ON THE THEORY OF TESTING COMPOSITE HYPOTHESES WITH ONE CONSTRAINT

By Henry Scheffé

Princeton University

1. Introduction. Our purpose is to extend some of the Neyman-Pearson theory of testing hypotheses to cover certain cases of frequent interest which are complicated by the presence of nuisance parameters. Our results give methods of finding critical regions of types B and B_1. Type B regions were defined by Neyman [1] for the case of one nuisance parameter. Type B_1 regions are the natural generalization of the type A_1 regions of Neyman and Pearson [5] to permit the occurrence of nuisance parameters. The reader familiar with the work of these authors will recognize most of the notation and some of the methods.

We consider a joint distribution of n random variables $x_1, x_2, \cdots, x_n$, depending on l parameters $\theta_1, \theta_2, \cdots, \theta_l$, $l \leqq n$. The functional form of the distribution is given. The random variables may be regarded as the coordinates of a point E in an n-dimensional sample space W, the parameters, as the coordinates of a point Θ in an l-dimensional space Ω of admissible parameter values. Ω, unlike W, in general will not be a complete Euclidian space. Let ω denote the subspace of Ω defined by $\theta_1 = \theta_1^0$. The hypothesis we consider is

$$H_0 : \Theta \epsilon \omega.$$

Neyman and Pearson [4] call H_0 a hypothesis with $l - 1$ degrees of freedom; for our present purpose we shift the emphasis by saying it has one constraint.

It is clear that whenever we test whether a parameter has a given value, and other parameters occur in the distribution, we are testing a hypothesis with one constraint. Hypotheses of the type $\theta_1 = \theta_2$, in which we do not specify the common value of θ_1 and θ_2, nor the values of any other parameters, may always be transformed to H_0 by choosing new parameters. In general, the hypothesis that the parameter point Θ lies on some hypersurface in Ω, $g(\theta_1, \theta_2, \cdots, \theta_l) = g_0$, may be transformed to H_0 if the function g satisfies certain conditions,—say, g is continuous and monotone-increasing in one of the θ's for all Θ in Ω. Another circumstance lending importance to the theory of testing hypotheses with one constraint is its connection with the theory of confidence intervals, which we shall point out below.

The path which led Neyman to critical regions of type B is the following: Every Borel-measurable region w of sample space determines a test of H_0, which consists of rejecting H_0 if and only if E falls in w. In deciding which is a most efficient test, one may limit the competition to similar[1] regions, if such exist. Because of the general non-existence [2, p. 372] of uniformly most

[1] Defined by condition (a) of definition 1.

powerful tests, one is led to consider common best critical regions [4] if he is interested only in alternatives $\theta_1 < \theta_1^0$ (or $\theta_1 > \theta_1^0$), or else regions giving an unbiased test [1, p. 251]. Narrowing the competition further to the latter class of regions, one is led to regions of type B if he seeks tests which are most powerful for θ_1 very near to θ_1^0, and to type B_1 regions if he is not content with this. These types of regions are defined in section 2.

We may now state the relationship of hypotheses with one constraint to the theory of confidence intervals [2]. To find confidence intervals for θ_1, we must first find similar regions $w(\theta_1^0)$ for testing H_0. If with every admissible θ_1 we can associate a $w(\theta_1)$, then confidence regions for θ_1 are determined, and if these be intervals, they are confidence intervals. Every class of similar regions mentioned above is intimately related to a category of confidence intervals. In particular, to find Neyman's short unbiased confidence intervals we must first solve the problem of type B regions. Likewise, if we define shortest unbiased confidence intervals in the obvious way along the lines laid down by Neyman, their discovery rests on the solution of the problem of type B_1 regions.

While the assumptions of section 3, especially 3^0, are unpleasantly restrictive—they are obviously tailored to fit the proof rather than the problem—they are nevertheless satisfied in many sampling problems associated with normal distributions. An application of the theorems of section 4 will be given in another paper *On the ratio of the variances of two normal populations*. The present theory was needed to round out that paper and was originally planned as a section thereof. However, it seems desirable for the convenience of other workers who might have use for the theory not to bury it under the preceding title.

Section 5 consists of an appendix on the moment problem raised by assumption 5^0.

2. Definitions. The symbols w, w_0, w_1 will always be understood to denote Borel-measurable regions in W. We shall symbolize $\partial^i Pr\{E \epsilon w \mid \Theta\}/\partial\theta_1^i$ for $i = 0, 1, 2$ by $P(w \mid \Theta)$, $P'(w \mid \Theta)$, $P''(w \mid \Theta)$, respectively. Since θ_1 plays a distinguished rôle, it will often be convenient to write $\Theta = (\theta_1, \vartheta)$, where the nuisance parameters are denoted by $\vartheta = (\theta_2, \theta_3, \cdots, \theta_l)$.

Definition 1: *w_0 is said to be a type B region for testing H_0 if for all Θ in ω*

(a) $P(w_0 \mid \theta_1^0, \vartheta) = \alpha$, *where α is independent of ϑ,*

(b) $P'(w_0 \mid \theta_1^0, \vartheta)$, $P''(w_0 \mid \theta_1^0, \vartheta)$ *exist,*

(c) $P'(w_0 \mid \theta_1^0, \vartheta) = 0$,

(d) $P''(w_0 \mid \theta_1^0, \vartheta) \geqq P''(w_1 \mid \theta_1^0, \vartheta)$ *for all w_1 satisfying* (a), (b), (c).

Definition 2: *w_0 is said to be of type B_1 if the conditions* (a), (b′), (c), (d′) *are satisfied. The conditions* (a), (c) *are given in definition* 1, *the other two are*

(b′) $P'(w_0 \mid \theta_1, \vartheta)$ *is continuous in θ_1 at $\theta_1 = \theta_1^0$ for all Θ in ω,*

(d′) $P(w_0 \mid \theta_1, \vartheta) \geqq P(w_1 \mid \theta_1, \vartheta)$ *for all w_1 satisfying* (a), (b′), (c), *and all Θ in Ω.*

3. Assumptions. $p(z_1, z_2, \cdots, z_m \mid \Theta)$ will be a generic notation for the p.d.f. (probability density function) of random variables $z_1, z_2, \cdots, z_m$ whose distribution depends on Θ. The numbering of the following assumptions follows that of Neyman elsewhere [1].

1^0. (a) There exists a p.d.f. $p(E \mid \Theta)$ such that for any w, and any $\Theta \in \Omega$,

$$P(w \mid \Theta) = \int_w p(E \mid \Theta) dW \tag{1}$$

where dW denotes the volume element $dx_1 dx_2 \cdots dx_n$.

(b) The region W_+ in W defined by $p(E \mid \Theta) > 0$ is independent of Θ for $\Theta \in \omega$.

(c) The connectivity of ω is such that it is possible to pass from any point Θ' in ω to any other point Θ'' in ω by a path lying entirely in ω and consisting of a finite number of segments on each of which all but one of $\theta_2, \theta_3, \cdots, \theta_l$ are constant.

2^0. For all $E \in W_+$ and $\Theta \in \omega$, $p(E \mid \Theta)$ is differentiable twice with respect to θ_1 and indefinitely with respect to $\theta_2, \theta_3, \cdots, \theta_l$. For any w, and any $\Theta \in \omega$, the corresponding derivatives of $P(w \mid \Theta)$ exist and may be obtained by differentiating under the integral sign in (1).

We now define

$$\phi_i = \partial \log p(E \mid \Theta)/\partial\theta_i, \qquad \phi_{ij} = \partial\phi_i/\partial\theta_j, \quad i, j = 1, 2, \cdots, l.$$

3^0. For all $E \in W_+$ and $\Theta \in \omega$, $\phi_i = \phi_i(E, \Theta)$ is continuous in E, $i = 1, 2, \cdots, l$, and

$$\phi_{ij} = A_{ij} + \sum_{k=2}^{l} B_{ijk}\phi_k, \qquad i, j = 2, 3, \cdots, l, \tag{2}$$

$$\phi_{i1} = A_{i1} + \sum_{k=1}^{l} B_{i1k}\phi_k, \qquad i = 1, 2, \cdots, l, \tag{3}$$

where $A_{ij} = A_{ij}(\theta_1^0, \vartheta)$, $B_{ijk} = B_{ijk}(\theta_1^0, \vartheta)$ are continuous in each of $\theta_2, \theta_3, \cdots, \theta_l$.

4^0. The matrix $(\partial\phi_i/\partial x_j)$, $i = 1, 2, \cdots, l$; $j = 1, 2, \cdots, n$, contains an $l \times l$ minor which is non-singular[2] for all $E \in W_+$ and $\Theta \in \omega$, and whose elements are continuous in E.

Write $\Phi = (\phi_2, \phi_3, \cdots, \phi_l)$, and denote by $p(\phi_1, \Phi \mid w, \Theta)$ the p.d.f. of (ϕ_1, Φ) calculated under the assumption that $E \in w$, i.e., that the p.d.f. of E is $p(E \mid \Theta)/P(w \mid \Theta)$ for $E \in w$ and zero for $E \in W - w$. Define

[2] If for each $\Theta \in \omega$, 4^0 is violated on an exceptional set $U(\Theta)$ for which $P(U(\Theta) \mid \Theta) = 0$, the theorems 1 and 2 may still be valid. What is essential is the existence of the p.d.f. $p(\phi_1, \phi_2, \cdots, \phi_l \mid \Theta)$ for all $\Theta \in \omega$. On reconsidering the theorems and their proofs, the reader will see that if the set $U(\Theta)$ is deleted from W_+, then $1^0(b)$ may be violated, but not seriously, and no essential changes are necessary. The addition of the necessary qualifying clauses to our statements, regarding sets of probability zero, would encumber the developments.

$$Q_s(\Phi \mid w, \Theta) = \int_{-\infty}^{+\infty} \phi_1^s p(\phi_1, \Phi \mid w, \Theta)\, d\phi_1. \tag{4}$$

Let w_1 be any region satisfying condition (a) of definition 1:

5^0. We assume, for each $\Theta \in \omega$, that if the moments[3] of $Q_s(\Phi \mid w_1, \Theta)$ and $Q_s(\Phi \mid W, \Theta)$ are the same then these functions are equal for almost all Φ

(a) for $s = 0$,

(b) for $s = 1$.

Note that Q_0 is p.d.f., Q_1 is not.

4. Theorems. A result of Neyman's [1] for $l = 2$ is generalized in the following[4]

THEOREM 1: *Under the assumptions 1^0 to 5^0, consider the existence of functions $k_i(\Phi, \theta_1^0, \vartheta)$, $i = 1, 2$, such that $k_1 < k_2$ and*

$$\int_{k_1(\Phi,\theta_1^0,\vartheta)}^{k_2(\Phi,\theta_1^0,\vartheta)} \phi_1^s p(\phi_1, \Phi \mid \theta_1^0, \vartheta)\, d\phi_1 = (1 - \alpha) \int_{-\infty}^{+\infty} \phi_1^s p(\phi_1, \Phi \mid \theta_1^0, \vartheta)\, d\phi_1, \qquad s = 0, 1, \tag{5}$$

for all $\Phi = (\phi_2, \phi_3, \cdots, \phi_l)$. If such functions exist for some $\Theta = \Theta' \in \omega$, they exist for all $\Theta \in \omega$. Then the region w_0 in W defined by

$$\phi_1(E, \theta_1^0, \vartheta) < k_1(\Phi, \theta_1^0, \vartheta) \quad \textit{and} \quad \phi_1(E, \theta_1^0, \vartheta) > k_2(\Phi, \theta_1^0, \vartheta) \tag{6}$$

is independent of ϑ, and is a region of type B for testing the hypothesis H_0.

Since throughout the proof $\Theta = (\theta_1^0, \vartheta)$, we shall write Θ in place of these symbols to simplify the printing. It is to be understood that *every statement in the proof involving the symbol Θ is asserted for all Θ in ω.*

We suppose first that a type B region w_0 exists in W_+. Then from (a), (c) of definition 1 and assumptions 1^0(a) and 2^0,

$$\int_{w_0} p(E \mid \Theta)\, dW = \alpha, \tag{7}$$

$$\int_{w_0} \phi_1 p(E \mid \Theta)\, dW = 0. \tag{8}$$

Since the value of the integral (7) is independent of ϑ, all its derivatives with respect to $\theta_2, \theta_3, \cdots, \theta_l$ must vanish. This leads [3, pp. 50, 51. Insert k_t before ϕ_t^{-1} in (15)] to

---

[3] By this term we include "product moments."

[4] When I communicated this theorem to Professor Neyman, he informed me it was among the results of a thesis by R. Satô, *Contributions to the theory of testing statistical composite hypotheses*, University of London, 1937, and he kindly sent me a copy of the MS. I decided nevertheless to publish my version of theorem and proof, since for the reasons indicated in section 1 this theory should be available in the literature.

$$(9) \qquad \alpha^{-1} \int_{w_0} \prod_{i=2}^{l} \phi_i^{k_i} p(E \mid \Theta)\, dW = M(k_2, k_3, \cdots, k_l \mid \Theta), \qquad k_i = 0, 1, 2, \cdots,$$

where M is independent of w_0, and thus has the value obtained from (9) by putting $w_0 = W$ and $\alpha = 1$. In particular,

$$(10) \qquad \alpha^{-1} \int_{w_0} \phi_i p(E \mid \Theta)\, dW = 0, \qquad i = 2, 3, \cdots, l.$$

The necessary condition (9) for (7) is also sufficient. Denoting by $\mathcal{E}(f \mid w, \Theta)$ the expected value of a function $f(E, \Theta)$ calculated under the assumption that $E \epsilon w$, equation (9) may be written

$$(11) \qquad \mathcal{E}\left(\prod_{i=2}^{l} \phi_i^{k_i} \mid w_0, \Theta\right) = \mathcal{E}\left(\prod_{i=2}^{l} \phi_i^{k_i} \mid W, \Theta\right).$$

From assumption 5^0(a) it then follows that

$$(12) \qquad Q_0(\Phi \mid w_0, \Theta) = Q_0(\Phi \mid W, \Theta)$$

for almost all Φ. Conversely, (12) implies (11).

In a similar manner we get from (8) with the aid of (9),

$$(13) \qquad \mathcal{E}\left(\phi_1 \prod_{i=2}^{l} \phi_i^{k_i} \mid w_0, \Theta\right) = \mathcal{E}\left(\phi_1 \prod_{i=2}^{l} \phi_i^{k_i} \mid W, \Theta\right).$$

We calculate the moments of the function $Q_1(\Phi \mid w, \Theta)$ to be

$$\mathcal{E}\left(\phi_1 \prod_{i=2}^{l} \phi_i^{k_i} \mid w, \Theta\right),$$

and hence because of 5^0(b), (13) implies

$$(14) \qquad Q_1(\Phi \mid w_0, \Theta) = Q_1(\Phi \mid W, \Theta)$$

almost everywhere in the Φ-space. The pair of conditions (12), (14) are equivalent to the pair (7), (8).

In order that w_0 be a type B region, it is necessary and sufficient that it satisfy (12) and (14) and that

$$P''(w_0 \mid \Theta) \geqq P''(w_1 \mid \Theta)$$

for all w_1 satisfying (12) and (14). The inequality may be transformed with the help of 1^0(a), 2^0, (3), (7), (8), and (10) to

$$\int_{w_0} \phi_1^2 p(E \mid \Theta)\, dW \geqq \int_{w_1} \phi_1^2 p(E \mid \Theta)\, dW,$$

which is equivalent to

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \phi_1^2 p(\phi_1, \Phi \mid w_0, \Theta)\, d\phi_1\, d\phi_2 \cdots d\phi_l$$
$$\geqq \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \phi_1^2 p(\phi_1, \Phi \mid w_1, \Theta)\, d\phi_1\, d\phi_2 \cdots d\phi_l.$$

Sufficient for this is

$$Q_2(\Phi \mid w_0, \Theta) \geqq Q_2(\Phi \mid w_1, \Theta). \tag{15}$$

We note the functions in (12), (14), and (15) are all of the form (4) with $s = 0, 1, 2$, and propose to transform these to integrals over certain portions of the sample space W. First, we write (4) in the form

$$Q_0(\Phi \mid w, \Theta) \int_{-\infty}^{+\infty} \phi_1^s p(\phi_1 \mid \Phi, w, \Theta)\, d\phi_1 = Q_0(\Phi \mid w, \Theta) \mathcal{E}(\phi_1^s \mid \Phi, w, \Theta). \tag{16}$$

Next, we consider "surfaces" $S(\Phi, \Theta)$ in W_+, constructed as follows: For any fixed Θ let $D(\Theta)$ be the $l - 1$ dimensional domain of values of $\phi_i(E, \Theta)$, $i = 2, 3, \cdots, l$, for $E \in W_+$. A "surface" $S(\Phi, \Theta)$ is the locus of points E for which

$$\phi_i(E, \Theta) = \phi_i', \text{ a constant}, \qquad i = 2, 3, \cdots, l, \tag{17}$$

the set of constants being in $D(\Theta)$. Over every "surface" we now define a density ρ: Without loss of generality, and to simplify the notation, we shall assume that the non-singular minor postulated in 4^0 contains the minor $(\partial\phi_i/\partial x_j)$, $i = 2, 3, \cdots, l; j = 1, 2, \cdots, l - 1$, and denote by $J(E, \Theta)$ its determinant. For E on $S(\Phi, \Theta)$ we define the density

$$\rho(E \mid \Theta) = p(E \mid \Theta)/\mid J(E, \Theta) \mid, \tag{18}$$

and consider "surface" integrals

$$\int_{wS(\Phi,\Theta)} F_s(E, \Theta)\, dx_l\, dx_{l+1} \cdots dx_n, \tag{19}$$

where

$$F_s(E, \Theta) = \phi_1^s(E, \Theta)\rho(E \mid \Theta). \tag{20}$$

A "surface" integral (19) is to be distinguished from an ordinary multiple integral, in that the integrand is not merely a function of $x_l, x_{l+1}, \cdots, x_n$; there may be several points E on the surface with the same values for these coordinates, but different values for the integrand. The integral is to be thought of as follows: The part $wS(\Phi, \Theta)$ of the "surface" $S(\Phi, \Theta)$ is partitioned into pieces ΔS, on each a point E is chosen, and the value of the integrand at E is multiplied by the "area" of the projection (taken non-negative) of ΔS on the $x_l, x_{l+1}, \cdots, x_n$-space. The "surface" integral is the limit of the sum of such products as the norm of the partition approaches zero.

Denoting the integral (19) by $I(s)$ for the moment, we may calculate that for $\Phi \in D(\Theta)$

$$I(s) = I(0)\mathcal{E}(\phi_1^s \mid \Phi, w, \Theta), \qquad I(0) = Q_0(\Phi \mid w, \Theta)P(w \mid \Theta),$$

and hence we see that the right member of (16) is equal to the integral (19) divided by $P(w \mid \Theta)$. The desired relationship between the ordinary integrals (4) and the "surface" integrals (19) is thus

$$Q_s(\Phi \mid w, \Theta) = \int_{wS(\Phi,\Theta)} F_s(E, \Theta) \prod_{j=l}^{n} dx_j / P(w \mid \Theta). \tag{21}$$

The conditions (12), (14), (15) may now be written

$$\int_{w_0S(\Phi,\Theta)} F_s(E, \Theta) \prod_{j=l}^{n} dx_j = \alpha \int_{S(\Phi,\Theta)} F_s(E, \Theta) \prod_{j=l}^{n} dx_j, \qquad s = 0, 1, \tag{22}$$

$$\int_{w_0S(\Phi,\Theta)} F_2(E, \Theta) \prod_{j=l}^{n} dx_j \geqq \int_{w_1S(\Phi,\Theta)} F_2(E, \Theta) \prod_{j=l}^{n} dx_j, \tag{23}$$

if Φ is in the domain $D(\Theta)$, else they are satisfied trivially. w_0 will be a type B region if equations (22) are satisfied for almost all $\Phi \in D(\Theta)$, and if (23) is valid for all w_1 satisfying (22).

We now hold Θ fixed in ω and Φ fixed in $D(\Theta)$, so that $S(\Phi, \Theta)$ is fixed, and the right members of equations (22) have constant values. The proof [5, p. 11] of the lemma of Neyman and Pearson giving sufficient conditions that a region maximize an integral, subject to integral side-conditions, is easily seen to be valid for our "surface" integrals, and a sufficient condition that $w_0S(\Phi, \Theta)$ have the desired property is then that it be defined by

$$\phi_1^2(E, \Theta) > a_0 + a_1\phi_1(E, \Theta), \tag{24}$$

where a_0, a_1 are independent of E on $S(\Phi, \Theta)$, and are such that equations (22) are satisfied. Since Θ and Φ are fixed, we may permit a_i to be of the nature $a_i = a_i(\Phi, \Theta)$, $i = 1, 2$. Introducing functions $k_1 < k_2$, $k_i = k_i(\Phi, \Theta)$, and defining a_0, a_1 from

$$a_0 = -k_1k_2, \qquad a_1 = k_1 + k_2,$$

the inequality (24) is satisfied if (6) is. Still holding Θ fixed, suppose that k_1, k_2 can be determined for all Φ (hence almost all Φ) in $D(\Theta)$ so that for the part $w_0S(\Phi, \Theta)$ of $S(\Phi, \Theta)$, defined by (6), the equations (22) are satisfied. The parts $w_0S(\Phi, \Theta)$ of "surfaces" then sweep out a "solid" $w_0(\Theta)$ in W_+, defined by (6). If we can similarly determine k_1 and k_2, and hence $w_0(\Theta)$, for every Θ in ω, and if furthermore $w_0(\Theta)$ is independent of Θ, then it is the type B region we seek.

The equations (22) have now served their main purpose, and we return to their equivalents, (12) and (14). For $w_0(\Theta)$ defined by (6)

$$p(\phi_1, \Phi \mid w_0, \Theta) = p(\phi_1, \Phi \mid W, \Theta)/\alpha \quad \text{if} \quad \phi_1 < k_1 \quad \text{or} \quad \phi_1 > k_2,$$

and vanishes otherwise, and hence equations (12) and (14) are equivalent to (5).

The remainder of the proof consists of deducing that k_1, k_2 exist, and that the associated region $w_0(\Theta)$ is independent of Θ, for all $\Theta \in \omega$, from the hypothesis of our theorem that k_1, k_2 exist for some $\Theta = \Theta'$. By 1^0(c), Θ' lies on a line segment L entirely in ω, on which all but one of the nuisance parameters, say θ_2, are constant. Let us vary Θ over L. Then θ_3, θ_4, $\cdots$, θ_l remain fixed and θ_2 varies over an interval I. The equations (2) for $j = 2$ now become ordinary differential equations in which the independent variable is θ_2, the dependent variables are $\phi_2, \phi_3, \cdots, \phi_l$, and $\theta_1^0, \theta_3, \cdots, \theta_l$ are parameters. A well known existence theorem assures us of the existence of particular solutions u_i and a non-singular (for all θ_2 in I) matrix (u_{ij}) of complementary solutions, $i, j = 2, 3, \cdots, l$, such that the general solution is

$$\phi_i = u_i + \sum_{j=2}^{l} u_{ij} c_j .$$

The u_i are determined by initial conditions for the system (2) with $j = 2$, and the u_{ij} by sets of initial conditions for the corresponding complementary system. Clearly, if these initial conditions are all chosen independent of E, then since the coefficients of the differential equations are all independent of E, the solutions u_i and u_{ij} enjoy the same property. On the other hand, the c_j are independent of θ_2. Hence

$$\phi_i(E, \theta_2) = u_i(\theta_2) + \sum_{j=2}^{l} u_{ij}(\theta_2) c_j(E), \qquad i = 2, 3, \cdots, l. \tag{25}$$

The dependence of the ϕ's, u's and c's on the parameters $\theta_1^0, \theta_3, \cdots, \theta_l$ has not been indicated, since these remain fixed throughout the present calculations.

Let $\mathfrak{D}$ be the $l - 1$ dimensional domain of the values of $c_j(E)$ for $E \epsilon W_+$, and C: $(c_2', c_3', \cdots, c_l')$ be a point in $\mathfrak{D}$, and denote by $S(C)$ the "surface" $c_j(E) = c_j'$. Denote the surface $S(\Phi, \Theta)$ defined in (17) by $S(\Phi, \theta_2)$, and the domain $D(\Theta)$ of Φ by $D(\theta_2)$. Then since $| u_{ij} | \neq 0$, therefore for every $\theta_2 \epsilon I$, every $S(C)$ with $C \epsilon \mathfrak{D}$ is identical with some $S(\Phi, \theta_2)$ with $\Phi \epsilon D(\theta_2)$, and vice versa. From this we conclude for later reference: (A) the functions $c_j(E)$ are constant on every $S(\Phi, \theta_2)$; (B) if θ_2', θ_2'' are any two values in I, then for every $\Phi = \Phi'' \epsilon D(\theta_2'')$ there exists a $\Phi' \epsilon D(\theta_2')$ such that $S(\Phi', \theta_2')$ is identical with $S(\Phi'', \theta_2'')$, and vice versa.

Now let us integrate with respect to θ_2 the equation

$$\partial \log p(E \mid \theta_2)/\partial\theta_2 = \phi_2 = u_2(\theta_2) + \sum_{j=2}^{l} u_{2j}(\theta_2) c_j(E).$$

$$\log p(E \mid \theta_2) = v(\theta_2) + \sum_{j=2}^{l} v_j(\theta_2) c_j(E) + f(E),$$

where $v(\theta_2)$, $v_j(\theta_2)$, $f(E)$, and all new undefined symbols in the sequel have obvious meanings. We get

$$p(E \mid \theta_2) = \bar{v}(\theta_2)\bar{f}(E) \exp\left[\sum_{j=2}^{l} v_j(\theta_2) c_j(E)\right]. \tag{26}$$

Next we differentiate the equations (25) with respect to x_k, and write the result in matrix form,

$$(\partial\phi_i/\partial x_k) = (u_{ij})(\partial c_j/\partial x_k), \qquad i, j = 2, 3, \cdots, l; k = 1, 2, \cdots, l - 1.$$

Taking determinants, we have

$$J(E, \theta_2) = J_1(\theta_2) J_2(E). \tag{27}$$

Finally, we shall need to know the nature of the dependence of ϕ_1 on θ_2 and E: From (3),

$$\partial\phi_1/\partial\theta_2 = A_{12}(\theta_2) + B_{121}(\theta_2)\phi_1 + \sum_{j=2}^{l} B_{12k}(\theta_2)\phi_k.$$

Substituting from (25), we get

$$\partial\phi_1/\partial\theta_2 = B_{121}(\theta_2)\phi_1 + A(\theta_2) + \sum_{j=2}^{l} B_j(\theta_2)c_j(E),$$

and integrating,

$$\phi_1(E, \theta_2) = B(\theta_2)\left[\int^{\theta_2} \frac{A(\xi) + \sum_{j=2}^{l} B_j(\xi)c_j(E)}{B(\xi)}\, d\xi + g(E)\right],$$

where

$$B(\theta_2) = \exp\left[\int^{\theta_2} B_{121}(\eta)\, d\eta\right]. \tag{28}$$

Thus

$$\phi_1(E, \theta_2) = \bar{A}(\theta_2) + \sum_{j=2}^{l} \bar{B}_j(\theta_2)c_j(E) + B(\theta_2)g(E). \tag{29}$$

In equations (22) we now use the definitions (20), (18) for the integrands and then substitute (26), (27), (29). As a result we obtain the equality of

$$\int_{w_0 S(\Phi,\theta_2)} \frac{\left[\bar{A}(\theta_2) + \sum_{j=2}^{l} \bar{B}_j(\theta_2)c_j(E) + B(\theta_2)g(E)\right]^s \bar{v}(\theta_2)\bar{f}(E) \cdot \exp\left[\sum_{j=l}^{l} v_j(\theta_2)c_j(E)\right]}{|J_1(\theta_2)J_2(E)|} \prod_{j=l}^{n} dx_j$$

and α times the "surface" integral of the same integrand over $S(\Phi, \theta_2)$. Putting first $s = 0$ and then $s = 1$, and employing the previous conclusion (A), we find that the equations (22) are equivalent to

$$\int_{w_0 S(\Phi,\theta_2)} \{g^s(E)\bar{f}(E)/|J_2(E)|\} \prod_{j=l}^{n} dx_j = \alpha \int_{S(\Phi,\theta_2)} \{g^s(E)\bar{f}(E)/|J_2(E)|\} \prod_{j=l}^{n} dx_j, \qquad s = 0, 1. \tag{30}$$

Again using the expression (29) for ϕ_1, and noting from (28) that $B(\theta_2) > 0$, we may write the inequality (6) in the form

$$g(E) < \kappa_1(\Phi, \theta_2) \quad \text{and} \quad g(E) > \kappa_2(\Phi, \theta_2), \tag{31}$$

where

$$\kappa_i(\Phi, \theta_2) = \left[k_i(\Phi, \theta_1^0, \vartheta) - \bar{A}(\theta_2) - \sum_{j=2}^{l} \bar{B}_j(\theta_2)c_j(E)\right] \Big/ B(\theta_2). \tag{32}$$

It follows from our hypothesis that for $\theta_2 = \theta_2'$ (the θ_2 coordinate of Θ') and any $\Phi \epsilon D(\theta_2')$, functions $\kappa_i(\Phi, \theta_2')$ exist such that for the part $w_0S(\Phi, \theta_2')$ of $S(\Phi, \theta_2')$, defined by (31), equations (30) are satisfied. The region $w_0(\Theta')$ is "swept out" by $w_0S(\Phi, \theta_2')$ as Φ ranges over $D(\theta_2')$. Now let Θ'' be any other $\Theta \epsilon L$, call its θ_2 coordinate θ_2'', let Φ'' be any $\Phi \epsilon D(\theta_2'')$, and consider the possibility of finding $\kappa_i(\Phi'', \theta_2'')$ such that on the part $w_0S(\Phi'', \theta_2'')$ of $S(\Phi'', \theta_2'')$, defined by (31), equations (30) are satisfied. From the conclusion (B), $S(\Phi'', \theta_2'')$ is identical with $S(\Phi', \theta_2')$ for a suitably chosen $\Phi' \epsilon D(\theta_2')$. Hence if we take $\kappa_i(\Phi'', \theta_2'') = \kappa_i(\Phi', \theta_2')$, then $w_0S(\Phi'', \theta_2'')$ becomes identical with $w_0S(\Phi', \theta_2')$ where equations (30) are already satisfied. Letting Φ'' range over $D(\theta_2'')$, every $w_0S(\Phi'', \theta_2'')$ thus determined becomes identical with some $w_0S(\Phi', \theta_2')$, and vice versa, by (B). Thus the region $w_0(\Theta'')$ "swept out" is identical with $w_0(\Theta')$. This process defines $\kappa_i(\Phi, \theta_2)$ for all $\theta_2 \epsilon I$ and $\Phi \epsilon D(\theta_2)$, and hence determines $k_i(\Phi, \theta_1^0, \vartheta)$ from (32). We now have functions $k_i(\Phi, \theta_1^0, \vartheta)$, $k_1 < k_2$, satisfying (5), and corresponding regions $w_0(\Theta)$ independent of Θ, for all $\Theta \epsilon L$. To conclude the proof, we use $1^0(c)$ to reach any point Θ in ω from Θ' by a path consisting of a finite number of segments like L on which only one of the nuisance parameters varies. The definitions of $k_i(\Phi, \theta_1^0, \vartheta)$ are continued along this path as above and the region $w_0(\Theta)$ is seen to be independent of Θ for all Θ in ω.

The following theorem may be regarded as a generalization of one by Neyman [6, p. 33] giving sufficient conditions that a type A region be also of type A_1:

THEOREM 2. *Suppose the assumption $1^0(b)$ holds for all $\Theta \epsilon \Omega$. Denote $\phi_i(E, \theta_1^0, \vartheta)$ by ϕ_i^0 and let $R(\vartheta)$ be the domain of values of $\phi_1^0, \phi_2^0, \cdots, \phi_l^0$ for $E \epsilon W_+$ and $\Theta \epsilon \omega$. Then a sufficient condition that a region w_0 of type B, found by application of theorem 1, be also of type B_1 is that for all $\Theta \epsilon \Omega$ and all $E \epsilon W_+$*

$$p(E \mid \theta_1, \vartheta) = p(E \mid \theta_1^0, \vartheta)g(\phi_1^0, \phi_2^0, \cdots, \phi_l^0; \theta_1^0; \theta_1, \vartheta), \tag{33}$$

where $g(y_1, y_2, \cdots, y_l; \theta_1^0; \theta_1, \vartheta)$ is a function such that $\partial^2 g/\partial y_1^2 > 0$ for all $y_1, y_2, \cdots, y_l$ in $R(\vartheta)$ and $\Theta \epsilon \Omega - \omega$.

For the w_0 satisfying the sufficient conditions of theorem 1, the conditions (a), (b′), (c) of definition 2 are satisfied, and it remains only to verify the condition (d′). The regions w_1 admitted for comparison in (d), as well as w_0, must satisfy the equations (22) since these are equivalent to the conditions (a), (c). We recall that $\Theta = (\theta_1^0, \vartheta)$ in equations (22) and rewrite them in a notation better adapted to our present considerations:

$$\int_{w_0S(\Phi^0,\theta_1^0,\vartheta)} [\phi_1^0]^s\{p(E \mid \theta_1^0, \vartheta)/|J(E, \theta_1^0, \vartheta)|\} \prod_{j=l}^{n} dx_j = \alpha \int_{S(\Phi^0,\theta_1^0,\vartheta)} [\phi_1^0]^s\{p(E \mid \theta_1^0, \vartheta)/|J(E, \theta_1^0, \vartheta)|\} \prod_{j=l}^{n} dx_j, \quad s = 0, 1 \tag{34}$$

where $\Phi^0 = (\phi_2^0, \phi_3^0, \cdots, \phi_l^0) \epsilon D(\theta_1^0, \vartheta)$.

To express the condition (d) in a convenient way, we now "shred" the regions w_0, w_1 of (d) for every θ_1 by means of the same "surfaces" we have been using for $\theta_1 = \theta_1^0$: For any w in W_+, $\Theta \epsilon \Omega$, and $\Phi^0 \epsilon D(\theta_1^0, \vartheta)$ we define a "surface" integral

$$I(\Phi^0, w \mid \theta_1, \vartheta) = \int_{wS(\Phi^0, \theta_1^0, \vartheta)} \{p(E \mid \theta_1, \vartheta)/|J(E, \theta_1^0, \vartheta)|\} \prod_{j=l}^{n} dx_j .$$

Then

$$P(w \mid \theta_1, \vartheta) = \int \cdots \int_{D(\theta_1^0, \vartheta)} I(\Phi^0, w \mid \theta_1, \vartheta)\, d\phi_2^0\, d\phi_3^0 \cdots d\phi_l^0,$$

and a sufficient condition for (d) is

$$(35) \qquad I(\Phi^0, w_0 \mid \theta_1, \vartheta) \geqq I(\Phi^0, w_1 \mid \theta_1, \vartheta)$$

for all $\Theta \epsilon \Omega$ and all $\Phi^0 \epsilon D(\theta_1^0, \vartheta)$.

Again applying the lemma of Neyman and Pearson to the integrands of the "surface" integrals in (34) and (35), we find that a sufficient condition that our region w_0 be of type B_1 is that there exist functions $b_i(\Phi^0, \theta_1^0, \theta_1, \vartheta)$, $i = 1, 2$, such that

$$p(E \mid \theta_1, \vartheta) > p(E \mid \theta_1^0, \vartheta)[b_0 + b_1\phi_1^0(E, \theta_1^0, \vartheta)]$$

if and only if $E \epsilon w_0$. Employing (33), we may replace this inequality by

$$(36) \qquad g(\phi_1^0, \Phi^0; \theta_1^0; \theta_1, \vartheta) > b_0 + b_1\phi_1^0 .$$

Define b_0, b_1 from

$$g(k_i, \Phi^0; \theta_1^0; \theta_1, \vartheta) = b_0 + b_1 k_i, \qquad i = 1, 2,$$

where $k_i = k_i(\Phi^0, \theta_1^0, \vartheta)$. Since $k_1 < k_2$, these equations have unique solutions b_0, b_1. Now hold Φ^0, θ_1, ϑ all fixed ($\theta_1 \neq \theta_1^0$) and consider the graphs of the members of (36) as functions of ϕ_1^0. From our definition of b_0, b_1, these graphs intersect at k_1, k_2. But by hypothesis, the graph of the left member is everywhere concave up, and hence for $k_1 < \phi_1^0 < k_2$, it lies below the linear graph of the right member, and for $\phi_1^0 < k_1$ and $\phi_1^0 > k_2$, it lies above. That is (36) is true if and only if $E \epsilon w_0$.

5. Appendix on the moment problem. Easily applied criteria [8] are available for the moment problem of assumption $5^0(a)$. The moment problem $5^0(b)$ is much more difficult, however, because the function to be determined by its moments is not of constant sign. Below we offer a proof that the solutions of both problems $5^0(a)$ and $5^0(b)$ are unique in the important case where $p(E \mid \Theta)$ is a multivariate normal p.d.f. and $\phi_1, \phi_2, \cdots, \phi_l$ are polynomials in x_1, $x_2, \cdots, x_n$ of degree $\leqq 2$ and not necessarily homogeneous. Since Θ is held fixed, we will not indicate dependence on Θ, nor will the dependence of various functions on s be indicated, since $s = 0$ or else 1 throughout.

Let w_1, w_2 be any two regions, $\alpha_j = P(w_j) \neq 0$, for which the moments of $Q_s(\Phi \mid w_1)$ and $Q_s(\Phi \mid w_2)$ are the same. To prove the equality (almost everywhere) of these two functions it suffices to prove that their Fourier transforms are identical [7, theorem 61]. Suppressing the customary multiple of $\sqrt{2\pi}$, the Fourier transform of $Q_s(\Phi \mid w_j)$ is

$$\Psi_j(\mathbf{t}) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{i\mathbf{t}\cdot\Phi} Q_s(\Phi \mid w_j)\, d\phi_2 \cdots d\phi_l,$$

where $\mathbf{t}$ is the vector $(t_2, t_3, \cdots, t_l)$ and $\mathbf{t}\cdot\Phi = t_2\phi_2 + \cdots + t_l\phi_l$. From (4) we get

$$\begin{aligned}\Psi_j(\mathbf{t}) &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{i\mathbf{t}\cdot\Phi} \phi_1^s p(\phi_1, \Phi \mid w_j)\, d\phi_1\, d\phi_2 \cdots d\phi_l \\ &= \mathcal{E}(e^{i\mathbf{t}\cdot\Phi} \phi_1^s \mid w_j) \\ &= \frac{1}{\alpha_j} \int_{w_j} e^{i\mathbf{t}\cdot\Phi} \phi_1^s(E) p(E)\, dW.\end{aligned}$$

A device of Cramér and Wold [8] for reducing the dimensionality of the problem now suggests itself. Let z be a scalar variable and consider $\psi_j(z \mid \mathbf{t}) = \Psi_j(z\mathbf{t})$ for fixed $\mathbf{t}$ as a function of z. Obviously if for every fixed $\mathbf{t}$, $\psi_1(z \mid \mathbf{t}) = \psi_2(z \mid \mathbf{t})$, then $\Psi_1(\mathbf{t}) = \Psi_2(\mathbf{t})$, and we are through. We propose to prove the former equality by showing first that ψ_j is an analytic function of z for all real z and secondly that the coefficients of the power series for ψ_1 and ψ_2 in powers of z are equal. Holding $\mathbf{t}$ fixed now, $\xi = \mathbf{t}\cdot\Phi$ is a polynomial of degree $\leqq 2$, and

$$\psi_j(z \mid \mathbf{t}) = \frac{1}{\alpha_j} \int_{w_j} e^{iz\xi} \phi_1^s p\, dW. \tag{37}$$

By our assumption of normality,

$$p = C \exp\left[- \sum_{\kappa, \nu = 1}^{n} a_{\kappa\nu} y_\kappa y_\nu \right], \qquad y_\kappa = x_\kappa - \mu_\kappa,$$

where the matrix $(a_{\kappa\nu})$ is positive definite. To prove the analyticity of ψ_j for any real $z = z_0$, let $z = z_0 + \zeta$, and restrict ζ to real values. Substitute in (37)

$$e^{i\zeta\xi} = \sum_{q=0}^{m-1} \frac{(i\zeta\xi)^q}{q!} + \frac{(i\zeta\xi)^m}{m!} f_m(\zeta\xi),$$

where $|f_m(\zeta\xi)| \leqq 1$. Then

$$\psi_j(z_0 + \zeta \mid \mathbf{t}) = \sum_{q=0}^{m-1} \frac{(i\zeta)^q}{q!\, \alpha_j} \int_{w_j} e^{iz_0\xi} \xi^q \phi_1^s p\, dW + R_{jm}(z_0, \zeta),$$

where

$$R_{jm} = \frac{(i\zeta)^m}{m!\, \alpha_j} \int_{w_j} e^{iz_0\xi} f_m(\zeta\xi) \xi^m \phi_1^s p\, dW,$$

and all integrands are absolutely integrable over W. Let σ be the sphere of unit radius with center at $(\mu_1, \mu_2, \cdots, \mu_n)$ in W and write

$$R_{jm} = \frac{(i\zeta)^m}{m!\,\alpha_j}\left[\int_{w_j\sigma} + \int_{w_j - w_j\sigma}\right].$$

Call the two terms of the right member R'_{jm} and R''_{jm},

$$R_{jm} = R'_{jm} + R''_{jm}.$$

$$|R'_{jm}| \leqq \frac{|\zeta|^m}{m!\,\alpha_j}\int_{w_j\sigma} |\xi^m \phi_1^s|\, p\, dW.$$

Let $M = \max |\xi|$, $M_1 = \max |\phi_1^s|$, for $E \epsilon \sigma$. Then

$$|R'_{jm}| \leqq \frac{M_1 |M\zeta|^m}{m!\,\alpha_j}\int_\sigma p\, dW \leqq M_1 |M\zeta|^m/m!\,\alpha_j.$$

Hence $R'_{jm} \to 0$ for all real ζ as $m \to \infty$.

$$|R''_{jm}| \leqq \frac{|\zeta|^m}{m!\,\alpha_j}\int_{W-\sigma} |\xi^m \phi_1^s|\, p\, dW.$$

Let $r = \left(\sum_{\kappa=1}^{n} y_\kappa^2\right)^{\frac{1}{2}}$, and M_2, M_3 be the sums of the absolute values of the coefficients of the polynomials ϕ_1^s, ξ, respectively, when expanded in powers of y_κ. Then for $E \epsilon W - \sigma$, $|\phi_1^s| \leqq M_2 r^2$, $|\xi| \leqq M_3 r^2$, $p \leqq C \exp(-\lambda r^2)$, where $\lambda > 0$ is the smallest characteristic root of $(a_{\kappa\nu})$. Hence

$$|R''_{jm}| \leqq \frac{CM_2 |M_3\zeta|^m}{m!\,\alpha_j}\int_{W-\sigma} r^{2m+2} e^{-\lambda r^2}\, dW.$$

Integrating over spherical shells concentric with σ, $dW = M_4 r^{n-1}\, dr$, and

$$|R''_{jm}| \leqq \frac{CM_2 M_4 |M_3\zeta|^m}{m!\,\alpha_j}\int_1^\infty r^{2m+n+1} e^{-\lambda r^2}\, dr \leqq \frac{CM_2 M_4 |M_3\zeta|^m}{m!\,\alpha_j}\int_0^\infty.$$

If we evaluate the last integral in terms of a Gamma function and employ Stirling's formula we easily find that for $M_3 |\zeta| < \lambda$, $R''_{jm} \to 0$. The convergence of R_{jm} to zero for *real* ζ, $|\zeta| < \lambda/M_3$, is sufficient to insure the analyticity of ψ_j.

Now let $z_0 = 0$. Then the coefficient of ζ^q in the power series for ψ_j is

$$\frac{i^q}{q!\,\alpha_j}\int_{w_j} (t_2\phi_2 + \cdots + t_l\phi_l)^q \phi_1^s\, p\, dW,$$

a linear combination (the same for $j = 1, 2$) of the q-th order moments of $Q_s(\Phi \mid w_j)$, and hence corresponding coefficients for ψ_1 and ψ_2 are equal.

REFERENCES

[1] J. Neyman, "Sur la vérification des hypothèses statistiques composées," *Bull. Soc. Math. France*, Vol. 63 (1935), pp. 246–266.
[2] J. Neyman, "Outline of a theory of statistical estimation based on the classical theory of probability," *Phil. Trans. Roy. Soc. London*, A, Vol. 236 (1937), pp. 333–380.
[3] J. Neyman, "On a statistical problem arising in routine analyses and in sampling inspections of mass production," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 46–76.
[4] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. Roy. Soc. London*, A, Vol. 231 (1933), pp. 289–337.
[5] J. Neyman and E. S. Pearson, "Contributions to the theory of testing statistical hypotheses: Part I," *Stat. Res. Mem.*, Vol. 1 (1936), pp. 1–37.
[6] J. Neyman and E. S. Pearson, "Contributions to the theory of testing statistical hypotheses: Part II," *Stat. Res. Mem.*, Vol. 2 (1938), pp. 25–36.
[7] S. Bochner, *Vorlesungen über Fouriersche Integrale*, Leipzig, 1932.
[8] H. Cramér and H. Wold, "Some theorems on distribution functions," *Jour. London Math. Soc.*, Vol. 11 (1936), pp. 290–294.