Journal of Statistical Physics, Vol. 11, No. 1, 1974

Quantum Theory and the Bayesian Inference Problems

Stanislav Sýkora[1,2]

Received December 28, 1973

Many physicists take it for granted that their theories can be either refuted or verified by comparison with experimental data. In order to evaluate such data, however, one must employ statistical estimation and inference methods which, unfortunately, always involve an ad hoc proposition. The nature of the latter depends upon the statistical method adopted; in the Bayesian approach, for example, one must use *some* Lebesgue measure in the "set of all possible distributions." The ad hoc proposition has usually nothing in common with the physical theory in question, thus subjecting its verification (or refutation) to further doubt. This paper points out one notable exception to this rule. It turns out that in the case of the quantum mechanical systems associated with finite-dimensional Hilbert spaces the proposition is completely determined by the premises of the quantum theory itself.

KEY WORDS: Quantum theory; statistical inference; Bayesian inference.

1. INTRODUCTION

1.1. A Sketch of the Bayes Method of Inference

Consider a finite probability field[(1)] F generated by r elementary events e_i, $i = 1, 2,..., r$. The simplest way of defining the probability function in F is to specify the probabilities $p_1, p_2,..., p_r$ associated with the events $e_1, e_2,..., e_r$,

Supported by the National Science Foundation (USA).

[1] Physikalisches Institut der Universitaet Basel, Basel, Switzerland.

[2] This work was completed at the Department of Chemistry, University of Illinois, Urbana, Illinois.

respectively. This will be done by means of an *empirical* selection process σ: Suppose that n selections (trials) have been made and write $e(\sigma_k)$ for the outcome of the kth trial. By definition, $e(\sigma_k) \in \{e_i; i = 1, 2, \ldots, r\}$. We assume that the outcomes of the individual trials are statistically independent so that the order in which they are performed is irrelevant. The only quantities of interest are then the frequencies n_i, $i = 1, 2, \ldots, r$, where n_i specifies how many times the event e_i occurs within the set $\{e(\sigma_k); k = 1, 2, \ldots, n\}$. We must also assume that the following reproducibility criterion holds true: For $n \to \infty$ and for any $i = 1, 2, \ldots, r$ the ratio n_i/n converges and its limit is the same for any two infinite sequences of trials. In this case we put $p_i = \lim(n_i/n)$.

The principal aim of the Bayes theory[2] is the estimation of the probabilities p_i on the basis of the empirical frequencies $n_1, n_2, \ldots, n_r$. Because of the limit transition involved, the *exact* values of the probabilities are, however, inaccessible by any empirical method. What follows is a modern-language transcription of the Bayes approach to this problem.

Let E_r be an r-dimensional, real Euclidean space and let $\bar{p} \in E_r$, $\bar{p} \equiv (p_1, p_2, \ldots, p_r)$, where the components p_i are taken with respect to a fixed orthonormal basis in E_r. We are interested in the $(r - 1)$-dimensional simplex $B \in E_r$ defined by the conditions $p_i \geqslant 0$ for all i and $p_1 + p_2 + \cdots + p_r = 1$. In principle, the components $p_1, p_2, \ldots, p_r$ of any element $\bar{p} \in B$ represent an acceptable set of the probabilities of the events $e_1, e_2, \ldots, e_r$, respectively.

Given the empirical frequencies $\bar{n} \equiv \{n_1, n_2, \ldots, n_r\}$, we assign to every $\bar{p} \in B$ the weight

$$w_{\bar{n}}(\bar{p}) = (n!/n_1!\, n_2! \cdots n_r!)\bar{p}^{\bar{n}} \tag{1}$$

where $n = n_1 + n_2 + \cdots + n_r$ and $\bar{p}^{\bar{n}} = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$. Evidently, $w_{\bar{n}}(\bar{p})$ is the conditional probability that n random trials will lead to the frequencies $n_1, n_2, \ldots, n_r$ provided that the probability of the event e_i is actually equal to p_i for all i. Equation (1) must be modified in the case that, for some reason, different elements of B are not supposed to be equally likely even prior to any empirical verification. Then

$$w_{\bar{n}}(\bar{p}) = (n!/n_1!\, n_2! \cdots n_r!)\bar{p}^{\bar{n}}\mu(\bar{p}) \tag{2}$$

where $\mu(\bar{p})$ is the so-called a priori weight of the element $\bar{p} \in B$. Since B is measurable, the weight $w_{\bar{n}}(\bar{p})$ can be normalized. This leads to a probability density $\rho(\bar{p})$ in B:

$$\rho_{\bar{n}}(\bar{p}) = \bar{p}^{\bar{n}}\mu(\bar{p}) \Big/ \int_B \bar{p}^{\bar{n}}\mu(\bar{p})\, d\sigma \tag{3}$$

where $d\sigma$ is the Cartesian surface element of B.

From the last equation it is evident that the a priori weight $\mu(\bar{p})$ is intimately connected with the measure in B. Indeed, a slightly more rigorous treatment would show that μ has to be *identified* with a measure. Then Eq. (3) can be rewritten by means of the Lebesgue integral

$$\rho_{\bar{n}}(\bar{p}) = \bar{p}^{\bar{n}} \Big/ \int_B \bar{p}^{\bar{n}} \mu\{d\sigma\} \tag{4}$$

Advancing one step further, we ask the question: Given the empirical frequencies $\bar{n}$, what is the probability $P_{\bar{n}}(\bar{m})$ that m new trials will lead to the frequencies $\bar{m} \equiv \{m_1, m_2, \ldots, m_r\}$? Combining Eqs. (1) and (4), we obtain

$$P_{\bar{n}}(\bar{m}) = \frac{m\,!}{m_1\,!\, m_2\,! \cdots m_r\,!} \frac{\int_B \bar{p}^{(\bar{n}+\bar{m})} \mu\{d\sigma\}}{\int_B \bar{p}^{\bar{n}} \mu\{d\sigma\}} \tag{5}$$

This is the fundamental Bayesian inference formula for finite probability fields.

1.2 The Central Impass of Inference Theories

The sore point of the Bayes method is the choice of the measure $\mu\{d\sigma\}$. Since the following sections of this paper must be viewed in the light of this basic problem, I will discuss briefly the historical attempts at its solution. Before doing so, however, it might be useful to demonstrate the depth of the problem by means of a simple example.

Let us first put $\mu\{d\sigma\} \equiv \mu^{\mathrm{B}}\{d\sigma\} = d\sigma$. The measure $\mu^{\mathrm{B}}\{d\sigma\}$ will be called "Bayesian" since it coincides with Bayes' own choice[3] in the special case he considered. Explicit evaluation shows (see Appendix A) that in this case

$$\rho_{\bar{n}}^{B}(\bar{p}) = r^{-1/2} \bar{p}^{\bar{n}} \Gamma(n + r) \Big/ \prod_{i=1}^{r} \Gamma(n_i + 1) \tag{6}$$

and

$$P_{\bar{n}}^{B}(\bar{m}) = \left[\prod_{i=1}^{r} \binom{n_i + m_i}{m_i} \right] \Big/ \binom{n + m + r - 1}{m} \tag{7}$$

Our second choice will seem a bit more complicated. Consider the unit sphere S_r in E_r and denote its Cartesian surface element $d\tau$. Let us define a regular map $\mathscr{T}_r$ of S_r on the whole of B by putting $\mathscr{T}_r \bar{x} = \bar{p}$, where $p_i = x_i^2$ for all $i = 1, 2, \ldots, r$. Given any measure $\nu\{d\tau\}$ on S_r, the mapping $\mathscr{T}_r$ defines a corresponding measure $\mu\{d\sigma\} = \nu\{\mathscr{T}_r^{-1}\, d\sigma\}$ on B. We will choose $\nu\{d\tau\} = \nu^{\mathrm{S}}\{d\tau\} = d\tau$ and denote the corresponding measure on B as $\mu^{\mathrm{S}}\{d\sigma\}$. This leads

[3] It is also known as the Bayes postulate. See Ref. 3.

to (see Appendix B)

$$\rho_{\bar{n}}^{S}(\bar{p})\mu^{S}\{d\sigma\} = \rho_{\bar{n}}'(\bar{x})\, d\tau \tag{8}$$

where $\bar{x} \in S_r$, $\bar{p} = \mathscr{T}_r \bar{x}$, $d\tau = \mathscr{T}_r^{-1}\, d\sigma$, and

$$\rho_{\bar{n}}'(\bar{x}) = \tfrac{1}{2}\bar{x}^{2\bar{n}}\Gamma(n + \tfrac{1}{2}r)\Big/\prod_{i=1}^{r}\Gamma(n_i + \tfrac{1}{2}) \tag{9}$$

Consequently,

$$P_{\bar{n}}^{S}(\bar{m}) = \left[\prod_{i=1}^{r}\binom{n_i + m_i - \frac{1}{2}}{m_i}\right]\Big/\binom{n + m + \frac{1}{2}r - 1}{m} \tag{10}$$

The fact that the two measures lead to two different inference formulas eliminates the argument that the lack of *any* a priori knowledge justifies the use of a particular measure. Historically, this argument has often been used[4] to justify the Bayes choice. The intuitively appealing feature of such reasoning is the homogeneity of $\mu^{B}\{d\sigma\}$ with respect to translations along B (notice that B is a section of a hyperplane). But, on the other hand, the measure $\nu^{S}\{d\tau\}$ is isotropic on the sphere S_r. Considering the fact that the symmetry of S_r is substantially higher than that of B, one's intuition might easily be tempted to switch sides. The problem is that the symmetry operations involved are purely abstract concepts with no empirical meaning.

Von Mises proved that under certain conditions the effect of the choice of the measure on the inference formula becomes negligible if $n \to \infty$.[5] This theorem, however, does not really solve anything since it simply refers back to the limit transition on which the definition of the p_i has been based.

Von Mises' later proposal[(6)] was to look at the measure $\mu\{d\sigma\}$ as an empirical *hypothesis* based on the sum of our previous experience (history). There are a few facts which seem to support this idea. Thus, for example, the everyday statistical practice is based on the Bayesian measure $\mu^{B}\{d\sigma\}$. If normal distributions are supposed to appear only as limit cases and/or accumulation points of Bernoullian distributions, then any measure on B determines a corresponding measure $\varphi\{da, ds\}$ in the space of their location (a) and scale (s) parameters. The particular measure generated by $\mu^{B}\{d\sigma\}$ turns out to be homogeneous in the coordinates a and s^2, i.e., $\varphi\{da, ds\} \propto s\, da\, ds$, and it is exactly this choice which leads to the universally used chi-square test. Our whole statistical experience (including the overall success of the insurance business) therefore indicates that, as a hypothesis, the Bayes choice is not unreasonable. It is nevertheless evident that while this approach may satisfy a pragmatist, it does not lead to any absolute certainty either. After all, our

[4] A discussion of this problem appears in almost every monograph dealing with statistical inference. See, e.g., Refs. 4, 6–8.

[5] Most of von Mises' excellent work on Bayesian problems appears in Ref. 7. Also see Ref. 6.

historical experience is limited and, even worse, we must sometimes face problems which are brand new to us.

Mathematicians have for some time recognized the elusive nature of the problem and have solved it by simply ignoring it.[8] The measure $\mu\{d\sigma\}$ may vary from one class of empirical problems to another. Its estimation therefore belongs to the domain of empirical science. Since there does not exist any rigorous algorithm on which such an estimation could be based, the empiricist is *forced* to take refuge in an ad hoc proposition and to see whether his choice is satisfactory in the light of his growing experience. He must keep in mind, however, that "satisfactory" is by no means identical with "true" or even "optimal." There exists at least one extremely successful example of this kind of approach in science. Statistical mechanics of many-body systems is based on the *assumption* that the density operator of a system which is in a thermodynamic equilibrium equals $\exp(-\theta\mathscr{H})$, where $\mathscr{H}$ is the Hamiltonian and θ is a real number. This Boltzmann "law" is in reality nothing more than an ad hoc choice of a weight function in the associated Liouville space.

In the last decade and a half there have appeared a number of new mathematical methods[9,10] which either reduce the sensitivity of the inference formulas to the choice of the measure or substitute the measure by substantially weaker concepts. These efforts are undoubtedly very valuable but they still lead to the same conclusion: No mathematical inference from a set of empirical data is possible unless one is given some knowledge *prior* to and independent of the data.

2. THE PHYSICIST'S POINT OF VIEW

Physical methods apply to special classes of phenomena. Within each class there exists a model (theory) which implicitly determines the outcomes of all relevant observations and/or the correlations between such observations. The physical theory is, however, just a hypothesis. In principle, many different and even conflicting theories may be successfully applied to the same class of phenomena.

The outcomes of observations (data) are presumably independent of the physical model adopted. It does not follow, however, that the way the data are evaluated is also independent of the model. Since observations are never exact, statistical methods must be used for their evaluation. We have seen that this is possible only in connection with a special proposition such as the choice of the measure $\mu\{d\sigma\}$ in the case of the Bayesian methods. This proposition becomes an integral part of the physical model. Vice versa, a physical theory is complete only if it includes a prescription for the evaluation of empirical data.

The relevant statistical proposition is most often just an independent appendix to the physical theory. In some cases (as in statistical physics) it may rely on the knowledge of quantities which have a meaning only within the particular theory. Finally—and in the next section we are going to discuss a particular example of this case—it may become a *consequence* of the very premises of the physical theory and thus acquire the character of a law.

3. BAYES' METHOD AND QUANTUM THEORY

We will consider a quantum mechanical system with a finite-dimensional Hilbert space H. By the postulates of quantum theory, a measurement process $\mathcal{O}$ is described by means of an associated Hermitian operator (observable) $\mathcal{O}$. Two measurement processes are said to be compatible if the corresponding observables commute. Two different measurements can be carried out simultaneously if and only if they are compatible. The set of all possible outcomes of the measurement process $\mathcal{O}$ coincides with the set of all distinct eigenvalues of $\mathcal{O}$. The measurement is said to *specify* the linear manifold in H belonging to the particular eigenvalue of $\mathcal{O}$. A set $\{\mathcal{O}_\alpha; \alpha = 1, 2, \ldots, k\}$ of measurements is said to be complete if any two of them are compatible and if their outcomes specify a unique state, i.e., a unique element on the unit sphere C_r in H. Denoting the dimension of H as r, it follows that the set $\{\mathcal{O}_\alpha; \alpha = 1, 2, \ldots, k\}$ defines an empirical selection process on the ensemble $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_r$, where the "events' $\bar{e}_i$ represent the orthonormalized simultaneous eigenvectors of the operators $\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_k$.

If the system is in a state $\bar{e} \in C_r$, then, in general, the outcome of the measurements $\{\mathcal{O}_\alpha\}$ may be any of the events $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_r$. Quantum theory predicts only the probability p_i that the result of the "trial" will be $\bar{e}_i$: $p_i = |(\bar{e}, \bar{e}_i)|^2$, where $(.,.)$ is the scalar product in H. This prescription defines a regular mapping $\mathcal{T}_c$ of C_r on B.

Let $\varphi\{d\omega\}$ be a measure on C_r ($d\omega$ is the Cartesian surface element of the r-dimensional complex sphere). The mapping $\bar{e}' = \mathcal{U}\bar{e}$ defined by a unitary operator $\mathcal{U}$ transforms $\varphi\{d\omega\}$ into a new measure $\varphi'\{d\omega\} = \varphi\{\mathcal{U}^\dagger\, d\omega\}$ on C_r. From the physical point of view, however, it makes no difference whether we transform the Hilbert space H while leaving unchanged the operators or leave H unchanged and transform all observables according to the formula $\mathcal{O}' = \mathcal{U}\mathcal{O}\mathcal{U}^\dagger$. Since the transformed observables $\{\mathcal{O}_\alpha'\}$ define a new complete set of measurements, and since no such set is preferred to any other, this implies that $\varphi\{\mathcal{U}\, d\omega\} = \varphi\{d\omega\}$ for any unitary $\mathcal{U}$. But there is only one fully isotropic measure and that is $\varphi\{d\omega\} \equiv \varphi^C\{d\omega\} = d\omega$.

Notice that our present ability to specify the measure is based on two facts: (i) The rotations in H now have a precise empirical interpretation—they describe the transitions from one complete set of measurements to another; and (ii) all sets of measurements are considered equivalent in the sense that it is our choice which one will be adopted. If the second condition is modified, the whole argument falls. For example, if the knowledge of the energy of the system is made mandatory, then the set $\{\mathcal{O}_\alpha\}$ must include the Hamiltonian. The symmetry therefore disappears and the isotropic measure has to be replaced by another one.

Any measure $\varphi\{d\omega\}$ on C_r generates a measure $\mu\{d\sigma\} = \varphi\{\mathcal{T}_c^{-1}\, d\sigma\}$ on B. The isotropic case leads to (see Appendix C)

$$\rho_{\bar{n}}{}^C(\bar{p})\mu^C\{d\sigma\} = \rho_{\bar{n}}''(\bar{e})\, d\omega \tag{11}$$

where $d\omega = \mathcal{T}_c^{-1}\, d\sigma$, $\bar{p} = \mathcal{T}_c\bar{e}$, and

$$\rho_{\bar{n}}''(\bar{e}) = \tfrac{1}{2}\pi^{-r}|\bar{e}|^{2\bar{n}}\Gamma(n + r)\Big/\prod_{i=1}^{r}\Gamma(n_i + 1) \tag{12}$$

Here $|\bar{e}|^{2\bar{n}} = |z_1|^{2n_1}|z_2|^{2n_2}\cdots|z_r|^{2n_r}$, where $z_i = (\bar{e}, \bar{e}_i)$. Surprisingly, the inference formula resulting from this relation is identical with formula (7), i.e.,

$$P_{\bar{n}}{}^C(\bar{m}) = P_{\bar{n}}{}^B(\bar{m}) \tag{13}$$

4. FINAL REMARKS

If we accept the thesis that experiment is the ultimate way of weighing the truth of a theory, then physical theories can be verified or refuted only in the statistical sense (that is, up to a certain significance level). From what has been said in Sections 1 and 2 it follows that, strictly speaking, no verification is possible unless the relevant inference method is an integral part of the theory. The merit of Section 3 consists primarily in the demonstration that such "complete" theories can be constructed.

Formulas (11)–(13) represent the solution of the inference problem within the framework of quantum mechanics for the special case characterized by finite-dimensional Hilbert spaces. The fact that, after much calculation, the results are identical with the classical Bayes choice is very intriguing but it hardly provides sufficient ground for any far-reaching speculation.

The generalization of the above results to the cases with infinite-dimensional Hilbert spaces presents the same kind of problems as the generalization of the Bayes method to infinite probability fields. The main obstacle is the nonexistence of any Lebesgue measure in the spaces involved. This results in the necessity of introducing further ad hoc restrictions on the a priori weight function. It can not be excluded, however, that this problem will be eventually solved by quite different methods.

APPENDIX A

Write

$$B_r(\bar{n}; R) = \int_{B(R)} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}\, d\sigma \tag{A.1}$$

where $B(R)$ is the simplex lying in the $(r-1)$-dimensional hyperplane $Q(R) \equiv \{\bar{p}; p_1 + p_2 + \cdots + p_r = R\}$ of the r-dimensional Euclidean space E_r and delimited by the inequalities $p_i \geqslant 0$, $i = 1, 2, \ldots, r$. We are especially interested in the quantity $B_r(\bar{n}) = B_r(\bar{n}; 1)$. Applying the scaling transformation $\bar{p}' = R^{-1}\bar{p}$, we find

$$B_r(\bar{n}; R) = R^{n+r-1} B_r(\bar{n}) \tag{A.2}$$

where $n = n_1 + n_2 + \cdots + n_r$.

Let $\bar{u}$ be the vector perpendicular to $Q(R)$, $\bar{u} \equiv (r^{-1/2}, r^{-1/2}, \ldots, r^{-1/2})$, and θ the angle between $\bar{u}$ and the rth coordinate axis. Then, putting $p_r = x$, one can write (see Fig. 1A)

$$B_r(n_1, n_2, \ldots, n_r) = \frac{1}{\sin\theta} \int_0^1 x^{n_r} B_{r-1}(n_1, n_2, \ldots, n_{r-1}; 1 - x)\, dx \tag{A.3}$$

Since $\sin\theta = [(r-1)/r]^{1/2}$, Eqs. (A.2) and (A.3) combine to give

$$\begin{aligned} B_r(n_1, n_2, \ldots, n_r) &= \left(\frac{r}{r-1}\right)^{1/2} B_{r-1}(n_1, n_2, \ldots, n_{r-1}) \int_0^1 x^{n_r}(1-x)^{n-n_r+r-2}\, dx \\ &= \left(\frac{r}{r-1}\right)^{1/2} B_{r-1}(n_1, n_2, \ldots, n_{r-1}) \beta(n_r + 1, n - n_r + r - 1) \end{aligned} \tag{A.4}$$

where $\beta(x, y)$ is the beta function.

Putting $B_1(n) = 1$, one obtains by induction

$$B_r(\bar{n}) = r^{1/2} \left[\prod_{i=1}^{r} \Gamma(n_i + 1)\right] \Big/ \Gamma(n + r) \tag{A.5}$$

Combining this formula with Eqs. (4) and (5), one obtains the desired Eqs. (6) and (7), respectively.

It is of some interest to notice that Eq. (A.5) is also valid for noninteger n_i as long as $n_i \geqslant 0$.

APPENDIX B

Let

$$V_r(\bar{n}; R) = \int_{S(R)} x_1^{2n_1} x_2^{2n_2} \cdots x_r^{2n_r}\, d\tau \tag{B.1}$$

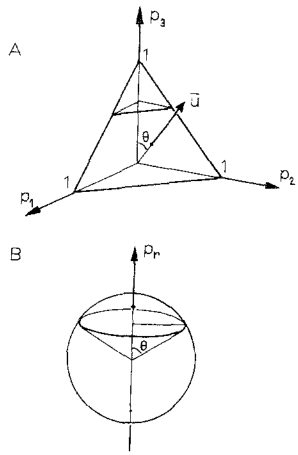

Fig. 1. (A) The simplex B for $r = 3$. The cut for $p_3 = x =$ const is identical with the simplex $B(R = 1 - x)$ in the two-dimensional space of the cut. $\bar{u}$ is the unit vector perpendicular to B. (B) A schematic drawing of the unit sphere S_r in an r-dimensional Euclidean space E_r. The cut for $p_r = \cos\theta =$ const is a sphere of radius $\sin\theta$ in E_{r-1}.

where $\bar{x} \equiv (x_1, x_2,..., x_r)$ is an element of the r-dimensional Euclidean space and $S(R)$ is the sphere $\{\bar{x};\, x_1{}^2 + x_2{}^2 + \cdots + x_r{}^2 = R^2\}$. We actually need to determine only the quantity $V_r(\bar{n}) = V_r(\bar{n}; 1)$, which is related to $V_r(\bar{n}; R)$ by means of the scaling transformation $\bar{x}' = R^{-1}\bar{x}$,

$$V_r(\bar{n}; R) = R^{2n+r-1}V_r(\bar{n}) \tag{B.2}$$

where $n = n_1 + n_2 + \cdots + n_r$.

Using Eq. (B.2), one can easily derive an induction formula for $V_r(\bar{n})$ Putting $x_n = \cos\theta$, $V_r(\bar{n})$ can be rewritten as (see Fig. 1B):

$$\begin{aligned} V_r(n_1, n_2,..., n_r) &= \int_0^\pi V_{r-1}(n_1, n_2,..., n_{r-1};\, |\sin\theta|)(\cos\theta)^{2n_{r-1}}\, d\theta \\ &= \beta(n_r + \tfrac{1}{2}, n - n_r + \tfrac{1}{2}r - \tfrac{1}{2})V_{r-1}(n_1, n_2,..., n_{r-1}) \end{aligned} \tag{B.3}$$

where $\beta(x, y)$ is the beta function.

The degenerate case $r = 1$ requires $V_1(n) = 2$ so that, by induction,

$$V_r(\bar{n}) = 2\left[\prod_{i=1}^{r} \Gamma(n_r + \tfrac{1}{2})\right] \Big/ \Gamma(n + \tfrac{1}{2}r) \tag{B.4}$$

Equations (9) and (10) then follow straightforwardly from Eqs. (B.4), (4), and (5). Similarly as in the case of Eq. (A.5), formula (B.4) is valid for all non-negative n_i.

APPENDIX C

We want to determine the value of the integral

$$W_r(\bar{n}) = \int_{C_r} |z_1|^{2n_1}|z_2|^{2n_2}\cdots|z_r|^{2n_r}\, d\omega \tag{C.1}$$

where $\bar{z} \equiv (z_1, z_2, \ldots, z_r)$ is an element of the r-dimensional complex Euclidean space H_r, $C_r \equiv \{\bar{z}; |z_1|^2 + |z_2|^2 + \cdots + |z_r|^2 = 1\}$ is the unit sphere in H_r, and $d\omega$ is the Cartesian surface element on C_r.

H_r is isomorphic to the $2r$-dimensional, real Euclidean space E_{2r}; a particular isomorphism Φ can be defined, for example, by the relations

$$x_{2i-1} = \operatorname{Re} z_i, \qquad x_{2i} = \operatorname{Im} z_i, \qquad i = 1, 2, \ldots, r \tag{C.2}$$

where $\bar{z} \in H_r$, $\bar{x} \in E_{2r}$, and Re and Im denote the real and imaginary parts, respectively. Since Φ maps C_r on the unit sphere S_{2r} in E_{2r} and, by definition, $\Phi(d\omega) = d\tau$, where $d\tau$ is the Cartesian surface element on S_{2r}, Eq. (C.1) can be written as

$$\begin{aligned} W_r(\bar{n}) &= \int_{S_{2r}} (x_1{}^2 + x_2{}^2)^{n_1}(x_3{}^2 + x_4{}^2)^{n_2} \cdots (x_{2r-1}^2 + x_{2r}^2)^{n_r}\, d\tau \\ &= \sum_{l_1=0}^{n_1} \sum_{l_2=0}^{n_2} \cdots \sum_{l_r=0}^{n_r} \left[\prod_{i=1}^{r} \binom{n_i}{l_i}\right] \\ &\quad \times V_{2r}(2l_1, 2n_1 - 2l_1, 2l_2, 2n_2 - 2l_2, \ldots, 2l_r, 2n_r - 2l_r) \end{aligned} \tag{C.3}$$

Using Eq. (B.4), we have, further,

$$\begin{aligned} W_r(\bar{n}) &= [2/\Gamma(n + r)] \sum_{l_1=0}^{n_1} \cdots \sum_{l_r=0}^{n_r} \\ &\quad \times \prod_{i=0}^{r} \begin{bmatrix} n_i \\ l_i \end{bmatrix} \Gamma(n_i + \tfrac{1}{2})\Gamma(n_i - l_i + \tfrac{1}{2}) \end{aligned} \tag{C.4}$$

Employing the identity

$$\sum_{i=0}^{n} \binom{n}{i} \Gamma(a + i)\Gamma(n + b - i) = \beta(a, b)\Gamma(n + a + b) \tag{C.5}$$

we arrive at the formula

$$W_r(\bar{n}) = 2\pi^r \left[\prod_{i=1}^{r} \Gamma(n_i + 1)\right] \Big/ \Gamma(n + r) \tag{C.6}$$

REFERENCES

1. A. N. Kolmogorov, *Foundations of the Theory of Probability*, Chelsea, New York (1956), Chapter I.
2. T. Bayes, *Phil. Trans.* **53**:376 (1763); reproduced, accompanied by a biographical note by G. A. Barnard, in *Biometrica* **45**:293 (1958); also in E. S. Pearson and M. G. Kendall, *Studies in the History of Statistics and Probability*, Hafner, Darien, Connecticut (1970).
3. M. G. Kendall and W. R. Buckland, *A Dictionary of Statistical Terms*, Oliver and Boyd, Edinburgh (1957).
4. H. Freeman, *Introduction to Statistical Inference*, Addison-Wesley, Reading, Massachusetts (1963).
5. R. Von Mises, *Math. Z.* **4**:1 (1919).
6. R. Von Mises, *Ann. Math. Stat.* **13**:156 (1942).
7. R. Von Mises, *Mathematical Theory of Probability and Statistics* (collection edited by H. Geiringer), Academic Press, New York (1964).
8. W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley, New York (1968), Vol. I, Chapter V.
9. M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, C. Griffin and Co., London (1961), Chapter 31.
10. P. J. Huber, *Ann. Math. Stat.* **43**:1041 (1972).