{\bar{g}\theta}(y)\frac{d\mu}{d\mu g^{-1}}(y)d\mu g^{-1}(y) =$$

$$= \int_{gA} p_{\bar{g}\theta}(y)d\mu(y) = P_{\bar{g}\theta}(gA) = P_\theta(A) = \int_A p_\theta(y)d\mu(y).$$

Hence $p_\theta(x) = p_{\bar{g}\theta}(gx)\frac{d\mu}{d\mu g^{-1}}(gx)$ a.e.(μ).

(ii) In order to avoid difficulties with division by zero we use as a definition for the likelihood ratio

$$\lambda(x) = \begin{cases} \sup_{\theta\in\omega} p_\theta(x)/\sup_{\theta\in\Omega} p_\theta(x), & \text{when } \sup_{\theta\in\Omega} p_\theta(x) > 0, \\ 0, \text{ otherwise.} \end{cases}$$

(This definition is commonly used, cf footnote on p. 15; moreover, if λ is almost invariant then so is $1/\lambda$.)

We first note that $\sup_{\theta\in\omega} p_\theta(x)$ and $\sup_{\theta\in\Omega} p_\theta(x)$ are measurable functions. This follows from the countability of ω and Ω and can be seen as follows

$$\text{for all } c \in \mathbb{R} : \{x : \sup_{\theta\in\Omega} p_\theta(x) \le c\} = \bigcap_{\theta\in\Omega} \{x : p_\theta(x) \le c\}.$$

Define the measurable set A by

$$A = \{x : \sup_{x\in\Omega} p_\theta(x) \le 0\}.$$

Then we have for all $\theta \in \Omega$

$$P_\theta(A) = \int_A p_\theta(x)d\mu(x) \le 0; \text{ i.e. } P_\theta(A) = 0.$$

For any fixed $g \in G$, let h_g be a fixed version of $\frac{d\mu}{d\mu g^{-1}}$, and define for all $\theta \in \Omega$

$$N_g^{(\theta)} = \{x : p_\theta(x) \ne p_{\bar{g}\theta}(gx)h_g(gx)\} \cup \{x : h_g(gx) \le 0\}.$$

It follows from (i) that $\mu(N_g^{(\theta)}) = 0$ for all θ. Hence $N_g^{(\theta)}$ is a $\mathcal{P}$-null set for all $\theta \in \Omega$. Put $N_g = A \cup \{\bigcup_{\theta\in\Omega} N_g^{(\theta)}\}$. N_g is a $\mathcal{P}$-null set (by the countability of Ω).

Suppose $x \notin N_g$. Then by invariance

$$\sup_{\theta\in\Omega} p_\theta(x) = \sup_{\theta\in\Omega} [p_{\bar{g}\theta}(gx)h_g(gx)] = h_g(gx)\cdot\sup_{\theta\in\Omega} p_\theta(gx);$$

$$\sup_{\theta\in\omega} p_\theta(x) = h_g(gx)\cdot\sup_{\theta\in\omega} (gx).$$

It follows that $\lambda(x) = \lambda(gx)$. Hence λ is almost invariant.

(iii) Because Ω is separable there exists by definition a countable subset $D_0 \subset \Omega$ that is dense in Ω. Hence, by the continuity of $p_\theta(x)$ w.r.t. θ for all x

$$\sup_{\theta\in D_0} p_\theta(x) = \sup_{\theta\in\Omega} p_\theta(x).$$

Assume*) that there exists a countable subset $D_1 \subset \omega$ such that for all x

$$\sup_{\theta\in D_1} p_\theta(x) = \sup_{\theta\in\omega} p_\theta(x) \quad \text{a.e. } (\mu).$$

For any fixed $g \in G$, let h_g be a fixed version of $\frac{d\mu}{d\mu g^{-1}}$, and define

$$D_g = \bigcup_{n\in\mathbb{Z}} \bar{g}^n(D_o) \quad \text{and} \quad \tilde{D}_g = \bigcup_{n\in\mathbb{Z}} \bar{g}^n(D_1).$$

D_g and $\tilde{D}_g$ are subsets of Ω and ω respectively (by invariance of Ω and ω with respect to $\bar{G}$); moreover D_g and $\tilde{D}_g$ are both countable. Since $\Omega \supset D_g \supset D_0$ and since $\sup_{\theta\in\Omega} p_\theta(x) = \sup_{\theta\in D} p_\theta(x)$, for all x, we have

$$\sup_{\theta\in\Omega} p_\theta(x) = \sup_{\theta\in D_g} p_\theta(x).$$

Define N_{g1} and N_{g2} by

$$N_{g1} = \{x : \sup_{\theta\in\Omega} p_\theta(x) \le 0\} \cup \{x : h_g(gx) \le 0\},$$

$$N_{g2} = \bigcup_{\theta\in D_g} \{x : p_\theta(x) \ne p_{\bar{g}\theta}(gx)h_g(gx)\}.$$

Then for all x outside $N_{g1} \cup N_{g2}$

$$\sup_{\theta\in\Omega} p_\theta(x) = \sup_{\theta\in D_g} p_\theta(x) = h_g(gx) \cdot \sup_{\theta\in D_g} p_{\bar{g}\theta}(gx)$$

$$= h_g(gx) \cdot \sup_{\theta\in D_g} p_\theta(gx), \quad \text{(since } \theta \in D_g \Leftrightarrow \bar{g}\theta \in D_g\text{, by the definition of } D_g\text{);}$$

(a)

$$= h_g(gx) \cdot \sup_{\theta\in\Omega} p_\theta(gx).$$

Since $\omega \supset \tilde{D}_g \supset D_1$ we have for μ-almost all x

*) This assumption is satisfied if Ω is a separable pseudometric space (see Lemma 1 of the Appendix).

$$\sup_{\theta\in\Omega} p_\theta(x) = \sup_{\theta\in\tilde{D}_g} p_\theta(x).$$

Define N_{g3}, N_{g4}, N_{g5} by

$$N_{g3} = \{x : \sup_{\theta\in\omega} p_\theta(x) \neq \sup_{\theta\in\tilde{D}_g} p_\theta(x)\}$$

$$N_{g4} = \{x : \sup_{\theta\in\omega} p_\theta(gx) \neq \sup_{\theta\in\tilde{D}_g} p_\theta(gx)\}$$

$$N_{g5} = \bigcup_{\theta\in\tilde{D}_g} \{x : p_\theta(x) \neq p_{\bar{g}\theta}(gx)h_g(gx)\}.$$

Then for all x outside $N_{g1} \cup N_{g3} \cup N_{g4} \cup N_{g5}$

$$\sup_{\theta\in\omega} p_\theta(x) = \sup_{\theta\in\tilde{D}_g} p_\theta(x), \quad (\text{since } x \notin N_{g3});$$

$$= h_g(gx) \cdot \sup_{\theta\in\tilde{D}_g} p_{\bar{g}\theta}(gx), \quad (\text{since } x \notin N_{g1} \cup N_{g5});$$

(b)

$$= h_g(gx) \cdot \sup_{\theta\in\omega} p_\theta(gx), \quad (\text{since } x \notin N_{g4}).$$

It is easily seen that $\mu(\cup_{i=1}^5 N_{gi}) = 0$, Hence $\cup_{i=1}^5 N_{gi}$ is a $\mathcal{P}$-null set and the almost invariance of λ (and $1/\lambda$) follows from the equalities (a) and (b).

Problem 18.

Let Ω be the parameter space. $\Omega = \Omega_H \cup \Omega_K$, where

$$\Omega_H = \left\{\left(\frac{\alpha}{n},\ldots,\frac{\alpha}{n},\frac{1-2\alpha}{n},\ldots,\frac{1-2\alpha}{n},\alpha\right)\right\} \subset \mathbb{R}^{2n+1},$$

and

$$\Omega_K = \left\{\left(\frac{p_1}{n},\ldots,\frac{p_n}{n},0,\ldots,0,\frac{n-1}{n}\right) : p_i \geq 0,\ i = 1,\ldots,n,\ \sum_{i=1}^n p_i = 1\right\} \subset \mathbb{R}^{2n+1}.$$

Furthermore $G = \{g_0,\ldots,g_{n-1}\}$ where g_k is the rotation of the plane by the angle $2k\pi/n$. Since both $P_1,\ldots,P_n$ and $Q_1,\ldots,Q_n$ are equidistant points on a circle with centre 0, it follows that

$$P_\theta\{g_k(X,Y) = P_i\} = P_\theta\{(X,Y) = P_{i-k(\text{mod } n)}\},$$

$$P_\theta\{g_k(X,Y) = Q_i\} = P_\theta\{(X,Y) = Q_{i-k(\text{mod } n)}\},$$

$$P_\theta\{g_k(X,Y) = 0\} = P_\theta\{(X,Y) = 0\}.$$

Hence $\bar{g}_k\theta = \theta$ if $\theta \in \Omega_H$, while

$$\bar{g}_k\theta = \left(\frac{p_{1-k(\text{mod } n)}}{n},\ldots,\frac{p_{n-k(\text{mod } n)}}{n},0,\ldots,0,\frac{n-1}{n}\right)$$

if

$$\theta = \left(\frac{p_1}{n},\ldots,\frac{p_n}{n},0,\ldots,0,\frac{n-1}{n}\right) \in \Omega_K.$$

Therefore the testing problem remains invariant under G.

The rejection region R of the level α likelihood ratio test consists of these points (x,y) for which

$$L(x,y) = \frac{\sup_{\theta\in\Omega_K} P_\theta\{(X,Y) = (x,y)\}}{\sup_{\theta\in\Omega_H} P_\theta\{(X,Y) = (x,y)\}} > C,$$

where C is such that $P_\theta\{L(X,Y) > C\} \le \alpha$, for all $\theta \in \Omega_H$.

Since $L(P_i) = \frac{1}{\alpha}$, $L(Q_i) = 0$ and $L(0) = \frac{n-1}{n\alpha}$ we have $L(Q_i) \le L(0) \le L(P_j)$, for $i,j = 1,\ldots,n$. Since moreover $P_\theta\{(X,Y) \in \{P_1,\ldots,P_n\}\} = \alpha$, for all $\theta \in \Omega_H$, we see that $R = \{P_1,\ldots,P_n\}$. Hence the power of the level α likelihood ratio test equals (for all $\theta \in \Omega_K$)

$$P_\theta\{(X,Y) \in \{P_1,\ldots,P_n\}\} = \frac{1}{n}.$$

Now let φ be any invariant test. Then $\varphi(g_k(x,y)) = \varphi(x,y)$ and hence both $\varphi(P_i)$ and $\varphi(Q_i)$ are independent of i. Hence the power of φ equals

$$\sum_{i=1}^{n} \varphi(P_i)\cdot\frac{p_i}{n} + \varphi(0)\cdot\frac{n-1}{n} = \varphi(P_1)/n + \varphi(0)\cdot\frac{n-1}{n}.$$

To determine the level α UMP invariant test, we have to maximize this last expression subject to

$$\alpha \ge \sup_{\theta\in\Omega_H} E_\theta\varphi(X,Y) = \frac{\alpha}{n}\sum_{n=1}^{n}\varphi(P_i) + \frac{1-2\alpha}{n}\sum_{n=1}^{n}\varphi(Q_i) + \alpha\varphi(0)$$

$$= \alpha\varphi(P_1) + (1-2\alpha)\varphi(Q_1) + \alpha\varphi(0).$$

It is easily seen that for $n > 2$ the UMP invariant test is given by $\varphi(P_i) = \varphi(Q_i) = 0$ and $\varphi(0) = 1$, and that the power of this test equals $\frac{n-1}{n}$, for all $\theta \in \Omega_K$. For $n = 1$ the UMP invariant test is given by $\varphi(P_1) = 1$, $\varphi(Q_i) = \varphi(0) = 0$, with power equals to 1; while for $n = 2$, any test with

$\varphi(P_1) = \varphi(P_2) = 1 - \varphi(0) \in [0,1]$ and $\varphi(Q_i)$ is UMP invariant with power $\frac{1}{2}$.

(LEHMANN (1950)).

Problem 19.

(i) Let $A \in \mathcal{A}_0$. Then, by definition, $I_A(x) = I_A(gx)$, for any $g \in G$ and $x \in \mathcal{X} - N_g$, $\mu(N_g) = 0$. Thus $I_{\tilde{A}}(x) = 1 - I_A(x) = 1 - I_A(gx) = I_{\tilde{A}}(gx)$, which implies $\tilde{A} \in \mathcal{A}_0$. Consider any sequence $A_1, A_2, \ldots \in \mathcal{A}_0$. Put $A = \cup_{n=1}^{\infty} A_n$. Without loss of generality suppose that $A_1, A_2, \ldots$ are disjoint. For each n, $I_{A_n}(x) = I_{A_n}(gx)$, for all g, $x \in \mathcal{X} - N_g^{(n)}$, $\mu(N_g^{(n)}) = 0$. Hence $I_A(x) = \Sigma_{n=1}^{\infty} I_{A_n}(x) = \Sigma_{n=1}^{\infty} I_{A_n}(gx) = I_A(gx)$, for all g, $x \in \mathcal{X} - N_g$, $N_g = \cup_{n=1}^{\infty} N_g^{(n)}$, $\mu(N_g) = 0$, which implies $A \in \mathcal{A}_0$. Hence $\mathcal{A}_0$ is a σ-field.
Let f be an almost invariant critical function, that is $f(x) = f(gx)$, for all $g \in G$, $x \in \mathcal{X} - N_g$, $\mu(N_g) = 0$.
Define, for $r \in \mathbb{R}$, $A_r = \{x : f(x) < r\}$. Then $I_{A_r}(x) = I_{A_r}(gx)$, so $A_r \in \mathcal{A}_0$. Hence f is $\mathcal{A}_0$-measurable.
Conversely, let f be $\mathcal{A}_0$-measurable. For any rational number r in $[0,1]$, define $A_r = \{x : f(x) < r\}$. Then $A_r \in \mathcal{A}_0$, which means that I_{A_r} is almost invariant, that is, for any $g \in G$ there exists a set N_g^r with measure zero such that $I_{A_r}(x) = I_{A_r}(gx)$ for all $x \in \mathcal{X} - N_g^r$, which implies that $I_{A_r}(x) = I_{A_r}(gx)$ for all $x \in \mathcal{X} - \cup_r N_g^r$. This means that $f(x) < r$ iff $f(gx) < r$, for all r,g and $x \notin \cup_r N_g^r \stackrel{\text{def}}{=} N_g$, with $\mu(N_g) = 0$. But this means that $f(x) = f(gx)$ for any $x \notin N_g$. Hence f is almost invariant.

(ii) Lehmann does not define the concept of sufficiency for σ-fields.
A possible definition is:
A σ-field $\mathcal{A}_0$ is sufficient for θ iff the $\mathcal{A}_0$-measurable function $T : (X,\mathcal{A}) \to (X,\mathcal{A}_0)$, defined by $T(x) = x$, is sufficient for θ.
HALL, WIJSMAN and GHOSH (1965) give as definition:
A σ-field $\mathcal{A}_0$ is sufficient for θ iff for any $\mathcal{A}$-measurable, $\mathcal{P}$-integrable function f_1 there exists an $\mathcal{A}_0$-measurable function f_2 such that $E_{P_\theta}[f_1(X) \mid \mathcal{A}_0] = f_2$, a.e.-$\mathcal{P}$, where $E_{P_\theta}[f_1(X) \mid \mathcal{A}_0]$ is defined by the relation

$$\text{for all } A \in \mathcal{A}_0, \quad \int_A E_{P_\theta}[f_1(X) \mid \mathcal{A}_0](x)\,dP_\theta(x) = \int_A f_1(x)\,dP_\theta(x).$$

By Problem 10 of Chapter 2 both definitions are equivalent.
We use the definition of HALL, WIJSMAN and GHOSH (1965).

For any $A \in \mathcal{A}$, $g \in G$, $P_\theta(gA) = P_{\bar{g}\theta}(gA) = P_\theta(A)$, by $\bar{g}\theta = \theta$. By Theorem 2 of the Appendix there exists a measure $\lambda = \Sigma\, c_i P_{\theta_i}$ which is equivalent to $\mathcal{P}$. Hence $\lambda(gA) = \lambda(A)$ for all g and A. By Lemma 2 of Chapter 2

$$\int_A \frac{dP_\theta}{d\lambda}(gx)\,d\lambda(x) = \int_{gA} \frac{dP_\theta}{d\lambda}(y)\,d\lambda(g^{-1}y) = \int_{gA} \frac{dP_\theta}{d\lambda}(y)\,d\lambda(y) = P_\theta(gA) =$$

$$= P_\theta(A) = \int_A \frac{dP_\theta}{d\lambda}(x)\,d\lambda(x).$$

Thus $\frac{dP_\theta}{d\lambda}(gx) = \frac{dP_\theta}{d\lambda}(x)$, a.e.$\lambda$. So $\frac{dP_\theta}{d\lambda}$ is almost invariant and consequently $\mathcal{A}_0$-measurable. Consider any $\mathcal{A}$-measurable $\mathcal{P}$-integrable function f_1. Define $f_2 = E_\lambda[f_1(X) \mid \mathcal{A}_0]$. Then f_2 is $\mathcal{A}_0$-measurable and for all $A \in \mathcal{A}_0$

$$\int_A E_{P_\theta}[f_1(X) \mid \mathcal{A}_0](x)\,dP_\theta(x) = \int_A f_1(x)\,dP_\theta(x) =$$

$$= \int_A \frac{dP_\theta}{d\lambda}(x) f_1(x)\,d\lambda(x) = \int_A E_\lambda\left[\frac{dP_\theta}{d\lambda}(X) f_1(X) \mid \mathcal{A}_0\right](x)\,d\ (x) =$$

$$= \int_A \frac{dP_\theta}{d\lambda}(x) f_2(x)\,d\lambda(x) = \int_A f_2(x)\,dP_\theta(x),$$

by Lemma 3 (ii) of Chapter 2.
Hence $f_2 = E_{P_\theta}[f_1(X) \mid \mathcal{A}_0]$, a.e.-$\mathcal{P}$, which implies that $\mathcal{A}_0$ is sufficient for θ.

Section 8

Problem 20.

(i) Let $(j_1, j_2, \ldots, j_n)$ be the permutation of $\{1,2,\ldots,n\}$ for which $Y_{j_1} < Y_{j_2} < \ldots < Y_{j_n}$. Then we have

$$U = \sum_{j=1}^{m} \sum_{i=1}^{n} U_{ij} = \sum_{k=1}^{m} \sum_{i=1}^{n} U_{ij_k} = \sum_{k=1}^{m} \#\{X_i : X_i < Y_{j_k}; i = 1,\ldots,m\}.$$

Put $A = \{X_1,\ldots,X_m,Y_1,\ldots,Y_n\}$. Then the rank S_k satisfies

$$S_k = \#\{Z \in A : Z \le Y_{j_k}\}.$$

Since $\{Z \in A : Z \le Y_{j_k}\} = \{X_i : X_i < Y_{j_k},\ i = 1,\ldots,m\} \cup \{Y_{j_1},\ldots,Y_{j_k}\}$ we have

$$S_k = \#\{X_i : X_i < Y_{j_k},\ i = 1,\dots,m\} + k.$$

Hence $\Sigma_{k=1}^{n} S_k = U + \frac{n(n+1)}{2}$.

(ii) We will show that the total number of steps equals n.m−U. The arrangement x...xy....y can be achieved by steps in which a configuration ...yx... is replaced by ...xy... (if for a given arrangement no pair yx exists, this arrangement must be the final one; and conversely). Hence the total number of steps equals the number of steps of the form ...yx... → ...xy... .

Given n+m distinct points $(x_1,\dots,x_m, y_1,\dots,y_n)$ define $X = \{x_1,x_2,\dots,x_m\}$ and $Y = \{y_1,y_2,\dots,y_n\}$. Let $z = (z_1,z_2,\dots,z_{m+n})$ be any permutation of $(x_1,\dots,x_m,y_1,\dots,y_n)$. Define f(z) as the sum of the i for which $z_i \in Y$. If z_0 is the original arrangement, i.e. if $z_o = (z_1,\dots,z_{m+n})$ with $z_1 < z_2 < \dots < z_{m+n}$, then $f(z_0) = \Sigma_{i=1}^{n} S_i$, with S_i = the rank of y_i. If z_f is the final arrangement, i.e. if $z_f = (x_{i_1},\dots,x_{i_m},y_{j_1},\dots,y_{j_n})$ then $f(z_f) = \Sigma_{k=1}^{n} (m+k) = nm + \frac{n(n+1)}{2}$. Replacing a pair yx by xy increases the value of f by 1. No interchange of neighboring elements is such that the value of f is changed by more than 1. Hence the total number of steps equals $f(z_f) - f(z_0)$ which is equal to nm − U, by (i).

Problem 21.

Let $(X_1,\dots,X_m)$ and $(Y_1,\dots,Y_n)$ denote a sample from the continuous distribution functions F and G respectively. Since U_{ij} equals 1 or 0 whether $X_i < Y_j$ or $X_i > Y_j$ respectively, we have $E(U_{ij}) = P\{X_1 < Y_1\} = \int FdG$. Hence $E(U) = mn \int FdG$.

For the calculation of the variance we observe that

$$E(U_{ij}^2) = E(U_{ij}) = \int FdG,$$

(if $j \neq k$)

$$\begin{aligned} E(U_{ij}U_{ik}) &= P\{X_i < Y_j, X_i < Y_k\} \\ &= \int P\{X_i < Y_j, X_i < Y_k \mid X_i = x\}dF(x) \\ &= \int (1 - G)^2 dF, \end{aligned}$$

(if $i \neq k$)

$$E(U_{ij}U_{kj}) = \int F^2 dG \qquad \text{(by a similar argument)},$$

(if $i \neq k$, $j \neq k$)

$$E(U_{ij}U_{kl}) = [E(U_{ij})]^2 = [\int FdG]^2.$$

Hence

$$E(\Sigma\Sigma U_{ij})^2 = mn \int FdG + mn(n-1) \int (1-G)^2 dF + m(m-1) \int F^2 dG + m(m-1)n(n-1)[\int FdG]^2.$$

This implies relation (30) because mn Var $(U/mn) = [E(U^2) - (E(U))^2]/mn$. If $F = G$ then (31) follows from (29) and (30), since $\int FdF = \frac{1}{2}$ and $\int F^2 dF = \frac{1}{3}$.

Remark: the proof of these results can also be found on pp. 335-336 of LEHMANN (1975).

Problem 22.

(i) Let $(t_1,t_2,\ldots,t_N)$ be any permutation of $(1,2,\ldots,N)$. Then we have

$$P\{T_1 = t_1,\ldots,T_N = t_N\} = \int_A \cdots \int f_1(z_1) \ldots f_N(z_N) dz_1 \ldots dz_N,$$

where $A = \{(z_1,\ldots,z_N) \in \mathbb{R}^N : z_i \neq z_j, z_i$ is the t_i-th smallest of the set $\{z_1,\ldots,z_N\}$ for $i,j = 1,2,\ldots,N, i \neq j\}$. Application of the transformation $w_{t_i} = z_i$ for $i = 1,2,\ldots,N$ yields

$$P\{T_1 = t_1,\ldots,T_N = t_N\} = \int_{w_1<\ldots<w_N} \int f_1(w_{t_1})\ldots f_N(w_{t_N}) dw_1 \ldots dw_N$$

$$= \frac{1}{N!} \int_{w_1<\ldots<w_N} \int \frac{f_1(w_{t_1})}{f(w_{t_1})} \cdots \frac{f_N(w_{t_N})}{f(w_{t_N})} \cdot (N! f(w_1)\cdots f(w_N)) dw_1 \ldots dw_N$$

$$= \frac{1}{N!} E\left[\frac{f_1(V^{(t_1)})}{f(V^{(t_1)})} \cdots \frac{f_N(V^{(t_N)})}{f(V^{(t_N)})} \right],$$

since the density of $V^{(1)},\ldots,V^{(N)}$ equals $N! f(w_1)\ldots f(w_N)$ on the set $\{(w_1,\ldots,w_N) \in \mathbb{R}^N : w_1 < w_2 < \ldots < w_N\}$, and equals zero elsewhere (see Problem 26).

(ii) Applying equation (32) with $N = n+m$; $f_1 = \ldots = f_m = f$, $f_{m+1} = \ldots = f_N = g$; assuming that f is positive whenever g is, we find, for every permutation $(t_1,\ldots,t_m,t_{m+1},\ldots,t_{m+n})$

$$P\{T_1 = t_1,\ldots,T_m = t_m, T_{m+1} = t_{m+1},\ldots,T_{m+n} = t_{m+n}\}$$

$$= \frac{1}{N!} E\left[\frac{f(V^{(t_1)})}{f(V^{(t_1)})}\cdots\frac{f(V^{(t_m)})}{f(V^{(t_m)})}\cdot\frac{g(V^{(t_{m+1})})}{f(V^{(t_{m+1})})}\cdots\frac{g(V^{(t_{m+n})})}{f(V^{(t_{m+n})})}\right]$$

$$= \frac{1}{N!} E\left[\frac{g(V^{(t_{m+1})})}{f(V^{(t_{m+1})})}\cdots\frac{g(V^{(t_{m+n})})}{f(V^{(t_{m+n})})}\right] .$$

Now let $S_1,\ldots,S_n$ denote the ordered ranks of $Z_{m+1},\ldots,Z_{m+n}$ (among the Z's), and let $(s_1,\ldots,s_n)$ be an arbitrary but fixed value of $(S_1,\ldots,S_n)$. Then

$$P\{S_1 = s_1,\ldots,S_n = s_n\} =$$

$$\Sigma\ P\{T_1 = t_1,\ldots,T_m = t_m, T_{m+1} = t_{m+1},\ldots,T_N = t_N\},$$

where the summation extends over all permutations $(t_1,\ldots,t_N)$ of $(1,\ldots,N)$ which satisfy $\{t_{m+1},\ldots,t_{m+n}\} = \{s_1,\ldots,s_n\}$. Since there are precisely $n!(n-m)!$ permutations with this property, and since the probability of occurence for each of these permutations is equal to

$$\frac{1}{N!} E\left[\frac{g(V^{(t_{m+1})})}{f(V^{(t_{m+1})})}\cdots\frac{g(V^{(t_{m+n})})}{f(V^{(t_{m+n})})}\right] = \frac{1}{N!} E\left[\frac{g(V^{(s_1)})}{f(V^{(s_1)})}\cdots\frac{g(V^{(s_n)})}{f(V^{(s_n)})}\right] ,$$

the desired result follows.

Problem 23.

(i) Let $y \in (0,1]$ and $\{x : F(x) = Y\} \neq \emptyset$. Since F is continuous, the set $\{x : F(x) = y\}$ is (non empty and) closed. Hence it contains its infimum. This implies $F[F^{-1}(y)] = y$.

If $y \in (0,1]$ and $\{x : F(x) = y\} = \emptyset$ then $y = 1$ and $F(x) < 1$ for all $x \in \mathbb{R}$. In this case $F^{-1}(y) = +\infty$ and hence $F[F^{-1}(y)] = F(\infty) = 1 = y$.

Evidently $F[F^{-1}(0)] = F(-\infty) = 0$.

To show that $F^{-1}[F(y)]$ may be smaller then y, define F_0 as follows: $F_0(x) = 0$ if $x \le 0$, $F_0(x) = x$ if $0 < x \le \frac{1}{2}$, $F_0(x) = \frac{1}{2}$ if $\frac{1}{2} \le x \le 1$, $F_0(x) = \frac{1}{2}x$ if $1 \le x \le 2$, $F_0(x) = 1$ if $2 \le x$. Then $F_0^{-1}[F_0(\frac{3}{4})] = \frac{1}{2} < \frac{3}{4}$.

(ii) If $F^{-1}(y) \le x$ then $y = F[F^{-1}(y)] \le F(x)$. Conversely we see that (since the infimum of the empty set equals $+\infty$) $F^{-1}(y) = \inf\{x : F(x) \geq y\}$. Hence if $y \leq F(x)$ then $x \in \{x : F(x) \geq y\}$ and $F^{-1}(y) \leq x$. It follows that $F^{-1}(y) \leq x \Leftrightarrow y \leq F(x)$, or equivalently $F(x) < y \Leftrightarrow x < F^{-1}(y)$. Hence $P\{Y < y\} = P\{f(Z) < y\} = P\{Z < F^{-1}(y)\} = h[F[F^{-1}(y)]]$. This last expression is equal to $h(y)$, by (1).

(iii) Taking $h(t) = t$ for $0 \leq t \leq 1$, the desired result follows from (ii).

Remark.

For arbitrary (not necessarily continuous distribution functions F, the relations $F^{-1}(y) \leq x \Leftrightarrow y \leq F(x)$ and $F(x) < y \Leftrightarrow x < F^{-1}(y)$ are still true, provided that F^{-1} is defined by $F^{-1}(y) = \inf\{x : F(x) \geq y\}$.

Problem 24.

(i) Let Z have c.d.f. F. If F' is the c.d.f. of f(Z), then for all y $F'(y) = P\{f(Z) \leq y\} = P\{Z \leq f^{-1}(y)\} = F(f^{-1}(y))$, since f is continuous and strictly increasing.

(ii) Suppose $(F_i)_{i=1,\ldots,N}$ and $(F_i')_{i=1,\ldots,N}$ are on the same orbit. Then for $i = 1,\ldots,N$: $F_i' = F_i(f^{-1})$, with f^{-1} continuous and strictly increasing. Let F be any continuous strictly increasing c.d.f.. Put $h_i = F_i(F^{-1})$ and $F' = F(f^{-1})$. Then $F_i = h_i(F)$ and $F_i' = F_i(f^{-1}) = h_i(F(f^{-1})) = h_i(F')$.

Suppose conversely that $F_i = h_i(F)$ and $F_i' = h_i(F')$, for all i, with F and F' strictly increasing continuous c.d.f.'s. Put $f = (F')^{-1}(F)$. Then f is continuous and strictly increasing; moreover $F_i(f^{-1}) = F_i(F^{-1}(F')) = h_i(F(F^{-1}(F'))) = h_i(F') = F_i'$.

Problem 25.

In view of Problem 23 (ii) $F(Z_i)$ has distribution function h_i with density h_i'. Noting that the ranks of $Z_1,\ldots,Z_N$ are the same as the ranks of $F(Z_1),\ldots,F(Z_N)$ and applying Problem 22 of this chapter with $f_i = h_i'$, $f = I_{(0,1)}$ and $V^{(i)} = U^{(i)}$ we see that (33) holds.

Problem 26.

(i) Since $(Z^{(1)},\ldots,Z^{(N)})$ has density $N!\,\Pi_{i=1}^{N} f(z^{(i)})$ on $\{(z^{(1)},\ldots,z^{(N)}) \mid z^{(1)} \leq z^{(2)} \leq \ldots \leq z^{(N)}\}$, we obtain the result by intergrating

$$N! \prod_{\substack{i=1 \\ i\notin\{s_1,\dots,s_n\}}}^{N} f(z^{(i)}) \prod_{j=1}^{n} f(y_j)$$

over the set $\{(z^{(1)},\dots,z^{(s_1-1)},z^{(s_1+1)},\dots,z^{(N)}) \mid z^{(1)} \le z^{(2)} \le \dots \le z^{(s_1-1)} \le y_1 \le z^{(s_1+1)} \le \dots \le z^{(s_2-1)} \le y_2 \le z^{(s_2+1)} \le \dots \le z^{(N)}\}$, i.e. by computing

$$N! \prod_{j=1}^{N} f(y_j) \int_{-\infty}^{y_1} \int_{z^{(1)}}^{y_1} \int_{z^{(2)}}^{y_1} \dots \int_{z^{(s_1-2)}}^{y_1} \int_{y_1}^{y_2} \int_{z^{(s_1+1)}}^{y_2} \dots \int_{z^{(s_2-2)}}^{y_2} \dots$$

$$\int_{y_n}^{\infty} \int_{z^{(s_n+1)}}^{\infty} \dots \int_{z^{(N-1)}}^{\infty} \prod_{\substack{i=1 \\ i\notin\{s_1,\dots,s_n\}}}^{N} f(z^{(i)}) dz^{(N)} dz^{(N-1)} \dots$$

$$dz^{(s_n+1)} dz^{(s_n-1)} \dots dz^{(s_1+1)} dz^{(s_1-1)} \dots dz^{(1)}.$$

(cf. DAVID (1970) p.9, HÁJEK and ŠIDÁK (1967), Theorem II,1.2.c.).

(ii) needs no comment.

(iii) The Jacobian of the transformation $(v_1,\dots,v_n) \to (y_1,\dots,y_n)$, given by $y_i = \Pi_{j=i}^{n} v_j$, is the determinant of an $n \times n$ upper triangular matrix with diagonal elements $\Pi_{j=i+1}^{n} v_j$, $i = 1,\dots,n$, and consequently equals $\Pi_{j=2}^{n} v_j^{j-1}$. Hence the distribution of $(V_1,\dots,V_n)$ is given by

$$\frac{N!}{(s_1-1)!\dots(N-s_n)!} v_2 v_3^2 \dots v_n^{n-1} (v_1,\dots,v_n)^{s_1-1}$$

$$\cdot [(1-v_1)v_2\dots v_n]^{s_2-s_1-1} [(1-v_2)v_3\dots v_n]^{s_3-s_2-1} \dots (1-v_n)^{N-s_n}$$

which factorizes in

$$\prod_{i=1}^{n} \frac{(s_{i+1}-1)!}{(s_i-1)!(s_{i+1}-s_i-1)!} v_i^{s_i-1} (1-v_i)^{s_{i+1}-s_i-1}$$

with $s_{n+1} = N+1$.

<u>Problem 27.</u>

(i) Write $V_i = F(X_i)$ $(i = 1,2,\dots,m)$ and $W_j = F(Y_j)$ $(j = 1,2,\dots,n)$. From problem 23 of this chapter it follows that $V_1,\dots,V_m$ are i.i.d. and $R(0,1)$ distributed and that $W_1,\dots,W_n$ are i.i.d. with distribution function h concentrated on $(0,1)$. Since h is differentiable the density of each of the W_j equals h'. For any pair (i,j) with $i \neq j$:

$$P\{X_i = X_j\} = \int_{\mathbb{R}} P\{X_i = x_j \mid x_j\} dF(x_j) = 0,$$

$P\{Y_i = Y_j\} = 0$ (analogously), $P\{V_i = V_j\} = P\{U_1 = U_2\} = 0$ and $P\{W_i = W_j\} \le P\{G(Y_i) = G(Y_j)\} = P\{U_1 = U_2\} = 0$, where U_1 and U_2 are independent and R(0,1) distributed. Hence, with probability 1, rank (X_i) = rank (V_i) $(i = 1,2,\ldots,m)$ and rank (Y_j) = rank (W_j) $(j = 1,2,\ldots,n)$. Let S_j denote rank (Y_j), then applying Problem 22 (ii) we find

$$P\{S_1 = s_1,\ldots,S_n = s_n\} = E[h'(U^{(s_1)})\ldots h'(U^{(s_n)})] \Big/ \binom{m+n}{n},$$

where $U^{(1)} < \ldots < U^{(m+n)}$ is an ordered sample from the R(0,1) distribution.

(ii) Let $h(x) = x^k$ for $0 \le x < 1$ where k is a positive integer, then $E[h'(U^{(s_1)})\ldots h'(U^{(s_n)})] = k^n E[(U^{(s_1)})^{k-1}\ldots(U^{(s_n)})^{k-1}]$.
By Problem 26 (iii) there exist independent random variables $R_1,R_2,\ldots,R_n$ such that $U^{(s_i)} = R_i R_{i+1}\ldots R_n$ $(i = 1,2,\ldots,n)$, and R_i has the beta distribution $B_{s_i, s_{i+1}-s_i}$. Hence

$$E[(U^{(s_1)})^{k-1}\ldots(U^{(s_n)})^{k-1}] = E\left[\prod_{i=1}^{n}\prod_{j=1}^{n} R_j^{k-1}\right] = \prod_{j=1}^{n} ER_j^{j(k-1)} =$$

$$\prod_{j=1}^{n} \frac{\Gamma(s_j+j(k-1))}{\Gamma(s_j)} \cdot \frac{\Gamma(s_{j+1})}{\Gamma(s_{j+1}+j(k-1))} .$$

From this (37) easily follows.

(cf. DAVID (1970), pp. 27-28)

Problem 28.

Applying Problem 27 with $h(x) = (1-\theta)x + \theta x^2$ we see that the distribution of the ranks $S_1 < \ldots < S_n$ of the Y's is given by

$$P_\theta\{S_1 = s_1,\ldots,S_n = s_n\} = \binom{N}{n}^{-1} E \prod_{i=1}^{n} (1 + \theta(2U^{(s_i)} - 1)),$$

where $U^{(1)} < \ldots < U^{(N)}$ is an ordered sample from the R(0,1) distribution. For $\theta \to 0$ this means

$$P_\theta\{S_1 = s_1,\ldots,S_n = s_n\} = \binom{N}{n}^{-1}\left\{1 + \theta \sum_{i=1}^{n} E(2U^{(s_i)} - 1)\right.$$

$$+ \theta^2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E\left[(2U^{(S_i)} - 1)(2U^{(S_j)} - 1)\right] + O(\theta^3)\Big\} =$$

$$\binom{N}{n}^{-1}\Big\{1 + \theta \sum_{i=1}^{n} \left(\frac{2S_i}{N+1} - 1\right) + \theta^2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Big[4\frac{S_iS_j + S_i \wedge S_j}{(N+1)(N+2)} -$$

$$2\frac{S_i + S_j}{N+1} + 1\Big] + O(\theta^3)\Big\}.$$

By the Neyman-Pearson lemma we see that the derivative of the power function at $\theta = 0$ is maximized among rank tests by each test of the Wilcoxon type, i.e. each test with critical function $\varphi \in [0,1]$ satisfying

$$\varphi(S_1,\ldots,S_n) = \begin{cases} 1 \\ \text{something} \\ 0 \end{cases} \quad \text{if } \sum_{i=1}^{n} S_i \begin{matrix} > \\ = \\ < \end{matrix} C,$$

where C and "something" have to be such that $E_0\varphi(S_1,\ldots,S_n) = \alpha$.
We make two remarks.

1. For some significance levels α i.e. those α not equal to $k / \binom{N}{n}$ for any integer k, there exist infinitely many tests of the Wilcoxon type which maximize the derivative of the power function. For such a significance level we choose a test which maximizes the second derivative given the significance level and given that it maximizes the first derivative. However, there may again exist infinitely many of such tests. We may proceed in this way up to and including the n-th derivative.
2. By an argument analogous to the one given after formula (19) on p. 237 we see that for every α there exists a $\theta(\alpha) > 0$ such that, uniformly for $\theta \in (0,\theta(\alpha))$, every test of the Wilcoxon type has power higher than every test which is not of Wilcoxon type (which means that $\varphi(S_1,\ldots,S_n) > 0$ for some $(S_1,\ldots,S_n)$ with $\Sigma_{i=1}^{n} S_i < C$). Tests with this property are called locally most powerful, rank, LMPR, tests. (See Problem 2 of Chapter 8.)

We may draw the following conclusion.

Let α be a fixed significance level. We proceed along the lines of remark 1. If there exists a j such that considering the j-th derivative yields a unique test, then there exists (by an argument similar to the one given in remark 2) a $\theta(\alpha) > 0$ such that uniformly for $\theta \in (0,\theta(\alpha))$ this test has a power higher than every other test. If an analysis of all derivatives (including the n-th one) does not yield a unique optimal test, then the remaining optimal tests are equivalent in the sense that their power functions are equal.

Problem 29.

The result of Problem 28 can be generalized to alternatives (F,G) with $G = (1-\theta)F + \theta F^2 + \theta^2 M_1(F) + \theta^3 M_2(F) + \theta^4 M_3(F) + \dots$, where M_i $(i = 2,3,\dots)$ are differentiable (Application of Problem 27 with $h(x) = (1-\theta)x + \theta x^2 + \theta^2 M_1(x) + \theta^3 M_2(x) + \dots$ yields that the distribution of the ranks $S_1 < \dots < S_n$ of the Y's is given by

$$(16) \qquad P_\theta\{S_1 = s_1,\dots,S_n = s_n\} = \binom{N}{n}^{-1} E \prod_{i=1}^{n} (1-\theta+2\theta U^{(s_i)} + \theta^2 M_1'(U^{(s_i)}) + \dots).$$

Differentiation of (16) w.r.t. θ under the expectation sign will give the same expression for $\frac{\partial}{\partial\theta} P_\theta\{S_1 = s_1,\dots,S_n = s_n\}\Big|_{\theta=0}$ as in Problem 28. Hence in the problem of detecting a shift θ in a distribution F the Wilcoxon test is a LMPR test (see Problem 2 of Chapter 8) for those distributions F for which the alternatives $G(x) = F(x-\theta)$ are of the form $G(x) = (1-\theta)F(x) + \theta F^2(x) = \theta^3 M_1(F(x)) + \dots$. Of course the remarks in Problem 28 remain valid here.
If $F'(x) = F(x) - F^2(x)$, using the expression $F(x-\theta) = F(x) - \theta F'(x) + \frac{1}{2}\theta^2 F''(x) - \frac{\theta^3}{6} F'''(x) + \dots$, we see that $F(x-\theta)$ is of the form $(1-\theta)F(x) + \theta F^2(x) + \theta^2 M_1(F(x)) + \theta^3 M_2(F(x)) + \dots$ as above, provided $F''(x), F''(x),\dots$ can be expressed as differentiable functions of F, which is seen to be true by using $F'(x) = F(x) - F^2(x)$ and finding $F''(x) = F'(x) - 2F'(x)F(x) = F(x) - 3F^2(x) + 2F^3(x)$, etc.
The logistic distribution $F(x) = 1/(1+e^{-x})$ satisfies the differential equation $F' = F - F^2$.

Problem 30.

It is sufficient to prove $\forall F_1 \in \mathcal{F}_1 : E_{F_1}\varphi(X) \geq \alpha$. Let $F_1 \in \mathcal{F}_1$, $f \in \mathcal{C}$ and $F_0 \in \mathcal{F}_0$ be such that F_1 is the distribution function of the random variable $f(X)$, where X is assumed to have distribution F_0. Then by conditions (a) and (b)

$$E_{F_1}\varphi(X) = E_{F_0}\varphi(f(X)) \geq E_{F_0}\varphi(X) = \alpha.$$

Hence φ is unbiased for testing $\mathcal{F}_0$ against $\mathcal{F}_1$.

Problem 31.

Recall (Problem 20) that $U = \Sigma_{i=1}^{n} S_i - \frac{1}{2}n(n+1)$, where $S_1 < S_2 < \ldots < S_n$ denote the ranks of the Y's. Further we know that under the hypothesis the samples are from a common distribution

$$P\{S_1 = s_1,\ldots,S_n = s_n\} = \frac{1}{\binom{m+n}{n}}$$

Define S_i' (the inverse rank) as $(m+n+1) - S_i$ (i.e. the subject that held rank S_i now has rank S_i'). Then, under the hypothesis,

$$P\{S_1' = s_1,\ldots,S_n' = s_n\} = \frac{1}{\binom{m+n}{n}} ,$$

and therefore $\Sigma_{i=1}^{n} S_i$ and $\Sigma_{i=1}^{n} S_i'$ have the same distribution.
From

$$\sum_{i=1}^{n} S_i' = \{(m+n+1) - S_1\} + \ldots + \{(m+n+1) - S_n\} =$$
$$= n(m+n+1) - \sum_{i=1}^{n} S_i$$

it follows that, again under the hypothesis,

$$\sum_{i=1}^{n} S_i \quad \text{and} \quad n(m+n+1) - \sum_{i=1}^{n} S_i$$

have the same distribution. By substracting $\frac{1}{2}n(n+1) - \frac{1}{2}mn$ from these statistics we find that

$$U - \tfrac{1}{2}mn \text{ and } \tfrac{1}{2}mn - U$$

are identically distributed, and hence U is distributed symmetrically about $\frac{1}{2}mn$.

(LEHMANN (1975)).

Problem 32.

(i) Apply the two sided Wilcoxon test to the sample $X_1,\ldots,X_m$ and $Y_1 - \Delta_0,\ldots,Y_n - \Delta_0$. Let $S_1,\ldots,S_m$ and $R_1,\ldots,R_n$ denote the ranks of $X_1,\ldots,X_m$ and $Y_1 - \Delta_0,\ldots,Y_n - \Delta_0$ in the combined sample. The two sided Wilcoxon test rejects if

$$T = \left| \sum_{i=1}^{m} S_i - \sum_{i=1}^{n} R_i \right| > C$$

for some suitable chosen C.

This statistic attains its maximal value if the combined sample can be ordered in one of the following two ways

$$(17)\qquad \begin{array}{l} X^{(1)} < \ldots < X^{(m)} < Y^{(1)} - \Delta_0 < \ldots < Y^{(n)} - \Delta_0 \\ Y^{(1)} - \Delta_0 < \ldots < Y^{(n)} - \Delta_0 < X^{(1)} < \ldots < X^{(m)}. \end{array}$$

T attains its second largest value in case of the orderings

$$(18)\qquad \begin{array}{l} X^{(1)} < \ldots < X^{(m-1)} < Y^{(1)} - \Delta_0 < X^{(m)} < Y^{(2)} - \Delta_0 < \ldots < Y^{(n)} - \Delta_0 \\ Y^{(1)} - \Delta_0 < \ldots < Y^{(n-1)} - \Delta_0 < X^{(1)} < Y^{(n)} - \Delta_0 < X^{(2)} < \ldots < X^{(m)}. \end{array}$$

Under $H : \Delta = \Delta_0$ each ordering of the combined sample has equal probability $1/\binom{m+n}{n}$. Therefore the rejection region for $\alpha = 2/\binom{m+n}{n}$ is given by (17) and for $\alpha = 4/\binom{m+n}{n}$ by (17) and (18) together, i.e. reject H if (17) or (18) is satisfied.

Hence for $\alpha = 2/\binom{m+n}{n}$ we find the confidence region

$$X^{(m)} > Y^{(1)} - \Delta_0 \quad \text{and} \quad Y^{(n)} - \Delta_0 > X^{(1)},$$

and a confidence interval for Δ

$$Y^{(1)} - X^{(m)} < \Delta < Y^{(n)} - X^{(1)},$$

with confidence coefficient $1 - 2/\binom{m+n}{n}$.

The acceptance region for $\alpha = 4/\binom{m+n}{n}$ is given by the negation of ((1) or (2)) which is equivalent to

$$(19)\qquad \begin{array}{l} X^{(m)} > Y^{(1)} - \Delta_0, \\ Y^{(n)} - \Delta_0 > X^{(1)}, \\ (X^{(m-1)} > Y^{(1)} - \Delta_0 \quad \text{or} \quad X^{(m)} > Y^{(2)} - \Delta_0), \\ (Y^{(n-1)} - \Delta_0 > X^{(1)} \quad \text{or} \quad Y^{(n)} - \Delta_0 > X^{(2)}). \end{array}$$

Using

$$Y^{(1)} - X^{(m-1)} > Y^{(1)} - X^{(m)}$$
$$Y^{(2)} - X^{(m)} > Y^{(1)} - X^{(m)}$$
$$Y^{(n)} - X^{(2)} < Y^{(n)} - X^{(1)}$$
$$Y^{(n-1)} - X^{(1)} < Y^{(n)} - X^{(1)}$$

(19) is seen to be equivalent to

$$\min (Y^{(1)} - X^{(m-1)}, Y^{(2)} - X^{(m)}) < \Delta_0 < \max (Y^{(n)} - X^{(2)}, Y^{(n-1)} - X^{(1)}),$$

a confidence interval for Δ with confidence coefficient $1 - 4/\binom{m+n}{n}$.

(ii) We follow the construction of confidence intervals for a shift as proposed in LEHMANN (1975), pp. 91-95. As in Problem 20 U denotes the number of pairs i,j for which $Y_j > X_i$. Let $D_{(1)} < \ldots < D_{(mn)}$ denote the ordered differences $Y_j - X_i$. With $D_{(0)} = -\infty$ and $D_{(mn+1)} = \infty$ we have

$$(20) \qquad P_\Delta\{D_{(s)} \le \Delta < D_{(s+1)}\} = P_0\{U = s\}, \qquad s = 0,1,\ldots,m.$$

(for proof see LEHMANN (1975)), where the subscript on P indicates the value of Δ for which the probability is computed. Since the right hand side of (20) is tabled for small values of m and n we are able to construct a confidence interval for Δ.
Equation (20) implies

$$(21) \qquad P_\Delta\{\Delta \in [D_{(s)}, D_{(t)})\} = \sum_{i=s}^{t-1} P_0\{U = i\} \quad \text{for } 0 \le s < t \le mn+1.$$

If $m = n = 6$ and $\alpha = 1/21 = .048$ we have

$$P_0\{U \le 5\} = .0206 < \tfrac{1}{2}\alpha$$
$$P_0\{U \le 6\} = .0325 > \tfrac{1}{2}\alpha.$$

Since the distribution of U is symmetric about 18 (Problem 31) if $\Delta = 0$ the interval $[D_{(6)}, D_{(31)})$ is a confidence interval for Δ with confidence coefficient 20/21. $D_{(6)}$ and $D_{(31)}$ can be determined by writing down all differences $Y_i - X_j$, which leads to $D_{(6)} = -.089$ and $D_{(31)} = .804$.

Note that the construction of (ii) can be used to derive the confidence intervals of (i).

Problem 33.

(i) Since $\max(X,X')$ and $\min(Y,Y')$ have distribution functions F^2 and $1-(1-G)^2$, the probability $P\{\max(X,X') < \min(Y,Y')\}$ equals both

$$\int F^2 d[1-(1-G)^2] = 2\int F^2(1-G)dG$$

and

$$\int (1-G)^2 dF^2 = 2\int F(1-G)^2 dF.$$

We conclude that

$$\begin{aligned} p &= \int F^2(1-G)dG + \int F(1-G)^2 dF + \int (1-F)G^2 dF + \int (1-F)^2 G dG \\ &= \int (F-F^2)dF + \int (G-G^2)dG + \int (F-G)^2 d[F+G] \\ &= \tfrac{1}{3} + 2\Delta. \end{aligned}$$

(ii) If $F = G$ it immediately follows that $\Delta = 0$.
Let $\Delta = 0$. Suppose that $G(x_1) - F(x_1) = \eta > 0$ then there exists $x_0 < x_1$ such that $G(x_0) = F(x_1) + \frac{1}{2}\eta$. Hence for all $x \in [x_0, x_1]$ we have $G(x) - F(x) \geq G(x_0) - F(x_1) = \frac{1}{2}\eta$. Moreover, $G(x_1) - G(x_0) = \frac{1}{2}\eta$ implying $\Delta > 0$, a contradiction. Similarly the assumption $F(x_1) - G(x_1) < 0$ leads to a contradiction. This completes the proof.

(LEHMANN (1951)).

Problem 34.

(i) Let $\alpha \in (0,1)$ and let $X_i, X_i'; Y_i, Y_i'$, $(i = 1,2,\ldots,n)$, be samples of size n from F and G. Because of Problem 33, the given problem is equivalent to that of testing $p = \frac{1}{3}$ against $p > \frac{1}{3}$, with
$p = \frac{1}{3} + 2\int (F-G)^2 d[(F+G)/2]$.
If we put

$$V_i = \begin{cases} 1 & \text{if } \max(X_i,X_i') < \min(Y_i,Y_i') \text{ or } \max(Y_i,Y_i') < \min(X_i,X_i') \\ 0 & \text{otherwise,} \end{cases}$$

$(i = 1,2,\ldots,n)$, then $\Sigma_{i=1}^n V_i$ is a binomial(n,p) r.v. and the (randomized) level α test which rejects if $\Sigma_{i=1}^n V_i > c$, for some constant c, is strictly unbiased because of Theorem 2.(ii) of Chapter 3.

(ii) Consider samples $X_1,\ldots,X_m$ from a distribution F and $Y_1,\ldots,Y_n$ from a distribution G. We prove that there does not exist a nonrandomized unbiased rank test of H against all $F \neq G$ at level $\alpha = 1/\binom{m+n}{n}$. In fact, if we denote the ordered ranks of the X's in the combined sample by $(R_1,\ldots,R_m)$, then a nonrandomized rank test at level α has its rejection region equal to $\{R_1 = r_1^0,\ldots,R_m = r_m^0\}$ for some $(r_1^0,\ldots,r_m^0) \in \{(r_1,\ldots,r_m) \in \{1,2,\ldots,m+n\}^m : r_1 < r_2 < \ldots < r_m\}$. But now the power against at least one of the alternatives for which $P\{X < Y\} = 1$ or $P\{Y < X\} = 1$, is zero, because for these alternatives

$$P\{R_1 = 1,\ldots,R_m = m\} = 1, \text{ resp.}$$

$$P\{R_1 = n+1,\ldots,R_m = m+n\} = 1.$$

(LEHMANN (1951)).

Section 9.

Problem 35.

(i) Let K be the number of distinct sums $Z_i + Z_j$ that are positive. Then

$$K = \sum_{i=1}^{N} \sum_{j=1}^{N} U_{ij}, \text{ where } U_{ij} = \begin{cases} 1 \text{ if } Z_j > 0 \text{ and } |Z_i| \leq |Z_j|, \\ 0 \text{ otherwise.} \end{cases}$$

Furthermore, let

$$A_j = \begin{cases} 1 \text{ if } Z_j > 0 \\ 0 \text{ otherwise} \end{cases}, \quad V_{ij} = \begin{cases} 1 \text{ if } |Z_i| \leq |Z_j| \\ 0 \text{ otherwise.} \end{cases}$$

Let $R_1,\ldots,R_N$ be the ranks of $|Z_1|,\ldots,|Z_N|$. Then

$$U_{ij} = A_j V_{ij} \quad \text{and} \quad \sum_{j=1}^{N} V_{ij} = R_j.$$

It follows that

$$\sum_{j=1}^{N} S_j = \sum_{j=1}^{N} A_j R_j = \sum_{j=1}^{N} A_j \sum_{i=1}^{N} V_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} U_{ij} = K,$$

which is the desired result.

(ii) Let the function h be defined as

$$h(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Define

$$U_N^{(A)} = \frac{1}{N} \sum_{i=1}^{N} h(Z_i),$$

$$U_N^{(B)} = \binom{N}{2}^{-1} \sum\sum_{1 \le i < j \le N} h(Z_i + Z_j).$$

Now PURI and SEN ((1971), p. 172) remark that

$$\binom{N}{2}^{-1} \sum_{j=1}^{N} S_j = U_N^{(B)} + \frac{2}{N-1} U_N^{(A)},$$

or equivalently

$$\sum_{j=1}^{N} S_j = \sum\sum_{1 \le i < j \le N} h(Z_i + Z_j) + \sum_{i=1}^{N} h(Z_i).$$

Hence

$$E\left(\sum_{j=1}^{N} S_j\right) = \frac{N(N-1)}{2} P\{Z_1 + Z_2 > 0\} + NP\{Z_1 > 0\}.$$

Since

$$P\{Z_1 + Z_2 > 0\} = 1 - \int D(-z)dD(z)$$

and

$$P\{Z_1 > 0\} = 1 - D(0)$$

it follows that

$$E\left(\sum_{j=1}^{n} S_j\right) = \frac{N(N-1)}{2}\{1 - \int D(-z)dD(z)\} + N(1 - D(0)) =$$

$$= \tfrac{1}{2}N(N+1) - ND(0) - \tfrac{1}{2}N(N-1) \int D(-z)dD(z).$$

(WALSH (1949)).

Problem 36.

(i) As in Section 9 we characterize the problem by the triple (ρ,F,G) with $\rho = P\{Z \le 0\}$, $F(z) = P\{|Z| \le z \mid Z < 0\}$ and $G(z) = P\{Z \le z \mid Z > 0\}$. The density of F is $\rho^{-1}f(z+\theta)$ and the density of G is $(1-\rho)^{-1}f(z-\theta)$. Let $1 \le i_1 < \dots < i_n \le N$ and let $C_{i_1 \dots i_n}$ be the event "$Z_{i_1} > 0, \dots, Z_{i_n} > 0$, $Z_j < 0$ for $j \notin \{i_1,\dots,i_n\}$". So $i_1,\dots,i_n$ are the indices of the positive

Z's. Then by conditioning on $C_{i_1 \ldots i_n}$ and by applying Problem 22 (i) we find for any possible set of ranks $t_1, \ldots, t_N$ of $|Z_1|, \ldots, |Z_N|$

$$P\{T_1 = t_1, \ldots, T_N = t_N \mid C_{i_1 \ldots i_n}\} =$$

$$= \frac{1}{N!} E\Big\{ \prod_{j \in \{i_1,\ldots,i_n\}} (1-\rho)^{-1} f(V^{(t_j)} - \theta) \cdot$$

$$\prod_{j \notin \{i_1,\ldots,i_n\}} \rho^{-1} f(V^{(t_j)} + \theta)/2f(V^{(1)}) \cdots 2f(V^{(N)}) \Big\}$$

where T_j is the rank of $|Z_j|$ and $V^{(1)} < \ldots < V^{(N)}$ is an ordered sample from a distribution with density $2fI_{(0,\infty)}$.

Next let $r_1 < \ldots < r_m$ be the ordered ranks of the absolute values of the negative Z_i in the sample $|Z_1|, \ldots, |Z_N|$ and similar by $s_1 < \ldots < s_n$ the ordered ranks of the absolute values of the positive Z's, then there are m!n! possible sets of ranks $\{t_1, \ldots, t_N\}$ having these prescribed sets of ordered ranks and we find

(22)
$$P\{S_1 = s_1, \ldots, S_n = s_n \mid C_{i_1 \ldots i_n}\} = \frac{m!n!\rho^{-m}(1-\rho)^{-n}}{2^N N!} \cdot$$
$$E\left\{\frac{f(V^{(r_1)} + \theta) \cdots f(V^{(r_m)} + \theta) f(V^{(s_1)} - \theta) \cdots f(V^{(s_n)} - \theta)}{f(V^{(1)}) \cdots f(V^{(N)})}\right\} .$$

Since this probability is independent of $i_1, \ldots, i_n$ it also equals

(23) $$P\{S_1 = s_1, \ldots, S_n = s_n \mid \text{the number of positive Z's is n}\} .$$

Multiplying by $\binom{N}{n}\rho^m(1-\rho)^n$, the probability of n positive Z's gives formula (38).

(ii) Firstly we derive the rank test of the hypothesis of symmetry which maximizes the derivative of the conditional powerfunction given the number of positive Z's. The reasoning is analogous to the reasoning on p. 237.

Since $\rho = \int_{-\infty}^{-\theta} f(z)dz$ we have $\frac{d}{d\theta}\rho\Big|_{\theta=0} = -f(0)$ and therefore

$$\frac{d}{d\theta}\rho^{-m}\Big|_{\theta=0} = \frac{-mf(0)}{2^{-m-1}} \quad \text{and} \quad \frac{d}{d\theta}(1-\rho)^{-n}\Big|_{\theta=0} = \frac{-nf(0)}{2^{-n-1}} .$$

Now, under some regularity conditions, the derivative of (22) and (23) at $\theta = 0$ is equal to

$$M + \binom{N}{n}^{-1}\left[E \sum_{j=1}^{m} \frac{f'(V^{(r_j)})}{f(V^{(r_j)})} - E \sum_{j=1}^{n} \frac{f'(V^{(s_j)})}{f(V^{(s_j)})}\right] =$$

$$M + \binom{N}{n}^{-1}\left[E \sum_{j=1}^{N} \frac{f'(V^{(t_j)})}{f(V^{(t_j)})} - 2E \sum_{j=1}^{n} \frac{f'(V^{(s_j)})}{f(V^{(s_j)})}\right] =$$

$$M' - 2\binom{N}{n}^{-1} E \sum_{j=1}^{n} \frac{f'(V^{(s_j)})}{f(V^{(s_j)})},$$

where M and M' are constants depending on n, but not on $r_1,\ldots,r_m$ and $s_1,\ldots,s_n$.

Since under the hypothesis $H : \theta = 0$ the probability of any outcome is equal to 2^{-N} it follows from the generalization of the Neyman-Pearson lemma, Theorem 5 of Chapter 3, that the derivative of the conditional power function at $\theta = 0$ is maximized by a test which rejects if

$$M' - 2\binom{N}{n} E \sum_{j=1}^{n} \frac{f'(V^{(s_j)})}{f(V^{(s_j)})} > C'_n$$

which is equivalent to

$$(24) \qquad -E \sum_{j=1}^{n} \frac{f'(V^{(s_j)})}{f(V^{(s_j)})} > C_n,$$

where C_n and C'_n are constants depending on n, but not on $r_1,\ldots,r_m$ or $s_1,\ldots,s_n$.

This conditional test also maximizes the conditional power function for sufficiently small θ (see p. 237).

Now consider the unconditional test we get by rejecting if there are n positive Z's and (24) holds. This test has the desired properties.

(iii) If f is the normal density function with zero mean and variance σ^2 we have

$$\frac{f'(z)}{f(z)} = -\frac{z}{\sigma^2}, \text{ so (24) becomes } E \sum_{j=1}^{n} V^{(s_j)} > C_n$$

where $V^{(1)} < \ldots < V^{(N)}$ is an ordered sample from a distribution with density $2fI_{(0,\infty)}$, which is, in case f is the standard normal density function, the distribution of the square root of a χ_1^2 distributed r.v., which in its turn is equal to the distribution of the absolute value of a standard normal r.v..

(iv) Let $F(x) = 1/(1+e^{-x})$, $f(x) = e^{-x}/(1+e^{-x})^2$.
Then $-f'(x)/f(x) = 2F(x) - 1$ and the conditional rejection region (24) can be written as

$$E \sum_{j=1}^{n} F(V^{(s_j)}) > C_n.$$

If V has the distribution with density $2fI_{(0,\infty)}$, then

$$P\{F(V) \le y\} = P\{V \le F^{-1}(y)\} = 2y - 1 \text{ for } \tfrac{1}{2} < y < 1,$$

so $U = F(V)$ is uniformly distributed over $(\frac{1}{2},1)$. The conditional rejection region can therefore be written as

$$E \sum_{j=1}^{n} U^{(s_j)} > C_n,$$

where $U^{(1)} < \ldots < U^{(N)}$ is an ordered sample of size N from the uniform distribution $R(\frac{1}{2},1)$. Since

$$EU^{(s_j)} = \tfrac{1}{2} \frac{s_j}{N+1} + \tfrac{1}{2}$$

conditionally on the number of positive Z's the test is equivalent to the Wilcoxon one sample test.

Problem 37.

Apply Problem 25 in the same way Problem 22 (i) was applied in the previous problem to obtain the expression (38).

Problem 38.

We restrict ourselves to continuous distributions. Suppose Z has a continuous distribution function F, and let $S = \{G : G$ is a continuous distribution function with $G(x) + G(-x) = 1\}$. Since there belongs to every F in the alternative some $G \in S$ with $F(x) \le G(x)$ for all x, there exists in view of Lemma 1 of Chapter 3 nondecreasing functions f_0 and f_1 and a random variable V such that $f_0(v) \le f_1(v)$ and $f_0(V)$ and $f_1(V)$ have distribution function G and F respectively. Using the monotonicity hypothesis on φ it follows that, with $V_1,\ldots,V_N$ a sample form the distribution G,

$$(25)\qquad \begin{aligned} E_F\varphi(Z_1,\ldots,Z_N) &= E\varphi(f_1(V_1),\ldots,f_1(V_N)) \geq E\varphi(f_0(V_1),\ldots,f_0(V_N)) \\ &= E_G\varphi(Z_1,\ldots,Z_N) \qquad \text{(cf. Problem 11 of Chapter 3).} \end{aligned}$$

Since φ is a rank test it is symmetric in its N variables and therefore by Lemma 3, $E_G\varphi(Z_1,\ldots,Z_N)$ is constant on S as a function of G. Hence using (25)

$$E_F\varphi(Z_1,\ldots,Z_N) \geq \sup_{G\in S} E_G\varphi(Z_1,\ldots,Z_N),$$

for any F in the alternative, implying the unbiasedness of φ.

Problem 39.

We consider the problem of testing $H : F_1 = F_2 = \ldots = F_N$ against $K : i < j \Leftrightarrow \forall\, x\; F_i(x) \geq F_j(x)$ and $\exists\, x\; F_i(x) > F_j(x)$.
The joint distribution of the T_i under H is

$$(26)\qquad P\{T_1 = t_1, T_2 = t_2, \ldots, T_N = t_N\} = \frac{1}{N!}$$

for each permutation $t_1,\ldots,t_N$ of $1,\ldots,N$.
If $F_1,\ldots,F_N$ have densities $f_1,\ldots,f_N$, respectively, and f is any density which is positive whenever at least one of the f_i is positive, then the joint distribution of the T_i is given by

$$(27)\qquad P\{T_1 = t_1,\ldots,T_N = t_N\} = \frac{1}{N!}E\left[\frac{f_1(W^{(t_1)})}{f(W^{(t_1)})}\cdots\frac{f_N(W^{(t_N)})}{f(W^{(t_N)})}\right],$$

where $W^{(1)} < \ldots < W^{(N)}$ is an ordered sample from a distribution with density f. (See Problem 22).
Consider in particular the translation alternatives $f_i(y) = f(y - i\delta)$, $f_i > 0$ and the problem of maximizing the power for small values of δ. Suppose that f is differentiable and that the probability (27), which is now a function of δ, can be differentiated with respect to δ under the expectation sign.
The derivative of (27) at $\delta = 0$ is then

$$\frac{\partial}{\partial\delta} P_\delta\{T = t_1,\ldots,T = t_N\}\Big|_{\delta=0} = -\frac{1}{N!}E\left[\sum_{i=1}^{N} i\,\frac{f'(W^{(t_i)})}{f(W^{(t_i)})}\right].$$

Since under the hypothesis the probability of any ranking is given by (26) it follows from the Neyman-Pearson lemma in the extended form of Theorem 5 of Chapter 3, that the derivative of the power function at $\delta = 0$ is maximized by the rejection region:

$$(28) \qquad -\sum_{i=1}^{n} i\,E\left[\frac{f'(W^{(t_i)})}{f(W^{(t_i)})}\right] > C.$$

The same test maximizes the power itself for sufficiently small δ. To see this the same arguments as in Problem 28 can be used here. Hence (28) is a LMPR test.

(i) If $f(x)$ is the normal density $N(\gamma,\sigma^2)$, $-\frac{f'(x)}{f(x)} = -\frac{d}{dx}\log f(x) = \frac{x-\gamma}{\sigma^2}$ and the left hand side of (28) becomes:

$$\sum_{i=1}^{N} i\,E\left[\frac{W^{(t_i)}-\gamma}{\sigma^2}\right] = \frac{1}{\sigma}\sum_{i=1}^{N} i\,E\left[V^{(t_i)}\right]$$

where $V^{(1)} < V^{(2)} < \ldots < V^{(N)}$ is an ordered sample from a standard normal distribution. Hence (41) is most powerful among rank tests against normal alternatives $F = N(\gamma + i\delta,\sigma^2)$ for sufficiently small δ.

(ii) If $f(x)$ is the logistic distribution $f(x) = e^{-x}/(1+e^{-x})^2$, $-\frac{f'(x)}{f(x)} = 2F(x) - 1$ and the locally most powerful rank test (28) rejects when

$$\sum_{i=1}^{N} i\,E\left[F(W^{(t_i)})\right] > C$$

where W has the distribution F, so that F(W) is uniformly distributed over (0,1). Since $E[F(W^{(t_i)})] = \frac{t_i}{N+1}$ the rejection region of the locally most powerful rank test against logistic alternatives is (40).

(iii) Let $G_1(x) \geq G_2(x) \geq \ldots \geq G_N(x)$, $G_i \neq G_j$ if $i \neq j$ be an alternative, and assume that G_i is continuous $(i = 1,\ldots,N$.
Define $f_i(v) := G_i^{-1}(v) := \inf\{z : G_i(z) \geq v\}$. Then we have $f_1(v) \leq f_2(v) \leq \ldots \leq f_N(v)$ and $v_1 < v_2 \Rightarrow f_i(v_1) < f_i(v_2)$ $(i = 1,\ldots,N)$. For $v_i \in [0,1]$ $(i = 1,\ldots,N)$ we put $z_1 = f_1(v_1),\ldots,z_N = f_1(v_N)$ and $z_1' = f_1(v_1),\ldots,z_N' = f_N(v_N)$. Then $i < j$, $z_i < z_j$ implies $f_1(v_i) < f_1(v_j) \Rightarrow v_i < v_j \Rightarrow f_i(v_i) < f_i(v_j) \leq f_j(v_j)$ that is $z_i' < z_j'$, hence $\varphi(z_1,\ldots,z_N) \leq \varphi(z_1',\ldots,z_N')$. Thus for all $v_1,\ldots,v_N$ we have $\varphi(f_1(v_1),\ldots,f_N(v_N)) \geq \varphi(f_1(v_1),\ldots,f_1(v_N))$. Now

$$E_G\varphi(z) = \int\cdots\int \varphi(z_1,\ldots,z_N)\,dG_1(z_1)\cdots dG_N(z_N) =$$

$$\int_{[0,1]}\cdots\int_{[0,1]} \varphi(f_1(v_1),\ldots,f_N(v_N))\,dv_1\cdots dv_N \geq$$

$$\geq \int_{[0,1]} \cdots \int_{[0,1]} \varphi(f_1(v_1),\dots,f_1(v_N))dv_1\cdots dv_N =$$

$$= \int \cdots \int \varphi(z_1,\dots,z_N)dG_1(z_1)\cdots dG_1(z_N) = \alpha.$$

Hence any rank test satisfying $\varphi(z_1,\dots,z_N) \leq \varphi(z_1',\dots,z_N')$ for any two points for which $i < j$, $z_i < z_j$ implies $z_i' < z_j'$ for all i and j is unbiased against alternatives of an upward trend.

In order to show that the tests (40) and (41) are special cases of tests φ with

$$\varphi(z_1,\dots,z_N) \leq \varphi(z_1',\dots,z_N')$$

for any two points z and z' such that

(29) $$i < j,\quad z_i < z_j \Rightarrow z_i' < z_j',$$

we use a result of LEHMANN (1966). Note that when two points z and z' satisfy (29), the rank numbers $(i_1,\dots,i_n)$ of z' are better ordered (in the sence of formula (7.2) of Lehmann's paper, p. 1149) than the rank numbers $(j_1,\dots,j_n)$ of z; i.e.

(30) $$\alpha < \beta \text{ and } j_\alpha < j_\beta \Rightarrow i_\alpha < i_\beta.$$

It now follows from Lehmann's Corollary 2 to his Theorem 5 (p. 1150) that for any non-decreasing function h and $a_1 < a_1 < \dots < a_k$

$$(30) \Rightarrow \sum_{k=1}^{N} a_k h(i_k) \geq \sum_{k=1}^{N} a_k h(j_k).$$

Applications of this result to $h(t) = t$ and $h(t) = E(V^{(t)})$ (both non-decreasing) show that (40) and (41) are special cases of the described tests φ.

Problem 40.

We shall restrict attention to the case without ties.

(i) Define

$$V_{ij} = \begin{cases} 1 & \text{if } Z_i \leq Z_j \\ 0 & \text{if } Z_i > Z_j, \end{cases}$$

then for $i \neq j$

$$V_{ij} = 1 \Leftrightarrow Z_i \le Z_j \Leftrightarrow \begin{cases} \text{if } i < j : U_{ij} = 1 \\ \text{if } i > j : U_{ij} = 0, \end{cases}$$

and, again for $i \neq j$,

$$V_{ij} = 0 \Leftrightarrow Z_i > Z_j \Leftrightarrow \begin{cases} \text{if } i < j : U_{ij} = 0 \\ \text{if } i > j : U_{ij} = 1. \end{cases}$$

Since evidently $V_{ii} = 1 - U_{ii}$ it follows that

$$V_{ij} = \begin{cases} U_{ij} & \text{if } i < j \\ 1 - U_{ij} & \text{if } i \ge j. \end{cases}$$

Using $T_j = \#\{Z_i : Z_i \le Z_j,\ i = 1,2,\dots,N\} = \Sigma_{i=1}^N V_{ij}$ we find

$$\sum_{j=1}^N jT_j = \sum_{j=1}^N j \sum_{i=1}^N V_{ij} = \sum_{j=1}^N j \sum_{i=j}^N V_{ij} + \sum_{j=2}^N j \sum_{i=1}^{j-1} V_{ij} =$$

$$= \sum_{j=1}^N j \sum_{i=j}^N (1 - U_{ij}) + \sum_{j=2}^N j \sum_{i=1}^{j-1} U_{ij} =$$

$$= \sum_{j=1}^N j(N-j+1) + \sum_{j=1}^N j \sum_{i=j}^N U_{ij} + \sum_{j=2}^N j \sum_{i=1}^{j-1} U_{ij} =$$

$$= (N+1) \sum_{j=1}^N j - \sum_{j=1}^N j^2 - \sum_{i \ge j} jU_{ij} + \sum_{i<j} jU_{ij}$$

$$= (N+1)\frac{N(N+1)}{2} - \frac{N(N+1)(2N+1)}{6} - \sum_{i>j} jU_{ij} + \sum_{i<j} jU_{ij} \quad (U_{ii} = 0)$$

$$= \frac{N(N+1)(N+2)}{6} + \sum_{i<j} (j-i)U_{ij} \quad (\text{since } U_{ij} = U_{ji}).$$

(ii) For all $Z = (Z_1,\dots,Z_N)$ define $U(Z) = \Sigma_{i<j} U_{ij}$. Note that if $i < j$ $U_{ij} = 1$ if $Z_i < Z_j$ and 0 otherwise.
If $Z' = (Z'_1,\dots,Z'_N)$ with

$$\begin{cases} Z'_i = Z_i \text{ for all } i \neq k,k+1 \\ Z'_k = Z_{k+1} \\ Z'_{k+1} = Z_k \end{cases}$$

Then

$$U(Z') = \begin{cases} U(Z) + 1 & \text{if } Z_k > Z_{k+1} \\ U(Z) - 1 & \text{if } Z_k < Z_{k+1}. \end{cases}$$

If $Z_1 < Z_2 < \dots < Z_N$ then $U(Z) = N(N-1)/2$. Conversely if there is an index i such that $Z_i > Z_{i+1}$ then $U(Z) < N(N-1)/2$. Complete ordering of $(Z_1,\dots,Z_N)$ can be achieved by a sequence of steps, each of which consists of exchanging adjacent elements Z_k, Z_{k+1}, with $Z_k > Z_{k+1}$. Since in each of these steps the value of U(Z) is increased by one, and since it is impossible to increase the value of U(Z) by more than one by exchanging two adjacent elements of Z, the smallest number of steps equals $N(N-1)/2 - U$.

Problem 41.

As usual we suppose that the distribution of (X_i,Y_i) is continuous.

(i) We shall prove the following result: testing independence versus positive dependence is equivalent to testing, conditionally on $X^{(1)},\dots,X^{(N)}$, randomness of the Z's versus the alternative of an upward trend (see Remark 1 below). Let F_x be the distribution function of Y_i conditional on $X_i = x$. Define the anti-ranks $D_1,\dots,D_N$ by $X_{D_i} = X^{(i)}$; in words, D_i is the random index of the i'th smallest X. Note that $Z_i = Y_{D_i}$.
Let $x_1,\dots,x_N$ be distinct values of $X_1,\dots,X_N$ and let $x^{(1)},\dots,x^{(N)}$ and $d_1,\dots,d_N$ be the corresponding values of $X^{(1)},\dots,X^{(N)}$ and $D_1,\dots,D_N$. Then conditional on $X_1 = x_1,\dots,X_N = x_N$, we have that $Z_1,\dots,Z_N$ are independent with distribution functions $F_{x^{(1)}},\dots,F_{x^{(N)}}$. For

$$P\{Z_1 \le z_1,\dots,Z_N \le z_N \mid X_1 = x_1,\dots,X_N = x_N\} =$$
$$= P\{Y_{D_1} \le z_1,\dots,Y_{D_N} \le z_N \mid Z_1 = x_1,\dots,X_N = x_N\} =$$
$$= P\{Y_{d_1} \le z_1,\dots,Y_{d_N} \le z_N \mid X_{d_1} = x^{(1)},\dots,X_{d_N} = x^{(N)}\} =$$
$$= \prod_{i=1}^{N} F_{x^{(i)}}(z_i).$$

Since this joint distribution only depends on $x^{(1)},\dots,x^{(N)}$ (but not on $d_1,\dots,d_N$), it is also the distribution of $Z_1,\dots,Z_N$ given $X^{(1)} = x^{(1)}$, $\dots,X^{(N)} = x^{(N)}$.
Clearly independence of X_i and Y_i implies randomness, positive dependence implies an upward trend (for all $x^{(1)} < \dots < x^{(N)}$). Conversely, if $Z_1,\dots,Z_N$ are identically distributed (randomness) whatever $x^{(1)} < \dots < x^{(N)}$ then X_i and Y_i are independent for each i; if $Z_1,\dots,Z_N$ have an upward trend whatever $x^{(1)} < \dots < x^{(N)}$ then X_i and Y_i are positively dependent.

(ii) Since $\Sigma_{i=1}^{N} R_i = \Sigma_{i=1}^{N} S_i = \frac{1}{2}N(N+1)$, rejecting for large values of the rank correlation coefficient is equivalent to rejecting for large values of $\Sigma_{i=1}^{N} R_i S_i$. Now $\Sigma_{i=1}^{N} R_i S_i = \Sigma_{i=1}^{N} iT_i$. Under the null hypothesis, $(T_1,\ldots,T_N)$ is independent of $(X^{(1)},\ldots,X^{(N)})$. Therefore the critical value for a conditional test bases on $\Sigma_{i=1}^{N} iT_i$ (conditional on $X^{(1)},\ldots,X^{(N)}$) does not depend on $X^{(1)},\ldots,X^{(N)}$. Hence the conditional test based on (40) and the unconditional test based on the rank correlation coefficient are identical.

(iii) The mean and variance of the uniform distribution on $\{1,\ldots,N\}$ are $\mu = \frac{1}{2}(N+1)$ and $\sigma^2 = \frac{1}{12}(N^2-1)$. These are therefore the mean and variance of both the numbers $\{R_1,\ldots,R_N\}$ and $\{S_1,\ldots,S_N\}$.
Since $\operatorname{var}(R-S) = \operatorname{var} R - 2\sqrt{\operatorname{var} R \operatorname{var} S}\,\operatorname{corr}(R,S) + \operatorname{var}(S) = 2\sigma^2(1-\operatorname{corr}(R,S))$ we have $\operatorname{corr}(R,S) = 1 - \operatorname{var}(R-S)/2\sigma^2$; i.e.

$$\frac{\Sigma\,(R_i-\bar{R})(S_i-\bar{S})}{\sqrt{\Sigma\,(R_i-\bar{R})^2\,\Sigma\,(S_i-\bar{S})^2}} = 1 - \frac{\Sigma\,(R_i-S_i)^2}{N}\cdot\frac{6}{N^2-1}\,.$$

(iv) We say that the set of two distinct ordered pairs $\{(a,b),(c,d)\}$ is *concordant* if $(a-c)(b-d) > 0$; *discordant* if $(a-c)(b-d) < 0$. Note the any $\{(i,Z_i),(j,Z_j)\}$ (for $i \neq j$) is either concordant or discordant when there are no ties among the X's or Y's (which we assume from now on). The total number of such pairs is $\frac{1}{2}N(N-1)$. We have

$$\begin{aligned}\sum_{i<i} U_{ij} &= \#\{\text{concordant } \{(i,Z_i),(j,Z_j)\},\ i \neq j\}\\ &= \#\{\text{concordant } \{(i,T_i),(j,T_j)\},\ i \neq j\}\\ &= \#\{\text{concordant } \{(R_i,S_i),(R_j,S_j)\},\ i \neq j\}\\ &= \#\{\text{concordant } \{(X_i,Y_i),(X_j,Y_j)\},\ i \neq j\}.\end{aligned}$$

Therefore

$$\begin{aligned}\sum_{i<j} V_{ij} &= \#\{\text{concordant } \{(X_i,Y_i),(X_j,Y_j)\},\ i \neq j\}\\ &\quad - \#\{\text{discordant } \{(X_i,Y_i),(X_j,Y_j)\},\ i \neq j\}\\ &= 2\sum_{i<j} U_{ij} - \tfrac{1}{2}N(N-1).\end{aligned}$$

(v) Consider first the test (ii). Suppose it has size α. By part (i), conditionally on $X^{(1)},\ldots,X^{(N)}$ the test still has size α while by Problem 39 (iii) its conditional power at each alternative is at least α. Therefore

its unconditional power is also at least α.
Consider next the test (iv), and suppose its size is α. By the same arguments as in part (i), conditionally on $X^{(1)},\dots,X^{(N)}$ the test still has size α. It is easily seen that its critical function has the property specified in Problem 39 (iii). Its conditional power is therefore at least α, and hence its unconditional power is too.

Remark 1. Though (unconditionally) $Z_1,\dots,Z_N$ are stochastically increasing in the case of positive dependence, they are generally not independent. We were unable to prove the following interpretation of (i): the joint distribution of $T_1,\dots,T_N$ is the same as that of the ranks of some independent, stochastically ordered r.v.'s $Z_1',\dots,Z_N'$.

Remark 2. The unbiasedness of test (ii) also follows from Theorem 4, special case (i), or Corollary 2 (iii) in LEHMANN (1966). Similarly the unbiasedness of test (iv) follows from Corollary 2 (i) in the same paper. Actually part of the proof of the just mentioned Theorem 4 was used in the solution of Problem 39.

Section 10

Problem 42.

Let $\{S(x,y) \mid x,y \in \mathbb{R}\}$ be a class of invariant confidence sets and denote the rotation around (0,0) over the angle α be g_α and the translation over (x,y) by $g_{x,y}$. Then

$$g_\alpha S(0,0) = S(0,0) \qquad \text{for all } \alpha \tag{31}$$

and

$$g_{x,y} S(0,0) = S(x,y) \qquad \text{for all } x,y. \tag{32}$$

Now (31) implies that

$$S(0,0) = \bigcup_{r\in R} \{(\xi,\eta) \mid \xi^2+\eta^2 = r^2\}$$

with $R = \{r \mid r = (s^2+t^2)^{\frac{1}{2}}, (s,t) \in S(0,0)\}$. Together with (32) this yields

$$S(x,y) = \bigcup_{r\in R} \{(\xi,\eta) \mid (\xi-x)^2 + (\eta-y)^2 = r^2\}, \qquad x,y \in \mathbb{R}. \tag{33}$$

Conversely, it is easy to verify that (33) is a class of invariant confidence sets.

Problem 43.

Let $X_1,\ldots,X_n;Y_1,\ldots,Y_n$ be samples from $N(\xi,\sigma^2)$ and $N(\eta,\tau^2)$ respectively. Confidence intervals for $\Delta = \tau^2/\sigma^2$ are based on the hypotheses $H(\Delta_0)$: $\Delta = \Delta_0$, which are invariant under the groups G_{Δ_0} generated by the transformations $X_i' = aX_i + b$, $Y_i' = aY_i + c$ $(a \neq 0)$ and the transformstion $X_i' = \Delta_0^{-\frac{1}{2}}Y_i$, $Y_i' = \Delta_0^{\frac{1}{2}}X_i$. The UMP invariant test of $H(\Delta_0)$ has acceptance region

$$\max\{\Sigma(Y_i - \bar{Y})^2/\Sigma(\Delta_0^{\frac{1}{2}}X_i - \Delta_0^{\frac{1}{2}}\bar{X})^2, \Sigma(\Delta_0^{\frac{1}{2}}X_i - \Delta_0^{\frac{1}{2}}\bar{X})^2/\Sigma(Y_i - \bar{Y})^2\} < k$$

(cf. Problem 7 with X_i and Y_i replaced by $\Delta_0^{\frac{1}{2}}X_i$ and Y_i).
The associated confidence intervals are

$$(34) \qquad \Sigma(Y_i - \bar{Y})^2/k\Sigma(X_i - \bar{X})^2 < \Delta < k\Sigma(Y_i - \bar{Y})^2/\Sigma(X_i - \bar{X})^2.$$

The group G in the present case is the group generated by the transformations $X_i' = aX_i + b$, $Y_i' = aY_i + c$ $(a \neq 0)$ and the transformations $X_i' = dY_i$, $Y_i' = X_i/d$ $(d \neq 0)$.
The transformation g given by $X_i' = aX_i + b$, $Y_i' = aY_i + c$ $(a \neq 0)$ induces the transformation $\bar{g}\Delta = \Delta$. Since for such transformations the confidence intervals S remain unaltered, the confidence intervals are invariant under such transformations. The transformations g given by $X_i' = dY_i$, $Y_i' = X_i/d$ $(d \neq 0)$ induces the transformation $\bar{g}\Delta = (\Delta d^4)^{-1}$. Application of the associated transformation g^* to the confidence interval takes it into the interval

$$(\Sigma(X_i - \bar{X})^2/kd^4\Sigma(Y_i - \bar{Y})^2, k\Sigma(X_i - \bar{X})^2/d^4\Sigma(Y_i - \bar{Y})^2).$$

Since this coincides with the interval obtained by replacing X_i and Y_i in (34) by dY_i and X_i/d, respectively, the confidence intervals (34) are invariant also under these transformations. Hence they are uniformly most accurate invariant.

Problem 44.

Let θ be fixed and let $S'(x) = \{\theta : \underline{\theta}'(x) \leq \theta\}$ be any (other) G-invariant class of (one sided) confidence sets at level $1-\alpha$. If $A'(\theta)$ denote the associated acceptance regions, then for any $g \in G_\theta$,

$$gA'(\theta) = \{g(x) : \theta \in S'(x)\} = \{x : \theta \in S'(g^{-1}(x))\} =$$
$$= \{x : \theta \in g^{*-1}S'(x)\} = \{x : \bar{g}\theta \in S'(x)\} =$$
$$= \{x : \underline{\theta}'(x) \le \bar{g}\theta\} \subset \{x : \underline{\theta}'(x) \le \theta\},$$

since $S'(x)$ is invariant and since $g \in G_\theta$ implies $\bar{g}\theta \le \theta$. Hence for any $g \in G_\theta$

$$gA'(\theta) \subset A'(\theta)$$

and also

$$g^{-1}A'(\theta) \subset A'(\theta)$$

or

$$A'(\theta) \subset gA'(\theta).$$

Therefore $A'(\theta)$ is invariant under G_θ for each θ. It follows that these tests are at most as powerful as those with acceptance regions $A(\theta)$ and hence the associated lower confidence limits $\underline{\theta}(X)$ are uniformly most accurate invariant.

Note that this proof is an adaptation of the proof of Lemma 4.

Problem 45.

Consider the group of transformations $G_1 = \{g : g(x_1,\ldots,x_n) = (x_1+c,\ldots,x_n+c),\ c \in \mathbb{R}\}$. Then $\bar{G}_1 = \{\bar{g} : \bar{g}(\xi,\sigma^2) = (\xi+c,\sigma^2),\ c \in \mathbb{R}\}$. Hence, for all σ_0^2 the problem of testing $H_0 : \sigma^2 \ge \sigma_0^2$ against $K(\sigma_0^2) : \sigma^2 < \sigma_0^2$ remains invariant under G_1.

It follows from Example 5 of this chapter that for all σ_0^2 the acceptance region

$$A(\sigma_0^2) = \{x : \sum_{i=1}^{n} (x_i - \bar{x})^2 \ge C_0\sigma_0^2\}$$

is UMP G_1-invariant, where C_0 is determined by

$$\int_{C_0}^{\infty} \chi^2_{n-1}(t)dt = \alpha.$$

If $S(x) = \{(\xi,\sigma^2) : x \in A(\sigma^2)\}$ then $S(x) = \{(\xi,\sigma^2) : \sigma^2 \le t(x)\}$, where $t(x) = \frac{1}{C_0}\Sigma_{i=1}^n (x_i - \bar{x})^2$.

For all $g \in G_1$, $x \in \mathbb{R}^n$

$$g^*S(x) = \{\bar{g}(\xi,\sigma^2) : (\xi,\sigma^2) \in S(x)\} = \{(\xi',\sigma^2) : \sigma^2 \le t(x)\} =$$
$$= \{(\xi',\sigma^2) : \sigma^2 \le tg(x)\} = S(gx),$$

and therefore the confidence sets $S(x)$ are G_1-invariant. By an obvious modification of Problem 44 it follows that the upper confidence limits $t(x)$ are most accurate G_1-invariant.
Next consider the group $G_2 = \{g : g(x_1,\ldots,x_n) = (ax_1+c,\ldots,ax_n+c),$ $a \ne 0, -\infty < c < \infty\}$, then $\bar{G}_2 = \{\bar{g} : \bar{g}(\xi,\sigma^2) = (a\xi+c,a^2\sigma^2),\ a \ne 0,$ $-\infty < c < \infty\}$, and for all $g \in G_2$

$$g^*S(x) = \{(a\xi+c,a^2\sigma^2) : \sigma^2 \le t(x)\} =$$
$$= \{(\xi',a^2\sigma^2) : a^2\sigma^2 \le a^2t(x)\} =$$
$$= \{(\xi',\tau^2) : \tau^2 \le t(gx)\} = S(gx).$$

Hence the confidence sets $S(x)$ are also G_2-invariant.
Let $S^*(x)$ be any family of G_2-invariant confidence sets then these sets are also G_1-invariant. Hence, for all $(\xi,\sigma^2) \ne (\xi_0,\sigma_0^2)$,

$$P_{\xi_0,\sigma_0^2}\{(\xi,\sigma^2) \in S(X)\} \le P_{\xi_0,\sigma_0^2}\{(\xi,\sigma^2) \in S^*(X)\}.$$

Therefore the $t(x)$ are also uniformly most accurate G_2-invariant upper confidence limits at confidence level $1-\alpha$.

Problem 46

(i) Let $X_1,\ldots,X_n$ be independently distributed as $N(\xi,\sigma^2)$ and let $\theta = \xi/\sigma$. By Section 4, the UMP invariant test under transformations $X_i' = aX_i$ for $H : \theta \le \theta_0$ has acceptance region

$$A(\theta_0) = \{x : T(x) \le c_0\},$$

where

$$T(x) = \sqrt{n}\,\bar{x} \Big/ \left\{\sum_{i=1}^n (x_i-\bar{x})^2/(n-1)\right\}^{\frac{1}{2}}$$

and c_0 is determined by $P\{t_{n-1}(\sqrt{n}\,\theta_0) > c_0\} = \alpha$. Here $t_{n-1}(\sqrt{n}\,\theta_0)$ denotes a random variable with a noncentral student distribution with $n-1$ degrees of freedom and noncentrality parameter $\sqrt{n}\,\theta_0$.
Next let $S(x) = \{\theta : x \in A(\theta)\}$. Using the monotone likelihood ratio property of the noncentral student distribution (see Section 4) we prove

that $S(x) = \{\theta : C^{-1}[T(x)] \le \theta\}$ where $C(\cdot)$ is a strictly increasing continuous function determined by

$$P_\theta\{t_{n-1}(\sqrt{n}\ \theta) > C(\theta)\} = \alpha.$$

Denote the cumulative distribution function of $t_{n-1}(\sqrt{n}\ \theta)$ at the point x by $F(x;\sqrt{n}\ \theta)$. By Lemma 2 of Chapter 3 we have $\theta' > \theta \Rightarrow F(x;\sqrt{n}\ \theta') \le F(x;\sqrt{n}\ \theta)$ which in its turn implies that C is increasing. In fact we shall prove that C is strictly increasing. By dominated convergence $F(x;\sqrt{n}\ \theta)$ can be shown to be continuous in θ for fixed x. Using this we prove continuity of C. Suppose there is a $\tilde{\theta}$ such that C is not right continuous in $\tilde{\theta}$. Then there exists an $\varepsilon > 0$ and a sequence $\theta_1 \ge \theta_2 \ge \theta_3 \ge \dots$ converging to $\tilde{\theta}$ with $C(\theta_k) \ge C(\tilde{\theta}) + \varepsilon$. But then

$$F(C(\tilde{\theta}) + \varepsilon;\sqrt{n}\ \theta_k) \le F(C(\theta_k);\sqrt{n}\ \theta_k) = 1-\alpha$$

and by continuity

$$F(C(\tilde{\theta}) + \varepsilon;\sqrt{n}\ \tilde{\theta}) = \lim_{k\to\infty} F(C(\tilde{\theta}) + \varepsilon;\sqrt{n}\ \theta_k) \le 1-\alpha$$

which contradicts $F(C(\tilde{\theta});\sqrt{n}\ \tilde{\theta}) = 1-\alpha$. Therefore C is right continuous and, because the same argument can be used to prove left continuity, C is continuous.

C is strictly increasing since if $\theta_1 < \theta_2$ then by Theorem 2 (ii) of Chapter 3 we have

$$F(C(\theta_1);\sqrt{n}\ \theta_1) = 1-\alpha > 1 - P_{\theta_2}\{T(X) > C(\theta_1)\} = F(C(\theta_1);\sqrt{n}\ \theta_2).$$

Since C is strictly increasing and continuous it is invertible and

$$S(x) = \{\theta : x \in A(\theta)\} = \{\theta : T(x) \le C(\theta)\} =$$
$$= \{\theta : C^{-1}[T(x)] \le \theta\}$$

is a family of one sided intervals for θ.

Furthermore

$$g^*S(x) = \bar{g}\theta : C^{-1}[T(x)] \le \theta\} =$$
$$= \{\bar{g}\theta : C^{-1}[T(g(x))] \le \bar{g}\theta\} = S(gx),$$

so the family of confidence intervals S(x) is G-invariant. By Problem 44 it now follows that the most accurate lower confidence bound under G is given by

$$\underline{\theta}(X) = C^{-1}[T(X)].$$

(ii) For this sample T(x) equals 7.6. $C^{-1}(7.6)$ can be determined from table 27 of PEARSON and HARTLEY (1972). We obtain $\underline{\theta} = 3.8/\sqrt{8} = 1.3$.

Problem 47.

(i) Let $(X_1,Y_1),\ldots,(X_n,Y_n)$ be a sample from a $N(\mu,\eta,\sigma^2,\tau^2,\rho)$ distribution. Let G be the group considered in the problem, G_1 the subgroup of G considered in Problem 11, $x = (x_1,\ldots,x_n)$, $y = (y_1,\ldots,y_n)$, $X = (X_1,\ldots,X_n)$, $Y = (Y_1,\ldots,Y_n)$ and

$$R(x,y) = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \Big/ \left\{ \sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2 \right\}^{\frac{1}{2}}.$$

Let, for any $\rho_0 \in (-1,1)$, $A(\rho_0) = \{(x,y) : R(x,y) \leq C(\rho_0)\}$, where $C(\rho_0)$ is determined by $P_{\rho_0}\{R \leq C(\rho_0)\} = 1-\alpha$, and let $S(x,y) = \{(\Delta,\rho) : (x,y) \in A(\rho)\}$, where Δ is the nuisance parameter $(\mu,\eta,\sigma^2,\tau^2)$.
We shall chech the conditions of Problem 44.
For any ρ_0 the problem of testing $H(\rho_0) : \rho \leq \rho_0$ against $K(\rho_0) : \rho > \rho_0$ is G-invariant. This is an immediate consequence of

(35) $\qquad g \in G,\ (\Delta,\rho) \in \Omega \Rightarrow$ there exists Δ' such that $\bar{g}(\Delta,\rho) = (\Delta',\rho)$.

For any ρ_0, $A(\rho_0)$ is UMP G-invariant. This follows from the fact that R(x,y) is a maximal invariant w.r.t. G_1 and invariant w.r.t. G, which implies that R is a maximal invariant w.r.t. G. Using Problem 11 (i), where it was shown that, for any ρ_0, $A(\rho_0)$ is UMP G_1-invariant, it follows that $A(\rho_0)$ is UMP G-invariant.
By the same reasoning as in Problem 46 (i) we can prove that C is continuous and strictly increasing and that we can rewrite S(x,y) as $\{(\Delta,\rho) : \underline{\rho}(x,y) \leq \rho\}$, where $\underline{\rho}(x,y) = C^{-1}[R(x,y)]$ (see Problem 11 where it is shown that the density of R(X,Y) has a monotone likelihood ratio). S(x,y) is G-invariant since by (35) and the G-invariance of R(x,y) the set $g^*S(x,y)$ is equal to $S(g(x,y))$, for any $g \in G$.
It now follows by Problem 44 that $\underline{\rho}(x,y)$ is a uniformly most accurate G-invariant lower confidence limit at confidence level $1-\alpha$.

(ii) Let (x_0,y_0) be fixed and $r_0 = R(x_0,y_0)$. Suppose that the equation

$F(r_0;\rho) = 1-\alpha$ has a solution, $\hat{\rho}$ say. Then this solution is unique, because if $F(r_0;\rho_1) = F(r_0;\rho_2) = 1-\alpha$ then $F(r_0;\rho_i) = F(C(\rho_i);\rho_i)$, $i=1,2$ and hence $C(\rho_1) = C(\rho_2) = r_0$ which implies $\rho_1 = \rho_2$ since C is strictly increasing. We determine $\hat{\rho} = C^{-1}(r_0)$. In the problem $\alpha = .05$ and $r_0 = -.22$. The equation $F(r_0;\rho) = 1-\alpha$ is easily seen to be equivalent to $F(-r_0;-\rho) = \alpha$. Using DAVID (1938) we obtain $\underline{\rho} = -0.7$.

Section 11

Problem 48.

Let $S(x)$ be a collection of confidence sets, and $\bar{g} \in \bar{G}$. In view of (1) on p. 214 we have $P_{\bar{g}\theta}\{\bar{g}\theta \in S(x)\} = P_\theta\{\bar{g}\theta \in S(gX)\}$. Since $S(x)$ is G-invariant this equals $P_\theta\{\bar{g}\theta \in g^*S(X)\}$. Further, by definition $\bar{g}\theta \in g^*S(x)$ if and only if $\theta \in S(X)$, and so it follows that $P_{\bar{g}\theta}\{\bar{g}\theta \in S(X)\} = P_\theta\{\theta \in S(X)\}$, that is $P_\theta\{\theta \in S(X)\}$ is invariant under $\bar{G}$.

Problem 49.

(i) Note that compared to Section 11 we are considering a larger group of transformations.

Let $g : \mathbb{R} \to \mathbb{R}$ be continuous, strictly monotone and onto. We define the mapping $\bar{g}$ from the set of continuous distribution functions into itself by

$$\bar{g}F = \begin{cases} Fg^{-1} & \\ & \text{if } g \text{ is} \\ 1 - Fg^{-1} & \end{cases} \begin{matrix} \text{increasing} \\ \\ \text{decreasing.} \end{matrix}$$

Clearly $\bar{g}$ is onto and 1 : 1. Let L_x and M_x be nondecreasing functions, defined for all sample points $x = (x_1,\ldots,x_n)$ with no x_i's equal, and let

$$S(X) = \{F \mid L_x(y) \le F(y) \le M_x(y), y \in \mathbb{R}\}$$

be the corresponding confidence band. If g is increasing then

$$\begin{aligned} g^*S(x) &= \{\bar{g}F \mid L_x(y) \le F(y) \le M_x(y), y \in \mathbb{R}\} = \\ &= \{Fg^{-1} \mid L_x(g^{-1}(y)) \le F(g^{-1}(y)) \le M_x(g^{-1}(y)), y \in \mathbb{R}\} = \\ &= \{F \mid L_x(g^{-1}(y)) \le F(y) \le M_x(g^{-1}(y)), y \in \mathbb{R}\} \end{aligned}$$

and if g is decreasing then

$$g^*S(x) = \{1 - Fg^{-1} \mid L_x(y) \le F(y) \le M_x(y), y \in \mathbb{R}\} =$$

$$= \{1 - Fg^{-1} \mid 1 - M_x(g^{-1}(y)) \le 1 - F(g^{-1}(y)) \le 1 - L_x(g^{-1}(y)), y \in \mathbb{R}\} =$$

$$= \{F \mid 1 - M_x(g^{-1}(y)) \le F(y) \le 1 - L_x(g^{-1}(y)), y \in \mathbb{R}\}.$$

We define the loss function L by $L(F,S(x)) = 1 - I_{S(x)}(F)$. Since $L(\bar{g}F, g^*S(x)) = 1 - I_{g^*S(x)}(\bar{g}F) = 1 - I_{S(x)}(F)$, the problem is invariant under g (see Section 5 of Chapter 1).
Let $R(x) = (R_1,\dots,R_n)$ be the vector of ranks of $x = (x_1,\dots,x_n)$ and $R^*(x) = (n+1-R_1,\dots,n+1-R_n)$. It is easy to verify that $T(x) = \{R(x), R^*(x)\}$ is maximal invariant.
The totality of invariant confidence bands is now seen to be the confidence bands for which there exist numbers $a_0,\dots,a_n; a_0',\dots,a_n'$ such that

$$(36) \qquad L_x(u) = a_i,\ M_x(u) = a_i' \text{ for } x^{(i)} < u < x^{(i+1)}, \quad i = 0,\dots,n$$

where $x^{(1)} < \dots < x^{(n)}$ is the ordered sample ($x^{(0)} = -\infty$ and $x^{(n+1)} = \infty$) (see p. 246), and in addition

$$(37) \qquad a_i = 1 - a_{n-i}', \quad i = 0,1,\dots,n.$$

This is a consequence of the fact that every decreasing transformation is the composition of an increasing transformation and the transformation $h : x \to -x$. Condition (37) follows from the invariance of the confidence band, which we know by Section 11 to satisfy (36), under h.

(ii) For this transformation g

$$\bar{g}F = P\{g(X) \le y\} = \begin{cases} F(y) & |y| \le 1 \\ F(y+1) - F(0) + F(-1) \quad \text{if} & -1 < y \le 0 \\ F(y-1) + F(1) - F(0) & 0 \le y < 1, \end{cases}$$

and $\bar{g}\bar{g}F(y) = P\{gg(X) \le y\} = P\{X \le y\} = F(y)$.
Let

$$(38) \qquad S(x) = \{F : L_x(y) \le F(y) \le M_x(y), y \in \mathbb{R}\}$$

be a band an observe that this band remains equal if we use

$$L_x'(y) = \inf_{F \in S(x)} F(y) \quad \text{and} \quad M_x'(y) = \sup_{F \in S(x)} F(y)$$

in place of L_x and M_x. Then L_x' and M_x' are nondecreasing and between zero and one. By the same reasoning if $g^*S(x)$ were a band then its lower and upper bounds, $\tilde{g}L_x$ and $\tilde{g}M_x$ say, would be given by

$$\tilde{g}L_x(y) \inf_{F\in S(x)} \bar{g}F(y) \quad \text{and } \tilde{g}M_x(y) = \sup_{F\in S(x)} \bar{g}F(y).$$

By carrying out this process twice we find lower and upperbounds, $\tilde{g}\tilde{g}L_x$ and $\tilde{g}\tilde{g}M_x$ say, for $g^*g^*S(x)$.
However, since $\bar{g}\bar{g}F = F$, we have $g^*g^*S(x) = S(x)$ which contradicts (38) in general. To see this we note that if $M_x(-1) \geq L_x(1)$ then

$$\tilde{g}L_x(y) = \begin{cases} L_x'(y) & \\ L_x'(-1) & \text{if} \\ L_x'(y-1) & \end{cases} \quad \begin{matrix} |y| \geq 1 \\ -1 < y \leq 0 \\ 0 \leq y < 1 \end{matrix}$$

and

$$\tilde{g}M_x(y) = \begin{cases} M_x(y) & \\ M_x(y+1) & \text{if} \\ M_1(1) & \end{cases} \quad \begin{matrix} |y| \geq 1 \\ -1 < y \leq 0 \\ 0 \leq y < 1 \end{matrix}$$

and hence

$$\tilde{g}\tilde{g}L_x(y) = \begin{cases} L_x(y) & \\ & \text{if} \\ L_x(-1) & \end{cases} \quad \begin{matrix} |y| \geq 1 \\ \\ |y| < 1 \end{matrix}$$

and

$$\tilde{g}\tilde{g}M_x(y) = \begin{cases} M_x(y) & \\ & \text{if} \\ M_x(1) & \end{cases} \quad \begin{matrix} |y| \geq 1 \\ \\ |y| < 1. \end{matrix}$$

As an example consider the band with bounds $L_x(y) = \frac{1}{4} + \frac{1}{8}I_{[0,\infty)}(y)$ and $M_x(y) = \frac{3}{4}$, then $\tilde{g}L_x(y) = \frac{1}{4} + \frac{1}{8}I_{[1,\infty)}(y)$ and $\tilde{g}M_x(y) = \frac{3}{4}$. But since for $F \in S(x)$ we have

$$\bar{g}F(-\tfrac{1}{4}) - \bar{g}F(-\tfrac{3}{4}) = F(\tfrac{3}{4}) - F(\tfrac{1}{4}) \leq \tfrac{3}{4} - (\tfrac{1}{4} + \tfrac{1}{8}) = \tfrac{3}{8}$$

and therefore $g^*S(x)$ contains no distribution functions F with $F(-\frac{1}{4}) - F(-\frac{3}{4}) > \frac{3}{8}$.

REFERENCES

ANDERSON, T.W. (1958). *An Introduction to Multivariate Statistical Analysis.* John Wiley & Sons. New York.

ARNOLD, K.J. (1951). Tables of faciliate sequential t-tests. *Appl. Math. Ser. Nat. Bur. Standards (U.S.)* 7, v-viii.

DAVID, F.N. (1938). *Tables of the correlation coefficient.* Biometrika Office, University College, London.

DAVID, H.A. (1970). *Order Statistics.* John Wiley & Sons, New York.

HÁJEK, J. & Z. ŠIDÁK (1967). *Theory of Rank Tests.* Academic Press, New York.

HALL, W.J., R.A. WIJSMAN and J.K. GHOSH (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Stat.* 36, 575-614.

LEHMANN, E.L. (1950). Some principals of the theory of testing hypotheses. *Ann. Math. Stat.* 21, 162-165.

LEHMANN, E.L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Stat.* 22, 165-179.

LEHMANN, E.L. (1966). Some concepts of dependence. *Ann. Math. Stat.* 37, 1137-1153.

LEHMANN, E.L. (1975). *Nonparametrics: Statistical methods based on ranks.* Holden-Day. San Francisco.

PEARSON, E.S. & H.O. Harthy (1972). *Biometrika tables for statisticians.* Cambridge University Press. Cambridge.

PURI, M.L. & P.K. Sen (1971). *Nonparametric Methods in Multivariate Analysis.* John Wiley & Sons, New York.

RUSHTON, S. (1952). On a two sided sequential t-test. *Biometrika* 39, 302-308.

CHAPTER 7

Section 1

Problem 1.

As is indicated on p. 266 the variables $Y_1,\ldots,Y_n$ are again independently normally distributed with common variance σ^2 and means $E(Y_i) = \eta_i$ for $i = 1,2,\ldots,s$, and $E(Y_i) = 0$ for $i = s+1,\ldots,n$.
Hence $(r \le s)$

$$E(r^{-1}\sum_{j=1}^{r} Y_j^2) = r^{-1}\sum_{j=1}^{r} E(Y_j^2) = r^{-1}\sum_{j=1}^{r}(\sigma^2+\eta_j^2) =$$
$$= \sigma^2 + r^{-1}\sum_{j=1}^{r}\eta_j^2$$

and

$$E((n-s)^{-1}\sum_{j=s+1}^{n} Y_j^2) = (n-s)^{-1}\sum_{j=s+1}^{n} E(Y_j^2) = (n-s)^{-1}\sum_{j=s+1}^{n}\sigma^2 = \sigma^2$$

Problem 2.

(i) Since $P\{X \le t\} = \Phi(t-\psi)$ we have for $v \ge 0$

$$P\{V \le v\} = P\{X^2 \le v\} = P\{-\sqrt{v} \le X \le \sqrt{v}\} = \Phi(\sqrt{r} - \psi) - \Phi(-\sqrt{v} - \psi).$$

Hence

$$p_\psi^V(v) = \frac{1}{2\sqrt{v}}\left[\Phi'(\sqrt{v}-\psi) + \Phi'(-\sqrt{v}-\psi)\right] =$$
$$= \frac{1}{2\sqrt{2\pi v}}\, e^{-\frac{1}{2}(v+\psi^2)}\left[e^{\psi\sqrt{v}} + e^{-\psi\sqrt{v}}\right] =$$
$$= \frac{1}{\sqrt{2\pi v}}\, e^{-\frac{1}{2}(v+\psi^2)}\sum_{k=0}^{\infty}\frac{(\psi\sqrt{v})^{2k}}{(2k)!}\,.$$

Since $(2k)! = \Gamma(2k+1) = 2k\Gamma(2k) = 2k2^{2k-1}\Gamma(k)\Gamma(k+\frac{1}{2})\pi^{-\frac{1}{2}}$ the preceding density can be rewritten as

$$p_\psi^V(v) = e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} \frac{\psi^{2k}}{2^k k!} \frac{e^{-\frac{1}{2}v} v^{k-\frac{1}{2}}}{2^{k+\frac{1}{2}}\Gamma(k+\frac{1}{2})} =$$

$$= \sum_{k=0}^{\infty} \frac{e^{-\frac{1}{2}\psi^2}\psi^{2k}}{2^k k!} \frac{e^{-\frac{1}{2}v} v^{k-\frac{1}{2}}}{2^{k+\frac{1}{2}}\Gamma(k+\frac{1}{2})} =$$

$$= \sum_{k=0}^{\infty} P_k(\psi) f_{2k+1}(v)$$

(ii) Let $Y_1,\dots,Y_r$ be independently normally distributed with unit variance and means $\eta_1,\dots,\eta_r$. Define $\psi = \{\Sigma_{j=1}^r \eta_i^2\}^{\frac{1}{2}}$. Let C denote an orthogonal $(r \times r)$-matrix with $(\eta_1\psi^{-1},\dots,\eta_r\psi^{-1})$ as first row and consider the orthogonal transformation $Z = (Z_1,\dots,Z_r)' = CY$, $Y = (Y_1,\dots,Y_r)'$. Then

a. $Z_1 = \Sigma_{i=1}^r \eta_i\psi^{-1}Y_i$;

b. $E(Z_1) = \psi$;

c. for $i = 2,3,\dots,r$ $E(Z_i) = 0$ since

$$E(Z_i) = \sum_{j=1}^{r} c_{ij}E(Y_j) = \sum_{j=1}^{r} c_{ij}\eta_j = 0 \quad (i = 2,3,\dots,r)$$

by orthogonality;

d. $Z_1, Z_2,\dots,Z_r$ are independently normally distributed with unit variance by the orthogonality of C.

Now let $U_1 = Z_1^2$ and $U_2 = \Sigma_{i=2}^r Z_i^2$, then by part (i) of the problem

$$p_\psi^{U_1}(u_1) = \sum_{k=0}^{\infty} P_k(\psi)\cdot f_{2k+1}(u_1).$$

Further it is well known that U_2 has the (central) χ^2-distribution with (r-1) degrees of freedom.

Finally since $U = U_1 + U_2$ is the sum of independent random variables, the density function is given by

$$p_\psi^{U_1}(u) = (p_\psi^{U_1} * f_{r-1})(u) = \sum_{k=0}^{\infty} P_k(\psi)(f_{2k+1} * f_{r-1})(u) =$$

$$= \sum_{k=0}^{\infty} P_k(\psi) f_{r+2k}(u),$$

where "*" denotes convolution (cf. any textbook on probability theory). For an alternative proof using characteristic functions see SEBER (1966), pp. 5-6.

Problem 3.

(i) Define $U = \Sigma_{i=1}^{r} (Y_i/\sigma)^2$ and $V = \Sigma_{i=s+1}^{n} (Y_i/\sigma)^2$. Then $W = U/V$. Note that U and V are independent and have the noncentral χ^2-distribution with r degrees of freedom and noncentrality parameter $\psi^2 = \sigma^{-2} \Sigma_{i=1}^{r} \eta_i^2$ (cf. Problem 7.2) and the (central) χ^2-distribution with (n-s) degrees of freedom, respectively.

Therefore the density function $f^W(w)$ of W at w is given by

$$f^W(w) = \int_0^\infty f^V(v) f^U(vw) v\,dv,$$

where f^V and f^U denote the densities of V and U, respectively.

By (86) and Fubini's theorem it follows that

$$f^W(w) = e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} \frac{(\frac{1}{2}\psi^2)^k}{k!} \int_0^\infty \frac{e^{-\frac{1}{2}v(1+w)} v^{\frac{1}{2}(r+n-s)+k-1} w^{\frac{1}{2}r+k-1}}{\Gamma(\frac{1}{2}r+k)\Gamma(\frac{1}{2}(n-s))2^{(n-s+r)+k}}\,dv =$$

$$= e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} c_k \frac{(\frac{1}{2}\psi^2)^k}{k!} \frac{w^{\frac{1}{2}r+k-1}}{(1+w)^{\frac{1}{2}(r+n-s)+k}} \int_0^\infty \frac{e^{-\frac{1}{2}t} t^{\frac{1}{2}(r+n-s)+k-1}}{\Gamma(\frac{1}{2}(r+n-s)+k)2^{\frac{1}{2}(n-s+r)+k}}\,dt$$

$$= e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} c_k \frac{(\frac{1}{2}\psi^2)^k}{k!} \frac{w^{\frac{1}{2}r+k-1}}{(1+w)^{\frac{1}{2}(r+n-s)+k}},$$

where $c_k = \frac{\Gamma(\frac{1}{2}(r+n-s)+k)}{\Gamma(\frac{1}{2}r+k)\Gamma(\frac{1}{2}(n-s))}$, as was to be proved.

If the random variable A has the noncentral χ^2-distribution with r degrees of freedom and noncentrality parameter ψ^2 and the random variable B has the (central) χ^2-distribution with (n-s) degrees of freedom and if A and B are independent, then by definition the random variable $C = \frac{Ar^{-1}}{B(n-s)^{-1}}$ has the noncentral F-distribution with r and (n-s) degrees of freedom and noncentrality parameter ψ^2.

Substituting U for A and V for B it is immediate that $(n-s)r^{-1}W = \frac{Ur^{-1}}{V(n-s)^{-1}}$ has a noncentral F-distribution.

(ii) Since $P\{B \le b\} = P\{W \le b(1-b)^{-1}\}$ the density $f^B(b)$ of B at b is given by

$$f^B(b) = f^W[b(1-b)^{-1}](1-b)^{-2} = \sum_{k=0}^{\infty} P_k(\psi) g_{\frac{1}{2}r+k,\frac{1}{2}(n-s)}(b),$$

where $P_k(\psi) = e^{-\frac{1}{2}\psi^2} \frac{(\frac{1}{2}\psi)^k}{k!}$ and

$$g_{p,q}(b) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} b^{p-1}(1-b)^{q-1}, \quad 0 \le b \le 1, \quad p > 0, \quad q > 0.$$

Problem 4.

(i) In the case of noncentral χ^2-distributions with r degrees of freedom and noncentrality parameters ψ_0^2 and ψ_1^2, $\psi_0^2 < \psi_1^2$, we have for $x > 0$

$$f(x) \stackrel{\text{def}}{=} p_{\psi_1}(x) / p_{\psi_0}(x) =$$

$$= \frac{\sum_{k=0}^{\infty} e^{-\frac{1}{2}\psi_1^2} \frac{(\frac{1}{2}\psi_1^2)^k}{k!} \frac{1}{2^{\frac{1}{2}r+k}\Gamma(\frac{1}{2}r+k)} x^{\frac{1}{2}r-1} e^{-\frac{1}{2}x} x^k}{\sum_{k=0}^{\infty} e^{-\frac{1}{2}\psi_0^2} \frac{(\frac{1}{2}\psi_0^2)^k}{k!} \frac{1}{2^{\frac{1}{2}r+k}\Gamma(\frac{1}{2}r+k)} x^{\frac{1}{2}r-1} e^{-\frac{1}{2}x} x^k}$$

$$= \frac{\sum_{k=0}^{\infty} b_k x^k}{\sum_{k=0}^{\infty} a_k x^k},$$

where $b_k = e^{-\frac{1}{2}\psi_1^2} \frac{(\frac{1}{4}\psi_1^2)^k}{k!} \frac{1}{\Gamma(\frac{1}{2}r+k)}$ $(k \ge 0)$ and $a_k = e^{-\frac{1}{2}\psi_0^2} \frac{(\frac{1}{4}\psi_0^2)^k}{k!} \frac{1}{\Gamma(\frac{1}{2}r+k)}$ $(k \ge 0)$. Since $\frac{b_k}{a_k} < \frac{b_{k+1}}{a_{k+1}}$ $(k \ge 0)$ by $\psi_0^2 < \psi_1^2$, we have

$$f'(x) = \frac{\left(\sum_{n=0}^{\infty} n b_n x^{n-1}\right)\left(\sum_{k=0}^{\infty} a_k x^k\right) - \left(\sum_{n=0}^{\infty} b_n x^n\right)\left(\sum_{k=0}^{\infty} k a_k x^{k-1}\right)}{\left(\sum_{k=0}^{\infty} a_k x^k\right)^2} =$$

$$= \frac{\sum\sum_{k<n} (n-k)(a_k b_n + a_n b_k) x^{k+n-1}}{\left(\sum_{k=0}^{\infty} a_k x^k\right)^2} > 0.$$

Hence f is increasing.

In the case of noncentral F-distributions with r and (n-s) degrees of freedom and noncentrality parameters ψ_0^2 and ψ_1^2, $\psi_0^2 < \psi_1^2$, the density functions are for $w > 0$

$$g_{\psi_i}(w) = e^{-\frac{1}{2}\psi_i^2} \sum_{k=0}^{\infty} c_k \frac{(\frac{1}{2}\psi_i^2)^k}{k!} \frac{(\frac{r}{n-s}w)^{\frac{1}{2}r-1+k}}{(1+\frac{r}{n-s}w)^{\frac{1}{2}(r+n-s)+k}},$$

where $c_k = \frac{\Gamma(\frac{1}{2}(r+n-s)+k)}{\Gamma(\frac{1}{2}r+k)\Gamma(\frac{1}{2}(n-s))}$, $i = 0,1$.

Since $z = \frac{\frac{r}{n-s}w}{1+\frac{r}{n-s}w}$ is an increasing function of w, it suffices to show that for $\psi_0^2 < \psi_1^2$

$$g(z) = \frac{e^{-\frac{1}{2}\psi_1^2} \sum_{k=0}^{\infty} c_k \frac{(\frac{1}{2}\psi_1^2)^k}{k!} z^k}{e^{-\frac{1}{2}\psi_0^2} \sum_{k=0}^{\infty} c_k \frac{(\frac{1}{2}\psi_0^2)^k}{k!} z^k}$$

is an increasing function of z. This is done in the same way as in the case of noncentral χ^2-distributions.

(ii) The hypothesis $H' : \psi^2 \le \psi_0^2$ ($\psi_0 > 0$ given) remains invariant under the groups of transformations G_1, G_2 and G_3, since for all $g_i \in G_i$ $(i = 1,2,3)$ $\overline{g_i}(\psi^2) = \psi^2$.
By part (i) the family of densities $p_\psi(w)$, given by (6), has monotone likelihood ratio in w. By Theorem 2, Chapter 3 there exists a UMP invariant test with rejection region $W > C'$, where C' is determined by $P_{\psi_0}\{W > C'\} = \alpha$.

Problem 5.

(i) The random variables $Y_1,\ldots,Y_n$ are independently normally distributed with (unknown) common variance σ^2 and means $E(Y_i) = \eta_i$, $i = 1,\ldots,s$, $E(Y_i) = 0$ $i = s+1,\ldots,n$ $(s < n)$. We wish to test $H: \eta_1 = \ldots = \eta_r = 0$ $(r \le s < n)$. Let

$$\Omega_H = \{(0,\ldots,0,\alpha_{r+1},\ldots,\alpha_s,\tau) : \alpha_j \in \mathbb{R}\ (j \ge r+1), \tau > 0\},$$
$$\Omega = \{(\alpha_1,\ldots,\alpha_s,\tau) : \alpha_j \in \mathbb{R}\ (1 \le j \le s), \tau > 0\}$$

and

$$\Omega_K = \Omega - \Omega_H.$$

Furthermore let $U = \Sigma_{i=1}^r Y_i^2$, $V = \Sigma_{l=s+1}^n Y_i^2$, $U' = \Sigma_{i=r+1}^s Y_i^2$.
The joint density of $Y_1,\ldots,Y_n$ under $\theta = (\eta_1,\ldots,\eta_s,\sigma) \in \Omega$, can be written as

$$h_\theta(y_1,\ldots,y_n) = (\sigma\sqrt{2\pi})^{-n} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{s} (y_i - \eta_i)^2 + \sum_{i=s+1}^{n} y_i^2\right]\right\}.$$

Let $\varphi = \varphi(Y_1,\dots,Y_n)$ be a level α test for H. The power of the test against a $\theta = (\eta_1,\dots,\eta_s,\sigma) \in \Omega_K$ will be written as

$$\beta_\varphi(\theta) = E_\theta\ (Y_1,\dots,Y_n).$$

Let P be the class of distributions of $Y = (Y_1,\dots,Y_n)$ as θ ranges over Ω. We also write y for $(y_1,\dots,y_n)$. Now suppose that φ is *unbiased*. Since P has the property that the power of any test for H is continuous in θ, unbiasedness implies similarity (cf. p. 125). Therefore, from now on we suppose that φ is a *similar* test. If ω denotes the boundary between Ω_H and Ω_K then $\omega = \Omega_H$.

The next step is to prove that φ has Neyman structure. The statistics $Y_{r+1},\dots,Y_s$, $U+V+U'$ are sufficient with respect to the nuisance parameters $\eta_{r+1},\dots,\eta_s$ and σ on Ω_H, as is clear from

$$h_\theta(y) = (\sigma\sqrt{2\pi})^{-n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=r+1}^{s}\eta_i^2\right)\cdot$$
$$\cdot\exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{r}y_i^2+\sum_{i=s+1}^{n}y_i^2+\sum_{i=r+1}^{s}y_i^2\right)+\sum_{i=r+1}^{s}\eta_i\sigma^{-2}y_i\right],$$

for $\theta \in \Omega_H$.

This constitutes an exponential family of the form

$$dP_\theta^Y(y) = C(\theta)\exp\left[\sum_{j=1}^{k}\theta_j T_j^*(y)\right]d\mu(y),$$

with

a. $\theta_1 = -\frac{1}{2\sigma^2}$, $\theta_j = \sigma^{-2}\eta_{j+r-1}$, $j = 2,\dots,s-r+1$;

b. $k = s-r+1$;

c. $T_1^*(y) = \Sigma_{i=1}^n y_i^2$, $T_j^*(y) = y_{j+r-1}$, $j = 2,\dots,s-r+1$;

d. $C(\theta) = (\sigma\sqrt{2\pi})^{-n}\exp\left(-\frac{1}{2\sigma^2}\Sigma_{i=r+1}^s \eta_i^2\right)$;

e. μ n-dimensional Lebesgue-measure.

Let $T^* = T^*(y) = (T_1^*(y),\dots,T_{s-r+1}^*(y))$.

The set

$$\{(\theta_1,\dots,\theta_{s-r+1}) : \theta_1 = -\frac{1}{2\sigma^2};\ \theta_j = \sigma^{-2}\eta_{j+r-1},\ j = 2,\dots,s-r+1;\ \sigma > 0;$$
$$\eta_j \in \mathbb{R},\ j = r+1,\dots,s\} =$$
$$= (-\infty,0) \times \mathbb{R}^{s-r}$$

contains an (s-r+1)-dimensional rectangle. It then follows from Theorem 1, Chapter 4 that the family of distributions $\mathcal{P}_H$ of Y under H (or $h_\theta(y)$, $\theta \in \Omega_H$) is complete.
Theorem 2, Chapter 4 then gives that all similar tests have Neyman structure, i.e.

$$E[\varphi(Y) \mid t^*] = \alpha \qquad \text{a.e. } \mathcal{P}_H^{T^*}.$$

Notice now that once the values of $Y_{r+1},\dots,Y_s$ are given, the value of U' is completely determined. Therefore the above may be interpreted as: "the conditional probability of rejection given $Y_{r+1},\dots,Y_s$ and $U+V$ equals α a.e." as was to be proved.
The optimality of the unconditional test may now be proved through the optimality of each of the conditional tests that are obtained by conditioning on the outcomes of T^*.
Introducing $T(y_1,\dots,y_n) = (y_{r+1},\dots,y_s, \Sigma_{i=1}^r y_i^2 + \Sigma_{i=s+1}^n y_i^2)$, so that $T(Y_1,\dots,Y_n) = (Y_{r+1},\dots,Y_s,U+V)$, and writing (for convenience) $T(y_1,\dots,y_n) = (t_{r+1},\dots,t_s,w) = t$, it follows from the above that the optimality of the unconditional test may be proved by maximizing the power of the conditional tests given $T(Y_1,\dots,Y_n) = (t_{r+1},\dots,t_s,w)$ (or $T(Y_1,\dots,Y_n) = t$), for each t separately.
Now let $h_\theta^t(y,\dots,y_n)$ be the conditional joint desity of $Y_1,\dots,Y_n$ given $T(Y_1,\dots,Y_n) = t$, $\theta \in \Omega$. Let $k_\theta(y_{r+1},\dots,y_s)$ be the joint (marginal) density of $Y_{r+1},\dots,Y_s$, $\theta \in \Omega$. Notice that $k_\theta(y_{r+1},\dots,y_s)$ depends only on the parameters $\eta_{r+1},\dots,\eta_s$ and σ^2. Let $g_\theta(t) = g_\theta(t_{r+1},\dots,t_s,w)$ be the joint density of $T(Y_1,\dots,Y_n) = (Y_{r+1},\dots,Y_s,U+V)$ under $\theta \in \Omega$. Because $U+V$ is independent of $Y_{r+1},\dots,Y_s$ we have $g_\theta(t) = k_\theta(t_{r+1},\dots,t_s)f_\theta(w)$, where $f_\theta(w)$ is the density of $U+V$ under $\theta \in \Omega$.
We now have (cf. RAO (1973), p. 99)

$$h_\theta^t(y) = \begin{cases} \dfrac{h_\theta(w)}{g_\theta(t)} & \text{when } T(y) = t \\ 0 & \text{otherwise.} \end{cases}$$

Now let $\theta_0 = (0,\dots,0,\eta_{r+1},\dots,\eta_s,\sigma) \in \Omega_H$ be fixed.
Define the sphere S as a subset of Ω_K as follows

$$S = S(\eta_{r+1},\dots,\eta_s,\sigma;\rho) =$$

$$= \{(\alpha_1,..,\alpha_s,\tau) \in \Omega_K : \sum_{i=1}^{r} \frac{\alpha_i^2}{\sigma^2} = \rho^2;\ \alpha_j = \eta_j\ (j = r+1,..,s); \tau = \sigma\}.$$

Consider the reduced problem of testing $H' : \theta = \theta_0$ against $K' : \theta \in S$. We now obtain the conditional test for the reduced problem that maximizes the average power against S.

Let $\theta \in S$. Then provided $T(y_1,...,y_n) = (t_{r+1},...,t_s,w)$

$$\frac{h_\theta^t(y)}{h_{\theta_0}^t(y)} = \frac{h_\theta(y)}{g_\theta(t)} \frac{g_{\theta_0}(t)}{h_{\theta_0}(y)} = \frac{h_\theta(y)}{h_{\theta_0}(y)} \frac{k_{\theta_0}(t_{r+1},..,t_s)}{k_\theta(t_{r+1},..,t_s)} \frac{f_{\theta_0}(w)}{f_\theta(w)} =$$

$$= \frac{h_\theta(y)}{h_{\theta_0}(y)} \frac{f_{\theta_0}(w)}{f_\theta(w)},$$

because $k_{\theta_0}(t_{r+1},...,t_s) = k_\theta(t_{r+1},...,t_s)$ for $\theta \in S$.

Here $f_{\theta_0}(w)$ is the density of a central χ^2-variable and $f_\theta(w)$ the density of a noncentral χ^2-variable.

A straightforward application of Neyman and Pearson's lemma to $h_{\theta_0}^t(y_1,...,y_n)$ and the average of $h_\theta^t(y_1,...,y_n)$ over S, i.e.

$$\int_S h_\theta^t(y)dA \,/ \int_S dA,$$

where θ ranges over S and dA is the differential area on the surface of S, gives that the average power against S is maximized by the test that rejects when the ratio of the average density against S to the density under H' is larger that a suitable constant, i.e. when

$$[h_{\theta_0}^t(y)]^{-1} \int_S h_\theta^t(y)dA \,/ \int_S dA > C(t),$$

or, equivalently, when

$$\int_S \frac{h_\theta^t(y)}{h_{\theta_0}^t(y)} dA > C(t) \int_S dA$$

Now we have

$$\int_S \frac{h_\theta^t(y)}{h_{\theta_0}^t(y)} dA = \int_S \frac{h_\theta(y)}{h_{\theta_0}(y)} \frac{f_{\theta_0}(w)}{f_\theta(w)} dA.$$

In this expression the factor $f_{\theta_0}(w) / f_\theta(w)$ still depends on θ. However, the density $f_\theta(w)$ depends on θ only through its noncentrality parameter $\psi^2 = \Sigma_{i=1}^r \eta_i^2\sigma^{-2}$ $(\theta = (\eta_1,\dots,\eta_s,\sigma))$, *which is constant on* S: $\psi^2 = \rho^2$ for $\theta \in S$. This means that the factor $f_{\theta_0}(w) / f_\theta(w)$ is constant with respect to the integration over S and may thus be absorbed in the constant C(t).

Furthermore, we have, for $\theta \in S$,

$$\frac{h_\theta(y)}{h_{\theta_0}(y)} = \frac{(\sigma\sqrt{2\pi})^{-n}\exp\{-\frac{1}{2\sigma^2}[\Sigma_{i=1}^s (y_i-\eta_i)^2 + \Sigma_{i=s+1}^n y_i^2]\}}{(\sigma\sqrt{2\pi})^{-n}\exp\{-\frac{1}{2\sigma^2}[\Sigma_{i=1}^r y_i^2 + \Sigma_{i=r+1}^s (y_i-\eta_i)^2 + \Sigma_{i=s+1}^n y_i^2]\}} =$$

$$= \exp\left(\sum_{i=1}^r \frac{y_i\eta_i}{\sigma^2} - \tfrac{1}{2}\rho^2\right).$$

Again the factor $\exp(-\frac{1}{2}\rho^2)$ may be absorbed in the constant C(t). Thus we obtain that the optimal conditional test rejects when

$$\int_S \exp\left(\sum_{i=1}^r \frac{y_i\eta_i}{\sigma^2}\right) dA > C(t).$$

Keeping in mind that we still work with a fixed S we define

$$g(y_1,\dots,y_r) = \int_S \exp\left(\sum_{i=1}^r \frac{y_i\eta_i}{\sigma^2}\right) dA .$$

Considering $\eta = (\eta_1,\dots,\eta_r)'$ and $\tilde{y} = (y_1,\dots,y_r)'$ as vectors we have

$$(\eta,\tilde{y}) = \|\eta\|\,\|\tilde{y}\|\cos(\eta,\tilde{y}),$$

that is

$$\sum_{i=1}^r \eta_i y_i = \left(\sum_{i=1}^r \eta_i^2 \sum_{i=1}^r y_i^2\right)^{\frac{1}{2}} \cos(\eta,\tilde{y}),$$

or, equivalently,

$$\sum_{i=1}^r \frac{\eta_i y_i}{\sigma^2} = \rho\sqrt{u}\frac{1}{\sigma}\cos(\eta,\tilde{y}) \overset{\text{def}}{=} \rho\sqrt{u}\frac{1}{\sigma}\cos\beta.$$

By symmetry the average of $\exp(\rho\sqrt{u}\,\frac{1}{\sigma}\cos\beta)$ when η ranges over S is unchanged when $\tilde{y}$ is replaced by an arbitrary vector of the same length

as $\tilde{y}$. This means that $g(\tilde{y})$ depends on $\tilde{y} = (y_1,\ldots,y_r)'$ only through u. So write $g(\tilde{y}) = h(u)$.
Consider a fixed $\tilde{y} = (y_1,\ldots,y_r)'$. Let S' be the subset of S for which $0 \le \gamma \le \frac{1}{2}\pi$, where γ is the angle between $\tilde{y}$ and η. Then

$$h(u) = \int_S [\exp(\rho\sqrt{u}\ \frac{1}{\sigma}\cos\gamma) + \exp(-\rho\sqrt{u}\ \frac{1}{\sigma}\cos\gamma)]dA,$$

which is an increasing function of u.
The conditional test for the reduced testing problem: $H' : \theta = \theta_0$ against $K' : \theta \in S$, which maximizes the average power against S is thus given by the rejection region

(1) $\qquad h(u) > C(t).$

Since the test with rejection region (1) is independent of the particular θ_0 chosen and independent of the particular sphere S, it follows that this test may also serve as a test for the original testing problem.
This means that (1) is the conditional test which maximizes the average power against alternatives on a sphere S for every such sphere.
Finally, because under H $U/(U+V)$ is independent of $U+V$, it follows that the unconditional distribution of $U/(U+V)$ and the conditional distribution of $U/(U+V)$ "given $U+V$" are the same. Because h(u) is a monotone function of u this means that the rejection regions determined by either $h(u) > C$ or $u/(u+v) > C^*$ are the same. This completes the proof.

(ii) Consider now the class of all *similar* tests whose power depends only on $\rho^2 = \sigma^{-2}\Sigma_{i=1}^r \eta_i^2$. Hence the power of such tests is constant over the sphere S. The test given by (9), Section 1 or by (1), part (i) of this problem, now maximizes "the" power against alternatives on spheres S. Because each alternative is an element of some sphere S, the test (9), Section 1 is UMP among all those tests.

(WALD (1942), HSU (1945))

Section 2

Problem 6.

(i) Without loss of generality we may assume that $\theta = \beta_1$: consider any nonsingular matrix B with first row $(e_1,e_2,\ldots,e_s)$. Put $\beta^* = B\beta$ and $A^* = AB^{-1}$, $A = (a_{ij})$. Then $\beta_1^* = \Sigma_{i=1}^s e_i\beta_i = \theta$ and $\xi = (AB^{-1})(B\beta) = A^*\beta^*$

(it is implicitly assumed that $\Sigma_{i=1}^{s} e_i^2 > 0$).
Following the hint, using column vectors only we infer that $Y_1 = \lambda\hat{\beta}_1$, where $\lambda = c_1'a_1 \overset{\text{def}}{=} \Sigma_{j=1}^{n} c_{1j}a_{j1}$. Now $\mathrm{var}(Y_1) = \lambda^2 \mathrm{var}(\hat{\beta}_1) = \lambda^2 \mathrm{var}(\Sigma_i d_i X_i) = \lambda^2\sigma^2 \cdot \Sigma d_i^2$, where $\sigma^2 = \mathrm{var}(X_1) = \mathrm{var}(Y_1)$. Hence $|\lambda| = (\Sigma d_i^2)^{-\frac{1}{2}}$ and $|Y_1| = |\hat{\beta}_1| / \sqrt{\Sigma d_i^2}$. Thus, by (12) and (13), the rejection region for testing $H(0) : \beta_1 = 0$ against $K(0) : \beta_1 \neq 0$ is given by

$$\frac{|\hat{\beta}_1| / \sqrt{\Sigma d_i}}{\sqrt{\Sigma (X_i - \hat{\xi}_i)^2 / (n-s)}} > C_0.$$

For any β_1^0, the transformation $X^* = X - \beta_1^0 a_1$ reduces the problem of testing $H(\beta_1^0) : \beta_1 = \beta_1^0$ against $K(\beta_1^0) : \beta_1 \neq \beta_1^0$ to the previous one, for $E(X^*) = (\beta_1 - \beta_1^0)a_1 + \Sigma_{j=2}^{s} \beta_j a_j$. Now the desired result follows.

(ii) Again we may assume, without loss of generality, that $\theta = \beta_1$.
With $Y_1,\ldots,Y_n$, $c_1,\ldots,c_n$ as in (i) we have $Y_i = c_i'X$, $i = 1,\ldots,n$. Hence $\eta_1 = E(Y_1) = E(c_1'X) = c_1'\xi = \Sigma\, \beta_i c_1' a_i = \beta_1 c_1' a_1 = \lambda\beta_1$, because c_1 is orthogonal to each $a_2,\ldots,a_s$. Since $\eta_1,\ldots,\eta_s$ do not involve β_1, the hypothesis $\beta_1 = \beta_1^0$ is equivalent to $\eta_1 = \eta_1^0$ with $\eta_1^0 = \lambda\beta_1^0$.
Now consider any fixed η_1^0. Define the groups

$$G_1 = \{g : g(y_1,\ldots,y_n) = (y_1, y_2+k_2,\ldots,y_s+k_s, y_{s+1},\ldots,y_n), k_2,\ldots,k_s \in \mathbb{R}\},$$

$$G_2(\eta_1^0) = \{g : g(y_1,\ldots,y_n) = (\tau(y_1-\eta_1^0)+\eta_1^0, y_2,\ldots,y_n), \tau = -1 \text{ or } \tau = 1\},$$

$$G_3(\eta_1^0) = \{g : g(y_1,\ldots,y_n) = (c(y_1-\eta_1^0)+\eta_1^0,\ cy_2,\ldots,cy_n), c \neq 0\}.$$

These groups leave the testing problem invariant. In the same way as on p. 267 we find that the test that rejects when

$$\frac{|y_1 - \eta_1^0|}{\Sigma_{i=s+1}^{n} Y_i^2 / (n-s)} > C_0$$

is UMP invariant with respect to the group $G(\eta_1^0)$ generated by G_1, $G_2(\eta_1^0)$ and $G_3(\eta_1^0)$.
It can be easily verified that the group G obtained from the $G(\eta_1^0)$'s by varying η_1^0 consists of all transformations of the form

$$g(y_1,\ldots,y_n) = (\tau\alpha y_1+\delta_1, \alpha y_2+\delta_1,\ldots,\alpha y_s+\delta_s, \alpha y_{s+1},\ldots,\alpha y_n),$$

where $\alpha, \delta_1, \ldots, \delta_s \in \mathbb{R}$, $\alpha \neq 0$ and $\tau \in \{-1, 1\}$.
It remains to show that the confidence intervals for β_1 given by

$$S(y) = \{\beta_1 : \frac{|\frac{1}{\lambda} y_1 - \beta_1|}{\Sigma_{i=s+1}^{n} y_i^2 / (n-s)} \leq k\},$$

in terms of the y_i's are uniformly most accurate G-invariant. Consider any fixed $g \in G$. Then

$$S[g(y)] = \left\{\beta_1 : \frac{\left|\frac{1}{\lambda}(\tau\alpha y_1 + \delta_1) - \beta_1\right|}{\left\{\Sigma_{i=s+1}^{n} (\alpha y_i)^2 / (n-s)\right\}^{\frac{1}{2}}} \leq k\right\} =$$

$$= \left\{\beta_1 : \frac{\left|\frac{1}{\lambda} y - \frac{1}{\tau\alpha}(\beta_1 - \frac{\delta_1}{\lambda})\right|}{\left\{\Sigma_{i=s+1}^{n} y_i^2 / (n-s)\right\}^{\frac{1}{2}}} \leq k\right\} =$$

$$= \{\beta_1 : \frac{1}{\tau\alpha}(\beta_1 - \frac{\delta_1}{\lambda}) \in S(y)\} =$$

$$= \{\tau\alpha\beta_1 + \frac{\delta_1}{\lambda} : \beta_1 \in S(y)\} \stackrel{\text{def}}{=} g^* S(y).$$

Hence the confidence intervals are G-invariant.
Now by Lemma 4 (ii), Chapter 6 they are uniformly most accurate G-invariant and hence the intervals (90) are uniformly most accurate invariant with respect to the group G^* induced by the inverse canonical transformation $x = \Sigma_{i=1}^{n} y_i c_i$.
Its elements are of the form:

$$g^*(x) = (\tau\alpha y_1 + \delta_1)c_1 + \sum_{i=2}^{s} (\alpha y_i + \delta_i)c_i + \sum_{i=s+1}^{n} (\alpha y_i)c_i =$$

$$= \alpha \sum_{i=1}^{n} y_i c_i + \alpha(\tau - 1)y_1 c_1 + \sum_{i=1}^{s} \delta_i c_i =$$

$$= \alpha x + \alpha(\tau-1)c_1 c_1' x + \xi =$$

$$= \alpha[I + (\tau-1)c_1 c_1']x + \xi,$$

Where $\alpha \neq 0$, $\tau \in \{-1, 1\}$ and $\xi \in \Pi_\Omega$.

Problem 7.

First we change notation a little: let Z_{ij} $(j=1,\dots,m_i)$ and Y_{ij} $(j=1,\dots,n_i)$, $i=1,\dots,p$, be independently normally distributed with common variance σ^2 and means $E(Z_{ij}) = \zeta_i$ and $E(Y_{ij}) = \zeta_i + \Delta$.
We now have a situation as has been described in Problem 6 with $n = \Sigma n_i + \Sigma m_i$, $s = p+1$, $\beta = (\zeta_1,\dots,\zeta_p,\Delta)'$, $X = (Z_{11},\dots,Z_{1m_1},\dots,Z_{p1},\dots,Z_{pm_p},Y_{11},\dots,Y_{1n_1},\dots,Y_{p1},\dots,Y_{pn_p})'$, A an $n \times (p+1)$ matrix of rank $p+1$ and $\theta = \Delta$.
The problem becomes that of minimizing

$$(2) \qquad \sum_{i=1}^{p}\left[\sum_{j=1}^{m_i} (Z_{ij}-\zeta_i)^2 + \sum_{j=1}^{n_i} (Y_{ij}-\zeta_i-\Delta)^2\right]$$

over all possible values of $\zeta_1,\dots,\zeta_p$ and Δ, and substituting the minimizing values $\hat{\zeta}_1,\dots,\hat{\zeta}_p$ and $\hat{\Delta}$ in (89).
Differentiating the above sum of squares and setting the partial derivatives equal to zero we get the following system of linear equations in $\hat{\zeta}_1,\dots,\hat{\zeta}_p$, $\hat{\Delta}$

$$\sum_{j=1}^{m_i} (Z_{ij}-\hat{\zeta}_i) + \sum_{j=1}^{n_i} (Y_{ij}-\hat{\zeta}_i-\hat{\Delta}) = 0, \quad i=1,\dots,p$$

$$\sum_{i=1}^{p}\sum_{j=1}^{n_i} (Y_{ij}-\hat{\zeta}_i-\hat{\Delta}) = 0$$

with solutions

$$\hat{\Delta}(=\hat{\theta}) = \left(\sum_{i=1}^{p} \frac{m_i n_i}{N_i}\right)^{-1} \sum_{i=1}^{p} \frac{m_i n_i}{N_i}(Y_{i\cdot}-Z_{i\cdot})$$

$$\hat{\zeta}_i = N_i^{-1}(m_i Z_{i\cdot} + n_i Y_{i\cdot} - n_i\hat{\Delta}), \quad i=1,\dots,p,$$

where

$$N_i = n_i + m_i, \quad Y_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij} \text{ and } Z_{i\cdot} = \frac{1}{m_i}\sum_{j=1}^{m_i} Z_{ij}.$$

Since (2) is a sum of squares these values give a minimum for (2). Substitution into (89), Problem 6, yields the rejection region of the UMP invariant test for $H: \Delta = 0$ $(\theta_0 = 0)$. Notice that since

$$\hat{\Delta} = \sum_i \sum_{j=1}^{m_i}\left(-\frac{n_i}{N_i D} Z_{ij}\right) + \sum_i \sum_{j=1}^{n_i} \frac{m_i}{N_i D},$$

where $D = \sum_i \frac{m_i n_i}{N_i}$ the factor $\sqrt{\Sigma d_i^2}$ in (89) is equal to

$$\left\{D^{-2}\left[\sum_i \frac{m_i n_i^2}{N_i^2} + \sum_i \frac{n_i m_i^2}{N_i^2}\right]\right\}^{\frac{1}{2}} = \left\{D^{-2} \sum_i \frac{n_i m_i}{N_i^2}(n_i + m_i)\right\}^{\frac{1}{2}} = D^{-\frac{1}{2}}.$$

Problem 8.

As in Section 1, Chapter 7 we reduce the problem to a canonical form. This yields the independently normally distributed variables $Y_1,\dots,Y_n$ with common variance σ_0^2 and means $E(Y_i) = \eta_i$ for $i = 1,\dots,s$ and $E(Y_i) = 0$ for $i = s+1,\dots,n$. Hence their joint density is

$$(\sigma_0\sqrt{2\pi})^{-n} \exp\left[-\frac{1}{2\sigma_0^2} \sum_{i=1}^{s} (y_i - \eta_i)^2 - \frac{1}{2\sigma_0^2} \sum_{i=s+1}^{n} y_i^2\right].$$

The hypothesis to be tested reduces to $H : \eta_1 = \dots = \eta_r = 0$ $(r \le s \le n)$. By sufficiency we may restrict attention to $Y_1,\dots,Y_s$ (σ_0 known!).
Now consider the groups of transformations G_1 and G_2 from p. 267.
The hypothesis remains invariant under G_1, which leaves $Y_1,\dots,Y_r$ as maximal invariants.
G_2 also leaves the testing problem invariant. A maximal invariant is $U = \Sigma_{i=1}^r Y_i^2$. Hence a UMP invariant test can be chosen to depend only on U.
As has been shown on p. 267 $\psi^2 = \Sigma_{i=1}^r \eta_i^2$ is a maximal invariant with respect to the induced groups $\bar{G}_1$ and $\bar{G}_2$. By Theorem 3, Chapter 6 it follows that the distribution of U depends only on ψ^2. To be more precise, U has the noncentral χ^2-distribution with r degrees of freedom and non-centrality parameter ψ^2/σ_0^2. Hence the principles of sufficiency and invariance reduce the problem to that of testing $H' : \psi^2 = 0$ against $K' : \psi^2 > 0$.
Since by Problem 4 (i), the class of probability densities $p_\psi(u)$ of U has monotone likelihood ratio in u, there exists a UMP invariant test that rejects when U is too large, that is when

$$U = \sum_{i=1}^{r} Y_i^2 > C \cdot \sigma_0^2.$$

C is determined by $\int_C^\infty \chi_r^2(y)dy = \alpha$.
In the same way as in Section 2, p. 269 we find that

$$U = \sum_{i=1}^{n} (X_i - \hat{\hat{\xi}}_i)^2 - \sum_{i=1}^{n} (X_i - \hat{\xi}_i)^2 = \sum_{i=1}^{n} (\hat{\xi}_i - \hat{\hat{\xi}}_i)^2.$$

Section 3

Problem 9.

Since the variables $X_{i.}$ $(i = 1,\dots,s)$ are independently normally distributed with means $E(X_{i.}) = \mu_i$ and variances $\mathrm{var}(X_{i.}) = n_i^{-1}\sigma^2$ $(i = 1,\dots,s)$ and $X_{..}$ is distributed as $N(\mu_.,n^{-1}\sigma^2)$, we have

$$
\begin{aligned}
E[\Sigma\, n_i(X_{i.} - X_{..})^2] &= E[(\Sigma\, n_i X_{i.}^2) - nX_{i.}^2) - nX_{..}^2] = \\
&= [\Sigma\, n_i(n_i^{-1}\sigma^2 + \mu_i^2)] - n(n^{-1}\sigma^2 + \mu_.^2) = \\
&= (s-1)\sigma^2 + \Sigma\, n_i(\mu_i - \mu_.)^2
\end{aligned}
$$

and

$$
\begin{aligned}
E\left[\sum_i \sum_{j=1}^{n_i} (X_{ij} - X_{i.})^2\right] &= E\left\{\sum_i \left[\left(\sum_{j=1}^{n_i} X_{ij}^2\right) - n_i X_{i.}^2\right]\right\} = \\
&= \sum_i \sum_{j=1}^{n_i} (\sigma^2 + \mu_i^2) - \sum_i n_i(n_i^{-1}\sigma^2 + \mu_i^2) = \\
&= (n-s)\sigma^2.
\end{aligned}
$$

Problem 10.

(i) Transform into new variables $X_i = a_i^{-1} Z_i$ $(i = 1,\dots,s)$. Then $X_1,\dots,X_s$ are independently normally distributed as $N(\xi_i,1)$ with $\xi_i = a_i^{-1}\zeta_i$.

The hypothesis $H : \zeta_1 = \dots = \zeta_s$ becomes $H' : a_1\xi_1 = \dots = a_s\xi_s$.

This is a linear hypothesis with $r = s-1$.

With respect to a suitable group of linear transformations the UMP invariant test is given in Problem 8. Its rejection is

$$
(3) \qquad \sum_{i=1}^{s} (\hat{\xi}_i - \hat{\hat{\xi}}_i)^2 > C,
$$

where C is determined by $\int_C^\infty \chi^2_{s-1}(y)dy = \alpha$. Consider any fixed i.

Since $\hat{\xi}_i$ minimizes $\Sigma_j\,(X_j - \xi_j)^2$ it must be equal to X_i.

$\hat{\hat{\xi}}_i$ minimizes $\Sigma_j\,(X_j - \xi_j)^2$ subject to $a_1\xi_1 = \dots = a_s\xi_s$, that is it minimizes $\Sigma_j(X_j - a_j^{-1}a_i\xi_i)^2$ over all possible values of ξ_i. Hence

$$
\hat{\hat{\xi}}_i = \left(a_i \sum_j a_j^{-2}\right)^{-1}\left(\sum_j X_j a_j^{-1}\right).
$$

Now the left hand side of (3) becomes

$$U \overset{\text{def}}{=} \sum_{i=1}^{s}\left[X_i - \left(a_i \sum_j a_j^{-2}\right)^{-1}\left(\sum_j X_j a_j^{-1}\right)\right]^2 =$$

$$= \sum_i a_i^{-2}\left[Z_i - \left(\sum_j a_j^{-2}\right)^{-1}\left(\sum_j Z_j a_j^{-2}\right)\right]^2 =$$

$$= \sum_i (Z_i a_i^{-1})^2 - \left(\sum_i a_i^{-2}\right)^{-1}\left(\sum_i Z_i a_i^{-2}\right)^2 (\overset{\text{def}}{=} f(Z_1,\dots,Z_s)),$$

as was to be shown.

(ii) Following the theory on pp. 270-271 formula (16) and further we find that U has the noncentral χ^2-distribution with $r = s-1$ degrees of freedom and noncentrality parameter $\lambda^2 = f(\zeta_1,\dots,\zeta_s)$.
Now the assertion of this part of the problem easily follows.

Problem 11.

The solution of this problem is based on the following theorem, which is slightly stronger than that formulated in Lehmann, p. 274 (cf. RAO (1973), p. 385. Note that we only need existence of the derivative at θ).

THEOREM. *If* $\{T_n\}$ *is a sequence of real-valued statistics such that* $\sqrt{n}(T_n - \theta)$ *has the limiting distribution* $N(0,\tau^2)$, *then for any function* $f(\theta)$ *which is differentiable at* θ, *the limiting distribution of* $\sqrt{n}[f(T_n) - f(\theta)]$ *is normal with zero mean and variance* $\tau^2\left(\frac{df}{d\theta}\right)^2$.

(i) We first show that for large λ the quantity $\sqrt{\lambda}(\lambda^{-1}X - 1)$ is approximally standard normally distributed. This follows by the Lévi-Cramér theorem, since

$$E \exp\left[it\sqrt{\lambda}(\lambda^{-1}X - 1)\right] = e^{-it\sqrt{\lambda}} E\left[\exp\left(\frac{itX}{\sqrt{\lambda}}\right)\right] =$$

$$= e^{-it\sqrt{\lambda}} \exp\left[\lambda(e^{it/\sqrt{\lambda}} - 1)\right] =$$

$$= \exp\left\{-it\sqrt{\lambda} - \lambda + \lambda\left[1 + i\frac{1}{\sqrt{\lambda}} - \frac{t^2}{2\lambda} + o\left(\frac{t^2}{\sqrt{\lambda}}\right)\right]\right\} \to e^{-\frac{1}{2}t}$$

as $\lambda \to \infty$.
Now applying the above theorem with $f(u) = \sqrt{u}$ gives the required result.

(ii) By the Central Limit Theorem (or the De Moivre-Laplace Theorem) $\sqrt{n}(n^{-1}X - p)$ is approximately distributed as $N(0,p(1-p))$ for large n.
Now by the above theorem with $f(u) = \arcsin\sqrt{u}$ it follows that the limiting

distribution of $\sqrt{n}\,[\arcsin\sqrt{n^{-1}}$ - $\arcsin\sqrt{p}]$ is normal with zero mean and variance $\frac{1}{4}$. In other words $\arcsin\sqrt{n^{-1}X}$ is for large n approximately normally distributed with mean $\arcsin\sqrt{p}$ and variance $\frac{1}{4n}$.

Remark

A slightly different version of the theorem can be found in SERFLING (1980), Theorem 3.1.A.

Section 5

Problem 12.

We use the notation of Section 5

Since

$$\sum_i \sum_j \sum_k (X_{ijk} - \xi_{ij})^2 = \sum_i \sum_j \sum_k (X_{ijk} - X_{ij\cdot})^2 +$$

$$+ m \sum_i \sum_j (X_{ij\cdot} - X_{i\cdot\cdot} - X_{\cdot j\cdot} + X_{\cdot\cdot\cdot} - \gamma_{ij})^2 +$$

$$+ mb \sum_i (X_{i\cdot\cdot} - X_{\cdot\cdot\cdot} - \alpha_i)^2 + ma \sum_j (X_{\cdot j\cdot} - X_{\cdot\cdot\cdot} - \beta_j)^2 +$$

$$+ mab\,(X_{\cdot\cdot\cdot} - \mu)^2,$$

it follows that in this case $\hat{\alpha}_i = \hat{\hat{\alpha}}_i = X_{i\cdot\cdot} - X_{\cdot\cdot\cdot}$, $\hat{\beta}_j = \hat{\hat{\beta}}_j = X_{\cdot j\cdot} - X_{\cdot\cdot\cdot}$, $\hat{\mu} = \hat{\hat{\mu}} = X_{\cdot\cdot\cdot}$, $\hat{\gamma}_{ij} = X_{ij\cdot} - X_{i\cdot\cdot} - X_{\cdot j\cdot} + X_{\cdot\cdot\cdot}$ and $\hat{\hat{\gamma}}_{ij} = 0$. Hence substitution in (15) yields the rejection region (30).

Problem 13.

By Problem 4 (i) the family of densities $p_\lambda(x)$ of X_λ has monotone likelihood ratio in x.

Hence by Lemma 2 (ii), Chapter 3 we have for $\lambda < \lambda'$ and $x \in \mathbb{R}$

$$F_{\lambda'}(x) \leq F_\lambda(x),$$

where F_λ is the cumulative distribution function of X_λ.

Problem 14.

The hypothesis $H : \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$ is a linear hypothesis with $n = abm$, $r = a-1$, $s = 1 + (a-1) + (b-1) + (m-1)$ and $n-s = abm - a - b - m + 2$.

Let $\xi_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k$ $(\Sigma\,\alpha_i = \Sigma\,\beta_j = \Sigma\,\gamma_k = 0)$.
Then

$$\sum_i \sum_j \sum_k (X_{ijk} - \xi_{ijk})^2 = \sum_i \sum_j \sum_k (X_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k)^2 =$$

$$= \sum_i \sum_j \sum_k [(X_{ijk} - X_{i..} - X_{.j.} - X_{..k} + 2X_{...}) +$$

$$+ (X_{i..} - X_{...} - \alpha_i) + (X_{.j.} - X_{...} - \beta_j) +$$

$$+ (X_{..k} - X_{...} - \gamma_k) + (X_{...} - \mu)]^2 =$$

$$= \sum_i \sum_j \sum_k (X_{ijk} - X_{i..} - X_{.j.} - X_{..k} + 2X_{...})^2 +$$

$$+ bm \sum_i (X_{i..} - X_{...} - \alpha_i)^2 + am \sum_j (X_{.j.} - X_{...} - \beta_j)^2 +$$

$$+ ab \sum_k (X_{..k} - X_{...} - \gamma_k)^2 + abm\,(X_{...} - \mu)^2,$$

because the cross-product terms vanish.
The least-squares estimates are

$$\hat{\alpha}_i = X_{i..} - X_{...},\quad \hat{\hat{\alpha}}_i = 0,\quad \hat{\beta}_j = \hat{\hat{\beta}}_j = X_{.j.} - X_{...},$$

$$\hat{\gamma}_k = \hat{\hat{\gamma}}_k = X_{..k} - X_{...} \quad \text{and} \quad \hat{\mu} = \hat{\hat{\mu}} = X_{...}\ .$$

Thus we have

$$\sum_i \sum_j \sum_k (X_{ijk} - \hat{\xi}_{ijk})^2 = \sum_i \sum_j \sum_k (X_{ijk} - X_{i..} - X_{.j.} - X_{..k} + 2X_{...})^2,$$

and

$$\sum_i \sum_j \sum_k (\hat{\xi}_{ijk} - \hat{\hat{\xi}}_{ijk})^2 = bm \sum_i (X_{i..} - X_{...})^2.$$

Hence the UMP invariant test rejects when

$$W^* = \frac{(abm - a - b - m + 2)bm\ \Sigma_i\ (X_{i..} - X_{...})^2}{(a-1)\ \Sigma_i\ \Sigma_j\ \Sigma_k\ (X_{ijk} - X_{i..} - X_{.j.} - X_{..k} + 2X_{...})^2} > C.$$

W^* has the noncentral F distribution with $(a-1)$ and $(abm - a - b - m + 2)bm$ degrees of freedom. For the noncentrality parameter ψ^2 we find following the theory on pp. 270-271,

$$\psi^2 = mb\sigma^{-2} \sum_i \alpha_i^2 .$$

Problem 15.

First notice the following symmetry: as given the level k of the third factor is a function of the levels i and j of the other factors, $k = f(i,j)$, satisfying

(4) $$\{f(i,j) : j = 1,\ldots,m\} = \{1,\ldots,m\}, \text{ for all } i, \text{ and}$$
$$\{f(i,j) : i = 1,\ldots,m\} = \{1,\ldots,m\}, \text{ for all } j.$$

It follows that there exist functions $g(i,k)$ and $h(j,k)$ such that $j = g(i,k)$ and $i = h(j,k)$ when $k = f(i,j)$. These functions have similar properties as f (cf. (4)).
In the following we will indicate by parentheses which index is considered dependent of the other two:

$$\xi_{ij(k)} = \xi_{ijf(i,j)}; \quad \xi_{i(j)k} = \xi_{ig(i,k)k}; \quad \xi_{(i)jk} = \xi_{h(j,k)jk}$$

(i) For any i

$$\xi_{i\cdot(\cdot)} = \frac{1}{m}\sum_j (\mu + \alpha_i + \beta_j + \gamma_{(k)}) = \mu + \alpha_i + \frac{1}{m}\sum_j \beta_j + \frac{1}{m}\sum_k \gamma_k = \mu + \beta_i$$

and by symmetry

$$\xi_{\cdot j(\cdot)} = \mu + \beta_j, \quad \xi_{\cdot\cdot(k)} = \xi_{\cdot(\cdot)k} = \mu + \gamma_k \quad \text{for any } j \text{ and } k.$$

Finally

$$\xi_{\cdot\cdot(\cdot)} = \frac{1}{m}\sum_i \xi_{i\cdot(\cdot)} = \mu.$$

(ii) Notice that

(5) $$\sum_i\sum_j [X_{ij(k)} - \xi_{ij(k)}]^2 = \sum_i\sum_j [(X_{i\cdot(\cdot)} - X_{\cdot\cdot(\cdot)} - \alpha_i) +$$
$$+ (X_{\cdot j(\cdot)} - X_{\cdot\cdot(\cdot)} - \beta_j) + (X_{\cdot\cdot(k)} - X_{\cdot\cdot(\cdot)} - \gamma_{(k)}) +$$
$$+ (X_{\cdot\cdot(\cdot)} - \mu) + (X_{ij(k)} - X_{i\cdot(\cdot)} - X_{\cdot j(\cdot)} - X_{\cdot\cdot(k)} + 2X_{\cdot\cdot(\cdot)})]^2 =$$
$$= m\sum_i [X_{i\cdot(\cdot)} - X_{\cdot\cdot(\cdot)} - \alpha_i]^2 + m\sum_j [X_{\cdot j(\cdot)} - X_{\cdot\cdot(\cdot)} - \beta_j]^2 +$$

$$+ m \sum_k [X_{\cdot\cdot(k)} - X_{\cdot\cdot(\cdot)} - \gamma_k]^2 + m^2[X_{\cdot\cdot(\cdot)} - \mu]^2 +$$

$$+ \sum_i \sum_j [X_{ij(k)} - X_{i\cdot(\cdot)} - X_{\cdot j(\cdot)} - X_{\cdot\cdot(k)} + 2X_{\cdot\cdot(\cdot)}]^2,$$

because the cross-product terms vanish.
Hence the least-squares estimates are

$$\hat{\alpha}_i = X_{i\cdot(\cdot)} - X_{\cdot\cdot(\cdot)},\ \hat{\beta}_j = X_{\cdot j(\cdot)} - X_{\cdot\cdot(\cdot)},\ \hat{\gamma}_k = X_{\cdot\cdot(k)} - X_{\cdot\cdot(\cdot)}$$

and $\hat{\mu} = X_{\cdot\cdot(\cdot)}$.

(iii) First we will establish the dimension of the parameter space

$$\Pi_\Omega = \{\underline{\xi} : \xi_{ij(k)} = \mu + \alpha_i + \beta_j + \gamma_{(k)}, \Sigma\alpha_i = \Sigma\beta_j = \Sigma\gamma_k = 0\}.$$

By part (i) there is a 1-1 correspondence between $\xi \in \Pi_\Omega$ and $(\mu,\alpha_1,\ldots,\alpha_{m-1},\beta_1,\ldots,\beta_{m-1},\gamma_1,\ldots,\gamma_{m-1}) \in \mathbb{R}^{3m-2}$. Since this 1-1 correspondence is linear we have $s = \dim \Pi_\Omega = 3m-2$. In the same way we see that the dimension of the space

$$\Pi_\omega = \{\underline{\xi} : \xi_{ij(k)} = \mu + \beta_j + \gamma_{(k)}, \Sigma\beta_j = \Sigma\gamma_k = 0\} =$$

is $2m-1$.
Hence the hypothesis $H : \alpha_1 = \alpha_2 = \ldots = \alpha_m = 0$ is a linear hypothesis with $s = 3m-2$, $r = (3m-2) - (2m-1) = m-1$ and $n-s = m^2 - (3m-2) = (m-1)(m-2)$.
The least-squares estimates of μ,α_i,β_j and γ_k under H follow from (5): $\hat{\hat{\alpha}}_i = 0$, $\hat{\hat{\beta}}_j = \hat{\beta}_j$, $\hat{\hat{\gamma}}_k = \hat{\gamma}_k$ and $\hat{\hat{\mu}} = \hat{\mu}$.
Hence

$$\sum_i \sum_j [\hat{\xi}_{ij(k)} - \hat{\hat{\xi}}_{ij(k)}]^2 = \sum_i \sum_j \hat{\alpha}_i^2 = m \sum_i [X_{i\cdot(\cdot)} - X_{\cdot\cdot(\cdot)}]^2$$

and

$$\sum_i \sum_j [X_{ij(k)} - \hat{\xi}_{ij(k)}]^2 = \sum_i \sum_j [X_{ij(k)} - X_{i\cdot(\cdot)} - X_{\cdot j(\cdot)} - X_{\cdot\cdot(k)} + 2X_{\cdot\cdot(\cdot)}]^2.$$

Substitution in (15) now gives the desired result.
The noncentrality parameter ψ^2 is given by

$$\psi^2 = m\sigma^{-2} \Sigma\, [\xi_{i\cdot(\cdot)} - \xi_{\cdot\cdot(\cdot)}]^2 = m\sigma^{-2} \sum_i \alpha_i^2 .$$

Remark.

For the practical use of latin squares see COCHRAN and COX (1957) or JOHN (1971).

A theoretical approach can be found in SCHEFFÉ (1959).

Section 6

Problem 16.

In this situation $Y_1,\ldots,Y_n$ are independently normally distributed with common variance $\sigma^2 = \sigma_v^2 + \beta^2\sigma_u^2$ and means $\xi_i = E(Y_i) = \alpha + \beta x_i$. For testing the hypothesis $H : \beta = \beta_0$ we have the rejection region (35):

$$\frac{|\hat{\beta} - \beta_0|\ \sqrt{\Sigma\,(x_i - \overline{x})^2}}{\sqrt{\Sigma\,(y_i - \hat{\alpha} - \hat{\beta}x_i)^2/(n-2)}} > C_0,$$

where $\hat{\beta} = [\Sigma\,(x_i - \overline{x})^2]^{-1}\,\Sigma\,(Y_i - \overline{Y})(x_i - \overline{x})$ and $\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{x}$.
For testing the hypothesis $H' : \alpha + \beta x_0 = \rho_0$ write $\xi_i = \alpha + \beta x_0 + \beta(x_i - x_0)$ and we have the rejection region (34):

$$\frac{|\hat{\rho} - \rho_0|\,\{n\,\Sigma\,(x_i - \overline{x})^2/\,\Sigma\,(x_i - x\ \ )^2\}^{\frac{1}{2}}}{\{\Sigma\,[Y_i - \hat{\rho} - \hat{\beta}(x_i - x_0)]^2\,/\,(n-2)\}^{\frac{1}{2}}} > C_1,$$

where $\hat{\beta} = [\Sigma\,(x_i - \overline{x})^2]^{-1}\,\Sigma\,(Y_i - \overline{Y})(x_i - \overline{x})$ and $\hat{\rho} = \overline{y} - \hat{\beta}(\overline{x} - x_0)$.
Each test statistic is distributed under the corresponding null hypothesis as the absolute value of Student's t with n-2 degrees of freedom.

(SCHEFFÉ (1958)).

Problem 17.

(i) The hypothesis $H : \beta = \delta$ is a linear hypothesis with $r = 1$ and $s = 4$. By Section 6 the minimum value of

$$(7) \qquad \sum_i\,[X_i - \alpha - \beta(u_i - \overline{u})]^2 + \sum_j\,[Y_j - \gamma - \delta(v_j - \overline{v})]^2$$

is attained at $\hat{\alpha} = \overline{X}$, $\hat{\beta} = [\Sigma_i\ (u_i - \overline{u})^2]^{-1}\,\Sigma_i\,(X_i - \overline{X})(u_i - \overline{u})$, $\hat{\gamma} = \overline{Y}$ and $\hat{\delta} = [\Sigma_j\ (v_j - \overline{v})^2]^{-1}\,\Sigma_j\,(Y_j - \overline{Y})(v_j - \overline{v})$.

Define

$$U = \sum_i (u_i - \overline{u})^2, \; V = \sum_j (v_j - \overline{v})^2, \; S_{xu} = \sum_i (X_i - \overline{X})(u_i - \overline{u}) \text{ and}$$

$$S_{yv} = \sum_j (Y_j - \overline{Y})(v_j - \overline{v}),$$

then

$$\hat{\beta} = S_{xu}U^{-1} \quad \text{and} \quad \hat{\delta} = S_{yv}V^{-1}.$$

To minimize (7) under H, first minimize over all possible values of α and γ for fixed β. This yields a minimum

$$\text{(8)} \qquad \sum_i [X_i - \overline{X} - \beta(u_i - \overline{u})]^2 + \sum_j [Y_j - \overline{Y} - \beta(v_j - \overline{v})]^2 =$$

$$= \beta^2(U + V) - 2\beta(S_{xu} + S_{yv}) + \sum_i (X_i - \overline{X})^2 + \sum_j (Y_j - \overline{Y})^2$$

at $\hat{\hat{\alpha}} = \overline{X} = \hat{\alpha}$ and $\hat{\hat{\gamma}} = \overline{Y} = \hat{\gamma}$.
Minimizing (8) over all possible values of β (= δ) we find

$$\hat{\hat{\beta}} \, (= \hat{\hat{\delta}}) = (S_{xu} + S_{yv})(U + V)^{-1} = \frac{U\hat{\beta} + V\hat{\delta}}{U + V}$$

Now the UMP invariant test is given by the rejection region $W = A/B > C$, where A is equal to

$$\sum_i \{[\hat{\alpha} + \hat{\beta}(u_i - \overline{u})] - [\hat{\hat{\alpha}} + \hat{\hat{\beta}}(u_i - \overline{u})]\}^2 +$$

$$+ \sum_j \{[\hat{\gamma} + \hat{\delta}(v_j - \overline{v})] - [\hat{\hat{\gamma}} + \hat{\hat{\delta}}(v_j - \overline{v})]\}^2 = (\hat{\beta} - \hat{\hat{\beta}})^2 U + (\hat{\delta} - \hat{\hat{\beta}})^2 V.$$

Observing

$$(\hat{\beta} - \hat{\hat{\beta}})^2 = \left[\hat{\beta} - \frac{U\hat{\beta} + V\hat{\delta}}{U + V}\right]^2 = \left[\frac{V(\hat{\beta} - \hat{\delta})}{U + V}\right]^2 = \frac{V^2(\hat{\beta} - \hat{\delta})^2}{(U + V)^2}$$

and by symmetry

$$(\hat{\delta} - \hat{\hat{\beta}})^2 = \frac{U \; (\hat{\beta} - \hat{\delta})^2}{(U + V)^2} \, ,$$

we have

$$A = (\hat\beta - \hat\delta)\,\frac{V^2U + U^2V}{(U+V)^2} = (\hat\beta - \hat\delta)^2\,\frac{UV}{(U+V)^2}\,(V+U) = (\hat\beta - \hat\delta)^2(U^{-1} + V^{-1}).$$

B equals

$$(9) \qquad (m+n-4)^{-1}\left\{\sum_i [X_i - \bar X - \hat\beta(u_i - \bar u)]^2 + \sum_j [Y_j - \bar Y - \hat\delta(v_j - \bar v)]^2\right\}.$$

W has the noncentral F-distribution with 1 and (m+n-4) degrees of freedom and noncentrality parameter $\psi^2 = \sigma^{-2}(\beta - \delta)^2(U^{-1} + V^{-1})$.

(ii) The hypothesis H' : $\alpha = \gamma$, $\beta = \delta$ is a linear hypothesis with $r = 2$ and $s = 4$.

The minimum of (7) under H', that is the minimum of

$$\sum_i [X_i - \alpha - \beta(u_i - \bar u)]^2 + \sum_j [Y_j - \alpha - \beta(v_j - \bar v)]^2$$

is attained at $\hat{\hat\alpha}\,(=\hat{\hat\gamma}) = (m+n)^{-1}[\Sigma_i X_i + \Sigma_j Y_j] = \frac{n}{m+n}\bar X + \frac{n}{m+n}\bar Y$ and

$$\tilde\beta\,(=\tilde\delta) = (U+V)^{-1}[\sum_i (X_i - \hat{\hat\alpha})(u_i - \bar u) + \sum_j (Y_j - \hat{\hat\gamma})(v_j - \bar v)] =$$

$$= (U+V)^{-1}[\sum_i (X_i - \bar X)(u_i - \bar u) + \sum_j (Y_j - \bar Y)(v_j - \bar v)] = \hat{\hat\beta}\,(=\hat{\hat\delta}),$$

since $\Sigma_i (X_i - \hat{\hat\alpha})(u_i - \bar u) = \Sigma_i [(X_i - \bar X) + (\bar X - \hat{\hat\alpha})](u_i - \bar u) =$
$= \Sigma_i (X_i - \bar X)(u_i - \bar u) + (\bar X - \hat{\hat\alpha})\Sigma_i (u_i - \bar u) = \Sigma_i (X_i - \bar X)(u_i - \bar u)$ and of course $\Sigma_j (Y_j - \hat{\hat\gamma})(v_j - \bar v) = \Sigma_j (Y_j - \bar Y)(v_j - \bar v)$.

The UMP invariant test is given by the rejection region $W' = P/Q > C'$, where

$$2P = \sum_i \{[\hat\alpha + \hat\beta(u_i - \bar u)] - [\hat{\hat\alpha} - \hat{\hat\beta}(u_i - \bar u)]\}^2 +$$

$$+ \sum_j \{[\hat\gamma + \hat\delta(v_j - \bar v)] - [\hat{\hat\alpha} - \hat{\hat\beta}(v_j - \bar v)]\}^2$$

and Q = B (9).

Observing

$$\sum_i \left\{\left[\hat\alpha + \hat\beta(u_i - \bar u)\right] - \left[\hat{\hat\alpha} + \hat{\hat\beta}(u_i - \bar u)\right]\right\}^2$$

$$= \sum_{i=1}^{m} \left[\frac{n}{m+n}(\bar X - \bar Y) + (\hat\beta - \hat{\hat\beta})(u_i - \bar u)\right]^2 =$$

$$= \frac{mn^2}{(m+n)^2}(\overline{X}-\overline{Y})^2 + (\hat{\beta} - \hat{\hat{\beta}})^2 U$$

and analogously

$$\sum_j \{[\hat{\gamma} + \hat{\delta}(v_j - \overline{v})] - [\hat{\hat{\gamma}} + \hat{\hat{\delta}}(v_j - \overline{v})]\}^2 =$$

$$= \frac{m^2 n}{(m+n)^2}(\overline{X}-\overline{Y})^2 + (\hat{\delta} - \hat{\hat{\delta}})^2 V$$

it follows that

$$2P = (\overline{X}-\overline{Y})^2 \cdot \frac{mn}{(m+n)^2}(m+n) + (\hat{\beta}-\hat{\hat{\beta}})^2 U + (\hat{\delta}-\hat{\hat{\delta}})^2 V =$$

$$= \frac{mn}{m+n}(\overline{X}-\overline{Y})^2 + A = (\hat{\alpha}-\hat{\gamma})^2(m^{-1}+n^{-1}) + (\hat{\beta}-\hat{\delta})^2(U^{-1}+V^{-1})$$

W' has the noncentral F-distribution with 2 and (m+n-4) degrees of freedom and noncentrality parameter $\lambda^2 = \sigma^{-2}[(\alpha-\gamma)^2(m^{-1}+n^{-1}) + (\beta-\delta)^2(U^{-1}+V^{-1})]$.

Problem 18.

Differentiating $\Sigma\,(X_i - \alpha - \beta t_i - \gamma t_i^2)^2$ with respect to α, β and γ we see that the least-squares estimates must satisfy the system of equations given in the problem. If $(1,1,\ldots,1)'$, $(t_1,\ldots,t_n)'$ and $(t_1^2,\ldots,t_n^2)'$ are linearly independent, then the matrix

$$T = \begin{pmatrix} 1 & t_1 & t_1^2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & t_n & t_n^2 \end{pmatrix}$$

has rank 3. By Lemma 1, Chapter 7 this implies that

$$\overline{T} \overset{\text{def}}{=} \frac{1}{n}T'T = \begin{pmatrix} 1 & \overline{t} & \overline{t^2} \\ \overline{t} & \overline{t^2} & \overline{t^3} \\ \overline{t^2} & \overline{t^3} & \overline{t^4} \end{pmatrix}, \text{ where } \overline{t^k} = \frac{1}{n}\sum_{i=1}^{n} t_i^k,$$

has also rank 3 and hence is nonsingular.

It follows that the system of equations has a unique solution, i.e.

$$\begin{pmatrix} \hat\alpha \\ \hat\beta \\ \hat\gamma \end{pmatrix} = \bar{T}^{-1} \begin{pmatrix} n^{-1}\Sigma X_i \\ n^{-1}\Sigma t_i X_i \\ n^{-1}\Sigma t_i^2 X_i \end{pmatrix}$$

With $r = 1$, $s = 3$ it follows by Problem 6 that the UMP invariant test for $H : \gamma = 0$ rejects when

$$\frac{|\hat\gamma| / \sqrt{\Sigma c_i^2}}{\{\Sigma (x_i - \hat\alpha - \hat\beta t_i - \hat\gamma t_i^2)^2/(n-3)\}^{\frac{1}{2}}} > C_0 .$$

Section 7

Problem 19.

(i) The joint density of $Z_{11},\ldots,Z_{sn}$ is given by (4.2) on p. 288. Let $\tau^2 = \sigma^2 + n\sigma_A^2$ and $\Delta_0' = 1 + n\Delta_0$ and

$$\theta = \left(-\frac{1}{2\tau^2} + \frac{1}{2\Delta_0'\sigma^2}\right), \ \vartheta_1 = \frac{\sqrt{sn}\,\mu}{\tau^2}, \ \vartheta_2 = -\frac{1}{2\sigma^2}$$

and

$$U = \sum_{i=1}^{s} Z_{i1}^2, \quad T_1 = Z_{11}, \quad T_2 = \sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2 + \frac{1}{\Delta_0'}\sum_{i=1}^{s} Z_{i1}^2 .$$

Then this density can be written as

$$C(\theta,\vartheta)\exp(\theta U + \vartheta_1 T_1 + \vartheta_2 T_2)$$

The hypothesis $H_1 : \sigma_A^2/\sigma^2 \le \Delta_0$ is equivalent to $H_1 : \theta \le 0$, so by Theorem 3 of Chapter 4 there exists a UMP unbiased test.
When $\theta = 0$ the distribution of the statistic

$$V = \frac{(U - T_1^2)/\Delta_0'}{T_2 - U/\Delta_0'} = \frac{\left(\sum_{i=2}^{s} Z_{i1}^2\right)\Big/\tau^2}{\left(\sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2\right)\Big/\sigma^2}$$

does not depend on μ, σ or τ and hence by Corollary 1 of Chapter 5 V is independent of (T_1,T_2). By Theorem 1 of Chapter 5 the UMP unbiased test has rejection region given by (43)

$$W^* = \frac{1}{\Delta_0'} \frac{\left(\sum_{i=2}^{s} Z_{i1}^2\right)\Big/(s-1)}{\left(\sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2\right)\Big/(n-1)s} > C.$$

This is (when $\theta = 0$) the ratio of two independent variables which are chi-square distributed with $s-1$ and $(n-1)s$ degrees of freedom, respectively each divided by its number of degrees of freedom; so the constant C can be determined by means of an F-distribution:

$$\int_C^\infty F_{s-1,(n-1)s}(y)dy = \alpha.$$

(ii) Let $W = \dfrac{(U - T_1^2)/\Delta_0'}{T_2 - T_1^2/\Delta_0'} = \dfrac{\left(\sum_{i=2}^{s} Z_{i1}^2\right)\Big/\tau^2}{\left(\sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2\right)\Big/\sigma^2 + \left(\sum_{i=2}^{s} Z_{i1}^2\right)\Big/\tau^2}$.

Under $H_2 : \Delta = \Delta_0$ which is equivalent to $H_2 : \theta = 0$ this statistic has a distribution that does not depend on μ, σ or τ and hence W is independent of (T_1,T_2). Furthermore it is linear in U. So by Theorem 1 of Chapter 5 the UMP unbiased acceptance region is given by $C_1 < W < C_2$.
The distribution of W is seen to be a beta-distribution with $\frac{1}{2}(s-1)$ and $\frac{1}{2}(n-1)s$ degrees of freedom. So C_1 and C_2 are determined by

$$\int_{C_1}^{C_2} B_{\frac{1}{2}(s-1),\frac{1}{2}(n-1)s}(y)dy = 1-\alpha,$$

$$\int_{C_1}^{C_2} yB_{\frac{1}{2}(s-1),\frac{1}{2}(n-1)s}(y)dy = (1-\alpha)\int_{\mathbb{R}} yB_{\frac{1}{2}(s-1),\frac{1}{2}(n-1)s}(y)dy.$$

The second equation here can equivalently be written as

$$\int_{C_1}^{C_2} B_{\frac{1}{2}(s+1),\frac{1}{2}(n-1)s}(y)dy = 1-\alpha.$$

The uniformly most accurate unbiased confidence sets for Δ can be derived from the acceptance region $C_1 < W < C_2$

$$\frac{1-C_2}{C_2}\,\frac{\sum_{i=2}^{s} Z_{i1}^2}{\sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2} < \frac{\tau^2}{\sigma^2} < \frac{1-C_1}{C_1}\,\frac{\sum_{i=2}^{s} Z_{i1}^2}{\sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2}$$

or, with $\Delta = \frac{\sigma_A^2}{\sigma^2} = \frac{1}{n}\left(\frac{\tau^2}{\sigma^2} - 1\right)$

$$\frac{1}{n}\left\{\frac{1-C_2}{C_2}\frac{\sum_{i=2}^{s} Z_{i1}^2}{\sum_{i=1}^{s}\sum_{j=2}^{n} Z_{ij}^2} - 1\right\} < \frac{\sigma_A^2}{\sigma^2} < \frac{1}{n}\left\{\frac{1-C_1}{C_1}\frac{\sum_{i=2}^{s} Z_{i1}^2}{\sum_{i=1}^{n}\sum_{j=2}^{n} Z_{ij}^2} - 1\right\}.$$

Problem 20.

Because $E(A_i) = E(U_{ij}) = E(U_{ik}) = 0$ the intraclass correlation coefficient $\rho(X_{ij},X_{ik})$ equals

$$\frac{E(A_i + U_{ij})(A_i + U_{ik})}{\sqrt{\operatorname{Var}(X_{ij})\cdot\operatorname{Var}(X_{ik})}} .$$

Since the A's and the U's are independent we have

$$\operatorname{Var}(X_{ij}) = \operatorname{Var}(X_{ik}) = \sigma^2 + \sigma_A^2$$

and

$$E(A_i + U_{ij})(A_i + U_{ik}) = \sigma_A^2.$$

Hence

$$\rho = \frac{\sigma_A^2}{\sigma^2 + \sigma_A^2} .$$

Section 8

Problem 21.

Except for a constant the joint distribution of Z_{ijk} ($i = 1,\ldots,a$; $j = 1,\ldots,b$; $k = 1,\ldots,n$) is given by (47) on p. 290. This density can be written in the following form

$$C(\theta,\vartheta)\exp\{\theta S_1^2(1+bn\Delta_0)^{-1}+\vartheta_1 Z_{111}+\vartheta_2(S_B^2 + S_1^2(1+bn\Delta_0)^{-1}) + \vartheta_3 S^2\}$$

where θ and $\vartheta = (\vartheta_1,\vartheta_2,\vartheta_3)$ are given by

$$\theta = -\tfrac{1}{2}(1+bn\Delta_0)(\sigma^2+n\sigma_B^2+bn\sigma_A^2)^{-1} + \tfrac{1}{2}(\sigma^2+n\sigma_B^2)^{-1}$$

$$\vartheta_1 = (abn)^{\frac{1}{2}}\mu(\sigma^2+n\sigma_B^2+bn\sigma_A^2)^{-1},$$

$$\vartheta_2 = -\tfrac{1}{2}(\sigma^2+n\sigma_B^2)^{-1}, \qquad \vartheta_3 = -\tfrac{1}{2}\sigma^{-2},$$

$C(\theta,\vartheta)$ is a normalizing constant and

$$S_1^2 = \sum_{i=1}^{a} Z_{i11}^2 = Z_{111}^2 + S_A^2, \quad S_B^2 = \sum_{i=1}^{a}\sum_{j=2}^{b} Z_{ij1}^2,$$

$$S^2 = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{k=2}^{n} Z_{ijk}^2.$$

The testing problem $H_1 : \sigma_A^2(\sigma^2+n\sigma_B^2)^{-1} \le \Delta_0$ against $K : \sigma_A^2(\sigma^2+n\sigma_B^2)^{-1} > \Delta_0$ is equivalent to $H_1 : \theta \le 0$ against $K_1 : \theta > 0$

Defining

$$W_1^* = h(S_1^2(1+bn\Delta_0)^{-1}, Z_{111}, S_B^2 + S_1^2(1+bn\Delta_0)^{-1}, S^2)$$
$$= \frac{1}{1+bn\Delta_0}\,\frac{(S_1^2 - Z_{111}^2)\,/\,(a-1)}{S_B^2\Big/(b-1)a} = \frac{1}{1+bn\Delta_0}\,\frac{S_A^2\Big/(a-1)}{S_B^2\Big/(b-1)a},$$

it is immediately seen that h is strictly increasing in the first argument. If $\theta = 0$, that is $\sigma_A^2 = \Delta_0(\sigma^2+n\sigma_B^2)$, then

$$W_1^* = \frac{S_A^2\Big/(\sigma^2+n\sigma_B^2+bn\sigma_A^2)(a-1)}{S_B^2\Big/(\sigma^2+n\sigma_B^2)(b-1)a}$$

is distributed as $F_{a-1,(b-1)a}$ and hence in particular the distribution of V does not depend on ϑ. Moreover $\{\vartheta : (0,\vartheta) \in \Omega\}$ (with Ω the whole parameter space) contains a 3-dimensional rectangle. Therefore by Corollary 1 on p. 162 W_1 is independent of $T = (Z_{111}^2, S_B^2 + S_1^2(1+bn\Delta_0)^{-1}, S^2)$ and it follows by Theorem 1 on p. 161 that the test (48) is UMP unbiased.

The density (47) on p. 290 can also be written as

$$K(\theta,\vartheta)\exp\{\theta_1 S_B^2(1+n\Delta_0)^{-1} + \vartheta_1 S_1^2 + \vartheta_2 Z_{111} + \vartheta_3(S^2 + S_B^2(1+n\Delta_0)^{-1})\}$$

where θ and $\vartheta = (\vartheta_1,\vartheta_2,\vartheta_3)$ are now given by

$$\theta = -\tfrac{1}{2}(1+n\Delta_0)(\sigma^2+n\sigma_B^2)^{-1} + \tfrac{1}{2}\sigma^{-2}$$

$$\vartheta_1 = -\tfrac{1}{2}(\sigma^2 + n\sigma_B^2 + nb\sigma_A^2)^{-1},$$
$$\vartheta_2 = (abn)^{\frac{1}{2}}\mu(\sigma^2 + n\sigma_B^2 + nb\sigma_A^2)^{-1}, \quad \vartheta_3 = -\tfrac{1}{2}\sigma^{-2}.$$

$K(\theta,\vartheta)$ is a normalizing constant and S_1^2, S_B^2, S^2 are the same as above. The testing problem $H_2 : \sigma_B^2\sigma^{-2} \le \Delta_0$ against $K_2 : \sigma_B^2\sigma^{-2} > \Delta_0$ is equivalent to $H_2 : \theta \le 0$ against $K_2 : \theta > 0$.
Defining

$$W_2^* = h(S_B^2(1+n\Delta_0)^{-1},\ S_1^2,\ Z_{111}^2,\ S^2 + S_B^2(1+n\Delta_0)^{-1})$$
$$= \frac{1}{1+n\Delta_0}\,\frac{S_B^2\big/(b-1)a}{S^2/(n-1)ab}$$

it is immediately seen that h is strictly increasing in the first argument. If $\theta = 0$, that is $\sigma_B^2 = \Delta_0\sigma^2$, then

$$W_2^* = \frac{S_B^2\big/(\sigma^2+n\sigma_B^2)(b-1)a}{S^2/\sigma^2(n-1)ab}$$

is distributed as $F_{(b-1)a,(n-1)ab}$ and hence the distribution of W_2 does not depend on ϑ. Moreover $\{\vartheta;(0,\vartheta) \in \Omega\}$ contains a 3-dimensional rectangle. Therefore by Corollary 1 on p. 162 W_2^* is independent of $T = (S_1^2, Z_{111}^2, S^2+S_B^2(1+n\Delta_0)^{-1})$ and it follows by Theorem 1 on p. 161 that the test (49) is UMP unbiased.

Problem 22.

If we put $(X_{i1},\ldots,X_{in_i}) = \underline{X}_i$, $i = 1,2,\ldots,s$, then the $\underline{X}_i$ are independent by assumption.
Each $\underline{X}_i$, $1 \le i \le s$ is subjected to an orthogonal transformation

$$\underline{Y}_i = (Y_{i1},Y_{i2},\ldots,Y_{in_i}) = \underline{X}_i\begin{pmatrix} \frac{1}{\sqrt{n_i}} & c_{12}^{(i)} & \cdots & c_{1n_i}^{(i)} \\ \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{n_i}} & c_{n_i2}^{(i)} & \cdots & c_{n_in_i}^{(i)} \end{pmatrix} = \underline{X}_i C^{(i)}$$

so that $Y_{i1} = \sqrt{n_i}\,X_{i.} = \sqrt{n_i}\,(\mu + A_i + U_{i.})$.
Furthermore for $2 \le j \le n_i$, $1 \le i \le s$

$$Y_{ij} = \sum_{k=1}^{n_i} C_{kj}^{(i)} X_{ik} = \sum_{k=1}^{n_i} C_{kj}^{(i)} (\mu + A_i + U_{ik}) = \sum_{k=1}^{n_i} C_{kj}^{(i)} U_{ik}$$

since by orthogonality $\sum_{k=1}^{n_i} C_{kj}^{(i)} = 0,\ 2 \le j \le n_i,\ 1 \le i \le s.$
Hence the Y_{ij} with $2 \le j \le n_i$, $1 \le i \le s$ are independently normally distributed with zero mean and variance σ^2. They are also independent of $U_{i.}$ since $(\sqrt{n_i}\,U_{i.}, Y_{i2}, \ldots, Y_{in})' = C^{(i)}(U_{i1}U_{i2}\ldots U_{in})'$ and thus independent of Y_{i1}, $1 \le i \le s$. Since the $\underline{X}_i$ are independent, the variables $\underline{Y}_{i1}$, $1 \le i \le s$ are independent with mean $\sqrt{n_i}\,\mu$ and variance $\sigma^2 + n_i\sigma_A^2$.
The joint density function of the Y's is then

$$\prod_{i=1}^{s} \frac{1}{(2\pi)^{n_i/2}} \sigma^{-(n_i-1)} (\sigma^2 + n_i\sigma_A^2)^{-\frac{1}{2}}$$

$$\exp\left\{- \sum_{i=1}^{s} \frac{(Y_{i1} - \sqrt{n_i}\,\mu)^2}{2(\sigma^2 + n_i\sigma_A^2)} - \sum_{i=1}^{s} \sum_{j=2}^{n_i} \frac{Y_{ij}^2}{2\sigma^2}\right\}.$$

<u>Problem 23.</u>

(i) First of all we establish a more convenient notation for the nested classification with a constant number of observations per cell. We write for m factors

$$X_{i_1 i_2 \ldots i_m i_{m+1}} = \mu + F_{i_1}^{(1)} + F_{i_1 i_2}^{(2)} + \ldots + F_{i_1 i_2 \ldots i_m}^{(m)} + U_{i_1 i_2 \ldots i_m i_{m+1}}$$

with $i_j = 1, \ldots, \nu_j$ for $j = 1, \ldots, m+1$. (I.e. ν_{m+1} observations per cell.)
All these variable are assumed to be independently normally distributed with zero means and with variances $\sigma_1^2, \sigma_1^2, \ldots, \sigma_m^2$ and σ^2, respectively.
We now proceed to prove the existence of an orthogonal transformation to variables $Z_{i_1 i_2 \ldots i_m i_{m+1}}$, the joint density of which, except for a constant, is equal to

$$(10) \qquad \exp\left\{-\frac{1}{2(\sigma^2+\nu_{m+1}\sigma_m^2+\nu_{m+1}\nu_m\sigma_{m-1}^2+\ldots+\nu_{m+1}\nu_m\ldots\nu_2\sigma_1^2)} \times \right.$$

$$\times \left((Z_{11\ldots1} - \sqrt{\nu_{m+1}\nu_m\ldots\nu_1}\,\mu)^2 + \sum_{i_1=2}^{\nu_1} Z_{i_1 11\ldots 1} \right) +$$

$$- \frac{1}{2(\sigma^2+\nu_{m+1}\sigma_m^2+\nu_{m+1}\nu_m\sigma_{m-1}^2+\ldots+\nu_{m+1}\nu_m\ldots\nu_3\sigma_2^2)} \times$$

$$\times \sum_{i_1=1}^{\nu_1} \sum_{i_2=2}^{\nu_2} Z^2_{i_1 i_2 1 \ldots 1}$$

$$- \frac{1}{2(\sigma^2+\nu_{m+1}\sigma^2+\nu_{m+1}\nu_m\sigma^2_{m-1}+\ldots+\nu_{m+1}\nu_m\ldots\nu_4\sigma^2_3)} \times$$

$$\times \sum_{i_1=1}^{\nu_1} \sum_{i_2=1}^{\nu_2} \sum_{i_3=2}^{\nu_3} Z^2_{i_1 i_2 i_3 1 \ldots 1} \quad + \ldots +$$

$$- \frac{1}{(\sigma^2+\nu_{m+1}\sigma^2_m)} \sum_{i_1=1}^{\nu_1} \sum_{i_2=1}^{\nu_2} \ldots \sum_{i_m=2}^{\nu_m} Z^2_{i_1 i_2 \ldots i_m 1} \; +$$

$$- \frac{1}{2\sigma^2} \sum_{i_1=1}^{\nu_1} \sum_{i_2=1}^{\nu_2} \ldots \sum_{i_{m+1}=2}^{\nu_{m+1}} Z^2_{i_1 i_2 \ldots i_{m+1}} \Big\} .$$

The number of indices of the Z's is m+1 in this formula.
The proof is by induction with respect to the number of factors. For $m = 1$ and $m = 2$, the existence of the orthogonal transformation is proved by Lehmann in Section 7.7 and 7.8, respectively. So consider the (m+1)-way nested classification,

$$X_{i_1 i_2 \ldots i_{m+1} i_{m+2}} =$$
$$\mu + F^{(1)}_{i_1} + F^{(2)}_{i_1 i_2} + \ldots + F^{(m+1)}_{i_1 i_2 \ldots i_{m+1}} + U_{i_1 i_2 \ldots i_{m+1} i_{m+2}} .$$

For fixed $i_1, i_2, \ldots, i_{m+1}$, consider the ν_{m+2} independent variables $X_{i_1 i_2 \ldots i_{m+1} 1}, \ldots, X_{i_1 i_2 \ldots i_{m+1} \nu_{m+2}}$.
In the usual way there exists an orthogonal transformation to variables $Y_{i_1 i_2 \ldots i_{m+1} 1}, \ldots, Y_{i_1 i_2 \ldots i_{m+1} \nu_{m+2}}$ such that

$$Y_{i_1 i_2 \ldots i_{m+1} 1} = \sqrt{\nu_{m+2}}\, X_{i_1 i_2 \ldots i_{m+1} \cdot} =$$
$$= \sqrt{\nu_{m+2}}\, \mu + \sqrt{\nu_{m+2}} \left(F^{(1)}_{i_1} + \ldots + F^{(m+1)}_{i_1 i_2 \ldots i_{m+1}} + U_{i_1 i_2 \ldots i_{m+1} \cdot} \right)$$

and the $Y_{i_1 i_2 \ldots i_{m+1} i_{m+2}}$ are independent variables with zero expectation and variance σ for $i_{m+2} > 1$.
On the other hand, the variables $Y_{i_1 i_2 \ldots i_{m+2} 1}$ have exactly the structure of the $X_{i_1 i_2 \ldots i_{m+1}}$ in the m-way nested classification, i.e.

$$Y_{i_1 i_2 \ldots i_{m+1} 1} = \xi + G^{(1)}_{i_1} + G^{(2)}_{i_1 i_2} + \ldots + G^{(m)}_{i_1 i_2 \ldots i_m} + V_{i_1 i_2 \ldots i_m i_{m+1}}$$

$$\xi = \sqrt{\nu_{m+2}}\,\mu;$$

$$G^{(j)}_{i_1 i_2 \dots i_j} = \sqrt{\nu_{m+2}}\,F^{(j)}_{i_1 i_2 \dots i_j} \qquad j = 1,\dots,m;$$

$$V_{i_1 i_2 \dots i_m i_{m+1}} = \sqrt{\nu_{m+2}}\left(F^{(m+1)}_{i_1 i_2 \dots i_{m+1}} + U_{i_1 i_2 \dots i_{m+1}\cdot}\right).$$

The expectations of the variables G and V are zero and the variances $\tau_1^2,\tau_2^2,\dots,\tau_m^2$ of $G^{(1)}_{i_1},\dots,G^{(m)}_{i_1 i_2 \dots i_m}$ are equal to $\tau_j^2 = \nu_{m+2}\sigma_j^2$, $j = 1,\dots,m$ and the variance τ of $V_{i_1 i_2 \dots i_m i_{m+1}}$ is equal to $\tau^2 = \sigma^2 + \nu_{m+2}\sigma^2_{m+1}$. Assuming that the induction hypothesis (the existence of an orthogonal transformation leading to (10) holds in the m-way nested classification, it follows that the variables $Y_{i_1 i_2 \dots i_{m+1} 1}$ may further be transformed by an orthogonal transformation to variables $Z_{i_1 i_2 \dots i_{m+1} 1}$, the joint density of which, except for a constant, is given by (10), provided that the proper replacement of variable and parameters has been made.
After completion of the transformation by putting

$$Z_{i_1 i_2 \dots i_{m+1} i_{m+2}} = Y_{i_1 i_2 \dots i_{m+1} i_{m+2}} \quad \text{for } i_{m+2} > 1,$$

we find that the joint distribution of the $Z_{i_1 i_2 \dots i_{m+1} i_{m+2}}$ is given by the equivalent of (10) for the (m+1)-way nested classification.
Herewith, the existence of the orthogonal transformation leading to (10) has been proved.

(ii) Two hypotheses of interest can now be tested on the basis of (*)

$$H_1: \qquad \sigma_1^2 \Big/ \left(\sigma^2 + \nu_{m+1}\sigma_m^2 + \nu_{m+1}\nu_m\sigma_{m-1}^2 + \dots + \nu_{m+1}\nu_m \cdots \nu_3\sigma_2^2\right) \le \Delta_0$$

$$H_2: \qquad \sigma_2^2 \Big/ \left(\sigma^2 + \nu_{m+1}\sigma_m^2 + \nu_{m+1}\nu_m\sigma_{m-1}^2 + \dots + \nu_{m+1}\nu_m \cdots \nu_4\sigma_3^2\right) \le \Delta_0$$

Let

$$S_j^2 = \sum_{i_1=1}^{\nu_1} \sum_{i_2=1}^{\nu_2} \dots \sum_{i_j=2}^{\nu_j} Z^2_{i_1 i_2 \dots i_j} \qquad j = 1,\dots,m$$

and

$$S^2 = \sum_{i_1=1}^{\nu_1} \sum_{i_2=1}^{\nu_2} \dots \sum_{i_{m+1}=2}^{\nu_{m+1}} Z^2_{i_1 i_2 \dots i_{m+1}}.$$

Following the same reasoning as in Section 8 for the case of two factors, but now conditioning on the event $S_3^2 = s_3^2 \wedge \ldots \wedge S_m^2 = s_m^2 \wedge S_2 = s^2$ we obtain the following (conditional) tests:
for H_1 we have the rejection region

$$W_1^* = \frac{1}{1+\nu_2\nu_3\Delta_0} \cdot \frac{S_1^2 / (\nu_1-1)}{S_2^2 / (\nu_2-1)\nu_1} \geq C_1 ;$$

for H_2 we obtain

$$W_2^* = \frac{1}{1+\nu_3\Delta_0} \cdot \frac{S_2^2 / (\nu_2-1)\nu_1}{S_3^2 / (\nu_3-1)\nu_2\nu_1} \geq C_2 .$$

These tests are in fact the same as the tests (48) and (49) on p. 291 because H_1 and H_2 only concern the first two factors in the nested classification.

Section 10

Problem 24.

First we prove the following lemma.

LEMMA. *Let* $Y_1,\ldots,Y_m$ *be independent* p*-dimensional random vectors whose probability distributions are absolutely continuous with respect to Lebesgue measure on* $\mathbb{R}^p$. *Then*

$$(11) \qquad P[Y_1,\ldots,Y_m \text{ are linearly independent}] = \begin{cases} 1 \text{ when } m \leq p \\ 0 \text{ when } m > p. \end{cases}$$

PROOF. For $m > p$ and for $m = 1$ (11) trivially holds. Let $2 \leq m \leq p$. Define

$$R_i(y_1,\ldots,y_m) = R_i(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_m) = \left\{\sum_{\substack{j=1\\j\neq i}}^{m} \alpha_j y_j : \alpha_j \in \mathbb{R}\right\},$$

that is, the space spanned by the vectors $y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_m$. Now for any $y_1,\ldots,y_m$, $R_i(y_1,\ldots,y_m)$ has dimension $\leq m-1 < p$. Hence its Lebesgue measure is zero. Thus

$$P(Y_i \in R_i(Y_1,\ldots,Y_m) \mid Y_1=y_1,\ldots,Y_{i-1}=y_{i-1},Y_{i+1}=y_{i+1},\ldots,Y_m=y_m)$$

$$= P(Y_i \in R_i(y_1,\ldots,y_m)) = 0$$

for any $y_1,\dots,y_m$ and therefore $P[Y_i \in R_i(Y_1,\dots,Y_m)] = 0$ too. So

$$P[Y_1,\dots,Y_m \text{ are linearly independent}] \geq 1 - \sum_{i=1}^{m} P(Y_i \in R_i(Y_1,\dots,Y_m))$$
$$= 1.$$

Throughout this problem $\theta = (\xi, A^{-1})$.

i) Because
$$E_\theta(\tfrac{1}{m} S) = E_\theta\left(\frac{Z'Z}{m}\right) = \frac{E_\theta(\Sigma_{k=1}^{m} Z_{ki}Z_{kj})_{ij}}{m} = \frac{\Sigma_{k=1}^{m} E_\theta(Z_{ki}Z_{kj})_{ij}}{m} = (\sigma_{ij}^2),$$

$\frac{1}{m}S$ is an unbiased estimate of A^{-1}.
By Lemma 1 (ii): δ is nonsingular if and only if rank $Z = p$ and $m \geq p$.
If $m \geq p$, $P_\theta[\delta$ is nonsingular$] = P_\theta[\text{rank } Z = p] \geq$
$\geq P_\theta[Z_1,\dots,Z_p$ are linearly independent$] = 1$ for every θ by the lemma in this problem.
If $m < p$ then by Lemma 1 (ii) δ is singular. This completes the proof.

ii) The U's are eliminated through G_1. Since the r+m row vectors of the matrices Y and Z may be assumed to be linearly independent, any such set of vectors can be transformed into any other through an element of G_3. Hence the Y's and Z's are eliminated. The only test that is invariant under the groups G_1 and G_3 is $\varphi(Y,U,Z) = \alpha$.

Problem 25.

First we prove the following extension of the lemma in Problem 24.

LEMMA. *Let* $\mathcal{P} = \{P_\theta : \theta \in \Omega\}$ *be a class of absolutely continuous (w.r.t. Lebesgue-measure) distributions on* $\mathbb{R}^{p(r+m)}$. *If* $\mathcal{P}$ *is a class of joint distributions of (the elements of)* $Y^{(r\times p)}$ *and* $Z^{(m\times p)}$, *then for any* $\mu \neq 0$

$$|Y'Y + \mu Z'Z| \neq 0, \; \mathcal{P}\text{-a.c.},$$

provided $p \leq r+m$.

PROOF. The following result is used:

(*) if $f(x_1,\dots,x_n)$ is a polynomial in real variables $x_1,\dots,x_n$ which is not identically zero then the subset $N = \{(X_1,\ldots,x_n) : f(x_1,\ldots,x_n) = 0\}$ of $\mathbb{R}^n$ has Lebesque-measure zero.

(cf. OKAMOTO (1973)).

Let $\mu \neq 0$ be fixed. Define $f_\mu(Y,Z) = |Y'Y + \mu Z'Z|$. Because the elements of $Y'Y$ and $Z'Z$ are polynomials in the elements of Y and Z, respectively, f_μ is a polynomial in the elements of Y and Z. By (*) and absolute continuity it suffices to show that f_μ is not identically zero. Because $p \leq r+m$ there exist Y_0 and Z_0 such that

$$\underset{p}{\begin{bmatrix} Y_0 \\ Z_0 \end{bmatrix}}\begin{matrix} r \\ m \end{matrix} = \underset{p}{\begin{bmatrix} I_p \\ 0 \end{bmatrix}}\begin{matrix} p \\ m+r-p \end{matrix}$$

Then $f_\mu(Y_0,Z_0) = \begin{cases} 1 & \text{when } p \leq r \\ \mu^{p-r} & \text{when } p > r. \end{cases}$

Hence $f_\mu(Y_0,Z_0) \neq 0$, which completes the proof.

Note. The result follows easily from the lemma in Problem 24 when $\mu > 0$. However for the present problem it is essential that we also consider the case $\mu < 0$.

Now let P be the class of joint normal distributions of (Y,Z).

(i) If $p < r+m$ and $V = Y'Y$, $S = Z'Z$, the lemma implies that $V + S$ is nonsingular, P-a.e.

Consider the roots $\lambda_1(V,S),\ldots,\lambda_p(V,S)$ of the equation (94) on p. 318

$$|V - \lambda(V + S)| = 0.$$

We will show that they constitute a maximal set of invariants w.r.t. the groups generated by G_1,G_2 and G_3. Since (V,S) is a maximal invariant w.r.t. the group generated by G_1 and G_3 (p. 297), it suffices to show, by Chapter 6, Theorem 2, that $\lambda_1(V,S),\ldots,\lambda_p(V,S)$ is a maximal set of invariants w.r.t. the group $G_3^* = \{g^* : g^*(V,S) = B'VB,B'SB)$, B nonsingular$\}$. Since $B'VB + B'SB = B'(V + S)B$ this can be shown in the same way as on p. 298, which completes the proof.

(ii) In the same way as on p. 299, first paragraph we find that $p-\min(p,r)$ roots of (94) equal zero (P-a.e.). As on p. 298 there exists a nonsingular matrix B such that $B'VB = \Lambda$ and $B'(V + S)B = I$, where Λ is a diagonal matrix whose elements are the roots of (94) and I is the identity matrix. Hence $B'SB = I - \Lambda$. Thus the multiplicity of the root $\lambda = 1$ is equal to $p - \mathrm{rank}(S) = p - \min(m,p)$ ($\mathcal{P}$-a.e.), by Lemma 1 and the lemma in Problem 24.

Applying the lemma in this problem we have, for any constant $\lambda \neq 0,1$, $|V - \lambda(V+S)| = (1-\lambda)^p \left|V - \frac{\lambda}{1-\lambda}S\right| \neq 0$ ($\mathcal{P}$-a.e.). Hence there are no other ($\mathcal{P}$-a.e.) constant roots, so that the number of ($\mathcal{P}$-a.e.) variable roots, which constitute a maximal invariant set, is $p - (p - \min(r,p)) + -(p - \min(m,p)) = \min(r,p + \min(m,p) - p$.

(OKAMOTO, (1973)).

Problem 26.

i) If x is a non zero solution of the equation $ABx = \lambda x$ with $\lambda \neq 0$, then $y = Bx$ is a non zero solution of $BAy = \lambda y$.

ii) Applying i) to the $p \times 1$ matrix Y' and the $1 \times p$ matrix $Y \cdot S^{-1}$ then $VS^{-1} = Y'YS^{-1}$ has the same characteristic non zero roots as $YS^{-1}Y'$. The only non zero root of the 1×1 matrix $YS^{-1}Y'$ is $W = \Sigma_{i=1}^{p} \Sigma_{j=1}^{p} S^{ij} Y_i Y_j$.

Problem 27.

The assertion can be proved according to the argument given in Section 10 with the invariance under G_2 omitten. However, this argument uses "the theory of the simultaneous reduction to diagonal form of two quadratic forms" in order to show that there exists a nonsingular matrix B such that $B'SB = I$ and $B'Y'YB$ is of diagonal form with rank $\min(p,r)$. For $r = 1$ this can also be seen as follows. There exists a nonsingular matrix B_1 such that $B_1'SB_1 = I$. Then YB_1 is a row vector and there exists an orthogonal matrix Q such that only the first coordinalte of YB_1Q is nonzero. With $B = B_1Q$ we now have $B'SB = Q'B_1'SB_1Q = Q'IQ = I$ and only the upper left element of $B'Y'YB = (YB_1Q)'(YB_1Q)$ is nonzero.

Problem 28.

Let $Z = (Z_{11},\ldots,Z_{m1},Z_{12},\ldots,Z_{m2},\ldots,Z_{mp})$ and $Z^* = (Z^*_{11},\ldots,Z^*_{m1},Z^*_{12},\ldots,Z^*_{m2},\ldots,Z^*_{mp})$. Then $Z^* = Z\tilde{Q}'$ where $\tilde{Q} = I_p \otimes Q$, the Kronecker product of the $p \times p$ identity matrix I_p and Q, is orthogonal. Given $Y = y$ the matrix $\tilde{Q}$ is a constant orthogonal matrix. By independence of Y and the Z's we can apply Problem 5.6 in order to prove that given

$Y = y$ the Z^*'s are independently distributed as $N(0,1)$. Since this conditional distribution does not depend on the value of y the result also holds for the unconditional distribution of the $Z^*_{\alpha i}$ and Z^* is independent of Y.

Problem 29.

The solution is given in the comments following the problem.
We only remark that the first row of Q has to be taken as

$$\left(\frac{Z_{11}}{R}, \frac{Z_{21}}{R}, \ldots, \frac{Z_{m1}}{R}\right)$$

in order to have $(Z_{11},\ldots,Z_{m1})Q' = (R,0,\ldots,0)$. The other rows can be determined by the classical Gram-Schmidt orthogonalisation procedure.
As to the last three lines of the comments we remark that, after stepwise reduction, the determinant of S is $|S| = R^2 \Sigma_{\alpha=1}^{m-p+1} \tilde{Z}^2_{\alpha 1}$ (with $\tilde{Z}_{\alpha 1}$, $\alpha = 1,\ldots,m-p+1$ independently distributed as $N(0,1)$) and $|S_1|$ is replaced by R^2.
Hence $\dfrac{|S|}{|S_1|} = \Sigma_{\alpha=1}^{m-p+1} \tilde{Z}^2_{\alpha 1}$, which has a $\chi^2 m-p+1$ distribution.

(WIJSMAN (1957))

Problem 30.

Let B be a nonsingular $p \times p$ matrix. Then $W = YS^{-1}Y' = Y(Z'Z)^{-1}Y' = YB((ZB)'ZB)^{-1}(YB)'$. We therefore may assume that the common covariance matrix of the vectors $Y = (Y_1,\ldots,Y_p),(Z_{11},\ldots,Z_{1p}),\ldots,(Z_{m1},\ldots,Z_{mp})$ is equal to the identity matrix.
Let Q be an orthogonal $p \times p$ matrix (depending on the Y's) such that

$$(Y_1,\ldots,Y_p)Q = (0,\ldots,0,T)$$

where $T^2 = \Sigma_{i=1}^{p} Y_i^2$. Since QQ' is the identity matrix one has

$$\begin{aligned} W &= (YQ)(Q'S^{-1}Q)(Q'Y') = (0,\ldots,0,T)(Q'S^{-1}Q)(0,\ldots,0,T)' \\ &= U_{pp}T^2 \end{aligned}$$

where U_{pp} is the element which lies in the p-th row and the p-th column of the matrix $Q'S^{-1}Q = (Q'S^{-1}Q)^{-1}$. Let $V = Q'Z'ZQ$. Then U_{pp} is equal to the ratio of determinants $|V_1| / |V|$, where V_1 is the matrix obtained by omitting the last row and column of V. Exchanging the role of m and p in Problem 28 it follows that the $Z^*_{\alpha i}$, defined by

$$(Z^*_{\alpha 1},\dots,Z^*_{\alpha p}) = (Z_{\alpha 1},\dots,Z_{\alpha p})Q$$

are independently N(0,1) distributed and independent of Y $(\alpha = 1,\dots,m;\ i = 1,\dots,p)$. Since $V = (Z^*)'Z^*$ Problem 29 implies that U^{-1}_{pp} is independent of the Y's (and hence independent of T^2) and distributed as χ^2_{m-p+1}.

Note that T^2 has to be read as $YB(YB)'$, where B satisfies $B'A^{-1}B = I$, that is $BB' = A$. Hence $ET^2 = \psi^2$. Since T^2 is a sum of p independent, normally distributed random variables each with variance 1, T^2 has a noncentral χ^2-distribution with p degrees of freedom and noncentrality parameter ψ^2.

(WIJSMAN (1957))

Problem 31.

For $i = 1,2,\dots,p$ we must minimize $\Sigma_\nu\ [X_{\nu i} - \alpha_i - \beta_i(u_\nu - u_.)]^2$ over all possible values of α_i and β_i.
Let i be fixed. Then the problem reduces to the regression problem that was studied in Section 7.6, Chapter 7. The Hypothesis $H : \beta_1 = \dots = \beta_p = 0$ obviously is a multivariate hypothesis with $r = 1$ and $s = 2$.
In the second paragraph of Section 7.6 is is shown that $\alpha_i = X_{.i}$ minimizes $\Sigma_\nu\ [X_{\nu i} - \alpha_i - \beta_i(u_\nu - u_.)]^2$ for every fixed value of β_i. Hence

$$\hat{\alpha}_i = \hat{\hat{\alpha}}_i = X_{.i}\ .$$

By (33) on p. 283

$$\hat{\beta}_i = \frac{\Sigma_\nu\ (u_\nu - u_.)(X_{\nu i} - X_{.i})}{\Sigma_\nu\ (u_\nu - u_.)^2}$$

holds.
P. 284 first formula yields $Y_i = \hat{\beta}_i\sqrt{\Sigma_\nu\ (u_\nu - u_.)^2}$.
Hence for all i and j $Y_iY_j = \hat{\alpha}_i\hat{\beta}_j\,\Sigma_\nu\,(u_\nu - u_.)^2$.
Finally

$$S_{ij} = \Sigma_\nu[X_{\nu i} - \hat{\alpha}_i - \hat{\beta}_i(u_\nu - u_.)][X_{\nu j} - \hat{\alpha}_i - \hat{\beta}_j(u_\nu - u_.)]$$

by (55) on p. 296.

Problem 32.

Note. Because we only use column vectors, there are some slight changes in notation.

We consider a sample $X^{(1)},X^{(2)},\dots,X^{(N)}$ from a p-variate normal distribution with covariance matrix Σ. Let $q < p$, $\max(q,p-q) \le N$.

Partition the matrix Σ as follows

(12) $$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where Σ_{11} is $q \times q$, $\Sigma_{12} = q \times (p-q)$, Σ_{21} is $(p-q) \times q$ and $\Sigma_{22}^{(p-q)\times(p-q)}$.

We test

(13) $$H : \Sigma_{12} = 0.$$

It is easy to see that the problem of testing H remains invariant under the transformations $X^* = X + B$, with B a p-vector of constants, and $X^* = CX$, where C is any nonsingular $p \times p$-matrix of the structure

(14) $$C = \begin{pmatrix} C_{11} & 0 \\ 0 & C_{22} \end{pmatrix},$$

where the order of the submatrices is as in (12).

[Let $X^* = CX + D$ be the result of a combination of both transformations. The covariance matrix Σ^* of X^* is equal to $\Sigma^* = C\Sigma C'$ which is equal to

$$\Sigma^* = \begin{pmatrix} C_{11} & 0 \\ 0 & C_{22} \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} C'_{11} & 0 \\ 0 & C'_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} C_{11}\Sigma_{11}C'_{11} & C_{11}\Sigma_{12}C'_{22} \\ C_{22}\Sigma_{21}C'_{11} & C_{22}\Sigma_{22}C'_{22} \end{pmatrix} = \begin{pmatrix} \Sigma^*_{11} & \Sigma^*_{12} \\ \Sigma^*_{21} & \Sigma^*_{22} \end{pmatrix}.$$

Because $\Sigma^*_{12} = 0 \Leftrightarrow C_{11}\Sigma_{12}C'_{22} = 0 \Leftrightarrow \Sigma_{12} = 0$, the problem of testing H is invariant. (C_{11} and C_{22} are nonsingular, so C_{11}^{-1} and C_{22}^{-1} exist. Pre and post multiplication of $C_{11}\Sigma_{12}C'_{22}$ by C_{11}^{-1} and $(C'_{22})^{-1}$ respectively then gives Σ_{12} back.)]

Next we prove the invariance of the roots of the equation

(15) $$|S_{12}S_{22}^{-1}S_{21} - \lambda S_{11}| = 0.$$

Suppose that x is a sample point with covariance matrix S. Let y be a sample point with covariance matrix T, such that $y = Cx + D$. Then

$$T = \begin{pmatrix} C_{11} & 0 \\ 0 & C_{22} \end{pmatrix} \begin{pmatrix} S_{12} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} C'_{11} & 0 \\ 0 & C'_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} C_{11}S_{11}C'_{11} & C_{11}S_{12}C'_{22} \\ C_{22}S_{21}C'_{11} & C_{22}S_{22}C'_{22} \end{pmatrix} .$$

We have

$$|T_{12}T_{22}^{-1}T_{21} - \lambda T_{11}| = 0 \quad \Leftrightarrow$$

$$|C_{11}S_{12}C'_{22}(C'_{22})^{-1}S_{22}^{-1}C_{22}^{-1}C_{22}S_{21}C_{11} - \lambda C_{11}S_{11}C'_{11}| = 0 \quad \Leftrightarrow$$

$$|C_{11}S_{12}S_{22}^{-1}S_{21}C'_{11} - \lambda C_{11}S_{11}C'_{11}| = 0 \quad \Leftrightarrow$$

$$|C_{11}|^2 |S_{12}S_{22}^{-1}S_{12} - \lambda S_{11}| = 0.$$

Because C_{11} is non-singular, it follows that the roots of $|T_{12}T_{22}^{-1}T_{21} - \lambda T_{11}| = 0$ and of (15) are the same. Hence these roots are invariant.

The roots are maximal invariant when the converse is also true. So suppose that $|S_{12}S_{22}^{-1}S_{21} - \lambda S_{11}| = 0$ and $|T_{12}T_{22}^{-1}T_{21} - \lambda T_{11}| = 0$ have the same roots. Then there exist matrices B and C such that $BS_{11}B' = I = CT_{11}C'$ and $BS_{12}S_{22}^{-1}S_{21}B' = CT_{12}T_{22}^{-1}T_{21}C' = \Lambda$. where Λ is the diagonal matrix whose diagonal elements are the roots λ. Since S_{22}^{-1} and T_{22}^{-1} are positive definite there exist nonsingular matrices E and F such that $S_{22}^{-1} = EE'$ and $T_{22}^{-1} = FF'$. Then

$$(BS_{12}E)(BS_{12}E)' = (CT_{12}F)(CT_{12}F)'$$

and it follows from the argument given in Section 10 in connection with G_2 that there exists an orthogonal matrix Q such that $BS_{12}EQ = CT_{12}F$, so that

$$C_{11} = C^{-1}B \quad \text{and} \quad C'_{22} = EQF^{-1}.$$

This proves the existence of the required transformation.

(ii) For the case $q = 1$, the solution of (15) is, trivially,

$\lambda = S_{12}S_{22}^{-1}S_{21}/S_{11} = R^2$, which is the square of the multiple correlation coefficient between X_{11} and $(X_{12},\dots,X_{1p})$ (ANDERSON, 1958). That the distribution of R^2 only depends on ρ^2 is readily seen from the formula given in (iii).

(iii) Denote the density of R^2 under ρ^2 by $p_{\rho^2}(R^2)$. Then, for $\rho_1^2 > \rho_0^2$, we have

$$\frac{p_{\rho_1^2}(R^2)}{p_{\rho_0^2}(R^2)} = \frac{(1-\rho_1^2)^{\frac{1}{2}(N-1)}}{(1-\rho_0^2)^{\frac{1}{2}(N-1)}} \frac{\sum_{h=0}^{\infty} (\rho_1^2)^h (R^2)^h \Gamma^2(\frac{1}{2}(N-1)+h)}{\sum_{h=0}^{\infty} (\rho_0^2)^h (R^2)^h \Gamma^2(\frac{1}{2}(N-1)+h)},$$

which is an increasing function of R^2. Furthermore ρ^2 is the maximal invariant in the parameter space and the distribution of R^2 depends only on ρ. The UMP invariant test therefore rejects $H : q = 1$ $H : \rho = 0$ when $R^2 > C_0$, using the Neyman-Pearson lemma.

(iv) When $\rho = 0$, R^2 has a beta distribution with parameter $\frac{1}{2}(N-1)$ and $\frac{1}{2}(N-p)$. The required result follows directly from this fact.

(SIMAIKA (1941))

Section 12

Problem 33.

There exists a nonsingular linear transformation B such that $BA^{-1}B' = I$ (= identity matrix). If Y has the $N_q(\eta, A^{-1})$-distribution, given by (62) on p. 304, then $Z = BY$ has the $N_q(\xi, I)$-distribution, where $\xi = (\xi_1,\dots,\xi_q)' = B\eta$. That is the components $Z_1,\dots,Z_q$ of Z are independently normally distributed with means $\xi_1,\dots,\xi_q$ and unit variance. The model assumption $\eta \in \Pi_\Omega$ becomes $\xi \in \Pi_\Omega = \{B\eta : \eta \in \Pi_\Omega\}$ while the hypothesis $\eta \in \Pi_\omega$ becomes $\xi \in \Pi_\omega^* = \{B\eta : \eta \in \Pi_\omega\}$. Performing the canonical transformation (1) we get the variable $X = (X_1,\dots,X_q)' = CZ$. The model assumption becomes $\beta_{s+1} = \dots = \beta_q = 0$ and the hypothesis turns into $\beta_1 = \dots = \beta_r = 0$ where $\beta_i = EX_i$.
This problem is invariant under the group G_1 of transformations $X_i' = X_i + c_i$ for $i = r+1,\dots,s$ and $X_i' = X_i$ for $i = 1,\dots,r;s+1,\dots,q$. This leaves $X_1,\dots,X_r,X_{s+1},X_q$ as maximal invariants. Another group leaving the problem invariant is the group G of all orthogonal transformations of $X_1,\dots,X_r$. A maximal invariant under G_2 is $U = \Sigma_{i=1}^r X_i^2, X_{s+1},\dots,X_q$. This reduces to U by sufficiency. In the parameter space this reduces to $\psi = \Sigma_{i=1}^r \beta_i^2$ as a maximal invariant.
It follows from Theorem 3 of Chapter 6 that the distribution of U depends

only of ψ, so that the principle of invariance reduces the problem to that of testing the simple hypothesis $\psi = 0$. Since $X_1,\dots,X_q$ are independently normally distributed with common variance 1 and mean $E(X_i) = \beta_i$ the distribution of the statistic $U = \Sigma_{i=1}^r X_i^2$ is the noncentral χ^2 with noncentrallity parameter ψ. By Problem 4 the class of noncentral χ^2 distributions has monotone likelihood ratio in ψ. Hence the UMP invariant test rejects when $U > C$. The cutoff point C is determined so that the probability of rejection is α when $\psi = 0$. Since in this case W is the central χ_r^2 distribution, C is determined by

$$\int_C^\infty \chi_r(y)\,dy = \alpha.$$

As in the case of an unknown common variance we find that

$$U = \sum_{j=1}^r X_j^2 = (Z-\hat{\hat{\xi}})'(Z-\hat{\hat{\xi}}) + (Z-\hat{\xi})'(Z-\xi) =$$

$$= (\hat{\xi}-\hat{\hat{\xi}})'(B^{-1})'AB^{-1}(\hat{\xi}-\hat{\hat{\xi}}) = (\hat{\eta}-\hat{\hat{\eta}})'A(\hat{\eta}-\hat{\hat{\eta}}) =$$

$$= \sum_{i=1}^q \sum_{j=1}^q a_{ij}(\hat{\eta}_i-\hat{\hat{\eta}}_i)(\hat{\eta}_j-\hat{\hat{\eta}}_j)$$

with obvious definitions of $\hat{\xi},\hat{\hat{\xi}},\hat{\eta},\hat{\hat{\eta}}$ (cf. p. 304).

Section 13.

Problem 34.

Consider the restricted class of alternatives $K : p \in S$, $p \neq \pi$ to the hypothesis $H : p = \pi$. The surface S is contained in the plane $M = \{(x_1,\dots,x_m) \mid \Sigma x_i = 1, x_i \in \mathbb{R}, i = 1,\dots,m\}$. Furthermore we have that $\overline{S}$, the tangent plane at π to S is in M, is of the form

$$p_i = \pi_i(1 + a_{i1}\xi_1 + \dots + a_{is}\xi_s) \quad i = 1,\dots,m$$

where

(16) $$\Sigma\, a_{ik}a_{il}\pi_i = 0 \text{ for all } k \neq l;\ k,l = 1,\dots,s.$$

We first introduce some notation. Let A denote the matrix with elements a_{ij} $(i = 1,\dots,m;\ j = 1,\dots,s)$. Note that A can be interpreted as the

Jacobian matrix $\left(\frac{\partial f_i}{\partial \theta_i}\right)$ of Section 13. Let D denote the diagonal matrix with positive diagonal elements $d_j = \Sigma_{i=1}^m a_{ij}^2 \pi_i$ $(j = 1,\dots,s)$, and let Π denote the $m \times m$ diagonal matrix with diagonal elements π_i $(i = 1,\dots,m)$. Furthermore let v be the vector $v = (v_1,\dots,v_m)'$, π be the vector $\pi = (\pi_1,\dots,\pi_m)'$ and ξ be the vector $\xi = (\xi_1,\dots,\xi_s)'$. The orthogonality relation (16) can be written as

$$A'\Pi A = D.$$

i) Taking the point $\hat{\xi}$ such that

$$\frac{\partial}{\partial \xi}(v - \pi - \Pi A)' \Pi^{-1}(v - \pi - \Pi A\xi) = 0$$

yields the normal equation

$$2A'\Pi\Pi^{-1}(v - \pi - \Pi A\hat{\xi}) = 0.$$

Hence

$$A'(v - \pi) = D\hat{\xi}.$$

Since $\overline{S} \in M$, we have that $A'\pi = 0$ and it follows that

$$\hat{\xi}_j = \sum_{i=1}^m a_{ij} v_i \Big/ \sum_{i=1}^m a_{ij}^2 \pi_i .$$

The vector of the second derivates is positive in $\hat{\xi}$ therefore $\hat{S}$ minimizes $\Sigma\,(v_i - p_i)^2/\pi_i$. The solution $\hat{p}$ need not satisfy $0 \le \hat{p}_i \le 1$.

ii) The test statistic (76) of p. 308 can be written as

$$n \sum_{i=1}^m \frac{(\hat{p}_i - \pi_i)^2}{\pi_i} = n\hat{\xi}'A'\Pi A\hat{\xi} = n\hat{\xi}'D\hat{\xi}$$

$$= n \sum_{i=1}^s \left(\sum_{i=1}^m a_{ij} v_i\right)^2 \Big/ \sum_{i=1}^m a_{ij}^2 \pi_i .$$

<u>Problem 35.</u>

<u>LEMMA</u>. *The likelihood of a multinomial sample* $x_1,\dots,x_m$ *with* m *classes is proportional to* $p_1^{x_1}\dots p_m^{x_m}$ *which has as maximum value* $\left(\frac{x_1}{n}\right)^{x_1}\dots\left(\frac{x_m}{n}\right)^{n}$.

PROOF. We have $P(x_1,\ldots,x_m) = \frac{n!}{x_1!\ldots x_m!}\, p_1^{x_1}\ldots p_m^{x_m}$. Because the geometric mean is less than or equal to the arithmetic mean

$$\prod_{i=1}^{m} \left(\frac{p_i}{x_i/n}\right)^{x_i} \le \sum_{i=1}^{m} \frac{x_i}{n}\,\frac{p_i}{(x_i/n)} = 1.$$

Thus

$$\prod_{i=1}^{m} p_i^{x_i} \le \prod_{i=1}^{m} \left(\frac{x_i}{n}\right)^{x_i}.$$

Applying the lemma to the multinomial situation under Ω with parameter p_{ij} ($i=1,\ldots,a$, $j=1,\ldots,b$; $\Sigma_{i,j}\, p_{ij} = 1$) the maximized likelihood is proportional to $\Pi_{i,j} \left(\frac{n_{ij}}{n}\right)^{n_{ij}}$.
Under ω when $p_{ij} = p_i p_j'$ ($\Sigma_i\, p_i = \Sigma_j\, p_j' = 1$) the likelihood function is proportional to

$$\prod_{i,j} (p_i p_j')^{n_{ij}} = \prod_i p_i^{N_{i\cdot}} \prod_j p_j'^{N_{\cdot j}}$$

where $N_{i\cdot} = \Sigma_j\, n_{ij}$, $N_{\cdot j} = \Sigma_i\, n_{ij}$. Applying the lemma, $\Pi_i\, p_i^{N_{i\cdot}}$ is maximal equal to $\Pi_i \left(\frac{N_{i\cdot}}{n}\right)^{N_{i\cdot}}$ and $\Pi_j\, p_j'^{N_{\cdot j}}$ to $\Pi_j \left(\frac{N_{\cdot j}}{n}\right)^{N_{\cdot j}}$. The likelihood ratio test therefore rejects when

$$\Lambda = \frac{\Pi_i \left(\frac{N_{i\cdot}}{n}\right)^{N_{i\cdot}} \Pi_j \left(\frac{N_{\cdot j}}{n}\right)^{N_{\cdot j}}}{\Pi_{ij} \left(\frac{n_{ij}}{n}\right)^{n_{ij}}} = \frac{\Pi_i\, N_{i\cdot}^{N_{i\cdot}}\ \Pi_j\, N_{\cdot j}^{N_{\cdot j}}}{n^n\, \Pi_{i,j}\, n_{ij}^{n_{ij}}} < k.$$

Under the null-hypothesis and under alternatives of the form (81), p. 310, $-2\log\Lambda$ is asymptotically equivalent to the test of the form (83). In the notation of p. 311, last paragraph we have $s = ab-1$ and $s-v = a-1+b-1$. Thus $r = (a-1)(b-1)$ and under H, $-2\log\Lambda$ has asymptotically a χ^2-distribution with $(a-1)(b-1)$ degrees of freedom.

Note. For a more detailed discussion see WITTING and NÖLLE (1970) Section 2.7.3, especially Example 2.32.

REFERENCES

ANDERSON, T.W. (1958). *An Introduction to Multivariate Analysis*. Wiley, New York.

COCHRAN, W.G. & G.M. COX (1957). *Experimental Designs*. Wiley, New York.

HSU, P.L. (1945). On the power function of the E^2-test and the T^2-test. *Ann. Math. Stat.* 16, 278-286.

JOHN, P.W. (1971). *Statistical Design and Analysis of Experiments*. The Macmillan Company, New York.

OKAMOTA, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* 1, 763-765.

RAO, C.R. (1973). *Linear Statistical Inference and its Applications*. Wiley, New York.

SCHEFFÉ, H. (1958). Fitting straight lines when one variable is controlled. *J. Amer. Statist. Assoc.* 53, 106-117.

SCHEFFÉ, H. (1959). *The Analysis of Variance*. Wiley, New York.

SEBER, G.A.F. (1966). *The Linear Hypothesis: a General Theory*. Griffin, London.

SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York

SIMAIKA, J.B. (1941). An optimum property of two statistical tests. *Biometrika* 32, 70-80.

WALD, A. (1942). On the power function of the analysis of variance test. *Ann. Math. Stat.* 13, 434-439.

WIJSMAN, R.A. (1957). Random orthogonal transformations and their use in some classical distribution problems in multivariate analysis. *Ann. Math. Stat.* 28, 415-423.

WITTING, H. & G. NÖLLE (1970). *Angewandte Mathematische Statistik*. Teubner, Stuttgart.

CHAPTER 8

Section 1

Problem 1.

Let $\beta = \sup \left[\inf_{\theta \in \Omega_K} E_\theta \varphi(X) \right]$, where the supremum is taken over all level α tests φ of $H : \theta \in \Omega_H$. Let $\{\varphi_n\}_{n=1}^{\infty}$ be a sequence of level α tests such that

$$\lim_{n\to\infty} \inf_{\theta\in\Omega_K} E_\theta \varphi_n(X) = \beta.$$

In view of the weak compactness theorem (Theorem 3 of the Appendix) there exists a subsequence $\{\varphi_{n'}\}$ which weakly converges to φ, say. This implies that for all $\theta \in \Omega_H$

$$E_\theta \varphi(X) = \lim_{n'\to\infty} E_\theta \varphi_{n'}(X) \le \alpha$$

and for all $\theta \in \Omega_K$

$$E_\theta \varphi(X) = \lim_{n'\to\infty} E_\theta \varphi_{n'}(X) \ge \lim_{n'\to\infty} \inf_{\theta'\in\Omega_K} E_{\theta'} \varphi_{n'}(X) = \beta.$$

Thus φ is a test with the desired property.

Problem 2.

(i) The assertion stated in (i) is not correct as is shown by the following example.
Let the distribution of X be given by

$$P_\theta\{X = 0\} = \tfrac{1}{2} + \theta^3$$

$$P_\theta\{X = 1\} = \tfrac{1}{2} - \theta^3$$

where $\theta \in \Omega = \{\theta : -\frac{1}{2} \le \theta^3 \le \frac{1}{2}\}$. Note that $P_\theta \ne P_{\theta'}$ if $\theta \ne \theta'$. A test φ

for testing the hypothesis H : $\theta = 0$ against K : $\theta > 0$ is of level α if

$$\tfrac{1}{2}(\varphi(0) + \varphi(1)) \leq \alpha.$$

The power of φ is given by

$$\beta_\varphi(\theta) = (\tfrac{1}{2} + \theta^3)\varphi(0) + (\tfrac{1}{2} - \theta^3)\varphi(1)$$

and has derivative

$$\beta'_\varphi(\theta) = 3\theta^2(\varphi(0)-\varphi(1)).$$

Since $\beta'_\varphi(0) = 0$ we have that all level α tests maximize the derivative of the power function at $\theta = 0$, but it is clear that not all level α tests are LMP.

To avoid this kind of counter examples, we make the extra assumption that the test which maximizes the derivative of the power function at $\theta = \theta_0$ is the unique such test. Furthermore, we need that the test is of exact size α.

Assume that the power function $\beta_\varphi(\theta)$ of any test φ is continuously differentiable at $\theta = \theta_0$, where differentiation may be taken under the integral sign. Then the test φ_0, which maximizes $\beta'(\theta_0)$ among all size α tests of the hypothesis H, can be found by applying the lemma of Neyman & Pearson in the extended from and, hence, is given by

$$\varphi_0(x) = \begin{cases} 1 & \\ & \text{if } \frac{d}{d\theta} p_\theta(x)\Big|_{\theta=\theta_0} \gtrless k\, p_{\theta_0}(x). \\ 0 & \end{cases}$$

Assume that φ_0 is the unique such test. We first show that φ_0 is LMP. Let φ be any size α test, then Taylor expansion yields for θ near θ_0

$$\beta_{\varphi_0}(\theta) = \alpha + (\theta - \theta_0)\beta'_{\varphi_0}(\theta_0 + \eta_0(\theta - \theta_0)) \qquad \text{for some } \eta_0 \in [0,1]$$

and

$$\beta_\varphi(\theta) = \alpha + (\theta - \theta_0)\beta'_\varphi(\theta_0 + \eta(\theta - \theta_0)) \qquad \text{for some } \eta \in [0,1].$$

Since β'_{φ_0} and β'_φ are continuous at θ_0, it follows that for θ sufficiently near θ_0

$$\beta'_{\varphi_0}(\theta_0 + \eta_0(\theta - \theta_0)) > \beta'_\varphi(\theta_0 + \eta(\theta - \theta_0)),$$

and hence, since $\theta > \theta_0$

$$\beta_{\varphi_0}(\theta) > \beta_{\varphi}(\theta).$$

It is clear that φ_0 is locally more powerful than any other level α test. To prove that a LMP level α test maximizes $\beta'(\theta_0)$ among all size α tests, we first note that a LMP level α test is of size α. The result now follows from the Taylor expansions written above.
A reference to this problem is FERGUSON (1967), pp. 235-237.

(ii) Let φ_0 be a LMP level α test. Then, given any other level α test φ, there exists Δ_1 such that

$$\beta_{\varphi_0}(\theta) \geq \beta_{\varphi}(\theta) \quad \text{for all } \theta \text{ with } 0 < d(\theta) < \Delta_1.$$

Since β_{φ_0} is bounded away from α for every set of alternatives which is bounded away from H, there exists $\varepsilon > 0$ such that

$$\beta_{\varphi_0}(\theta) > \alpha+\varepsilon \qquad \text{for all } \theta \text{ with } d(\theta) > \Delta_1.$$

Moreover, by continuity of $\beta_{\varphi_0}(\theta)$, there exists a $\Delta_0 < \Delta_1$ such that

$$\alpha \leq \beta_{\varphi_0}(\theta) \leq \alpha+\varepsilon \quad \text{for all } \theta \text{ with } d(\theta) < \Delta_0.$$

Hence for all $\Delta < \Delta_0$ there exists θ_Δ with $\Delta \leq d(\theta_\Delta) < \Delta_1$ such that

$$\inf_{\omega_\Delta} \beta_{\varphi_0}(\theta) = \beta_{\varphi_0}(\theta_\Delta) \geq \beta_{\varphi}(\theta_\Delta) \geq \inf_{\omega_\Delta} \beta_{\varphi}(\theta),$$

for any other test φ. Here ω_Δ is the set of θ's for which $d(\theta) > \Delta$.

(iii) By (i) the acceptance region of the LMP level α test φ_0 is

$$p_0'(x)/p_0(x) < k,$$

i.e. in the present case

$$2 \sum_{j=1}^{n} \frac{x_j}{1+x_j^2} < k,$$

where $k > 0$ because $\alpha < \frac{1}{2}$. (If $k \leq 0$ then $\alpha = P_{\theta=0}[\text{rejection of H}] \geq$
$\geq P_{\theta=0}\{\Sigma_{j=1}^{n} X_j/(1+X_j^2) > 0\} = \frac{1}{2}$.)
Note that the test which maximizes $\beta'(0)$ is unique in the case that $p_\theta(x)$ is the density of a Cauchy distribution with location parameter θ.
Since $x_j/(1+x_j^2) \to 0$ as $x_j \to \infty$, there exists M such that any point with $x_j \geq M$ for all $j = 1,\ldots,n$ lies in the acceptance region. Hence the power

of φ_0,

$$\beta_{\varphi_0}(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \varphi_0(x)\pi^{-n} \prod_{j=1}^{n} (1+(x_j-\theta)^2)^{-1} dx_1 \cdot \ldots \cdot dx_n \leq$$

$$\leq \int_{-\infty}^{M} \cdots \int_{-\infty}^{M} \pi^{-n} \prod_{j=1}^{n} (1+(x_j-\theta)^2)^{-1} dx_1 \cdot \ldots \cdot dx_n$$

$$= \pi^{-n} \prod_{j=1}^{n} (\arctan(M-\theta) + \tfrac{\pi}{2}),$$

tends to zero as θ tends to infinity (cf. also FERGUSON (1967), p. 237). It follows that the LMP test is not unbiased and hence does not maximize the minimum power locally. (Compare the power of φ_0 with the power of the test $\varphi(x) \equiv \alpha$).

(LEHMANN (1955))

Problem 3.

The assertion stated in this problem is not correct. A counter example similar to the example in the solution of Problem 2 (i) can easily be given. Again we need the extra condition that the test which maximizes $\beta''(\theta_0)$ among all unbiased level α tests is unique in order to show that it is also the LMP test.

Assume that the power function $\beta_\varphi(\theta)$ of any test φ is twice continuously differentiable at $\theta = \theta_0$, where differentiation may be taken under the integral sign. Then the test φ_0, which maximizes $\beta''(\theta_0)$ among all unbiased level α tests of the hypothesis H, can be constructed using the lemma of Neyman & Pearson in the extended form. Assume that φ_0 is the unique such test. We first show that φ_0 is LMP. Note that all unbiased level α tests are of size α and have $\beta'(\theta_0) = 0$. Let φ be any unbiased level α test, then Taylor expansion for θ near θ_0 yields

$$\beta_{\varphi_0}(\theta) = \alpha + \tfrac{1}{2}(\theta-\theta_0)^2 \beta''_{\varphi_0}(\theta_0 + \eta_0(\theta-\theta_0)) \quad \text{for some } \eta_0 \in [0,1]$$

and

$$\beta_{\varphi}(\theta) = \alpha + \tfrac{1}{2}(\theta-\theta_0)^2 \beta''_{\varphi}(\theta_0 + \eta(\theta-\theta_0)) \quad \text{for some } \eta \in [0,1].$$

Since β''_{φ_0} and β''_φ are continuous at θ_0, it follows that for θ sufficiently near θ_0

$$\beta''_{\varphi_0}(\theta_0 + \eta_0(\theta-\theta_0)) > \beta''_{\varphi}(\theta_0 + \eta(\theta-\theta_0))$$

and hence

$$\beta_{\varphi_0}(\theta) > \beta_{\varphi}(\theta).$$

The prove that a LMP unbiased level α test maximizes $\beta''(\theta_0)$ among all unbiased level α tests is along the same lines as above.
A reference to this problem is FERGUSON (1967), pp. 237-238.

Section 2

Problem 4.

(i) The statement is not true, as is seen by the following example. Let $0 < \alpha < 1$. The distribution of X is given by

$$P_\theta\{X = 1\} = \alpha + \theta^4,$$

$$P_\theta\{X = 0\} = \tfrac{1}{4}(1-\alpha) + \theta,$$

$$P_\theta\{X = -1\} = \tfrac{3}{4}(1-\alpha) - \theta - \theta^4,$$

where $\theta \in \Omega = \{\theta : -\frac{1}{4}(1-\alpha) < \theta < \frac{1}{4}(1-\alpha)\}$. Note that $P_\theta \neq P_{\theta'}$ if $\theta \neq \theta'$. The test

$$\varphi(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{if } x \neq 1 \end{cases}$$

of $H : \theta = 0$ against $\theta \neq 0$ is locally strictly unbiased. However,

$$\beta''_\varphi = \frac{d^2}{d\theta^2}\beta_\varphi(\theta)\Big|_{\theta=0} = 0.$$

Let f be a function whose domain contains an open set $A \subset \mathbb{R}^n$. Suppose that the second order derivatives of f exist at every $x = (x_1,\ldots,x_n) \in A$. Write

$$f_i(x) = \frac{\partial}{\partial x_i} f(x) \qquad i = 1,\ldots,n$$

and

$$f_{ij}(x) = \frac{\partial^2}{\partial x_i \partial x_j} f(x) \qquad i,j = 1,\ldots,n,$$

for $x \in A$. Suppose furthermore that f_{ij} is a continuous function. Let $x_0 \in A$ and denote by M the matrix $(f_{ij}(x_0))$. By Proposition 10 and Theorem 6 on pp. 60-62 in FLEMMING (1965) we have

$f_i(x_0) = 0$, $i = 1,\ldots,n$, M *is positive definite* $\Rightarrow$
f *has a strict relative minimum at* x_0,

and

f *has a relative minimum at* $x_0 \Rightarrow$
$f_i(x_0) = 0$, $i = 1,\ldots,n$, M *is non-negative definite.*

In the setting of Problem 4 application of the above result yields the following statement.

Suppose that for all critical functions φ *and all fixed* ϑ *the first and second derivatives*

$$\beta_\varphi^i(\theta,\vartheta) = \frac{\partial}{\partial\theta_i}\beta_\varphi(\theta,\vartheta), \qquad \beta_\varphi^{ij}(\theta,\vartheta) = \frac{\partial^2}{\partial\theta_i\partial\theta_j}\beta_\varphi(\theta,\vartheta)$$

exist and are continuous (*w.r.t.* θ) *in a* θ*-neighborhood of* θ^0. *If for each* ϑ, $\beta_\varphi(\theta^0,\vartheta) = \alpha$, $\beta_\varphi^i(\theta^0,\vartheta) = \alpha$ $(i = 1,\ldots,r)$, *and the matrix* $(\beta_\varphi^{ij}(\theta^0,\vartheta))$ *is positive definite, then* φ *is a locally strictly unbiased test of* $H : \theta = \theta^0$ *against* $\theta \neq \theta^0$ *at level* α. *If* φ *is a locally strictly unbiased test of* $H : \theta = \theta^0$ *against* $\theta \neq \theta^0$ *at level* α, *then for each* ϑ, $\beta_\varphi(\theta^0,\vartheta) = \alpha$, $\beta_\varphi^i(\theta^0,\vartheta) = 0$ $(i = 1,\ldots,r)$, *and the matrix* $(\beta_\varphi^{ij}(\theta^0,\vartheta))$ *is non-negative definite.*

(ii) The Gaussian curvature of the power surface at θ^0 is given by

$$\frac{|(\beta_\varphi^{ij}(\theta^0,\vartheta))|}{\left\{1 + \Sigma_{i=1}^r \{\beta_\varphi^i(\theta^0,\vartheta)\}^2\right\}^{r/2+1}} = |(\beta_\varphi^{ij}(\theta^0,\vartheta))|$$

because of the locally strictly unbiasedness of φ.
Now consider the set-up of Chapter 7, Section 1. So let $\theta = (\eta_1,\ldots,\eta_r)$, $\vartheta = (\eta_{r+1},\ldots,\eta_s,\sigma)$, $\theta^0 = (0,\ldots,0)$ and $\psi^2 = \sigma^{-2}\,\Sigma_{i=1}^r \eta_i^2$.
Let $(\theta',\vartheta') = (\eta_1',\ldots,\eta_s',\sigma')$ be any alternative and let $\psi_1^2 = \sigma'^{-2}\,\Sigma_{i=1}^r \eta_i'^2$.
The test φ_0 is MP for testing $\psi = 0$ against $\psi = \psi_1$ (that is for testing $p_0(w)$ against $p_{\psi_1}(w)$). In view of Corollary 1 on p. 67 of the book if follows that $\beta_{\varphi_0}(\theta',\vartheta') > \alpha$. Noting that for each ϑ, $\beta_{\varphi_0}(0,\vartheta) = \alpha$ by (7) and (8) on p. 268 of the book it is seen that φ_0 is (locally) strictly unbiased.
Let φ_1 be any locally strictly unbiased test. Then φ_1 is similar and hence (cf. Problem 5 of Chapter 7)

$$(1)\qquad \int_S [\beta_{\varphi_0}(\eta,\sigma^2) - \alpha]dA \geq \int_S [\beta_{\varphi_1}(\eta,\sigma^2) - \alpha]dA,$$

where $S = S(\eta_{r+1},\ldots,\eta_s,\sigma;\rho) = \{(t_1,\ldots,t_s,\sigma) : \Sigma_{i=1}^{r} t_i^2\sigma^{-2} = \rho^2,\ t_{r+1} = \eta_{r+1},\ldots,t_s = \eta_s\}$ and A is the Lebesgue measure on S. In the following ϑ will be kept fixed. For short $\eta_{r+1},\ldots,\eta_s,\sigma$ is dropped out of our notation. Theorem 9 on p. 52 of the book ensures that for any test φ its power function $\beta_\varphi(\eta_1,\ldots,\eta_r)$ is an analytic function of $\eta_1,\ldots,\eta_r$. Denote the matrix $(\beta^{ij}_{\varphi_k}(0))$ by M_k, $k = 0,1$. In view of the locally strictly unbiasedness Taylor expansion yields

$$|\eta|^{-2}|\beta_{\varphi_k}(\eta) - \alpha - \tfrac{1}{2}\eta' M_k \eta| \le D_k(C)\ |\eta|$$

for all $\eta = (\eta_1,\ldots,\eta_r)$ with $|\eta| \le C$ and some constant $D_k(C)$, only depending on C, $k = 0,1$. Hence

$$\lim_{\rho\to 0}\int_S \frac{[\beta_{\varphi_k}(\eta) - \alpha]}{|\eta|^2 \int_S dA} dA = \lim_{\rho\to 0}\int_S \frac{\tfrac{1}{2}\eta' M_k \eta}{\rho^2\sigma^2 \int_S dA} dA = \tfrac{1}{2}\sum_{j=1}^{r} \beta^{jj}_{\varphi_k}(0)/r, \tag{2}$$

because

$$\int_S \eta_i\eta_j dA = 0 \text{ for all } i \neq j$$

and

$$\int_S \eta_j^2 dA = \frac{1}{r}\sum_{j=1}^{r}\int_S \eta_j^2 dA = \frac{\rho^2\sigma^2}{r}\int_S dA.$$

Since for any non-negative definite matrix (b^{ij}), $|(b^{ij})| \le \prod_j b^{jj}$ (cf. RAO (1973), p. 56), it follows by (1) and (2) that

$$|(\beta^{ij}_{\varphi_1})| \le \prod_j \beta^{jj}_{\varphi_1} \le \left\{\sum_j \beta^{jj}_{\varphi_1}/r\right\}^r \le \left\{\sum_j \beta^{jj}_{\varphi_0}/r\right\}^r, \tag{3}$$

where the well-known inequality of the geometric and arithmic mean is used. By the first lines on p. 269 of the book the power function of φ_0 is of the form

$$\beta_{\varphi_0}(\eta) = e^{-\frac{1}{2}\psi^2}\sum_{k=0}^{\infty} a_k(\psi^2)^k = \sum_{k=0}^{\infty} b_k(\psi^2)^k.$$

Hence

$$\beta^{ij}_{\varphi_0}(0) = \begin{cases} 0 & \text{if } i \neq j \\ 2\sigma^{-2}b_1 & \text{if } i = j, \end{cases}$$

implying

$$\left\{\sum_{j=1}^{r} \beta^{jj}_{\varphi_0}/r\right\}^r = \{\beta^{''}_{\varphi_0}\}^r = |(\beta^{ij}_{\varphi_0})|. \tag{4}$$

Combination of (3) and (4) completes the proof.

(KIEFER (1958))

Problem 5.

The result stated in this problem follows directly from application of Theorem 1 of Chapter 8 with $\omega = \{D : D(0) = \frac{1}{2}\}$, $\omega' = \{D : D(0) \le q\}$ where $q < \frac{1}{2}$. The conditions of this theorem are verified as follows. We first show that the densities f and g given in the hint correspond to distribution functions $F \in \omega$ and $G \in \omega'$ respectively. The density f defined by

$$f(x) = \frac{1-2q}{2(1-q)} \left(\frac{q}{1-q}\right)^{[|x|]}$$

is symmetric about zero. Furthermore

$$\int_0^\infty f(x)dx = \sum_{x=0}^\infty \frac{1-2q}{2(1-q)} \left(\frac{q}{1-q}\right)^x = \frac{1-2q}{2(1-q)} \left(1 - \frac{q}{1-q}\right)^{-1} = \tfrac{1}{2},$$

hence f *is* a density and $F(0) = \frac{1}{2}$. The density g defined by

$$g(x) = (1-2q)\left(\frac{q}{1-q}\right)^{|[x]|}$$

satisfies

$$\int_0^\infty g(x)dx = (1-2q) \sum_{x=0}^\infty \left(\frac{q}{1-q}\right)^x = (1-2q)/\left(1 - \frac{q}{1-q}\right) = 1-q$$

and

$$\int_{-\infty}^0 g(x)dx = (1-2q) \sum_{x=1}^\infty \left(\frac{q}{1-q}\right)^x = \frac{q}{1-q}(1-q) = q.$$

Thus g is a density also, and $G(0) = q$.
We now show that the MP size α test for testing f against g is the sign test. Let $X = \#\{i : Z_i > 0\}$ and consider the sign test of size

$$\alpha = \gamma\binom{N}{x} 2^{-N} + \sum_{y=x+1}^{N} \binom{N}{y} 2^{-N},$$

which rejects H when $X > x$ and rejects H with probability γ if $X = x$. By the lemma of Neyman and Pearson the MP size α test of f versus g, based on observations $Z_1 = z_1, \ldots, Z_N = z_N$, rejects when

$$\prod_{i=1}^N \frac{g(z_i)}{f(z_i)} > k.$$

Since D is continuous we can assume that for each i, $z_i \notin \{0,-1,-2,\ldots\}$. Then

$$\prod_{i=1}^N \frac{g(z_i)}{f(z_i)} = (2(1-q))^N \prod_{i=1}^N \left(\frac{q}{1-q}\right)^{|[z_i]|-[|z_i|]} = \left(\frac{1-q}{q}\right)^x (2q)^N,$$

where $x = \#\{i : z_i > 0\}$. Because $q < \frac{1}{2}$, $\log((1-q)/q) > 0$ and so $\Pi_1^N g(z_i)/f(z_i) > k$ is equivalent with $x > k'$.

It remains to show that this test is of size α for testing H against K, and that its minimal power over K is attained against g. The first requirement is trivially satisfied, and the second follows from the fact that the power

$$\gamma\binom{N}{x}(1-D(0))^x D(0)^{N-x} + \sum_{y=x+1}^{N} \binom{N}{y}(1-D(0))^y D(0)^{N-y}$$

is non-increasing in $D(0)$. This follows from Lemma 2 on p. 74 of the book and the fact that $\binom{N}{y}\theta^y(1-\theta)^{N-y}$ has monotone likelihood ratio in y with respect to $\theta = 1 - D(0)$ (cf. Example 2 on p. 70 of the book).
It follows that the sign test is maximin for testing $D(0) \geq \frac{1}{2}$ against $D(0) \leq q$.

(RUIST (1954))

Problem 6.

First note that $p_\theta(x)$ has monotone likelihood ratio in x iff

$$x_1 < x_2 \Rightarrow \frac{\partial}{\partial\theta} \log p_\theta(x_1) \leq \frac{\partial}{\partial\theta} \log p_\theta(x_2).$$

In the present case we have

$$f_\theta(x) = \theta g(x) + (1-\theta)h(x) = h(x)\{\theta G(x) + 1 - \theta\},$$

where $G(x) = g(x)/h(x)$. Since

$$\frac{\partial}{\partial\theta} \log f_\theta(x) = \frac{G(x)-1}{\theta G(x) + 1 - \theta},$$

we have that $f_\theta(x)$ has monotone likelihood ratio in x iff

$$x_1 < x_2 \Rightarrow \frac{G(x_1)-1}{\theta G(x_1)+1-\theta} \leq \frac{G(x_2)-1}{\theta G(x_2)+1-\theta},$$

i.e.

$$x_1 < x_2 \Rightarrow G(x_1) \leq G(x_2).$$

(It is assumed that g(x) and h(x) are densities.)

Problem 7.

Interpreting $g_\theta(x;\xi)$ as the conditional density of x given ξ, and $h_\theta(\xi)$ as the a priori density of ξ, let $\rho_\theta(\xi;x)$ denote the a posteriori density of ξ given x, i.e.

$$\rho_\theta(\xi;x) = g_\theta(x;\xi)h_\theta(\xi)/p_\theta(x).$$

We assume that the joint density of x and ξ,

$$f_\theta(x,\xi) = g_\theta(x;\xi)h_\theta(\xi),$$

is positive for all x, ξ and θ.
Consider any fixed $x \le x'$ and $\theta \le \theta'$, then we must show that

(5) $$p_\theta(x')/p_\theta(x) \le p_{\theta'}(x')/p_{\theta'}(x).$$

Since

$$p_\theta(x')/p_\theta(x) = \int \frac{g_\theta(x';\xi)}{g_\theta(x;\xi)} \rho_\theta(\xi;x)d\nu(\xi),$$

and a similar expression holds for θ', it follows that (5) is equivalent to

$$\int \frac{g_\theta(x';\xi)}{g_\theta(x;\xi)} \rho_\theta(\xi;x)d\nu(\xi) \le \int \frac{g_{\theta'}(x';\xi)}{g_{\theta'}(x;\xi)} \rho_{\theta'}(\xi;x)d\nu(\xi).$$

By assumption (a) it is enough to prove that

$$D = \int \frac{g_\theta(x';\xi)}{g_\theta(x;\xi)} [\rho_{\theta'}(\xi;x) - \rho_\theta(\xi;x)]d\nu(\xi) \ge 0.$$

Now define for x, θ and θ' the sets $S^- = \{\xi : \rho_{\theta'}(\xi;x) < \rho_\theta(\xi;x)\}$ and $S^+ = \{\xi : \rho_{\theta'}(\xi;x) \ge \rho_\theta(\xi;x)\}$. Then for any $\xi \in S^-$ and $\xi' \in S^+$ we have

$$\frac{\rho_{\theta'}(\xi;x)}{\rho_\theta(\xi;x)} < 1 \le \frac{\rho_{\theta'}(\xi';x)}{\rho_\theta(\xi';x)}$$

i.e.

$$\frac{g_{\theta'}(x;\xi)h_{\theta'}(\xi)}{g_\theta(x;\xi)h_\theta(\xi)} < \frac{g_{\theta'}(x;\xi')h_{\theta'}(\xi')}{g_\theta(x;\xi')h_\theta(\xi')},$$

and hence by assumption (b) we must have that

$$\xi \le \xi' \quad \text{for all } \xi \in S^-,\ \xi' \in S^+.$$

Thus by assumption (c) we have

$$(6) \qquad \xi \in S^-,\ \xi' \in S^+ \Rightarrow \frac{g_\theta(x';\xi)}{g_\theta(x;\xi)} \le \frac{g_\theta(x';\xi')}{g_\theta(x;\xi')}.$$

Now define

$$a = \left(\int_{S^-} [\rho_\theta(\xi;x) - \rho_{\theta'}(\xi;x)]d\nu(\xi)\right)^{-1} \int_{S^-} \frac{g_\theta(x';\xi)}{g_\theta(x;\xi)} [\rho_\theta(\xi;x)-\rho_\theta{}'(\xi;x)]d\nu(\xi),$$

and

$$b = \left(\int_{S^+} [\rho_{\theta'}(\xi;x)-\rho_\theta(\xi;x)]d\nu(\xi)\right)^{-1} \int_{S^+} \frac{g_\theta(x';\xi)}{g_\theta(x;\xi)} [\rho_{\theta'}(\xi;x)-\rho_\theta(\xi;x)]d\nu(\xi),$$

then by (6) it follows that $a \le b$.
Hence

$$D = -a \int_{S^-} [\rho_\theta(\xi;x) - \rho_{\theta'}(\xi;x)]d\nu(\xi) + b \int_{S^+} [\rho_{\theta'}(\xi;x)-\rho_\theta(\xi;x)]d\nu(\xi) =$$

$$= (b-a) \int_{S^+} [\rho_{\theta'}(\xi;x) - \rho_\theta(\xi;x)]d\nu(\xi) \ge 0,$$

which was to be proved.

(LEHMANN (1955); see also footnote on p. 346 of the book)

Problem 8.

We first prove a general result on exponential families. Let $X_1,\ldots,X_n$ be a random sample from the exponential family

$$p_\theta(x) = C(\theta) \exp[Q(\theta)T(x)]h(x)$$

(w.r.t. some σ-finite measure μ) in which $Q(\theta)$ is an increasing function of θ. Consider, for testing $H : \theta = \theta_0$ against the alternatives $K : \theta \le \theta_1$ or $\theta \ge \theta_2$ where $\theta_1 < \theta_0 < \theta_2$, the test of (exact) size α based on $T_n = \Sigma_{i=1}^n T(X_i)$ of the form

$$\varphi(t) = \begin{cases} 1 & \text{if } t < t_1 \text{ or } t > t_2 \\ \gamma_i & \text{if } t = t_i,\ i=1,2 \\ 0 & \text{if } t_1 < t < t_2 \end{cases} .$$

Choose $\gamma_i = 0$ if $P\{T_n = t_i\} = 0$, $i=1,2$, and choose t_1 and t_2 as small as possible if this gives an equivalent test. We call such tests "natural two-sided tests (of size α)". Note that by the size requirement, (t_2,γ_2) is determined uniquely by (t_1,γ_1), and the tests can be ordered lexicographically by (t_1,γ_1).

Then

(a) *For any test of size* $\leq \alpha$ *of* H *against* K, *there exists a natural two sided size* α *test which is uniformly at least as powerful.*

(b) *There exists a point* $\theta_3 \in (\theta_1, \theta_2)$ *such that the power function of a natural two-sided test is non-increasing in* θ *for* $\theta < \theta_3$ *and non-decreasing in* θ *for* $\theta > \theta_3$.

(c) *For increasing* (t_1, γ_1), *the power against alternatives* $< \theta_0$ *is non-decreasing and the power against alternatives* $> \theta_0$ *is non-increasing.*

(d) *If for sample size* n *a natural two-sided size* α *test exists with power* $\geq \beta > \alpha$ *at* θ_1 *and* θ_2, *while for sample size* n-1 *one exists with power* $< \beta$ *at* θ_1 *and* θ_2, *then* n *is the smallest size admitting a size* α *test of power* $\geq \beta$ *of* H *against* K.

PROOF: By sufficiency we may consider tests based on $T_n = \Sigma_{i=1}^n T(X_i)$. We note that T_n also has a distribution from an exponential family with the exponential part of the density of T_n equal to $\exp [Q(\theta)t_n]$. So without loss of generality we consider (for part (a) to (c)) the case n = 1. Now we see that (a) is a consequence of part (i) of Problem 8 of Chapter 4. Statement (b) follows from the proof of Theorem 6 (iii) of Chapter 3. Statement (c) follows by applying Lemma 2 of Chapter 3 in the same way as is done on p. 90 of the book, after the proof of Theorem 6, (or see solution of Problem 8 (ii) of Chapter 4). For part (d), note that by sufficiency we always prefer a test based on T_n to one based on T_{n-1}. Thus if no test of power β exists at sample size n-1, it does not exist for smaller sample sizes either. By (c), we can at sample size n-1 only attain power β at θ_1 at the cost of still lower power at θ_2 or vice-versa. Thus power $\geq \beta$ at all alternatives in K cannot be attained. By (b), at sample size n the test under consideration has power $\geq \beta$ at all alternatives in K. □

Now we turn to the cases described in the problem.

(i) The above result is directly applicable. Trial and error gives

p_0	p_1	p_2	n	t_1	t_2	γ_1	γ_2	$\beta(p_0)$	$\beta(p_1)$	$\beta(p_2)$
$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$	68	26	42	.39	.39	.05	.796	.796
			69	26	43	.86	.86	.05	.805	.805
$\frac{2}{5}$	$\frac{1}{4}$	$\frac{4}{7}$	71	21	37	.25	.42	.05	.795	.797
			72	21	38	.63	.92	.05	.803	.802
$\frac{3}{10}$	$\frac{3}{17}$	$\frac{6}{13}$	82	17	34	.82	.39	.05	.797	.794
			83	18	34	.09	.07	.05	.803	.803
$\frac{1}{5}$	$\frac{1}{9}$	$\frac{1}{3}$	111	15	33	.58	.68	.05	.798	.798
			112	15	33	.76	.45	.05	.803	.803
$\frac{1}{10}$	$\frac{1}{19}$	$\frac{2}{11}$	203	13	32	.82	.25	.05	.799	.799
			204	13	32	.92	.12	.05	.803	.802

The required minimal sample sizes are therefore 69,72,83,112,204; smaller p_0 requiring smaller sample size.

(ii) Application of Example 4 on p. 332 of the book shows that we may restrict attention to tests based on the statistic $\Sigma_{i=1}^{n} (X_i - \overline{X})^2$, which is distributed as $\sigma^2\chi^2_{n-1}$. Inspection of the table of the UMP unbiased test in the solution of Problem 5 of Chapter 4 shows that n = 46 is the minimal sample size.

In the references to Chapter 8 it is stated that LEHMANN (1955) is relevant to this problem. However here only one-sided alternatives are considered.

Problem 9.

We show that for every test the power function is continuously differentiable at $\theta = 0$ and that the derivative of the power function at $\theta = 0$ is maximized by the sign test. Since the sign test need not be uniquely defined for certain levels α it does not follow that the sign test is LMP (see Problem 2 (i)).
We first consider a somewhat more general situation. Let $X_1,\ldots,X_n$ be i.i.d. with density $f(\cdot - \theta)$ and consider testing $\theta \le 0$ against $\theta > 0$. Let f be absolutely continuous with respect to the Lebesgue measure with

derivative f' such that f'/f is continuous Lebesgue-a.e. and $\int |f'| < \infty$. Using Vitali's theorem (cf. KLAASSEN (1979)) one can see that for every critical function φ the power function is differentiable

$$\frac{d}{d\theta} \int \cdots \int \varphi(x_1,\ldots,x_n) \prod_{i=1}^{n} f(x_i - \theta)\,dx_1\cdot\ldots\cdot dx_n =$$

$$\int \cdots \int \varphi(x_1,\ldots,x_n) \sum_{i=1}^{n} -\frac{f'}{f}(x_i - \theta) \prod_{i=1}^{n} f(x_i - \theta)\,dx_1\cdot\ldots\cdot dx_n.$$

Since

$$\left|\int \cdots \int \varphi(x_1,\ldots,x_n) \sum_{i=1}^{n} -\frac{f'}{f}(x_i - \theta) \prod_{i=1}^{n} f(x_i - \theta)\,dx_1\cdot\ldots\cdot dx_n\right| =$$

$$\left|\int \cdots \int \varphi(x_1 + \theta,\ldots,x_n + \theta) \sum_{i=1}^{n} -\frac{f'}{f}(x_i) \prod_{i=1}^{n} f(x_i)\,dx_1\cdot\ldots\cdot dx_n\right|$$

$$\leq n \int |f'| < \infty,$$

the dominated convergence theorem yields the continuity of the derivative of the power function. Maximizing

$$\int \cdots \int \varphi(x_1,\ldots,x_n) \sum_{i=1}^{n} -\frac{f'}{f}(x_i) \prod_{i=1}^{n} f(x_i)\,dx_1\cdot\ldots\cdot dx_n$$

under

$$\int \cdots \int \varphi(x_1,\ldots,x_n) \prod_{i=1}^{n} f(x_i)\,dx_1\cdot\ldots\cdot dx_n = \alpha,$$

we see by the lemma of Neyman & Pearson that the derivative at $\theta = 0$ of the power function is maximized by tests which reject for large values of

$$\sum_{i=1}^{n} -\frac{f'}{f}(x_i);$$

and only by such tests.
Since $-\frac{f'}{f}(x) = \mathrm{sign}(x)$ for $f(x) = \frac{1}{2}e^{-|x|}$ our assertion has been proved.

Section 4

Problem 10.

In accordance with our convention that all vectors are column vectors, we consider here the transposes of X, Y and A; i.e. we write $X = (X_1,\ldots,X_p)'$ etc. We assume that X and Y are independent and multivariate normally distributed with means zero and covariance matrices Σ and $\Delta\,\Sigma$ respectively, where Σ is nonsingular and $\Delta > 0$. We shall continually identify members g of some group G of transformations on a Euclidian space with their concrete representations as a matrix or vector of real numbers.

(i) Consider the group $G^{(p)}$ of transformations on $(\mathbb{R}^p)^2$ defined by $(x,y) \underset{A}{\to} (Ax,Ay)$ where A is a $p \times p$ nonsingular matrix with $a_{ij} = 0$ for $i < j$. By nonsingularity $a_{ii} \neq 0$ for all i. Clearly any testing problem concerning Δ is invariant under this group. For such a matrix A let $A_1,\dots,A_p$ be $p \times p$ matrices such that the q'th row of A_q equals the q'th row of A, the other rows being equal to the corresponding rows of the $p \times p$ identity matrix. Since $Ax = A_1A_2\cdot\dots\cdot A_px$ we see that $G^{(p)}$ is generated by the groups $G_q^{(p)}$ of transformations $(x,y) \underset{A_q}{\to} (A_qx,A_qy)$; $q = 1,\dots,p$.
First we shall consider, for arbitrary $1 < q \leq p$, the group $G_q^{(q)}$ of transformations on $(\mathbb{R}^q)^2$. Note that any set of Lebesgue measure zero contained in $(\mathbb{R}^q)^2$ is also assigned probability zero by the distribution of $((X_1,\dots,X_q)',(Y_1,\dots,Y_q)')$, whatever the values of Δ and $\ddagger$. Now an element A of $G_q^{(q)}$ is equal to the $q \times q$ identity matrix with the bottom row replaced by the transpose of some vector a. Under A, $(x,y) \in (\mathbb{R}^q)^2$ is transformed into $((x_1,\dots,x_{q-1},a'x)',(y_1,\dots,y_{q-1},a'y)')$. Clearly $((x_1,\dots,x_{q-1})',(y_1,\dots,y_{q-1})')$ is invariant under $G_q^{(q)}$. Consider two elements (x,y) and (x^*,y^*) of $(\mathbb{R}^q)^2$ which are on the same orbit under $G_q^{(q)}$. Thus there exists a vector a such that

(7) $$a'x = x_q^*, \quad a'y = y_q^*$$

while we must also have

(8) $$x_i = x_i^*, \quad y_i = y_i^* \text{ for } i < q.$$

We shall show that for almost any (w.r.t. Lebesgue measure) (x,y) and (x^*,y^*) satisfying (8), a vector a can be found such that (7) holds too. Thus $((x_1,\dots,x_{q-1})',(y_1,\dots,y_{q-1})')$ is not only invariant under $G_q^{(q)}$, but also *equivalent* (in the sence of p. 225 of the book) to a maximal invariant. (In fact we could determine a maximal invariant T such that $T(x,y) = ((x_1,\dots,x_{q-1})',(y_1,\dots,y_{q-1})')$ for almost all $(x,y) \in (\mathbb{R}^q)^2$, and $T(x,y) = ((x_1,\dots,x_{q-1})',(y_1,\dots,y_{q-1})',S(x,y))$ for the remaining (x,y) for some function S.) Outside of a set with Lebesgue measure zero, we have $x_q \neq 0$ and $y_q \neq 0$, and for some $i < q$

$$x_i/x_q \neq y_i/y_q.$$

Solving for a_q in (7) we find

$$a_q = x_q^*/x_q - \sum_{i=1}^{q-1} a_i x_i/x_q = y_q^*/y_q - \sum_{i=1}^{q-1} a_i y_i/y_q \tag{9}$$

so that

$$x_q^*/x_q - y_q^*/y_q = \sum_{i=1}^{q-1} a_i(x_i/x_q - y_i/y_q). \tag{10}$$

Thus starting with almost any (x,y) and (x^*,y^*) satisfying (8), we can find $a_1,\ldots,a_{q-1}$ such that (10) holds, and we can define a_q by (9). Moreover, for almost any (x,y) and (x^*,y^*) we can ensure that $a_q \neq 0$. We have thus finally determined a such that (7) holds.

For $q = 1$, $(x_1,y_1) \underset{A}{\to} (a_1x_1,a_1y_1)$. Defining $y_1/x_1 = \infty$ if $x_1 = 0$, we have that y_1/x_1 is a maximal invariant under $G_1^{(1)}$.

We now combine these results by using the facts that $G^{(p)}$ is generated by $G_p^{(p)},\ldots,G_1^{(p)}$ and that $G_q^{(p)}$ can be identified with $G_q^{(q)}$; $q = 1,\ldots,p$.

Consider for fixed $p > 1$ the induction hypothesis for induction on k, that $((x_1,\ldots,x_{p-k})',(y_1,\ldots,y_{p-k})')$ is an invariant statistic, equivalent to a maximal invariant under the group of transformations on $(\mathbb{R}^p)^2$ generated by $G_p^{(p)},\ldots,G_{p-k+1}^{(p)}$; $k = 1,\ldots,p-1$. We can identify $G_{p-k}^{(p)}$ acting on this invariant statistic with $G_{p-k}^{(p-k)}$. Then by an easy modification of Theorem 2 of Chapter 6 and the result just proved for $G_q^{(q)}$ for $k < p-1$, it follows that $((x_1,\ldots,x_{p-k-1})',(y_1,\ldots,y_{p-k-1})')$ is an invariant statistic, equivalent to a maximal invariant, under the group generated by $G_p^{(p)},\ldots,G_{p-k}^{(p)}$. For $k = p-1$ it follows that y_1/x_1 possesses the same properties. Since the induction hypothesis holds for $k = 1$, it follows that, for $p > 1$, y_1/x_1 is an invariant statistic, equivalent to a maximal invariant under $G^{(p)}$. For $p = 1$ this result has already been proved.

By Theorem 1 of Chapter 6, any invariant test is equivalent to a test which is a function of the invariant statistic $Z = Y_1/X_1$ which has the Cauchy distribution with location parameter 0 and scale parameter $\sqrt{\Delta}$. The likelihood ratio for testing $\Delta = \Delta_0$ against $\Delta = \Delta_1$ is therefore

$$\sqrt{\Delta_1/\Delta_0}\ (1+\Delta_0 z^2)/(1+\Delta_1 z^2),$$

which is easily verified to be a monotone increasing function of z^2 for $0 < \Delta_0 < \Delta_1$. Thus an UMP invariant test is to reject $H : \Delta \leq \Delta_0$ in favour of $K : \Delta > \Delta_0$ (or $K : \Delta \geq \Delta_1 > \Delta_0$) when $Y_1^2/X_1^2 > c$, where $c\Delta_0$ is the $1-\alpha$ percentile of the $F_{1,1}$ distribution (see also Problem 33 of Chapter 3).

(ii) By Problem 1 a maximin test does exist. By Theorem 2 (Hunt-Stein) and Lemma 2 of Chapter 8, if the group $G^{(p)}$ satisfies the conditions of Theorem 2, then there exists an almost invariant maximin test for testing $H : \Delta \leq \Delta_0$ against $K : \Delta \geq \Delta_1$ $(0 < \Delta_0 < \Delta_1)$. By Corollary 1 to Theorem 4 of Chapter 6 (p. 226 of the book) the UMP invariant test constructed in (i) is also UMP almost invariant, and hence will be the required maximin test.
We shall verify the conditions of the Hunt-Stein theorem by verifying it for the groups $G_p^{(p)}, G_{p-1}^{(p-1)}, \ldots, G_1^{(1)}$ in turn. As indicated in the remarks on p. 338, of the book, following Example 7, this is a valid procedure. As for the group $G_q^{(q)}$, we shall show, by following the hint, that it is isomorphic to a scale-translation group, and therefore by an extension of Example 7 the Hunt-Stein theorem holds for it too (the measurability assumptions of the theorem are everywhere trivially satisfied since we are working with continuous mappings in Euclidian spaces). Of course Example 7 has to be applied in two steps corresponding to the scale group and the translation group respectively.
As in the hint, let a' and b' be the bottom rows of two matrices A and B corresponding to two elements of $G_q^{(q)}$. Then the bottom row of the matrix AB has elements

$$(a_1 + a_q b_1, a_2 + a_q b_1, \ldots, a_{q-1} + a_q b_{q-1}, a_q b_q).$$

Now consider $a \in \mathbb{R}^q$ as the transformation on $\mathbb{R}^{q-1}$ defined by

$$(x_1, \ldots, x_{q-1})' \underset{a}{\rightarrow} (a_q x_1 + a_1, \ldots, a_q x_{q-1} + a_{q-1})'$$

(i.e. a scale change on all coordinates by the amount a_q, followed by a translation by the vector $(a_1, \ldots, a_{q-1})'$). Then we see that under the transformation a followed by the transformation b we have

$$(x_1, \ldots, x_{q-1})' \xrightarrow[b_0 a]{} (b_q(a_q x_1 + a_1) + b_1, \ldots, b_q(a_q x_{q-1} + a_{q-1}) + b_{q-1})'$$

$$= (b_a a_q x_1 + b_1 + a_1 b_q, \ldots, b_q a_q x_{q-1} + b_{q-1} + b_q a_{q-1})',$$

i.e. $G_q^{(q)}$ is isomorphic to the "transpose" of the group of positive or negative scale changes and translations. Generalizing Example 7 to the group of arbitrary translations (not just translations by $(g,\ldots,g)'$) and to positive and negative scale transformations and applying the remarks following Example 7, gives the desired result.

In the references to Chapter 8, LEHMANN (1950) is supposed to be relevant to this problem. This article contains a statement of the Hunt-Stein theorem together with the statement that it is applicable to translation groups, scale groups, finite groups, or products of such groups. For proofs of these statements the reader is referred to HUNT and STEIN (1946).

A more natural group to consider here would be the group of transformations by all nonsingular matrices A, not just matrices with zero elements above the diagonal. However Example 10 on p. 231 of the book shows that when $p = 2$, the only invariant size α test is $\varphi \equiv \alpha$

Problem 11.

Let φ be any test of size $\le \alpha$. By Theorem 2 of Chapter 8, there exists an almost invariant test ψ such that

$$\inf_{\overline{G}} E_{\bar{g}\theta}\varphi(X) \le E_\theta\psi(X) \le \sup_{\overline{G}} E_{\bar{g}\theta}\varphi(X) \qquad \text{for all } \theta \in \Omega.$$

By the right hand inequality, we have for any $\theta \in \Omega_H$

$$E_\theta\psi(X) \le \sup_{\overline{G}} E_{\bar{g}\theta}\varphi(X) \le \sup_{\theta'\in\Omega_H} E_{\theta'}\varphi(X) \le \alpha,$$

so ψ is also of size $\le \alpha$.

Now since φ_0 is an UMP almost invariant test of size α with respect to G it follows that for any $\theta \in \Omega_K$

$$w(\theta)E_\theta\varphi_0(X) + u(\theta) \ge w(\theta)E_\theta\psi(X) + u(\theta)$$

$$\ge w(\theta)\inf_{\overline{G}} E_{\bar{g}\theta}\varphi(X) + u(\theta) = \inf_{\overline{G}} [w(\bar{g}\theta)E_{\bar{g}\theta}\varphi(X) + u(\bar{g}\theta)]$$

$$\ge \inf_{\theta'\in\Omega_K} [w(\theta')E_{\theta'}\varphi(X) + u(\theta')].$$

Therefore φ_0 maximizes

$$\inf_{\theta\in\Omega_K} [w(\theta)E_\theta\varphi(X) + u(\theta)]$$

as was to be proved.

Section 5

Problem 12.

Let

$$\delta = \inf \left[\sup_{\theta\in\Omega_K} (\beta^*_\alpha(\theta) - \beta_\varphi(\theta)) \right],$$

where the infimum is taken over all level α tests φ of $H : \theta \in \Omega_H$. Let $\{\varphi_n\}_{n=1}^\infty$ be a sequence of level α tests such that

$$\lim_{n\to\infty} \sup_{\theta\in\Omega_K} (\beta^*_\alpha(\theta) - \beta_{\varphi_n}(\theta)) = \delta.$$

In view of the weak compactness theorem (Theorem 3 of the Appendix) there exists a subsequence $\{\varphi_{n'}\}$ which weakly converges to φ, say. This implies that for all $\theta \in \Omega_H$

$$E_\theta\varphi(X) = \lim_{n'\to\infty} E_\theta\varphi_{n'}(X) \le \alpha$$

and for all $\theta \in \Omega_K$

$$\begin{aligned}\beta^*_\alpha(\theta) - \beta_\varphi(\theta) &= \lim_{n'\to\infty} \beta^*_\alpha(\theta) - \beta_{\varphi_{n'}}(\theta) \\ &\le \lim_{n'\to\infty} \sup_{\theta\in\Omega_K} (\beta^*_\alpha(\theta) - \beta_{\varphi_{n'}}(\theta)) = \delta.\end{aligned}$$

Thus φ is a most stringent level α test.

Problem 13.

Let $\Omega_K = \cup\, \Omega_\Delta$ and $\beta^*_\alpha(\theta) = \text{constant} = \Delta$, say, for $\theta \in \Omega_\Delta$. Since φ_Δ maximizes the minimum power over Ω_Δ, we have that φ_Δ minimizes over all φ

$$\max_{\theta\in\Omega_\Delta} (\Delta - \beta_\varphi(\theta)) = \max_{\theta\in\Omega_\Delta} (\beta^*_\alpha(\theta) - \beta_\varphi(\theta)).$$

But φ_Δ does not depend on Δ. Therefore φ_Δ minimizes over all φ

$$\max_\Delta \max_{\theta\in\Omega_\Delta} (\beta^*_\alpha(\theta) - \beta_\varphi(\theta)) = \max_{\theta\in\Omega_K} (\beta^*_\alpha(\theta) - \beta_\varphi(\theta)).$$

Thus $\varphi = \varphi_\Delta$ is most stringent for testing $\theta \in \Omega_H$.

Problem 14.

We first show that the envelope power function of the permutation test of the problem is constant on each of the two-point sets Ω_Δ specified in

the hint. Therefore, we first note that the two sample problem with joint density given by (56) of Chapter 5 (complete randomization) is a special case, $c = 1$, of randomization within each of c subgroups having joint density (62) of Chapter 5. By the italicized remarks on p. 196, the most powerful test of the null hypothesis H against a simple alternative with $\xi > \eta$ is the one-sided permutation test (54) of Chapter 5 based on large values of $\Sigma_{j=1}^{n} Y_j$. By the remarks on p. 188 concerning the case $c = 1$, this is exactly the same as the one-sided permutation test (55) based on large values of the Student's t-statistic or equivalently on large values of $\overline{Y} - \overline{X}$. By changing the sign of all Y_j's and X_j's, we see on the other hand that the most powerful test of H against a simple alternative with $\xi < \eta$ is based on small values of the same statistic. Furthermore, for given m, n and σ, the power of these two tests against the corresponding alternatives depends only on $|\xi - \eta|$.
It follows from Problem 13 that the two-sided version of this test is most stringent if it is maximin for testing H against Ω_Δ. We shall verify the maximin property by showing that the test is most powerful against the symmetric mixture h' of the two densities in Ω_Δ, which one would expect to be the least favourable mixture. We shall take no mixture over distributions in the null-hypothesis, and need therefore the following obvious modification of a part of Theorem 1 on p. 327:

> *For any distribution* λ' *over* B', *let* $\varphi_{\lambda'}$ *be the most powerful test of the composite null-hypothesis* ω *against*
>
> $$h'(x) = \int_{\omega'} p_\theta(x) d\lambda'(\theta),$$
>
> *and let* $\beta_{\lambda'}$ *be its power against the alternative* h'. *If there exists* λ' *such that*
>
> $$\inf_{\omega'} E_\theta \varphi_{\lambda'}(X) = \beta_{\lambda'}, \tag{11}$$
>
> *then* $\varphi_{\lambda'}$ *maximizes* $\inf_{\omega'} E_\theta \varphi(X)$ *among all level* α *tests of the hypothesis* $H : \theta \in \omega$.

Note that the test we hope to find in this way, the two sided version of test (55) of Chapter 5, does have property (11), since its power against each alternative in $\omega' = \Omega_\Delta$ and hence also against h' is the same.
Now, by the remarks at the bottom of p. 195, when λ' assings

probability $\frac{1}{2}$ to each of the alternatives in Ω_Δ, the test φ_λ, is given by (52) of p. 185 with h(z), i.e. h'(x) above, given by

$$h(z) = \tfrac{1}{2}(\sqrt{2\pi}\,\sigma)^{-N}\Big\{\exp\Big[-\frac{1}{2\sigma^2}\Big(\sum_{j=1}^{m}(X_j-\xi_1)^2+\sum_{j=1}^{n}(Y_j-\eta_1)^2\Big)\Big]+$$
$$+\exp\Big[-\frac{1}{2\sigma^2}\Big(\sum_{j=1}^{m}(X_j-\xi_2)^2+\sum_{j=1}^{n}(Y_j-\eta_2)^2\Big)\Big]\Big\}$$
$$= \tfrac{1}{2}(\sqrt{2\pi}\,\sigma)^{-N}\Big\{\exp\Big[-\frac{1}{2\sigma^2}\Big(\sum_{j=1}^{m}(X_j-\zeta+\frac{n}{m+n}\delta)^2+\sum_{j=1}^{n}(Y_j-\zeta-\frac{m}{m+n}\delta)^2\Big)\Big]+$$
$$+\exp\Big[-\frac{1}{2\sigma^2}\Big(\sum_{j=1}^{m}(X_j-\zeta-\frac{n}{m+n}\delta)^2+\sum_{j=1}^{n}(Y_j-\zeta+\frac{m}{m+n}\delta)^2\Big)\Big]\Big\}$$
$$= \tfrac{1}{2}(\sqrt{2\pi}\,\sigma)^{-N}\exp\Big[-\frac{1}{2\sigma^2}\Big(\sum_{j=1}^{N}(Z_j-\zeta)^2+\frac{mn}{m+n}\delta^2\Big)\Big]\times$$
$$\times\Big\{\exp\Big[-\frac{\delta}{(m+n)\sigma^2}\Big(n\sum_{j=1}^{m}(X_j-\zeta)-m\sum_{j=1}^{n}(Y_j-\zeta)\Big)\Big]+$$
$$+\exp\Big[\frac{\delta}{(m+n)\sigma^2}\Big(n\sum_{j=1}^{m}(X_j-\zeta)-m\sum_{j=1}^{n}(Y_j-\zeta)\Big)\Big]\Big\}$$
$$= \tfrac{1}{2}(\sqrt{2\pi}\,\sigma)^{-N}\exp\Big[-\frac{1}{2\sigma^2}\Big(\sum_{j=1}^{N}(Z_j-\zeta)^2+\frac{mn}{m+n}\delta^2\Big)\Big]\times$$
$$\times\Big\{\exp\Big[-\frac{\delta}{mn(m+n)\sigma^2}|\overline{Y}-\overline{X}|\Big]+\exp\Big[\frac{\delta}{mn(m+n)\sigma^2}|\overline{Y}-\overline{X}|\Big]\Big\}.$$

Thus we see that the most powerful test of ω against h' is the permutation test with rejection region

$$|\overline{Y}-\overline{X}| > C[T(Z)].$$

In the article by LEHMANN and STEIN (1949), the result of Problem 13 is also given as a theorem quoted from HUNT and STEIN (1946).

REFERENCES

FERGUSON, T.S. (1967). *Mathematical Statistics, a Decision Theoretic Approach.* Academic Press. New York.

FLEMMING, W.H. (1965). *Functions of Several Variables.* Addison-Wesley.

HUNT, G. & C. STEIN (1946). Most stringent tests of statistical hypotheses. unpublished.

KIEFER, J. (1958). On the nonrandomized optimality and randomized non-optimality of symmetrical designs. *Ann. Math. Stat.* 29, 675-699.

KLAASSEN, C.A.J. (1979). Nonuniformity of the convergence of location estimators. *Proc. Second Prague Symp. Asymp. Stat.* North-Holland, Amsterdam, pp. 251-258.

LEHMANN, E.L. (1950). Some priciples of the theory of testing hypotheses. *Ann. Math. Stat.* 21, 1-26.

LEHMANN, E.L. (1955). Ordered families of distributions, *Ann. Math. Stat.* 26, 399-419.

LEHMANN, E.L. & C. STEIN (1949). On the theory of some nonparametric hypotheses. *Ann. Math. Stat.* 20, 28-45.

RAO, C.R. (1973). *Linear Statistical Inference and its Applications.* Wiley, New york.

RUIST, E. (1954). Comparison of tests for nonparametric hypotheses. *Arkiv. Mat.* 3, 133-163.

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.

1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.

1.3 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 3: statistiek.* 1966.

1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.

1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.

1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.

1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977.

1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.

1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.

1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.

1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.

2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.

2.2 L. Dekker, T.J. Dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.

3.1 H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967.

3.2 H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968.

3.3 H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968.

4 H.A. Lauwerier. *Representaties van groepen.* 1968.

5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.

6 K.K. Koksma. *Cursus ALGOL 60.* 1969.

7.1 *Colloquium moderne rekenmachines, deel 1.* 1969.

7.2 *Colloquium moderne rekenmachines, deel 2.* 1969.

8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.

9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970.

9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970.

10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970.

11 H. Bart, M.A. Kaashoek. H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.

12 T.J. Dekker. *Numerieke algebra.* 1971.

13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971.

14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971.

15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen. *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.

15.2 P.A. Beentjes, K. Dekker, H.C. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.

15.3 P.A. Beentjes, K. Dekker, P.W. Hemker, M. van Veldhuizen. *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.

16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973.

16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973.

17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.

17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.

17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.

18 F. van der Blij, H. Freudenthal, J.J. de Iongh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.

19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.

20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.

21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.

22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.

23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.

23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.

24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1974.

25 *Colloquium structuur van programmeertalen.* 1976.

26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.

26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.

27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.

28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.

29.1 J.C.P. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.

29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.

30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.

31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.

32 L. Geurts (red.). *Colloquium bedrijfsystemen.* 1976.

33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.

34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.

35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.

36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.

37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.

38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.

38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.

39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.

40 J. van Tiel. *Convexe analyse.* 1979.

41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.

42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.

43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations.* 1980.

45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur.* 1980.

46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1.* 1981.

46.2 P.G.M. Apers (red.). *Colloquium databankorganisatie, deel 2.* 1981.

47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices.* 1981.

47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations.* 1981.

47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I.* 1981.

47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II.* 1981.

47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I.* 1981.

47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II.* 1981.

47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation.* 1981.

48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1.* 1982.

48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2.* 1982.

49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups.* 1982.

50 H. Nijmeijer. *Inleiding systeemtheorie.* 1982.

51 P.J. Hoogendoorn (red.). *Cursus cryptografie.* 1983.

CWI SYLLABI

1 Vacantiecursus 1984 *Hewet - plus wiskunde.* 1984.

2 E.M. de Jager, H.G.J. Pijls (eds). *Proceedings Seminar 1981-1982. Mathematical structures in field theories.* 1984.

3 W.C.M. Kallenberg, et.al. *Testing statistical hypotheses: worked solutions.* 1984.

Addendum of the second edition of:
Testing statistical hypotheses: worked solutions (1987), CWI Syllabus 3.

PREFACE

It was during the 1978/1979 course that we decided to hold a seminar on Professor Lehmann's fundamental book "Testing Statistical Hypotheses". The objective we set ourselves was to solve all the problems. At some stage we concluded that these solutions might be worth publishing and that this could be done with some extra effort. (We now feel, though, that "some extra effort" is something of an understatement.)

The present text is based on the problems as they appear in the first (1959) edition of the book. Though the second edition (1986) with extra problems has appeared, we decided to confine ourselves to the first edition. To accommodate readers of the second edition in the matter of changed problem numbers, we include a separate addendum with a cross-reference list (see pages 311-319).

We thank Professor Lehmann for his support of our project, K. Snel for his excellent typing of the manuscript and the Centre for Mathematics and Computer Science (CWI) for giving us the opportunity to publish this syllabus. Our thanks are due also to all others who have contributed towards its realization. Among the participants it is Wilbert Kallenberg who deserves our special gratitude for doing most of the editorial work.

The members of the Lehmann group:

J. Beirlant
P. van Blokland
J.J. Dik
R.J.M.M. Does
A.J. van Es
R.D. Gill
P.L.J. Janssen
P.J.M. Kallenberg
W.C.M. Kallenberg
C.A.J. Klaassen
E. Opperdoes
C. van Putten
B.F. Schriever
M. Vandemaele
C.J. Warmer

July 1987

CROSS-REFERENCE LIST.

Part 1: Changes in new edition with respect to the old one.
Only those problems, which have been changed or renumbered in the new edition of Professor Lehmann's book Testing Statistical Hypotheses are mentioned in this cross-reference list.

CHAPTER 1.

Section#	problem# in new edition	changes
Section 2	1	*
Section 5	2	*
	3	**
	4	**
	6	ok/o7
Section 6	7	N
Section 9	18	*

CHAPTER 2.

Section#	problem# in new edition	changes
Section 2	3	N
Section 3	4	o3
	.	.
	.	.
Section 7	14	o13
	15	o14/**
	16	o15

CHAPTER 3.

Section#	problem# in new edition	changes
Section 2	1	N
	2	o1/*
	3	o2/**/N
	4	N
	5	o3
	.	.
	.	.
	11	o9
Section 3	12	N
	13	N
	14	o10
	15	o11/**
Section 4	16	o12
	.	.
Section 7	26	o22
	27	N
	28	o23/*
	29	o24/*
	30	o25/**
	31	o26
		.
Section 9	38	o33
Add. Prob.	39	N
	53	N

CHAPTER 4.

Section #	problem # in new edition	changes
Section 1	2	*
Section 3	12	N
	13	N
Section 5	18	**
	19	N
	20	o19
	21	o20
	22	N
	23	o21
Section 8	24	N
	⋮	⋮
Add. Prob.	36	N

CHAPTER 5.

Section #	problem # in new edition	changes
Section 2	8	**
Section 4	14	N
	⋮	⋮
Section 5	24	N
Section 6	25	o14
	26	o15
	27	o16
	28	o17
	29	N
	30	N
Section 7	31	o18
	32	o19
Section 8	33	o20
Section 9	34	N
	⋮	⋮
	43	N
Section 10	44	o21
Section 11	45	o22
	46	o23/**
Section 12	47	o24
	⋮	⋮
Section 13	54	o31
Section 14	55	N
	⋮	⋮
	63	N
Section 15	64	o32
	65	o33
	66	o34
	67	o35
	68	o36/*
	69	N
	⋮	⋮
Add. Prob.	79	N

CHAPTER 6.

Section #	problem # in new edition	changes
Section 4	9	o10
	10	o11
	11	o12
Section 5	12	o13
Section 6	13	N
Section 7	20	N
	.	.
	.	.
	26	N
Section 9	27	o20
	.	.
	35	o28/**
	.	.
	39	o32/***
	.	.
Section 10	43	o36/**
	.	.
	48	o41/*
	.	.
Section 11	51	o44/o45
	52	N
	53	o46
	54	o47
Section 12	55	N
	.	.
	.	.
	65	N
Section 13	66	o48
	67	o49
Add. Prob.	68	N
	.	.
	.	.
	81	N

CHAPTER 7.

Section #	problem # in new edition	changes
Section 1	4	*
	6	N
	7	N
Section 2	8	o6
	9	o7
	10	o8
	11	N
	12	N
Section 3	13	o9
	14	o10
	15	o11/***
	16	N
	.	.
	.	.
	21	N
Section 6	22	o12
	23	N
	24	N
	25	o13
	.	.
	.	.
Section 7	30	o18
Section 8	31	N
	.	.
	.	.
Section 10	47	N
Section 11	48	o19
	.	.
	.	.
Section 12	52	o23
	53	N
	.	.
	.	.
Add. Prob.	71	N

CHAPTER 8.

Section #	problem # in new edition	changes
Section 2	1	o7.24
	.	.
	.	.
	7	o7.30/*
Section 3	8	N
	.	.
	.	.
Section 6	34	N
Section 7	35	o7.33/**
	36	N
Section 8	37	o7.34
	38	N
	.	.
	.	.
Add. Prob.	44	N
	45	o7.32/*
	46	N
	47	N

CHAPTER 9.

Section #	problem # in new edition	changes
Section 1	1	o8.1
	2	o8.2/***
	3	N
Section 2	4	o8.4
	.	.
	.	.
	9	o8.9/**
Section 3	10	N
	.	.
	.	.
Section 4	18	N
Section 5	19	o8.10
	.	.
	.	.
Section 6	23	o8.14
Add. Prob.	24	N
	.	.
	.	.
	34	N

CHAPTER 10.

Section #	problem # in new edition	changes
Section 1	1	N
	.	.
	.	.
Section 4	30	N

Explanation of the symbols used in the preceding tables.

Symbol	meaning
ok	Nothing has been changed
N	This problem is a new one
o6	This problem was problem 6 of the same chapter in the old edition
o7.6	This problem was problem 6 of chapter 7 in the old edition
*	Only small or irrelevant changes have been made
**	A few relevant changes have been made
***	The problem has been changed almost completely

Part 2: Notes about the changes mentioned in the preceding part of the cross-reference list.

Chapter # Section #	problem # in new edition	changes	Note
Section 1.5	3	**	3 (ii) : 'Let Ω be connected' has been skipped
Section 1.5	4	**	The general case has been skipped
Section 1.5	6	ok/o7	6 (iii) is problem 7 from old edition
Section 1.9	18	*	The density has been changed
Section 2.7	15	o14/**	'For any θ ... space' has been added
Section 3.2	2	o1/*	A hint for (i) has been added
Section 3.2	3	o2/**/N	The density has been changed; (iii) to (vi) have been added
Section 3.3	15	o11/**	'then' has been replaced by 'if and only if'
Section 3.7	27	N	This problem contains the definitions used in 28 and 29
Section 3.7	28	o23/*	'Polya type' has been replaced by 'STP'
Section 3.7	29	o24/*	'Polya type' has been replaced by 'STP'
Section 3.7	30	o25/**	The conditions have been changed
Section 4.1	2	*	'critical levels' has been changed in 'p-values'
Section 4.5	18	**	Part (iii) has been changed
Section 5.2	8	**	Part (ii) has been skipped
Section 5.11	46	o23/**	Part (ii) has been added and the hint has been extended
Section 5.15	68	o36/*	'in the sense of Section 11' has been skipped
Section 6.9	35	o28/**	'A level ... integer' has been added
Section 6.9	39	o32/***	Part (i) has been changed and part (iii) has been added
Section 6.10	43	o36/**	'under ... conditions' has been added
Section 6.10	48	o41/*	'regression' has been added
Section 6.11	51	o44/o45	Part (ii) is problem 45 in old edition
Section 7.1	4	*	Another formulation has been used
Section 7.3	15	o11/***	For part (i) and part (ii) another formulation has been used and part (iii) has been added
Section 8.2	7	o7.30/*	The title 'Null ... T^{2}' has been added
Section 8.7	35	o7.33/**	The hint has been skipped
Ad. Pr. Ch. 8	45	o7.32/*	The title 'Testing ... independence' has been added
Section 9.1	2	o8.2/***	Parts (i), (ii) and (iii) have been added and parts (ii) and (iii) in the old edition have become parts (iv) and (v) in the new on
Section 9.2	9	o8.9/**	'provided ... nonrandomized' has been added

Part 3: Changes in the numbering of the problems in the old edition with respect to the new one.

Those problems in the old edition that also appear in the new one under the same number are not mentioned in this part of the cross-reference.

CHAPTER 1.

problem # in old edition	problem # in new edition
7	6 (iii)

CHAPTER 2.

problem # in old edition	problem # in new edition
3	4
.	.
.	.
15	16

CHAPTER 3.

problem # in old edition	problem # in new edition
1	2
2	3
3	5
.	.
.	.
9	11
10	14
.	.
.	.
22	26
23	28
.	.
.	.
33	38
34	X
.	.
.	.
39	X

CHAPTER 4.

problem # in old edition	problem # in new edition
12	X
13	X
19	20
20	21
21	23

CHAPTER 5.

problem # in old edition	problem # in new edition
14	25
15	26
16	27
17	28
18	31
19	32
20	33
21	44
.	.
.	.
31	54
32	64
.	.
.	.
36	68

CHAPTER 6.

problem # in old edition	problem # in new edition
9	X
10	9
11	10
12	11
13	12
20	27
.	.
.	.
43	50
44	51 (i)
45	51 (ii)
46	53
47	54
48	66
49	67

CHAPTER 7.

problem # in old edition	problem # in new edition
6	8
7	9
8	10
9	13
10	14
11	15
12	22
13	25
.	.
.	.
18	30
19	48
.	.
.	.
23	52
24	8.1
	.
30	8.7
31	X
32	8.45
33	8.35
34	8.37
35	X

CHAPTER 8.

problem # in old edition	problem # in new edition
1	9.1
2	9.2
3	X
4	9.4
.	.
.	.
9	9.9
10	9.19
.	.
.	.
14	9.23

Explanation of the symbols used in the preceding tables.

Symbol	meaning
4	Problem 4 in new edition
8.1	Problem 1 of chapter 8 in new edition
X	This problem has been skipped in the new edition