J. Sci. I. R. Iran

Vol. 1, No. 2
Winter, 1990

THE EMPIRICAL BAYES METHOD OF ANALYSIS OF A SERIES OF EXPERIMENTS

M.R. Meshkani

Department of Mathematics, Faculty of Science, Shahid Beheshti University, Tehran, Iran

L. Billard

University of Georgia, Athens, Georgia, U.S.A.

Keywords: Series of Experiments, Prior, Posterior, Bayes

Abstract

The classical method of analysis of a series of experiments is somewhat involved in being conditional on various, occasionally unrealistic, assumptions such as homogeneity of variances of experimental error, lack of interactions of treatments and places, etc. In this work, we adopt a Bayesian view to account for such heterogeneities. Our appoach is illustrated by a real series of experiments regarding the effect of fertilizers on crops of sugarbeet in various places of Iran. Conjugate priors are considered for parameters of the model and posteriors are obtained. Using information from different places, the unknown priors are estimated and the empirical Bayes procedure is employed.

1. Introduction

In many research programs it is quite common to repeat the same experiment at a number of different locations and/or on a number of different occasions. Sometimes it is desired to produce a set of recommendations which is to apply to a population that is extensive either in space or in time, or in both. In agricultural field experimentation, examples of this practice abound. The conclusions drawn from these series, if they are to be applied, must be valid for at least several seasons in the future and over a reasonably large area of farm land. For example, it has been found that the effectiveness of the common fertilizers on an extensively cultivated crop, such as sugarbeet, varies from field to field and, even more markedly, from season to season. A single experiment, however well conducted, provides information about only one location and one season. Series of experiments conducted at several different locations in the area for which recommendations are sought, can remedy the situation in supplying information about the whole area.

Our particular motivations in this study is the analysis of a series of agronomic experiments conducted in ten provinces of Iran to investigate the effects of different fertilizers on the yield of sugarbeet. There were two fertilizers, four levels of nitrogen and three levels of phosphate. At each location, the experiment consisted of a randomized block design with three replicates of each of the twelve fertilizer treatment combinations. We confine our attention to the statistical analysis.

The classical method of analysis of a series of experiments is somewhat involved in being conditional on various, occasionally unrealistic, assumptions such as homogeneity of variances of experimental error across each experiment in the series, lack of interactions between treatments and places, treatments and times, etc., that is, there is no interaction between the treatments and the particular experiment in the series itself (be that a series across time, locations etc.). Two major problems arise. First the experimental error variances do in fact often differ from place to place. This can be tested via Bartlett's test of homogeneity of

variances. If the variances are heterogeneous, the usual F tests are not valid and suitable transformations or separate analyses for each location, or sets of homogeneous locations, are required. Secondly, the treatment and location interactions may not be homogeneous especially in factorial experiments. Some factors may behave in a stable way, while others are more erratic. In this case it is usually advisable to analyze the experiments separately for each homogoneous set [9]. As we shall see below both of these complications are present in our sugarbeet data, thus prompting us to find a more efficient mode of analysis.

These difficulties, most especially the problems associated with intrinsic heterogeneities in the variances of the underlying error variables, can be accommodated by utilizing the techniques of Bayesian analysis. A direct Bayes approach suffers in that it is dependent upon subjective choice of prior distributions which may not be acceptable to all experimenters interested in the results of the subsequent analysis. To avoid this problem we therefore select the natural conjugate prior and estimate its parameters from data gathered on previous experiments in the series, thus implementing the ideas of empirical Bayes techniques to draw inferences about the single experiment at hand.

The empirical Bayes approach to statistical decision problems is applicable when we encounter the same decision problem in the sequence repeatedly and independently, with a fixed but unknown prior distribution for the parameter. It is not expected that all decision problems will be embedded in such a sequence. However, when they are, the empirical Bayes approach offers certain advantages over any aproach which ignores the fact that the parameter is itself a random variable. This approach also has advantages over approaches which assume a personal prior not changing with experience, [6]. See also [3,5] for a nice review of empirical Bayes techniques.

In particular, the assumptions of empirical Bayes procedures are especially suitable to the framework of series of experiments.

Before presenting our analysis (in Section 3) we first provide, in Section 2, some basic results for Bayes estimators and empirical Bayes estimators resulting from the natural conjugate priors. We estimate the treatment effects and perform various contrasts. Our analysis includes a comparison of results obtained by using maximum likelihood methods. We conclude that the empirical Bayes approach has smaller Bayes risk than does the maximum likelihood estimator and therefore should be considered seriously for use in the analysis of data of this type.

2. Theoretical Background

We consider the problem as a Bayesian decision theoretic problem. Let $\lambda=(\theta,\Sigma)$ be the parameters of the distribution from which we have some observations. We consider the loss function L (d, λ), where d is the decision to be made about the λ. Here, the decision rule is an estimator to be used for estimating the mean of a set of multivariate normal vectors whose variances are unknown and presumably not the same. A joint prior distibution, $\pi(\lambda)$ with some metaparameters $\delta=(\mu,\Sigma)$ is assumed. the competing decision rules will be judged according to their overall expected loss (i.e., Bayes risk), denoted by w (d, π). In the case of squared error loss, the Bayes estimator is the mean of posterior distribution which depends on the metaparameters. These metaparameters can be estimated from the marginal distribution of the data, which would give the empirical Bayes (EB) estimator, d_n. Such an EB estimator would share the optimal property of the Bayes estimator, given the metaparameters are efficiently estimated.

2.1. Conjugate priors for the parameters of the multinormal distribution.

Following Ando and Kaufman [2] and Press and Rolph [5], we obtain the conjugate priors of the multinormal distribution. Suppose Y is a normally distributed p-vector, with all parameters unknown. For a random sample of N observations,

$$p(Y_1,\ldots,Y_N|\theta,\Sigma)\propto|\Sigma|^{-\frac{N}{2}}\exp(-Q), \qquad (2.1)$$

where

$$Q=\frac{1}{2}\sum_{i=1}^{N}(Y_i-\theta)'\Sigma^{-1}(Y_i-\theta).$$

or setting $\Delta=\Sigma^{-1}$ and expanding Q, we have

$$p(Y_1,\ldots,Y_N|\theta,\Delta)$$
$$\propto[|\Delta|^{\frac{1}{2}}\exp(-Q_1)][|\Delta|^{\frac{N-1}{2}}\exp(-Q_2)] \qquad (2.2)$$

where

$$Q_1=N(\theta-\bar{y})'\Delta(\theta-\bar{y})/2,\quad Q_2=[\sum_{i=1}^{N}Y_i'\Delta Y_i-N\bar{Y}'\Delta\bar{Y}]/2$$

Interchanging the role of Y and θ in Q_1, we can show the conditional conjugate prior of θ is

$$\pi_1(\theta|\Delta)=N(\mu,\Sigma)$$

where μ is an arbitrary p-vector, the mean of the prior distribution. The joint prior of (θ,Δ) is

$$\pi(\theta,\Delta)=\pi_1(\theta|\Delta)\pi_2(\Delta).$$

To specify $\pi_2(\Delta)$, we note that the second term in the right hand side of (2.2) can be written as

$$|\Delta|^{\frac{N-1}{2}}\exp\{-\mathrm{tr}[\sum_{i=1}^{N} Y'_i\Delta Y_i - NY'\Delta Y]/2\}$$
$$=|\Delta|^{\frac{N-1}{2}}\exp\{-\frac{1}{2}\mathrm{tr}\Delta V\}$$

where tr denotes the trace of a matrix and where

$$V=\sum_{i=1}^{N} Y_i Y'_i - N\bar{y}\bar{y}'.$$

Now, by considering Δ to be a variable matrix, we assume it follows a Wishart distribution W(G,p,m), i.e.,

$$\pi_2(\Delta)\propto|G|^{-\frac{m}{2}}|\Delta|^{\frac{m-p-1}{2}}\exp\{-\mathrm{tr}\Delta G^{-\frac{1}{2}}\},\ m\geq p.$$

Thus, we have an inverse Wishart prior with arbitrary parameters G and m for the distribution of the covariance matrix $\Sigma=\Delta^{-1}$.

After some algebra, we arrive at the following conjugate prior for (θ,Σ)

$$\pi(\theta,\Sigma)\propto|\Sigma|^{-\frac{v+1}{2}}\exp\{-(\theta-\mu)'\Sigma^{-1}(\theta-\mu)/2+\mathrm{tr}\Sigma^{-1}G^{-1}\} \quad (2.3)$$

with parameters $v=m+p+1>2p$, μ, and G. For details of this derivation, see [4].

2.2. Bayes estimators for the multinormal parameters.

We assume the squared error loss function. Then, using well established techniques we have the Bayes estimator for (θ,Σ) as (see [1])

$$\hat{\theta}_B=E(\theta|\bar{y},V)=\mu+N\bar{y})/(N+1),\ N+m-2p>1, \quad (2.4)$$

and for $N+m-2p-2>0$,

$$\hat{\Sigma}_B=E(\Sigma|\bar{y},V)=[V+G+N(N+1)^{-1}(\mu-\bar{y})(\mu-\bar{y})]/(N+m-2p-2), \quad (2.5)$$

where N is the sample size, and where

$$\bar{y}=N^{-1}\sum_{i=1}^{N} y_i,\ V=(V_{kh}),\ k,h,=1,\ldots,p,$$

with

$$V_{kh}=\sum_{i=1}^{N}(y_{jk}-\bar{y}_k)(y_{jh}-\bar{y}_h)$$

are the sample mean vector and convariance matrix, respectively.

It seems unreasonable to assume that the parameters G,μ, and m are all known or that occasionally all persons interested in this experiment can agree on what values these parameters might assume. The above estimators are only applicable when one can supply some sort of estimates for the prior distribution. Thus we estimate these parameters to give us empirical Bayes estimators.

2.3. Empirical Bayes estimators

It is assumed that the data consist of a p-vector of observations Y_{ij} from each of the (n+1) locations, each location having N replicates, i=1,2,..., N, j= 1,2,...,n; n+1. Each time, the observations from one location are considered for inference about their (θ,Σ) and the rest of observations so called «past data» are used to estimate G,μ, and m.

Thus, for the particular location of interest (i=n+1), there is a pair of sufficient statistics $(\bar{Y},V)$ whose marginal distribution involves the parameters G, μ, and m. We need to estimate this parameters using as our data, the sample of n observations obtained from the locations i= 1,...,n. Upon substituting the estimates G,μ, and m in (2.4) and (2.5), we would obtain the corresponding empirical Bayes estimators. Thus, we need to derive the marginal density of $(\bar{Y},V)$ and thence estimate its parameters. To this end, some results from Siegel [8] are helpful.

Siegel [8] introduces a class of integral identities as generalizations of beta and gamma functions. Using his results repeatedly, one can find the marginal probability density function of $(\bar{Y},V)$ as follows. Note that

$$P(\bar{Y},V,\theta,\Sigma)=P(\bar{Y}|\theta,\Sigma)p(V|\theta,\Sigma)\pi(\theta,\Sigma)$$
$$=N(\theta,N^{-1}\Sigma)W(\Sigma,p,N-1)\pi(\theta,\Sigma).$$

This leads to

$$p(\bar{Y},V)=\propto|V|^{(N-p-2)/2}.|V+G+N(N+1)^{-1}(\bar{Y}-\mu)(\bar{Y}-\mu)'|^{-(N+m)/2}$$

which yields the marginal probability density function of $\bar{Y}$ and of V, as, respectively,

where

$$K = (m-2p)^{(m-p)/2}\pi^{-p/2}[N/(N+1)]^{p/2}\Gamma\{(m-p)/2\}\,[\Gamma\{(m-2p)/2\}]^{-1}.$$

and

$$p(V) = |G|^{(m-p-1)/2}|V|^{(N-p-2)/2}|V+G|^{-(N+m-p-2)/2}[B_p\{(N-1)/2,\,(m-p-1)/2\}]^{-1}$$

where for an r×r matrix Z, with a, b> (r-1)/2,

$$B_r(a,b) = \Gamma_r(a)\;\Gamma_r(b)\;[\Gamma_r(a+b)]^{-1} = \int |Z|^{a-(r+1)/2}|I+Z|^{-(a+b)}dZ$$

and

$$\Gamma_r(a) = \pi^{r(r-1)}\Gamma(a)\Gamma(a-\tfrac{1}{2})\ldots\Gamma(a-(r-1)/2).$$

Similarly, it can be shown that, for m> p,

$$p(Y) = K[2N/(N+1)]^{-p/2}\,|G|^{-1/2}\,[(m-2p)\{1+2^{-1}(Y-\mu)'G^{-1}(Y-\mu)\}]^{-(m+p)/2}$$

which is a multivariate t-distrubution with m-2p degrees of freedom. Hence, for m-2p > 1 it is known that E(Y) $=\mu$ and for m-2p > 2,

$$\mathrm{Var}\,(Y) = 2(m-2p-2)^{-1}G \qquad (2.6)$$

Here, we adopt a simple approach. Taking m = 2p+ 3, to insure that the constraint of (2.6) holds, we estimate μ and G, respectively, by

$$\hat{\mu} = (nN)^{-1}\Sigma_{i=1}^{n} Y_{ij} = n^{-1}\Sigma_{j=1}^{n}(\bar{Y}_j = \bar{\bar{Y}}), \qquad (2.7)$$

$$\hat{G} = (2nN)^{-1}\Sigma_{j=1}^{n}(Y_{ij}-\bar{\bar{Y}})(Y_{ij}-\bar{\bar{Y}})'. \qquad (2.8)$$

We have the empirical Bayes (EB) estimators of θ and Σ respectively as

$$\hat{\theta}_{EB} = (\bar{\bar{Y}}+N\bar{Y})/(N+1) \qquad (2.9)$$

$$\hat{\Sigma}_{EB} = \{V+\hat{G}+\frac{N}{N+1}(\bar{Y}-\bar{\bar{Y}})(\bar{Y}-\bar{\bar{Y}})'\}/(N+1),\ N>1. \qquad (2.10)$$

Note that $\bar{Y}$ and $\bar{\bar{Y}}$ are independent statistics. The competing maximum likelihood (ML) estimators will be, respectively,

$$\hat{\theta}_{ML} = \bar{Y}, \qquad (2.11)$$

$$\hat{\Sigma}_{ML} = V/(N-1). \qquad (2.12)$$

2.4. Comparison of the EB and ML estimators.

The Bayes risk is the overall expected loss with respect to all random variables involved. Denote the Bayes risks of Bayes, EB, and ML estimators by $W(\hat{\theta}_{EB})$, and $W(\hat{\theta}_{ML})$, respectively. Then straightforward computation shows (using, for example, Searle [7] to evaluate the expectation of quadratic forms) that

$$W(\hat{\theta}_B) = N/(N+1)^2\,\mathrm{tr}\,E(\Sigma), \qquad (2.13)$$

$$W(\hat{\theta}_{EB}) = (Nn+1)/[Nn(N+1)]\,\mathrm{tr}\,E(\Sigma), \qquad (2.14)$$

and

$$W(\hat{\theta}_{ML}) = N^{-1}\,\mathrm{tr}\,E(\Sigma). \qquad (2.15)$$

From (2.13), (2.14) and (2.15), it is clear that for all n and N,

$$W(\hat{\theta}_B) < W(\hat{\theta}_{EB}) < W(\hat{\theta}_{ML}).$$

Furthermore,

$$W(\hat{\theta}_{EB}) = (Nn+1)\,W(\hat{\theta}_{ML})/[n(N+1)].$$

Since the diminishing factor (DMF),(Nn+1)/[n(N+1)] is always less than one, it follows that the EB estimator is always more precise than is the MLE. As n ∞, this factor aproaches N/(N+1) and when N$\to\infty$ it tends to 1. That is, in experiments with a large number of replications the amount of gain in the Bayes risk is not significant. However, in experiments with few such the 2 to 4 replications commonly seen in practice, the gain in risk is considerable suggesting that in these cases, the EB estimator should be sought in preference to the ML estimator. From Efron and Morris [3], these facts can be rephrased in terms of the relative savings risk (RSR) of the EB estimator as

$$RSR(\hat{\theta}_{EB}) = W(\hat{\theta}_{EB})\text{-}W(\hat{\theta}_B)/[W(\hat{\theta}_{ML})\text{-}W(\hat{\theta}_B)] = [N(n+1)+1]/n(2N+1) = 1\text{-}(N+1)(n-1)/n(2N+1).$$

That is , $0 \leqslant RSR\ (\hat{\theta}_B) < 1$. The same conclusions that the EB estimator is preferable are again drawn. Table 2.1 shows the DMF and RSR for a few values of n and N illustrating this fact. Hence, it is concluded that for small N and/or large n the empirical Bayes approach renders a «better» estimator than does the maximum likelihood. This is intuitively reasonable since in an experiment with few replications, one would be better off to incorporate external relevant information where possible. The fewer replicates one has, the more valuable the external information becomes.

Table 2.1. DMF and RSR for selected values of N and n.

	DMF					RSR				
n	N=2	3	4	5	∞	N=2	3	4	5	∞
2	.83	.88	.90	.92	1.00	.70	.71	.72	.73	.75
3	.78	.83	.87	.89	1.00	.60	.62	.63	.64	.67
4	.75	.81	.85	.88	1.00	.55	.57	.58	.59	.63
5	.73	.80	.84	.87	1.00	.52	.54	.56	.56	.60
9	.70	.78	.82	.85	1.00	.47	.49	.51	.52	.56
∞	.67	.75	.80	.83	1.00	.40	.43	.44	.45	.50

3. An Application

The Soil Institute of Iran routinely conducts various experiments either at provincial agricultural stations or on farmers' land to provide some set of recommendations for various crops. We illustrate the techniques described earlier by analyzing data collected to investigate the effects of the appliction of various fertilizers on the productivity of sugarbeet. The experiment involved twelve treatment combinations of four levels of nitrogen (N) with four levels of phosphate (P). The four levels of nitrogen (15,30,45,60 kilograms per hectare) are denoted by subscripts 1,2,3, and 4 on N and likewise four levels of phosphate (0, 20, 40, and 60 kilograms per hectare) by subscripts on P. Combinations N_1P_3, N_2P_3, N_3P_o, and N_4P_1 were excluded from the experiment.

The preliminary analysis, Table 3.2, consisted of the traditional analysis of variance at each location, according to the standard model

$$Y=X\beta+ e,\ e \sim N(0,\sigma^2 I).$$

In this model X is the usual design matrix and ß is the appropriate vector of parameters for a complete block design model. The mean sum of squares and the calculated F statistics for each case are summarized in Table 3.1. Since in the second location $8.23 > 2.79 = F_{11,22,.01}$ we observe that there is a significant difference at the 1% level of significance across the different treatments. Also, there is a significant difference at the 5% level in the different treatments in the third location. Likewise, there were significant differences at the 1% level of significance in the effect of blocking in locations 8,9 and 10.

We conduct a test of homeogeneity of variance using the Bartlett's criteria M/C where

$$M= 2.3026f\,(a \log \bar{s}^2-\Sigma \log s_i^2),$$

$$C= 1+(a+ 1)/(3af),$$

with

$$\bar{s}^2= a^{-1}\Sigma\, s_i^2$$

and a (=10) being the number of estimates of the error mean square (s_i^2) each with f (=22) degrees of freedom. Thus for our data, substituting the appropriate values from Table 3.1, we have $\Sigma \log s_i^2= 17.1251$, $\bar{s}^2= 60.94$, and hence M=18.4342, C=1.0167 and therefore M/C= 18.1314. Since $18.1314 > \chi^2_{9,.05}= 16.92$ and $18.1314 < \chi^2_{9,.01} =21.67$ we conclude there is a significant difference across the variances at the 5% but not at the 1% level of significance. Hence, with this indication of heterogeneity in the variances, a collective analysis of the experiments is not advisable. A better (relative to the Bayes risk) estimate of the vector of parameters θ can be obtained from the empirical Bayes procedure,

Table 3.1. Preliminary analyses of variance and Bartlett's test.

Place	df	1	2	3	4	5	6	7	8	9	10	Total
Ms Treat.	11	73.45	219.96	54.48	51.23	197.61	104.19	172.26	125.36	39.88	77.52	
Ms Block	2	48.15	39.01	3.95	9.84	96.11	186.10	293.44	657.21	3684.05	401.58	
Ms Error	22	39.52	26.72	20.04	28.70	99.24	57.46	61.11	61.11	60.20	72.98	609.42
F(T)		1.86	8.23	2.72	1.79	1.72	1.81	1.20	2.05	0.66	1.06	
F(B)		1.22	1.46	0.20	0.33	0.97	3.24	2.05	10.75	61.20	5.50	
Log MSE		1.5968	1.4268	1.3019	1.4579	1.9967	1.7594	2.1567	1.7861	1.7796	1.8632	17.1251

Table 3.2. The mean vectors of observations ($\bar{Y}_j$) from 3 replicates in 10 locations.

Treat	1	2	3	4	5	6	7	8	9	10
N_1P_0	70.18	53.43	45.00	44.27	70.87	64.70	72.97	51.33	76.53	72.63
N_1P_1	76.98	57.80	41.80	53.07	76.50	61.90	64.30	49.70	65.83	71.93
N_1P_2	71.58	61.26	38.63	50.27	60.33	60.33	77.07	50.10	67.67	74.27
N_2P_0	71.73	58.83	42.07	49.27	69.13	74.73	89.50	63.67	66.33	71.83
N_2P_1	78.53	72.05	41.37	48.27	72.20	67.63	77.67	62.63	74.10	80.90
N_2P_2	75.28	65.83	44.90	50.80	79.03	74.67	78.07	61.03	73.37	80.33
N_3P_0	75.48	62.63	41.73	55.67	68.80	63.70	81.27	57.73	75.57	75.93
N_3P_1	81.48	70.67	37.40	51.57	80.47	75.10	83.33	65.63	73.93	75.17
N_3P_2	79.88	68.32	34.53	49.97	76.23	59.93	69.50	64.90	72.93	83.33
N_3p_3	80.10	67.58	41.00	56.77	84.97	74.37	86.26	68.70	72.00	83.53
N_4P_2	84.17	70.37	43.67	57.53	78.60	69.97	82.10	62.47	72.00	83.20
N_4p_3	85.53	70.62	49.57	55.97	87.33	70.67	88.30	64.07	72.67	85.43

that is $\hat{\theta}_{EB}$.

Before embarking on our empirical Bayes estimation we first adjusted the data for block effects whenever necessary (locations 8, 9, 10) to reduce the parameters to those for treatment effects only. We then develop the empirical Bayes estimate, $\hat{\theta}_{EB,k}$, for each location k= 1,...,10 where in the terminology of Section 2, the kth location corresponds to the k= n+1 «current» data set and j= 1,..., n=9 «past» data sets are the data of the other nine locations. For each location there are N= 3 replicates. Table 3.2 provides the mean vector of observations $\bar{Y}_k$, k= 1,..., 10, that is, the maximum likelihood estimation $\hat{\theta}_{ML,k}$ (see equation (2.9)). Table 3.3 gives us the empirical Bayes estimate for the vector of parameters $\hat{\theta}_{EB,k}$ k= 1,..., 10, obtained from equation (2.7) where in (2.7) $\bar{Y}$ is replaced by $\bar{Y}_k$ for each k. As we compare the estimates of these treatment effects in Tables 3.2 and 3.3, we observe that shrinkage has resulted either positively or negatively depending on the particular location. For example, the empirical Bayes estimators in location 1 are uniformly less than are the maximum likelihood estimators, while in contrast in the third region, the empirical Bayes estimators are all larger than the corresponding maximum likelihood estimators.

In some locations the shift can be either positive or negative, for example, in location 5, when nitrogen is present in either 15 or 30 kilograms per hectare, the empirical Bayes estimator is larger than the maximum likelihood estimator but the reverse holds when nitrogen is present at 45 or 60 kilograms per acre.

It can be shown that an empirical Bayes estimate is a convex combination of the maximum likelihood estimate and the prior mean of the parameter being

Table 3.3. The empirical Bayes estimates of mean vector for 10 locations $\hat{\theta}_{EB}$

Treat	Place	1	2	3	4	5	6	7	8	9	10
1	N_1P_0	68.18	55.62	49.30	48.75	68.70	64.07	70.27	54.05	72.95	70.02
2	N_1P_1	73.23	58.85	46.85	55.30	72.87	61.92	63.27	52.77	64.87	69.45
3	N_1P_2	68.98	61.23	44.26	52.99	60.54	60.54	73.09	52.84	66.04	70.99
4	N_2P_0	70.23	60.56	47.99	53.39	68.29	72.49	83.56	64.31	66.19	70.31
5	N_2P_1	75.78	70.92	47.91	53.09	71.04	67.61	75.14	63.86	72.44	77.56
6	N_2P_2	73.54	66.46	50.76	55.18	76.36	73.08	75.63	62.86	72.11	77.33
7	N_3P_0	73.07	63.44	47.76	58.21	68.06	64.24	77.41	59.76	73.14	73.41
8	N_3P_1	78.48	70.37	45.42	56.04	77.72	73.69	79.87	66.52	72.02	73.74
9	N_3P_2	76.34	67.67	42.33	52.16	73.61	61.38	68.56	65.11	71.13	78.93
10	N_3P_3	77.96	68.57	48.63	60.46	81.61	73.66	62.58	69.41	71.88	80.53
11	N_4P_2	80.73	70.38	50.35	60.75	76.55	70.08	79.18	64.45	71.60	80.00
12	N_4P_3	82.48	71.30	55.51	60.31	83.83	71.33	84.56	66.38	75.08	82.41

Table 3.4. Sample Variance of $\hat{\theta}_{EB,1}$, $\hat{H}_1$.

$H_1 = .38$

112.32	76.67	73.50	69.99	85.00	81.29	88.75	64.45	79.37	80.40	80.29	94.17
76.67	87.23	67.95	50.80	75.27	69.56	70.83	85.39	79.10	78.03	78.76	81.28
73.50	67.95	105.76	64.43	86.45	65.22	83.39	98.29	71.20	79.95	86.01	91.78
69.99	50.80	64.43	106.80	73.56	65.93	79.54	87.46	77.85	82.33	74.64	95.12
85.00	75.27	86.45	73.56	104.41	72.68	88.23	97.98	91.16	92.32	84.85	96.00
81.29	69.65	65.22	65.93	72.68	100.39	68.01	90.04	79.29	76.65	86.32	82.45
88.75	00.83	83.39	79.54	88.23	68.01	107.44	102.34	87.34	91.02	89.45	99.75
64.45	85.39	8.29	87.46	97.98	90.04	102.34	154.49	115.40	109.29	114.38	116.41
79.37	79.10	71.20	77.85	91.16	79.26	87.34	115.49	151.10	97.86	92.37	114.03
80.40	78.03	79.95	82.33	92.32	76.65	91.02	109.29	97.86	118.46	90.98	101.12
80.29	78.76	86.01	74.64	84.85	86.32	89.45	114.38	92.37	90.98	112.72	108.65
94.17	81.28	91.78	95.12	96.00	82.45	99.75	116.41	114.03	101.12	108.65	146.38

estimated; in this case,

$$\hat{\theta}_{EB} = \hat{\theta}_{ML} - (\bar{Y} - \bar{\bar{Y}})/(N+1),$$

so that if $\bar{Y}$ is above (below) the prior mean estimated by $\bar{\bar{Y}}$ the shrinkage is downward (upward).

We can show, from (2.6), (2.9) and the fact that $\bar{Y}$ and $\bar{\bar{Y}}$ are independent, empirical Bayes estimator has mean μ and variance H,

$$E(\hat{\theta}_{EB}) = \mu$$

$$Var(\hat{\theta}_{EB}) = 2(nN^2+1)[nN(N+1)^2]^{-1}G = H$$

where μ and G are estimated by $\hat{\mu}$ and $\hat{G}$ given in (2.7) and (2.8), respectively. In our case H is estimated by $\hat{H} = 0.38\,\hat{G}$. In Table 3.4 we have derived the variance-covariance matrix $\hat{G}_1$ for the first location. Notice from (2.8) $\hat{G}_k$ is the sample within location variation obtained from all locations other than the kth location.

The $\hat{G}_k, k = 2,\ldots,10$, can be calculated likewise, but are not reported here. The variance-covariance matrix of the maximum likelihood estimate $\hat{\theta}_{ML}$ is estimated by $\hat{L}_k$. $2\,\hat{L}_k = \sum_{r=1}^{3} (Y_{kr} - \bar{Y}_k)(Y_{kr} - \bar{Y}_k)'$, $k = 1,\ldots,10$. Thus, Table 3.5 gives the value of $\hat{L}_1$ for the first location.

Having specified the mean and variance of $\hat{\theta}_{EB}$, we can then undertake the component analysis desired as appropriate to the particular factorial experiment. This is performed using well known techniques. Without giving this detail, we provide the results of one component analysis, viz., the determination of the linear and quadratic components, for the first location only. We observe that the linear component of nitrogen when 40 kilograms per hectare of phosphate are present is statistically significant at 5% level of significance.

We have not demonstrated the calculation of $\hat{\Sigma}_{EB,j}$, $j = 1,\ldots,10$, here. These computations are not difficult to carry out although they are lengthy in detail. We simply use (2.10) replacing V, G and $\bar{Y}$ by V_k, G_k, $\bar{Y}_k$, where G_k is estimated by $\hat{G}_k$ as above and V_k is estimated by

$$\hat{V}_k = \sum_{i=1}^{N} (Y_{ik} - \bar{Y}_k)(Y_{ik} - \bar{Y}_{ik})'.$$

Table 3.5. Sample Variance of $\hat{\theta}_{ML,1}$, $\hat{L}_1$

	N_1P_0	N_1P_1	N_1P_2	N_2P_0	N_2P_1	N_2P_2	N_3P_0	N_3P_1	N_3P_2	N_3P_3	N_4P_2	N_4P_3
N_1P_0	36.0508											
N_1P_1	14.7458	6.5208										
N_1P_2	19.3108	5.5958	21.1808									
	34.0733	11.4333	30.0333	45.0133								
	14.1571	6.6521	3.5296	8.3667	7.0758							
	-19.2054	-15.0104	23.3821	18.4533	-20.1367	114.8408						
N_3P_0	68.4896	25.5396	48.3321	77.3933	22.5396	-0.3054	142.6308					
N_3P_1	-4.5529	-0.2979	-9.8004	-12.3067	0.9658	-20.4467	-16.5604	5.5758				
N_3P_2	-25.1304	-4.8854	-38.8429	-51.3467	-0.3742	-65.4717	-75.0179	20.4158	76.9658			
N_3P_3	37.7350	12.6250	33.4350	50.0400	9.8725	20.9775	85.8975	-13.7475	-57.2725	55.6300		
N_4P_2	-0.3833	1.1667	-6.4333	-7.1333	2.1792	-19.1458	-7.4208	4.2792	14.8542	-8.000	3.5833	
N_4P_3	-43.2592	-17.4792	-24.1842	-41.9867	-16.6908	19.9008	-83.2717	6.1508	32.5258	-46.5150	1.0417	52.0033

Table 3.6. The t-test for linear and quadratic components.

P \ N	1	2	3	4	Linear	$SE_{(L)}$	t_L	quadratic	$SE_{(Q)}$	t_Q
0	68.18	70.23	73.07	–	4.89	4101	1.22	0.79	9.27	0.09
1	73.23	75.78	78.48	–	5.25	5.19	1.01	0.15	7.23	0.02
2	68.98	73.54	76.34	80.73	38.04	13.94	2.73	0.17	8.04	-0.02
3	–	–	77.96	82.48	4.52	4.87	0.93	–	–	–
Linear(L)	0.80	3.31	12.53	1.75						
SE(L)	0.52	5.35	13.65	3.98						
t_L	1.54	0.62	0.92	0.44						
Quadratic	-9.30	-7.79	-3.79	–						
SE(Q)	8.33	8.08	7.56	–						
t_Q	-1.12	-0.96	-0.50	–						

Conclusion

By utilizing information contained in the data gathered from other locations in a series of experiments, we have been able to provide better estimates of the vector of parameters representing the treatment effects for a given specific location. These estimates have been obtained by exploiting the principles of empirical Bayes procedures and have been shown to be better (in that the Bayes risk is smaller) than are the conventional maximum likelihood estimates of these treatment effects. The advantages of this approach are particularly enhanced when, as in our series of experiments under study, there is heterogeneity of variances across the individual experiments.

Acknowledgement

Support from the Office of Naval Research is gratefully acknowledged.

References

1. Anderson, T.W. *An Introduction to Multivariate Statistical Analysis* (2nd.ed.). John Wiley, New York. (1984).
2. Ando, A. and Kuafman, G.M. Bayesian analysis of the independent multinormal process-Neither mean nor precision known. *JASA*, 60, 347-358 (1965).
3. Morris, C.N. Parametric empirical Bayes inference: theory and applications. *JASA*, 78, 42-53.(1983).
4. Press, S.J. *Applied Multivariate Analysis: Including Bayesian and Frequentist Methods of Inference*, Krieger Publishing Co., Malabar, Florida.(1982).
5. Press, S.J. and Rolph, J. Empirical Bayes estimation of multinormal mean vector. *Comm. Statist. A*, 15, 1122-1146. (1986).
6. Robbins, H. The empirical Bayes approach to Statistical decision problems, *Annal. of Math. Stat.* 31, 195-208(1964).
7. Searle, S.R., *Linear Models*, John Wiley, New York. (1974).
8. Siegel, C.L. Uherdie analytishe theorie der quadralischen foremen. *Annal. of Math.* 36, 527-606 (1935).
9. Snedecor, G.W. and Cochran, W.G. *Statistical Methods.* Iowa State Univ. Press. Ames, Iowa. (1967).