ON A STATISTICAL PROBLEM ARISING IN ROUTINE ANALYSES AND IN SAMPLING INSPECTIONS OF MASS PRODUCTION

By J. Neyman

University of California, Berkeley, Calif.

CONTENTS

1. **Introduction.** The words "routine analyses" are used to denote the analyses performed by laboratories, frequently attached to industrial plants, and distinguished by the following characteristics: (1) All the analyses or measurements are of the same kind, for example, are designed to measure the sugar content in beets or to determine the coordinate of a star. (2) The analyses are carried out day after day using the same methods and the same instruments. (3) While all the analyses are of the same kind, the quantity measured varies from time to time and each such quantity is measured repeatedly n times, where n represents some small number, 2, 3, 4, 5.

As an illustration we may consider the routine analyses of sugar beets performed in the process of selection and breeding. A small section is cut out of each of a great number of sugar beets expected to be suitable for further breeding. It is crushed and its juice extracted to determine ξ, the sugar content of each particular beet. From the juice available from each beet n samples are taken and a determination of the sugar content is made from each. Thus, if ξ_i represents the sugar content of the section from the ith beet and there are N beets, the laboratory will have to make nN analyses with their results $x_{i,1}$, $x_{i,2}, \cdots, x_{i,n}$, representing the measurements of the same quantity ξ_i. Obviously the sugar content ξ_i referring to the ith beet need have no relation to that of any other jth beet.

An essential point in the above description is that the number of measurements referring to the same quantity ξ_i is usually very small. For example, the quantitative analyses of urine in certain clinics are performed only twice for each patient, so that $n = 2$. Frequently, various practical considerations make it impossible to increase this number n of analyses intended to measure the same quantity ξ_i.

The smallness of n introduces difficulties in estiminating ξ_i. It is usual to consider $x_{i,1}, x_{i,2}, \cdots, x_{i,n}$ as independent variables, varying normally about ξ_i with an unknown standard error σ_i. If they have to be used to estimate ξ_i, then the confidence interval [1][1] for ξ_i will be determined by the familiar formula

$$(1) \qquad x_{i.} - s_i t_\alpha(n) \leq \xi_i \leq x_{i.} + s_i t_\alpha(n),$$

where $x_{i.}$ denotes the mean of the x_{ij},

$$(2) \qquad s_i^2 = \sum_{j=1}^{n} (x_{ij} - x_{i.})^2/n(n-1)$$

and $t_\alpha(n)$ is Fisher's t corresponding to the number of degrees of freedom $n - 1$ and to the chosen confidence coefficient α. It is known [2] that if the estimate of ξ_i is based only on its direct measurements $x_{i,1}, x_{i,2}, \cdots, x_{i,n}$, then the confidence interval (1) can not be made any smaller; in fact, formula (1) gives the shortest unbiased confidence interval for ξ_i. But if we try to substitute appropriate numbers in (1) we get disconcerting results. Namely, if $n = 2$ and $\alpha = .99$, then $t_\alpha(n) = 63.657$. If n is increased, the value of $t_\alpha(n)$ decreases rapidly but for $n = 5$ it is still very considerable, $t_\alpha(5) = 4.604$, and consequently the numerical confidence interval determined by (1) is frequently so broad that it is devoid of practical value.

The general conclusion is that, if n cannot be increased, satisfactory estimates of ξ_i can only be obtained when they are based on something else in addition to the direct measurements $x_{i,1}, x_{i,2}, \cdots, x_{i,n}$. This point was first noticed by "Student" [3]. His method of avoiding the difficulty consists in assuming that the accuracy of measurements performed in the same laboratory is constant in time, so that $\sigma_1 = \sigma_i = \cdots = \sigma_N = \sigma$. If this is true, then $s_0^2 = \Sigma s_i^2/N$ will be an unbiased estimate of the variance of x_{ij}, based on $N(n - 1)$ degrees of freedom. If the past experience of the laboratory is of any size, as measured by N, then the product $N(n - 1)$ will be of considerable size and the confidence interval for ξ_i

$$(3) \qquad x_{i.} - s_0 t_\alpha(N(n-1)+1) \leq \xi_i \leq x_{i.} + s_0 t_\alpha(N(n-1)+1)$$

will be much more satisfactory than (1).

The problem which arises is whether we are entitled to assume that $\sigma_1 = \sigma_2 = \cdots = \sigma_N$. The first study of this problem seems to have been made by Przyborowski [4] in a paper written in Polish. His findings, subsequently reported [5] in English, show that, at least in certain cases, the accuracy of routine analyses is quite difficult to keep constant. If it is not constant, then the relative frequency of the cases where formula (3) gives correct statements about ξ_i will generally be different from the expected α.

[1] Figures in square brackets refer to the literature quoted at the end of the paper.

The procedure employed by Przyborowski to test whether $\sigma_1 = \sigma_2 = \cdots = \sigma_N$ consisted in considering the quantities $v_i = (n - 1)s_i^2$ and applying the χ^2 test to see whether they follow the same χ^2 distribution with $n - 1$ degrees of freedom

$$p(v) = cv^{\frac{1}{2}(n-3)}e^{-\frac{1}{2}v/\sigma^2} \tag{4}$$

with an unknown σ.

Just this point is to be the main subject of this paper. The χ^2 test was devised by Karl Pearson with no particular set of alternative hypotheses in view. As a result we may expect that in many cases other tests may be devised which would be more powerful. A number of such cases are already on record [6], [7], [8].

2. **Statistical hypothesis H to be tested.** We shall consider the case where we can observe the particular values of Nn random variables $x_{i,j}$, $i = 1, 2, \cdots, N; j = 1, 2, \cdots, n$, and we know that $x_{i,j}$ is independent of x_{kl} for $i \neq k$ and that

$$p(x_{i,1}, x_{i,2}, \cdots x_{i,n}) = \left(\frac{1}{\sigma_i\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{j=1}^{n}(x_{i,j}-\xi_i)^2/\sigma_i^2} \tag{5}$$

with unknown values of ξ_i and $\sigma_i > 0$. The hypothesis H to be tested is that $\sigma_1 = \sigma_2 = \cdots = \sigma_N = \sigma$ without specifying, however, the actual value of σ. It will be noticed that this hypothesis has already been treated by a number of authors [9]–[17]. The need for considering it again arises from the fact that previously it was tested against the set of alternatives presuming that the σ_1, σ_2, $\cdots \sigma_N$, were positive constants having any values whatsoever. It seems to the author that, in the present case, the set of alternatives should be different. This will be explained in the next section. It follows that while the hypothesis tested is the same as in the papers quoted above, the problem of testing it is quite different.

Let us denote by E the whole set of Nn observable variables. If H is true then their elementary probability law will be

$$p(E \mid H) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{Nn} e^{-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n}(x_{i,j}-\xi_i)^2/\sigma^2} \tag{6}$$

3. **General problem of similar regions.** The development of the test will follow the general lines explained elsewhere [18], [19], [20]. Denoting by W the Nn dimensional space of the $x_{i,j}$'s, we want to determine a region w in W having the following properties: (a) if the hypothesis tested is true then the probability of E falling in w shall have some fixed value chosen in advance, e.g., $\epsilon = .05$ or $\epsilon = .01$. This probability is known as the probability of an error of the first kind. (b) If H is not true then the probability of E falling in w as determined by one of the alternative hypotheses (that we assume likely to be true when H is false) shall be as large as possible in a sense that requires further explanation.

The probability with which this condition is concerned is a complement of the probability of an error of the second kind. Once the region w is chosen it will be used to test H in this way: if E falls within w, then H will be rejected.

In the present section we shall deal only with ways of satisfying condition (a). The problem is similar to the one recently described by Hotelling [21]. The difficulty is that, if H is true, the probability law of E is given by (6) and contains $N + 1$ unspecified parameters, "nuisance" parameters as Hotelling very appropriately calls them. If we take just any region w then it is most likely that the probability of E falling in it will vary with different values of $\sigma, \xi_1, \cdots, \xi_N$. As a matter of fact, if we want the test to be absolutely most powerful, or at least relatively so, we must determine not just one single region satisfying (a) but actually *all* such regions or some broad family of them. From these we shall then select one which seems most satisfactory from the point of view of (b).

Systematic methods of determining regions of the above kind have already been considered [18], [20], [2]. In these publications they are called "similar" to the sample space W. The reason for this term is that the whole space W does possess the required properties with $\epsilon = 1$. In fact, whatever be the values of the nuisance parameters, $\sigma, \xi_1, \cdots, \xi_N$, the probability of E falling within W, as calculated from (6), is perfectly determined and equals 1. Our problem is to find a region w, part of W, with similar properties for $0 < \epsilon < 1$. However, in many cases no such regions exist [22].

The general methods in the above publications are applicable in the present case. However, a recent paper by Cramér and Wold [23] allows a slight improvement in presenting the matter. As this is a little involved, it seems desirable to take up the whole problem and present it anew.

Consider then the general case where the probability law of some m observable variables $y_1, y_2, \cdots, y_m$, say $p(E \mid \theta_1, \cdots, \theta_s)$, as specified by the hypothesis tested, depends on s nuisance parameters $\theta_1, \theta_2, \cdots, \theta_s$. Our problem will consist of determining the necessary and sufficient conditions for a region w to be similar to the sample space with respect to all these parameters. We shall assume that the probability law $p(E \mid \theta_1, \cdots, \theta_s)$ satisfies certain limiting conditions.

Let

$$\varphi_i = \frac{\partial \log p}{\partial \theta_i} \tag{7}$$

$$\varphi_{ij} = \frac{\partial^2 \log p}{\partial \theta_i \partial \theta_j} \tag{8}$$

Assume that for all values of i and $j = 1, 2, \cdots, s$

$$\varphi_{ij} = A_{i,j} + \sum_{k=1}^{s} B_{i,j,k} \varphi_k \tag{9}$$

where the coefficients $A_{i,j}$ and $B_{i,j,k}$ are independent of the observable variables E. Assume also that the probability law $p(E \mid \theta_1, \cdots, \theta_s)$ permits indefinite differentiation under the sign of the integral taken over any fixed region w in W. It is easy to check that the probability law (6) satisfies all of these conditions.

In order to find the necessary conditions for the region w to be similar to W with respect to $\theta_1, \theta_2, \cdots, \theta_s$, assume that w is actually similar and that, consequently,

$$(10) \qquad P\{E \epsilon w \mid \theta_1, \cdots, \theta_s\} = \int \cdots \int_w p(E \mid \theta_1, \cdots \theta_s)\, dy_1 \cdots dy_m \equiv \epsilon$$

for all possible values of $\theta_1, \theta_2, \cdots, \theta_s$. It follows that the derivatives of all orders with respect to $\theta_1, \theta_2, \cdots, \theta_s$ taken from the left side of (10) must be identically equal to zero. But we have

$$(11) \qquad \begin{aligned} &\frac{\partial}{\partial \theta_i} \int \cdots \int_w p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \\ &\qquad = \int \cdots \int_w \frac{\partial}{\partial \theta_i} p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \\ &\qquad = \int \cdots \int_w \varphi_i p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \equiv 0 \end{aligned}$$

for $i = 1, 2, \cdots, s$. Similarly, using (9)

$$(12) \qquad \begin{aligned} &\frac{\partial^2}{\partial \theta_i \partial \theta_j} \int \cdots \int p(E \mid \theta_1, \cdots \theta_s)\, dy_1 \cdots dy_m \\ &= \int \cdots \int_w \left(\varphi_i \varphi_j + A_{ij} + \sum_{k=1}^{s} B_{i,j,k} \varphi_k\right) p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \equiv 0. \end{aligned}$$

Using (10) and (11), the last identity will be reduced to

$$(13) \qquad \frac{1}{\epsilon} \int \cdots \int_w \varphi_i \varphi_j p(E \mid \theta_1, \cdots \theta_s)\, dy_1 \cdots dy_m \equiv -A_{i,j} \quad \text{for } i, j = 1, 2, \cdots, s$$

where the right side does not depend on the particular region w, provided that w is similar to the sample space. Considering the identities (11) and (13) which were obtained by differentiating (10) twice, we may guess what will happen if we differentiate (13) again and again. We may assume, in fact, that, whatever be the non-negative integers $k_1, k_2, \cdots, k_s$, we shall obtain

$$(14) \qquad \frac{1}{\epsilon} \int \cdots \int_w \prod_{i=1}^{s} \varphi_i^{k_i} p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \equiv M(k_1, k_2, \cdots, k_s),$$

where $M(k_1, \cdots, k_s)$ is independent of the particular region w, provided that w is similar to the sample space with respect to all of the θ's. Assume that this is found for all k's such that $\sum_{i=1}^{s} k_i \leq K$; also assume that the sum of the k's in (14) is exactly K. Differentiating with respect to θ_j, we obtain

$$(15) \qquad \begin{aligned} &\frac{1}{\epsilon} \int \cdots \int_w \left\{\varphi_j \prod_{i=1}^{s} \varphi_i^{k_i} + \prod_{i=1}^{s} \varphi_i^{k_i} \sum_{t=1}^{s} \varphi_t^{-1} \varphi_{t,j}\right\} p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \\ &\qquad\qquad \equiv \frac{\partial}{\partial \theta_j} M(k_1, \cdots, k_s). \end{aligned}$$

Because of the particular form of $\varphi_{t,j}$, the second expression in the curly brackets under the integral is a polynomial in the φ's of order not exceeding K. According to the assumption made, this expression multiplied by $p(E \mid \theta_1, \cdots, \theta_s)/\epsilon$ and integrated over w gives a result which is independent of w. As the right side of (15) is also independent of w, we conclude that

$$(16) \qquad \frac{1}{\epsilon}\int \cdots \int_w \varphi_j \prod_{i=1}^{s} \varphi_i^{k_i} p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \equiv M(k_1, \cdots, k_j + 1, \cdots, k_s)$$

is also independent of the particular similar region chosen. We have seen that (14) is true for $K \leq 2$ and that if it is true for K it is true for $K + 1$, that is, it is true in general.

We may now sum up our findings: if w is a region similar to the sample space with respect to all of the θ's and if ϵ denotes the value of the integral (10), then, whatever be the non-negative integers $k_1, k_2, \cdots, k_s$, the value of the integral on the left side of (14) is independent of the particular region w chosen.

As the whole sample space W is also "similar" with $\epsilon = 1$, it must satisfy this identity. This allows us to determine the M's, namely

$$(17) \qquad \int \cdots \int_W \prod_{i=1}^{s} \varphi_i^{k_i} p(E \mid \theta_1, \cdots, \theta_s)\, dy_1 \cdots dy_m \equiv M(k_1, \cdots, k_s).$$

It is obvious that the necessary condition above is also sufficient. If a region w is such that (14) holds for all systems of non-negative integers then all the derivatives of (10) must be identically zero; thus the left side of (10) is independent of $\theta_1, \theta_2, \cdots, \theta_s$.

It will be useful to interpret the above conditions as follows. We start by noticing that the left side of (17) represents the product moment of some specified order of the $\varphi_1, \varphi_2, \cdots, \varphi_s$ considered as random variables. We shall call it the absolute product moment. We will now interpret the left side of (14) as a product moment also. For this purpose we shall define a new elementary probability law of the y's to be denoted by $p(E \mid w, \theta_1, \cdots, \theta_s)$ and described as the relative probability law given w. We shall write it as

$$(18) \qquad p(E \mid w, \theta_1, \cdots, \theta_s) = \frac{1}{\epsilon} p(E \mid \theta_1, \cdots, \theta_s)$$

for all of the points E included in w and

$$(19) \qquad p(E \mid w, \theta_1, \cdots, \theta_s) = 0$$

for all other points. With this definition the left side of (14) appears to be the expectation of the product $\varphi_1^{k_1} \cdots \varphi_s^{k_s}$ calculated from the relative probability law of the y's given w. We will call it the relative product moment given w. The final result can now be stated as follows:

For a region w to be similar to the sample space with respect to $\theta_1, \theta_2, \cdots, \theta_s$ it is necessary and sufficient that all the relative moments and product moments of $\varphi_1, \varphi_2, \cdots, \varphi_s$ shall equal the corresponding absolute moments and product moments.

In order to make the method of constructing similar regions according to the above conditions clear we recall the procedure involved in the calculation of the probability laws of any given set of random variables.

Assume then that the elementary probability law of the original variables is given. Fix some values of the parameters $\theta_1, \theta_2, \cdots, \theta_s$, denote the resulting probability law by $p(E)$, and consider the problem of finding the elementary probability law of $\varphi_1, \varphi_2, \cdots, \varphi_s$ considered as functions of the y's. We shall assume that none of the φ's can be expressed as a function of the others not involving the y's explicitly so that the matrix

$$\left(\begin{array}{cccc} \frac{\partial \varphi_1}{\partial y_1} & \frac{\partial \varphi_1}{\partial y_2} & \cdots & \frac{\partial \varphi_1}{\partial y_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \varphi_s}{\partial y_1} & \frac{\partial \varphi_s}{\partial y_2} & \cdots & \frac{\partial \varphi_s}{\partial y_m} \end{array}\right) \tag{20}$$

is non-singular. In these circumstances it is possible to select $m - s$ functions of the y's say $\psi_{s+1}, \psi_{s+2}, \cdots, \psi_m$ which have continuous second derivatives such that the formulae

$$\begin{aligned} z_i &= \varphi_i \qquad & i &= 1, 2, \cdots, s \\ z_j &= \psi_j \qquad & j &= s + 1, \cdots, m \end{aligned} \tag{21}$$

determine a one-to-one transformation of the space W of the y's into the space W' of the z's. If w denotes any region in W then it will be transformed into a perfectly determined region w' in W'. If E' denotes a point in W' then the probability of E' falling in w' will be identical with that of E falling in w. Thus

$$P\{E' \epsilon w'\} \equiv P\{E \epsilon w\} = \int \cdots \int_w p(E)\, dy_1 \cdots dy_m. \tag{22}$$

Letting J be the Jacobian of the y's with respect to the z's in the transformation (21) and using the known formulae for transforming multiple integrals, we have

$$P\{E' \epsilon w'\} = \int \cdots \int_{w'} p(E)\Big]_{E'} |J|\, dz_1 \cdots dz_m, \tag{23}$$

where $p(E)]_{E'}$ denotes the result of substituting the expressions for the y's in terms of the z's as obtained from (21) into $p(E)$. It follows that, whatever be the region w' in W', the probability of E''s falling in it is obtained by integrating the function $p(E)]_{E'} \mid J \mid$ over w'. But this means, according to the usual definition, that the product $p(E)]_{E'} \mid J \mid$ is the elementary probability law of the z's. Denoting it by $p(E') = p(z_1, \cdots, z_m)$ we have

$$p(E') = p(E)]_{E'} \mid J \mid. \tag{24}$$

Now, to obtain the joint probability law of $\varphi_1, \varphi_2, \cdots, \varphi_s$ or that of $z_1, z_2, \cdots, z_s$ we must integrate $p(E')$ for all the other z's between their extreme limits, formally between $-\infty$ and $+\infty$ for each of the variables concerned,

$$(25) \qquad p(\varphi_1, \cdots, \varphi_s) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} p(E')\, dz_{s+1} \cdots dz_m.$$

This procedure will be applied when calculating the absolute probability law of the φ's and also the relative one given w. The only difference will be that in the latter case we shall have to start with (18) and (19) instead of the original probability law. The space W' and the transformation (21) will be the same in both cases. It is important to be clear about the difference between the two cases. This is connected with the difference between $p(E \mid \theta_1, \cdots, \theta_s)$ and $p(E \mid w, \theta_1, \cdots, \theta_s)$ of (18) and (19). The latter is proportional to the former at any point E within the region w but is zero outside of w. As mentioned above, the integrations for $z_{s+1}, z_{s+2}, \cdots, z_m$ in (25) should extend formally from $-\infty$ to $+\infty$ for each variable. However, the probability law $p(E')$ may equal zero within certain parts of this range. Fixing any system of values $z_i = \varphi_i$, for $i = 1, 2, \cdots, s$, is equivalent to fixing a hypersurface in the space W and considering the intersection of planes $z_i =$ constant in the space W'. Denote them by $W(\varphi)$ and $W'(\varphi)$, respectively. If we shift the point E or E' along $W(\varphi)$ or $W'(\varphi)$ respectively, the variables $z_j = \psi_j$, for $j = s + 1, s + 2, \cdots, m$ will assume a certain set $S(\varphi)$ of systems of values. When calculating the absolute probability law of $\varphi_1, \cdots, \varphi_s$ this set $S(\varphi)$ will be the real region of integration in (25); outside of it the function under the integral sign will be zero. On the other hand, when calculating the relative probability law of $\varphi_1, \cdots, \varphi_s$ given w, the function under the integral (25) is zero as soon as the point E moves outside of the region w. Denote by $w(\varphi)$ that part of $W(\varphi)$ which is included in w and by $w'(\varphi)$ the corresponding part of $W'(\varphi)$. So, the absolute and the relative, given w, probability laws of $\varphi_1, \cdots, \varphi_s$ can be obtained by using the formulae

$$(26) \qquad p(\varphi_1, \cdots, \varphi_s) = \int \cdots \int_{W'(\varphi)} p(E')\, dz_{s+1} \cdots dz_m$$

$$(27) \qquad p(\varphi_1, \cdots, \varphi_s \mid w) = \frac{1}{\epsilon} \int \cdots \int_{w'(\varphi)} p(E')\, dz_{s+1} \cdots dz_m.$$

Now the method of constructing regions similar to W with respect to $\theta_1, \theta_2, \cdots, \theta_s$ is clear: to construct any such region it is necessary and sufficient to select for each of all possible systems of values of $\varphi_1, \varphi_2, \cdots, \varphi_s$ a part $w(\varphi)$ of the hypersurface $W(\varphi)$ and to combine all these parts. The selection of $w(\varphi)$ is arbitrary save for the restriction that the probability law (27) have all its moments equal to those of (26), identically in the θ's. This last condition will certainly be satisfied if $w(\varphi)$ is so selected that for almost all systems of values of $\varphi_1, \varphi_2, \cdots, \varphi_s$

$$p(\varphi_1, \cdots, \varphi_s \mid w) \equiv p(\varphi_1, \cdots, \varphi_s) \tag{28}$$

for all values of the θ's.

By selecting $w(\varphi)$ in all possible ways that satisfy (28) we obtain an infinity of regions similar to W with respect to $\theta_1, \theta_2, \cdots, \theta_s$. They form a family which we shall denote by $F(\epsilon)$. However, it is known that in general all the moments of $p(\varphi_1, \cdots, \varphi_s \mid w)$ and $p(\varphi_1, \cdots, \varphi_s)$ may be identical without the two probability laws being equal almost everywhere. In such cases, the family $F(\epsilon)$ will not exhaust all the similar regions. It is important to be able to state whether or not $F(\epsilon)$ contains all the similar regions. To ascertain this we may use the conditions of Cramér and Wold [23] which are sufficient for the determinateness of the problem of moments, that is, for the uniqueness of a function having a given set of moments.

Let

$$\mu_\nu = M(\nu, 0, 0, \cdots, 0) + M(0, \nu, 0, \cdots, 0) + \cdots + M(0, 0, \cdots, 0, \nu). \tag{29}$$

With this notation the conditions of Cramér and Wold can be stated as follows: If any two probability laws, e.g., the probability laws $p(\varphi_1, \cdots, \varphi_s \mid w)$ and $p(\varphi_1, \cdots, \varphi_s)$, have all their moments and all their product moments identical and if the series

$$\sum_\nu \mu_{2\nu}^{-1/2\nu} \tag{30}$$

is divergent, then

$$p(\varphi_1, \cdots, \varphi_s \mid w) \equiv p(\varphi_1, \cdots, \varphi_s) \tag{31}$$

almost everywhere.

Therefore, to know whether the family $F(\epsilon)$ defined above exhausts all the regions similar to W, we must calculate the even moments of all the φ_i and see whether the series (30) depending on these moments is divergent. If it is, there is no similar region besides the family $F(\epsilon)$. Otherwise, there may be some others. These others will be constructed by selecting $w(\varphi)$'s such that the integral (27) equals any other probability law having the same moments as (26). In such cases, a region w selected, in one way or another, from the family $F(\epsilon)$ as the best from the point of view of controlling errors of the second kind will only be the relative best.

It should be mentioned that whether we can always, under the conditions considered, select a $w(\varphi)$ on any $W(\varphi)$ that satisfies the identity (28) has not yet been proved. However, it seems plausible that the differential equations (9) imply the existence of a sufficient set of statistics for $\theta_1, \theta_2, \cdots, \theta_s$. If this is so, the possibility of satisfying (28) is guaranteed (see [2], p. 366).

4. **Regions similar to the sample space with respect to $\sigma, \xi_1, \xi_2, \cdots, \xi_N$.** We may now return to the original problem and apply our theory to the probability law (6). We wish to construct the most general regions similar to the sample space with respect to the nuisance parameters $\sigma, \xi_1, \cdots, \xi_N$ unspecified by the hypothesis tested. We let

$$\varphi_\sigma = \frac{\partial \log p}{\partial \sigma} = -\frac{Nn}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{N} \sum_{j=1}^{n} (x_{ij} - \xi_i)^2, \tag{32}$$

$$\varphi_i = \frac{\partial \log p}{\partial \xi_i} = \frac{n(x_{i\cdot} - \xi_i)}{\sigma^2} \quad \text{with} \quad x_{i\cdot} = \frac{1}{n} \sum_{j=1}^{n} x_{ij}. \tag{33}$$

Then

$$\begin{aligned} \frac{\partial \varphi_\sigma}{\partial \sigma} &= -\frac{3}{\sigma}\varphi_\sigma - \frac{2Nn}{\sigma^2} \\ \frac{\partial \varphi_\sigma}{\partial \xi_i} &= -2\sigma\varphi_i \\ \frac{\partial \varphi_i}{\partial \xi_i} &= -\frac{n}{\sigma^2} \\ \frac{\partial \varphi_i}{\partial \xi_j} &= 0, \qquad i \neq j \end{aligned} \tag{34}$$

and we see that the probability law (6) satisfies the differential equations (9).

Now the hypersurfaces $W(\varphi)$ of the theory are the intersections of the hypersurfaces

$$\varphi_\sigma = \text{constant and} \qquad \varphi_i = \text{constant}, \qquad \text{for } i = 1, 2, \cdots, N. \tag{35}$$

The latter equations are clearly equivalent to

$$x_{i\cdot} = \text{constant}. \tag{36}$$

As to the former, we notice the identity

$$\sum_{i=1}^{N} \sum_{j=1}^{n} (x_{i,j} - \xi_i)^2 = n \sum_{i=1}^{N} (S_i^2 + (x_{i\cdot} - \xi_i)^2) = \chi^2, \text{ (say)} \tag{37}$$

where $nS_i^2 = \sum_{j=1}^{n} (x_{i,j} - x_{i\cdot})^2$. Therefore, $W(\varphi)$ denotes the intersection of the hypersurfaces (36) with, say,

$$T_1 = \sum_{i=1}^{N} S_i^2 = \text{constant}. \tag{38}$$

If we succeed in selecting from each hypersurface $W(\varphi)$ a part $w(\varphi)$ satisfying condition (28) identically then the sum of all such regions $w(\varphi)$ will form a region w similar to W with respect to all the unspecified parameters and belonging to the family $F(\epsilon)$. Before proceeding to this stage of the solution, let us see whether the family $F(\epsilon)$ exhausts all of the similar regions.

For this purpose notice first that instead of considering whether there is but one probability law with moments equal to those of φ_σ and the φ_i's, it is sufficient to concern ourselves with the moments of χ^2 and $x_{i.}$. In fact, all the φ's are functions of these variables and the problem of uniqueness of the distribution must have the same answer in both cases. The 2νth absolute moment of χ^2 as calculated from (6) equals

$$(39) \qquad (2\sigma^2)^{2\nu}\Gamma(\tfrac{1}{2}Nn + 2\nu)/\Gamma(\tfrac{1}{2}Nn).$$

The same order moment of $x_{i.}$ is

$$(40) \qquad \sigma^{2\nu}(2\nu)!/(2n)^{\nu}\nu!.$$

Thus, the quantity denoted by $\mu_{2\nu}$ in the theory becomes

$$(41) \qquad \mu_{2\nu} = \frac{(2\sigma^2)^{2\nu}\Gamma(\frac{1}{2}Nn + 2\nu)}{\Gamma(\frac{1}{2}Nn)} + N\left(\frac{\sigma^2}{n}\right)^{\nu}\frac{(2\nu)!}{2^{\nu}\nu!}.$$

We are interested in whether or not the series (30) is divergent. Since $\mu_{2\nu}$ satisfies the inequality

$$(42) \qquad \mu_{2\nu} < a^{2\nu}\Gamma(b + 2\nu) = C_{2\nu}^{-2\nu}, \qquad \text{(say)}$$

with $a = 2\sigma^2 + N$ and $2b = Nn$, if we prove that the series $\Sigma C_{2\nu}$ diverges, then (30) also diverges. To settle this conveniently we apply Stirling's formula to $\Gamma(b + 2\nu)$ and find that, as $\nu \to \infty$, the ratio $C_{2\nu}/\nu^{-1}$ tends to a finite limit. As the series $\Sigma\nu^{-1}$ is divergent, so is the series $\Sigma C_{2\nu}$ and thus the series $\Sigma\mu_{2\nu}^{-1/2\nu}$ is divergent. Therefore, there is but one probability law with moments identical to those of χ^2 and the $x_{i.}$'s and so the family $F(\epsilon)$ contains all the regions similar to the sample space with respect to $\sigma, \xi_1, \cdots, \xi_N$.

It may now be interesting to go into some details of the effective construction of any region similar to W with respect to $\sigma, \xi_1, \cdots, \xi_N$. For this purpose it is convenient to go back and express the identity (28), that the regions $w(\varphi)$ must satisfy, in terms of the relative probability law of $z_{s+1}, z_{s+2}, \cdots, z_m$ given $\varphi_1, \varphi_2, \cdots, \varphi_s$. This is denoted by $p(z_{s+1}, z_{s+2}, \cdots, z_m \mid \varphi_1, \cdots, \varphi_s)$ and defined for every system of values of the φ's for which $p(\varphi_1, \varphi_2, \cdots, \varphi_s) \neq 0$ as follows:

$$(43) \qquad \begin{aligned} &p(z_{s+1}, z_{s+2}, \cdots, z_m \mid \varphi_1, \varphi_2, \cdots, \varphi_s) \\ &\qquad = p(\varphi_1, \cdots, \varphi_s, z_{s+1}, \cdots, z_m)/p(\varphi_1, \cdots, \varphi_s). \end{aligned}$$

Using (26), (27), and (43), the identity (28) can be rewritten in the following form

$$(44) \qquad \int \cdots \int_{w'(\varphi)} p(z_{s+1}, \cdots, z_m \mid \varphi_1, \cdots, \varphi_s)\, dz_{s+1} \cdots dz_m \equiv \epsilon.$$

The function under this integral is the relative elementary probability law of $z_{s+1}, z_{s+2}, \cdots, z_m$ and it is integrated over the region $w'(\varphi)$. Therefore, the left side of (44) is nothing but the relative probability of the point E' falling in

$w'(\varphi)$ given that the first s of its coordinates have the fixed values $\varphi_1, \varphi_2, \cdots, \varphi_s$. In other words, and owing to the one-to-one correspondance between the spaces W and W', we have

$$(45) \qquad P\{E' \epsilon w'(\varphi) \mid E' \epsilon W'(\varphi)\} = P\{E \epsilon w(\varphi) \mid E \epsilon W(\varphi)\} \equiv \epsilon.$$

Now the general method of determining similar regions may be stated as follows:

1. Choose any system of variables $z_{s+1}, z_{s+2}, \cdots, z_m$ such that their values determine uniquely the position of the point E' on any fixed hypersurface $W'(\varphi)$. These z's considered as functions of the y's should be continuously differentiable twice.

2. Find the relative probability law of the z's given the φ's. This must be done for every possible set of values of the φ's.

3. In the space of $z_{s+1}, z_{s+2}, \cdots, z_m$ consider regions which satisfy the equality (44) identically in the θ's. Any such region could be taken to form a part of w', the region similar to the sample space, which we are trying to construct. If the assumption that the differential equations (9) imply the existence of a sufficient system of statistics for $\theta_1, \theta_2, \cdots, \theta_s$ is true, then (see [2], p. 366) the probability law $p(z_{s+1}, z_{s+2}, \cdots, z_m \mid \varphi_1, \cdots, \varphi_s)$ will be independent of the θ's and there will be an infinity of regions satisfying (44).

Obviously, instead of dealing directly with $\varphi_1, \varphi_2, \cdots, \varphi_s$ as described above, we may select any system of statistics $T_1, T_2, \cdots, T_s$ such that the system of equations T_i = constant is equivalent to φ_i = constant, for $i = 1, 2, \cdots, s$.

Returning to the particular problem of similar regions with respect to σ, $\xi_1, \cdots, \xi_N$, we notice that instead of the φ's we may consider

$$(46) \qquad T_1 = \sum_{i=1}^{N} S_i^2 \quad \text{and} \quad T_{i+1} = x_i. \qquad \text{for } i = 1, 2, \cdots, N.$$

Now we wish to select a convenient system of variables, denoted by z_{s+j}'s in the theory above, to determine the position of the point E' on any hypersurface $W'(\varphi)$ where all the functions (46) have fixed values. Obviously there is no unique choice and we shall use what we find convenient. But notice that the total number of these variables should be, in our case, $Nn - N - 1$. The following system may be suggested.

If the sum ΣS_i^2 has a fixed value T_1 then none of the S_i^2 can exceed T_1. Write

$$(47) \qquad \begin{aligned} S_i^2 &= u_i T_1 \\ S_N^2 &= \left(1 - \sum_{i=1}^{N-1} u_i\right) T_1 \qquad \text{for } i = 1, 2, \cdots, N-1 \end{aligned}$$

and consider $u_1, u_2, \cdots, u_{N-1}$ as belonging to the system of variables sought. The region of their variation is determined by the inequalities

$$(48) \qquad 0 \le u_i \quad \text{and} \quad \sum_{i=1}^{N-1} u_i \le 1$$

If the u's are fixed then they, together with the value of T_1, determine the values of $S_1, S_2, \cdots, S_N$. As the values of $x_{i.} = T_{i+1}$ are already fixed, we have to solve the problem of choosing for each $i = 1, 2, \cdots, N$ a system of $n - 2$ variables, say $z_{i,1}, z_{i,2}, \cdots, z_{i,n-2}$, which with $x_{i.}$ and S_i will completely determine the values of $x_{i,1}, x_{i,2}, \cdots, x_{i,n}$. However, this will only have to be done if $n > 2$. Following the now familiar method (see, for example, [5], pp. 33–43), we may determine the $z_{i,j}$ in two consecutive steps. First write

$$
\begin{aligned}
x_{i,1} &= x_{i.} + \sqrt{\frac{1}{1\cdot 2}}v_{i,1} + \sqrt{\frac{1}{2\cdot 3}}v_{i,2} + \cdots + \sqrt{\frac{1}{(n-1)n}}v_{i,n-1} \\
x_{i,2} &= x_{i.} - \sqrt{\frac{1}{1\cdot 2}}v_{i,1} + \sqrt{\frac{1}{2\cdot 3}}v_{i,2} + \cdots + \sqrt{\frac{1}{(n-1)n}}v_{i,n-1} \\
x_{i,3} &= x_{i.} \qquad\qquad - 2\sqrt{\frac{1}{2\cdot 3}}v_{i,2} + \cdots + \sqrt{\frac{1}{(n-1)n}}v_{i,n-1} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_{i,n} &= x_{i.} \qquad\qquad\qquad\qquad - (n-1)\sqrt{\frac{1}{(n-1)n}}v_{i,n-1}
\end{aligned} \tag{49}
$$

where $v_{i,1}, v_{i,2}, \cdots, v_{i,n-1}$ are new variables satisfying the identity

$$\sum_{j=1}^{n-1} v_{i,j}^2 = \sum_{j=1}^{n} (x_{i,j} - x_{i.})^2. \tag{50}$$

We transform them further by putting

$$
\begin{aligned}
v_{i,1} &= \sqrt{n}\, S_i \cos z_{i,n-2} \cos z_{i,n-3} \cdots \cos z_{i,2} \cos z_{i,1} \\
v_{i,2} &= \sqrt{n}\, S_i \cos z_{i,n-2} \cos z_{i,n-3} \cdots \cos z_{i,2} \sin z_{i,1} \\
v_{i,3} &= \sqrt{n}\, S_i \cos z_{i,n-2} \cos z_{i,n-3} \cdots \sin z_{i,2} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
v_{i,n-1} &= \sqrt{n}\, S_i \sin z_{i,n-2}
\end{aligned} \tag{51}
$$

with the z's varying as follows

$$
\begin{aligned}
0 &\le z_{i,1} < 2\pi \\
-\pi/2 &\le z_{i,j} \le \pi/2 \qquad \text{for } j = 2, 3, \cdots, n-2.
\end{aligned} \tag{52}
$$

Of course, instead of the S_i we should put their expressions in terms of T_1 and the u's into (51). With the exception of a set of measure zero, which can be ignored, the formulae above determine a one-to-one transformation of the original space W of the x's into the space W' of $T_1, T_2, \cdots, T_{N+1}, u_1, \cdots, u_{N-1}$, and $z_{i,1}, z_{i,2}, \cdots, z_{i,n-2}$ for $i = 1, 2, \cdots, N$.

In calculating the joint probability law of all the new variables, we notice that, on the hypothesis tested, all the Nn original variables are mutually independent. Consequently, the transformations (49) and (51), which refer to separate groups of the $x_{i,j}$'s, corresponding to fixed values of i, could be carried through separately. In doing so, we use formulae deduced elsewhere (see [5], pp. 38–39) directly and obtain

$$(53)\quad p(x_{i\cdot}, S_i, z_{i,1}, \cdots, z_{i,n-2}) = \left(\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\right)^n S_i^{n-2} e^{-\frac{1}{2}n(S_i^2+(x_{i\cdot}-\xi_i)^2)/\sigma^2} \prod_{j=2}^{n-2} \cos^{j-1} z_{i,1}.$$

It follows that

$$(54)\quad \begin{aligned} p(x_{1\cdot}, \cdots, x_{N\cdot}, S_1, \cdots, S_N, z_{i,1}, \cdots, z_{N,n-2}) &= \prod_{i=1}^{N} p(x_{i\cdot}, S_i, z_{i,1}, \cdots, z_{i,n-2}) \\ &= \left(\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\right)^{Nn} e^{-\frac{1}{2}n\sum_{i=1}^{N}(x_{i\cdot}-\xi_i)^2/\sigma^2} \prod_{i=1}^{N} S_i^{n-2} e^{-\frac{1}{2}nS_i^2/\sigma^2} \prod_{k=1}^{N}\prod_{j=2}^{n-2} \cos^{j-1} z_{k,j}. \end{aligned}$$

We now wish to introduce T_1 and the u_i instead of the S_i's. Since all other variables remain unchanged the Jacobian of this transformation reduces to that of (47). Simple calculations show that

$$(55)\quad \left|\frac{\partial(S_1, S_2, \cdots, S_N)}{\partial(T_1, u_1, \cdots, u_{N-1})}\right| = (\tfrac{1}{2})^N T_1^{\frac{1}{2}(N-2)} \left(1 - \sum_{i=1}^{N-1} u_i\right)^{-\frac{1}{2}} \prod_{i=1}^{N-1} u_i^{-\frac{1}{2}}.$$

Using this expression and substituting (47) in (54) we finally obtain

$$(56)\quad \begin{aligned} p(x_{1\cdot}, \cdots, x_{N\cdot}, T_1, u_1, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2}) &= (\tfrac{1}{2})^N \left(\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\right)^{Nn} e^{-\frac{1}{2}n\sum_{i=1}^{N}(x_{i\cdot}-\xi_i)^2/\sigma^2} T_1^{\frac{1}{2}N(n-1)-1} \\ &\cdot e^{-\frac{1}{2}nT_1/\sigma^2} \left(\left(1 - \sum_{i=1}^{N-1} u_i\right) \prod_{i=1}^{N-1} u_i\right)^{\frac{1}{2}(n-3)} \prod_{k=1}^{N}\prod_{j=2}^{n-2} \cos^{j-1} z_{k,j}. \end{aligned}$$

To obtain the relative probability law of $u_1, u_2, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2}$ given T_1 and the $T_{i+1} = x_{i\cdot}$, we must calculate $p(T_1, T_2, \cdots, T_{N+1})$ and divide expression (56) by it. Of course, $p(T_1, T_2, \cdots, T_{N+1})$ is obtained from (56) by integrating over the whole of $W'(\varphi)$, that is, for all other variables between the extreme limits of their variation. As these limits are independent of the values of $T_1, T_2, \cdots, T_{N+1}$, the result will be

$$(57)\quad p(T_1, T_2, \cdots, T_{n+1}) = ce^{-\frac{1}{2}n\sum_{i=1}^{N}(x_{i\cdot}-\xi_i)^2/\sigma^2} T_1^{\frac{1}{2}N(n-1)-1} e^{-\frac{1}{2}nT_1/\sigma^2}$$

where c denotes a constant. Thus

$$(58)\quad \begin{aligned} p(u_1, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2} \mid T_1, \cdots, T_{N+1}) &= c_1\left(\left(1 - \sum_{i=1}^{N-1} u_i\right) \prod_{i=1}^{N-1} u_i\right)^{\frac{1}{2}(n-3)} \prod_{k=1}^{N}\prod_{j=2}^{n-2} \cos^{j-1} z_{k,j} \end{aligned}$$

with the region of variation $W'(\varphi)$ limited by the following inequalities

$$
(59) \qquad \begin{aligned}
&0 \le u_i, \qquad \sum_{i=1}^{N-1} u_i \le 1 \\
&0 \le z_{k,1} < 2\pi \qquad\qquad \text{for } k = 1, 2, \cdots, N, \\
&-\pi/2 \le z_{k,1} \le \pi/2 \qquad\qquad j = 2, 3, \cdots, n-2.
\end{aligned}
$$

Since (58) integrated over $W'(\varphi)$ is identically unity, c_1 is a purely numerical constant.

Now to construct any region w similar to the sample space with respect to σ, $\xi_1, \cdots, \xi_N$, we must select, separately for each and all systems (φ) of values of $T_1, T_2, \cdots, T_{N+1}$, a region $w'(\varphi)$, part of $W'(\varphi)$ as defined by (59), with the sole restriction that

$$
(60) \qquad \int \cdots \int_{w'(\varphi)} p(u_1, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2} \mid T_1, \cdots, T_{N+1}) \cdot du_1, \cdots, dz_{N,n-2} = \epsilon.
$$

Obviously, there is an infinity of ways of selecting any single one of such regions. For example, we could let the u's vary as indicated in (59) and limit the z's by

$$
(61) \quad 0 \le z_{k,1} \le a, \quad -a \le z_{k,j} \le a \quad (k = 1, 2, \cdots, N; j = 2, 3, \cdots, n-2)
$$

where a is chosen so that (60) is satisfied. This choice of $w'(\varphi)$ may correspond to one particular system of values of $T_1, T_2, \cdots, T_{N+1}$ and no other. Again, the same region (61) may be chosen to serve for all systems of values of the T's. In this case, the region $w = \sum_{\varphi} w(\varphi)$ might be described as cylindrical. Any such region w will control errors of the first kind in testing H to the same level of significance ϵ and, as far as these errors alone are concerned, each of these regions is of equal value. Whatever the choice of regions $w'(\varphi)$ or $w(\varphi)$, the test of H will consist of (1) observing the values of the $x_{i,j}$'s, (2) calculating the corresponding value of $T_1, T_2, \cdots, T_{N+1}$, the u's, and the z's, and (3) noting whether the point with coordinates $u_1, u_2, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2}$ falls in the region $w(\varphi)$ chosen to correspond to the observed values of $T_1, T_2, \cdots, T_{N+1}$. Of course, in practical cases, the choice of $w'(\varphi)$ for one system of values of the T's will not be quite unconnected with that for others. On the contrary, there will probably be some more or less simple rule connecting $w'(\varphi)$ with the corresponding systems of the T's. As a result, the actual machinery of the test will be much simpler than that described above and will consist of the calculation of only a very few functions of the x's and in checking some simple inequalities.

Now our purpose is to select a region from the infinite family $F(\epsilon)$ of all regions similar to the sample space with respect to σ, $\xi_1, \cdots, \xi_N$ which we judge most satisfactory for controlling errors of the second kind. Roughly speaking, this region will have to be such that, if the hypothesis H is not true, the observed point E will fall in this particular region as frequently as possible, in general. Here we come to the necessity of specifying the ways in which we expect the hypothesis H to be untrue. It may be untrue in an infinite number of ways. For example, the values of the σ's may (1) be equally distributed over any given range, (2) may fall into just two groups $\sigma_i = 1$ and $\sigma_j = 2$, or (3) all σ_i's except the last may have the same value σ while the last is 10σ, and so forth. Any such assumption will be called an hypothesis alternative to H. It is obvious that the probability of E falling in any given region w will be different for each of them. Therefore, if we wish to deduce a test which will detect the falsehood of the hypothesis tested frequently, we must analyse the practical cases where the test is to be applied and guess the ways in which the hypothesis tested is usually wrong. Then we can deduce a test which will be, in one sense or another, most sensitive to the assumed deviations from the hypothesis tested. Needless to say, our guess may be right or wrong. In the latter case, an increased volume of observational material may demonstrate its fallacy and suggest the necessary modifications. In any case, it is important to know exactly the class of alternatives for which our test is, in some particular way, the best.

5. **The set of hypotheses alternative to** H. Let us consider the routine analyses made at some laboratory and try to discover the circumstances likely to cause variation in their accuracy. First of all, we may think of assignable causes such as a change in personnel, apparatus, or accommodation. These and similar causes are likely to produce lasting effects; the test of the hypothesis that they did not reduces to one of the equality of only two σ's. An easy application of known theory [20] shows that the familiar F or z test is unbiased of type B_1, which means that it is preferable to any other. Consequently, situations of this kind and also similar one for which the L_1 test is applicable [9], need not be considered here, so that we may concentrate on cases where there is no directly assignable cause of variation in the accuracy of the analyses. Assume then that the personnel, the apparatus, the accommodation, etc., remain the same. Now the accuracy of analyses depends on a multitude of causes evading identification, such as changes in the efficiency of the workers. In principle, they try to have the highest, and therefore a constant, level of accuracy. Uncontrollable circumstances cause some fluctuations about a certain average and we expect that small deviations from this average will occur more frequently than large ones. With this in mind, the author feels that it would be appropriate to expect that variations in accuracy, if any, will have a random character so that any σ_i referring to one particular group of analyses, or any monotonic function of that σ_i could be considered as an essentially positive random variable, having some unimodal probability law. To make the problem of the best test sufficiently specific, we must specify this law entirely. Here we face a somewhat embarassing freedom of choice. For lack of more precise information as to the random variability of σ_i, we guide ourselves by considerations of ease in

calculations. From this point of view it is convenient to consider the variable

$$h = \sigma^{-2} \tag{62}$$

and assume that, within a given period of time which is not too long, when the conditions in a laboratory are sensibly constant, it is varying according to the law

$$p(h) = \beta^{\alpha} h^{\alpha-1} e^{-\beta h}/\Gamma(\alpha) \quad \text{for} \quad 0 < h, \tag{63}$$

where α and β are unknown non-negative constants. It is useful to express these constants in terms of two new ones which have an obvious interpretation: h_0, the expectation of h, and ν, the square of the coefficient of variation of h. Easy calculations give

$$\alpha = 1/\nu, \qquad \beta = 1/h_0\nu. \tag{64}$$

Now $p(h)$ has the form

$$p(h) = \frac{1}{(h_0\,\nu)^{1/\nu}\Gamma(1/\nu)}\, h^{(1/\nu)-1} e^{-h/h_0\nu}. \tag{65}$$

We note that when $\nu \to 0$ the probability law (65) tends to a limiting discontinuous form with $P\{h = h_0\} = 1$. This corresponds to the hypothesis H that we wish to test. The type of law represented by (65) is known to be rather flexible. Consequently, we may easily assume that even though the true variability of h (or σ) does not exactly correspond to (65), there will be a system of values of h_0 and ν for which the difference between the true law and (65) will not be large. Therefore, a test which is particularly sensitive to deviations of ν from zero with law (65) will be reasonably sensitive in real practical cases. However, this is an assumption by the author. But it is subject to test and this will be done below.

Formula (63) represents the hypothetical probability law of the variable h which is not directly observable. We must use this formula to obtain the probability law of the observable x's alternative to (6), which corresponds to the hypothesis H being true. Using $h = 1/\sigma^2$, we write the relative probability law of $x_{i,1}, x_{i,2}, \cdots, x_{i,n}$ given h

$$p(x_{i,1}, \cdots, x_{i,n} \mid h) = \left(\frac{h}{2\pi}\right)^{n/2} e^{-\frac{1}{2}h\sum_{j=1}^{n}(x_{i,j}-\xi_i)^2}. \tag{66}$$

Multiplying (66) by (65) we obtain the joint probability law of h and the $x_{i,j}$'s referring to one group of analyses

$$p(h, x_{i,1}, \cdots, x_{i,n}) = \frac{1}{(2\pi)^{n/2}(h_0\nu)^{1/\nu}\Gamma(1/\nu)}\, h^{n/2+(1/\nu)-1} e^{-h\left(1/h_0\nu+\frac{1}{2}\sum_{j=1}^{n}(x_{i,j}-\xi_i)^2\right)}. \tag{67}$$

Integrating (67) with respect to h from zero to infinity, we obtain the absolute probability law of $x_{i,1}, x_{i,2}, \cdots, x_{i,n}$, all referring to the ith group of analyses. Assuming that the value of h in one group of analyses is independent of that in another, we obtain the joint probability law of all the Nn observable $x_{i,j}$'s by simply multiplying the probability laws referring to particular groups of n of them. The result will depend on $N + 2$ unknown parameters, $\xi_1, \xi_2, \cdots, \xi_N, h_0$, and ν. As the last two will play a more important role than the others we shall denote the probability law by $p(E \mid h_0, \nu)$. Easy calculations give

$$(68) \qquad p(E \mid h_0, \nu) = \left(\frac{\Gamma(n/2 + 1/\nu)}{(2\pi)^{n/2}\Gamma(1/\nu)}\right)^N \frac{(h_0\nu)^{\frac{1}{2}Nn}}{\prod_{i=1}^{N}\left(1 + \frac{h_0\nu}{2}\sum_{j=1}^{n}(x_{i,j} - \xi_i)^2\right)^{n/2+1/\nu}}.$$

We easily check that for $\nu \to 0$ (68) approaches the law (6) with $h_0 = \sigma^{-2}$. Therefore, the problem that we shall treat below will be to assume that the observable x's follow (68) with *some* $h_0 > 0$ and *some* $\nu \geq 0$ and to test the hypothesis H that $\nu = 0$. More specifically, we shall try to choose among all the regions of the family $F(\epsilon)$, found in the preceding section the one over which the integral of the function (68) is, in general, the largest.

Before doing so, it may be useful to exhibit some experimental evidence in favor of the assumption that, if σ is not constant in some conditions of analysis or measurement, then it varies in such a way that the variability of the x's has at least some characteristics appropriate to (68).

Introduce the notation

$$(69) \qquad \omega_i = nS_i^2 = \sum_{j=1}^{n}(x_{i,j} - x_{i.})^2.$$

Using transformations (49), (50), and (69), successively, we easily deduce the probability law of ω_i

$$(70) \qquad p(\omega_i) = \frac{(h_0\nu/2)^{\frac{1}{2}(n-1)}\Gamma(\frac{1}{2}(n-1) + 1/\nu)}{\Gamma(\frac{1}{2}(n-1))\Gamma(1/\nu)} \frac{\omega_i^{\frac{1}{2}(n-3)}}{(1 + \frac{1}{2}h_0\nu\omega_i)^{\frac{1}{2}(n-1)+1/\nu}}.$$

If the hypothesis we have made about the variability of h, as expressed by (65), is true in any particular case then the sums of squares (69), referring to each particular group of analyses, are distributed according to (70). The reverse is not necessarily true, of course, but it is comforting that a check of the above in a number of broadly divergent circumstances gives satisfactory results. By applying the transformation $1 + h_0\nu\omega_i/2 = t^{-1}$, the integral of (70) is easily reduced to an Incomplete Beta function whence Pearson's tables [24] provide an easy means of calculating the theoretical probability that ω_i is within any given limits.

Table I gives several observed distributions of the sums ω together with their expected ones, calculated from (70) with the values of h_0 and ν fitted by the method of moments. The last lines give particulars of the application of the χ^2 test for goodness of fit.

The origin of the data used to compile Table I is as follows:

For the data providing frequency distributions numbered 1 and 2, the author is deeply indebted to Professor Raymond T. Birge. The methods of measurement and their purpose are explained in the publications [25] and [26], respec-

TABLE I

Comparison of empirical distributions of ω with those calculated from (70)

Number	1		2		3		4		5	
Author or Source of Data	R. T. Birge		R. T. Birge		K. Buszczyński and Sons, Ltd.		A. A. Michelson, F. G. Pease, and F. Pearson		W. S. Svenson	
Kind of Measurement or Analysis	Strong Lines in the Band Spectra of Nitrogen		A Solar Spectrum Line		Sugar Content of Beets		Velocity of Light		Octane Rating	
	Frequency		Frequency		Frequency		Frequency		Frequency	
ω	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
0–1	29·38	29	15·10	17	15·56	16	3·50	2	14·90	17
1–2	19·30	20	13·14	11	12·67	17	7·73	10	18·88	16
2–3	13·11	17	11·39	15	10·70	13	9·37	13	16·83	14
3–4	9·16	7	9·84	5	8·98	2	9·66	8	13·93	12
4–5	6·56	6	8·46	9	7·53	11	9·28	17	11·20	10
5–6	4·80	1	7·24	9	6·34	4	8·60	7	8·91	7
6–7	3·59	4	6·17	11	5·36	3	7·80	7	7·04	10
7–8	} 4·80	1	5·23	4	4·54	7	6·99	7	5·58	9
8–9		3	4·40	2	3·86	4	6·22	4	4·43	7
9–10		2	3·69	2	} 6·09	0	5·52	4	3·52	7
10–11	} 3·94	0	} 5·63	2		5	4·88	3	} 5·08	3
11–12		0		1	} 4·45	0	4·32	5		1
12–13		4	} 3·76	3		0	3·82	3		0
13–14	} 5·36	1		1		5	} 6·37	2	} 4·51	1
14–15		0		1	} 4·61	1		5		0
15–16		0	} 5·95	3		0	} 5·03	1		1
16–17		1		1		0		0	} 6·18	1
17–18		0		1		3	} 4·00	3		1
18–19		0		1	} 4·37	1		2		2
19–20		1		1		0		2		0
20–21		1				1	} 4·55	0		0
21–22		0				0		1		0
22–23		0			} 4·94	1		2		0
23–24		0				0	} 4·23	1		0
24–25		0				0		1		0
25–26		0				1		1		0
26–32		2				4	3·94	3		1
32–43						1	3·58	3		1
>43							3·61	6		
Total	100·00	100	100·00	100	100·00	100	123·00	123	120·99	121
χ^2	9·63		12·67		18·75		18·09		13·35	
Degrees of Freedom	7		10		11		18		10	
$P(\chi^2)$	·21		·24		·066		·45		·21	

The symbols } are used to indicate the groupings used in the calculation of the χ^2. The groupings were made so as to have the expected frequency in a class at least equal to 3.5.

tively. These papers also contain various compilations of the results of the measurements. However, the original single measurements, necessary for the present paper, are naturally unpublished and Professor Birge was kind enough to find them for the author in his records.

Frequency distribution No. 3 was compiled from a book of records of sugar beet trials carried out by Messrs. K. Buszczyński and Sons, Ltd. in Górka Narodowa, Poland.

The 4th distribution was constructed from the original measurements of the velocity of light as published [27] by Michelson, Pease, and Pearson. The measurements made during single days were treated as forming separate groups.

Distribution No. 5 originated from repeated measurements of Octane Rating conducted by a refining company in California. They were made accessible by Mr. Walter S. Svenson and it is a pleasure to express the author's deep gratitude to him.

The number of observations in each column is not very large. It may be expected that if it were increased, the differences between the hypothetical distributions and the observed ones would become more apparent. It seems safe, however, to assume that in a number of instances the hypothesis as to the character of the variability of ω_i is not in very bad disagreement with the actual facts. It would be most interesting to have some more data on the subject.

6. **The best critical region for testing H against a particular alternative.** It seems unquestionable that the most desirable test of any hypothesis is the uniformly most powerful test (U. M. P. Test) with respect to the whole class of simple hypotheses alternative to the one which is being tested. Denote by H the hypothesis tested, by h any simple admissible hypothesis alternative to H, and by Ω the set of all h's. If w_0 is the critical region corresponding to the U. M. P. Test, then w_0 has these properties:

(71) (1) $$P\{E \epsilon w_0 \mid H\} \equiv \epsilon.$$

(2) If w is any other region such that $P\{E \epsilon w \mid H\} \equiv \epsilon$ then

(72) $$P\{E \epsilon w_0 \mid h\} \geq P\{E \epsilon w \mid h\},$$

whatever be $h \epsilon \Omega$.

Following the known method [18], we shall see whether a test of the hypothesis H considered in the preceding sections exists which is a U. M. P. Test with respect to the whole class of admissible hypotheses that specify the probability laws (68) with any $h_0 > 0$ and $\nu > 0$.

The method consists of considering one particular alternative hypothesis h', that is, one particular set of values of $h_0 > 0$ and $\nu > 0$ and finding the best critical region $w_{h_0,\nu}$ for testing H against h'. If this region appears to depend on ν and/or on h_0 then there is no U. M. P. Test. The region $w_{h_0,\nu}$ is found by determining, for each system (φ) of $T_1, T_2, \cdots, T_{N+1}$ separately, a part $w_{h_0,\nu}(\varphi)$ determined by the inequality

(73) $$p(E \mid h_0, \nu) \geq k(\varphi)p(E \mid H)$$

where $k(\varphi)$ is a function of $T_1, T_2, \cdots, T_{N+1}$ so determined that the relation (60) is satisfied. Substituting (6) and (68) in (73), taking the logarithm of both sides, and combining all terms which are constant or depend only on $T_1, T_2, \cdots, T_{N+1}$, we have

$$(74) \quad \sum_{i=1}^{N} \log\left(1 + \tfrac{1}{2}h_0 \nu n(S_i^2 + (T_{i+1} - \xi_i)^2)\right) \leq k_1(T_1, \cdots, T_{N+1}), \quad \text{(say)}.$$

Clearly, for $T_1, T_2, \cdots, T_{N+1}$ fixed, this inequality imposes a restriction on the variability of $u_1, u_2, \cdots, u_{N+1}$ while $z_{1,1}, \cdots, z_{N,n-2}$ are allowed to vary indiscriminately within the extreme limits (52). But the region $w_{h_0,\nu}(\varphi)$ determined by (74) also depends on the product $h_0\nu$. Therefore, there is no uniformly most powerful test for testing H against any and all simple alternatives specifying (68).

7. **A critical region of an unbiased type.** There seems to be no grounds for dissention that when a U. M. P. Test exists and is readily applicable, it is preferable to any other test, but the situation is quite different when there is no U. M. P. Test. In such cases, practical considerations may suggest a variety of requirements for a second best test of the hypothesis. Among these, we may suggest the following considerations:

Fix, for a moment, the values of $h_0, \xi_1, \cdots, \xi_N$, take any region w of the family $F(\epsilon)$, and consider the probability of E falling in w as a function of ν only. This is called the power function

$$(75) \qquad \beta(\nu \mid w) = \int \cdots \int_w p(E \mid h_0, \nu)\, dx_{1,1} \cdots dx_{N,n}$$

Here, of course, $\nu \geq 0$. Because of the properties of regions belonging to $F(\epsilon)$ we have $\beta(0 \mid w) \equiv \epsilon$. If $\nu > 0$, the value of $\beta(\nu \mid w)$ represents the corresponding probability of the test (based on w) discovering the falsehood of H. It is obviously desirable to have this probability as large as possible. In any case, it should be greater than ϵ. This last restriction is known as that of unbiasedness [19], [20], [28]. Further, since it is impossible to maximize $\beta(\nu \mid w)$ for *all values* of ν, we must choose those for which it is most desirable, in our opinion, to concentrate our efforts to increase $\beta(\nu \mid w)$. One possible point of view is that these values should be very close to the hypothetical value $\nu = 0$. For if ν is considerably larger than zero, we may argue that there will be no need to apply any refined statistical test to detect the falsehood of H. Of course, this argument has no mathematical character and its general acceptance is not suggested. In fact, we may argue that if ν is greater than zero but very small, it will be almost impossible to detect the falsehood of H by any test and, therefore, our efforts should be concentrated on values of ν which are of considerable size.

These are considerations of non-mathematical character; the role of mathematical statistics is limited to devising tests and elucidating their properties. If these last are understood by practical statisticians, each may choose according to his problem. Note that what could be termed the "properties" of a test are summarized in the power function $\beta(\nu \mid w)$ with its relation to the power functions of other possible tests of the same hypothesis.

In this paper we shall deal with tests particularly sensitive to small deviations of ν from its hypothetical value $\nu = 0$. In this respect, our first trial is to find a region w_0, belonging to the family $F(\epsilon)$ and satisfying the condition

$$\left.\frac{\partial\beta(\nu \mid w_0)}{\partial\nu}\right]_{\nu=0} \geq \left.\frac{\partial\beta(\nu \mid w)}{\partial\nu}\right]_{\nu=0}, \tag{76}$$

where w is any other region belonging to the same family $F(\epsilon)$.

Because of the peculiar structure of the regions belonging to $F(\epsilon)$, the problem is immediately reduced to finding regions $w_0(\varphi)$. According to theory explained elsewhere [18] these should satisfy the condition

$$\left.\frac{\partial p(E \mid h_0, \nu)}{\partial\nu}\right]_{\nu=0} \geq k(T)p(E \mid H), \tag{77}$$

where $k(T)$ depends on $T_1, T_2, \cdots, T_{N+1}$ only and is determined to satisfy the condition of similarity (60). Condition (77) is equivalent to

$$\left.\frac{\partial \log p(E \mid h_0, \nu)}{\partial\nu}\right]_{\nu=0} \geq k(T). \tag{78}$$

Taking the logarithm of (68), differentiating with respect to ν, putting ν equal to zero, substituting in (78), and combining all the terms which are constant on $W(\varphi)$ into a single term which we may write as $\frac{1}{8}h_0^2k_1(T)$, we have

$$\sum_{i=1}^{N} (S_i^2 + (T_{i+1} - \xi_i)^2)^2 \geq k_1(T). \tag{79}$$

We note that condition (79) determining, so to speak, the shape of the region $w_0(\varphi)$ does not imply any restriction on the variability of the z's but only on the u's. However, the region $w_0(\varphi)$ as determined by (79) has the disadvantage of being dependent on the values of the ξ_i. Since these are not specified by the hypothesis tested, we are not able to determine the critical regions belonging to the family $F(\epsilon)$ and maximizing the derivative $\partial\beta(\nu \mid w)/\partial\nu]_{\nu=0}$. The region which does so for some particular system $\xi_1', \xi_2', \cdots, \xi_N'$ of values of the ξ's will lose this property if the system of values of the ξ's is appropriately changed. Therefore, our choice of the region maximizing the derivative of the power function at $\nu = 0$ should be made not from the whole family $F(\epsilon)$ but from a subfamily $F_1(\epsilon)$ composed only of such regions which also possess the supplementary property that

$$\left.\frac{\partial\beta(\nu \mid w)}{\partial\nu}\right]_{\nu=0} = \text{constant} \tag{80}$$

has a value independent of $\xi_1, \xi_2, \cdots, \xi_N$. The determination of this subfamily $F_1(\epsilon)$ embracing all such regions is an interesting problem. Until it is solved, we use an obvious subfamily $F_2(\epsilon)$ of regions w which have the desired property, but we do not know whether or not $F_2(\epsilon)$ contains all such regions.[2]

The family $F_2(\epsilon)$ is defined as consisting of those regions belonging to $F(\epsilon)$ which could be described as cylindrical with their generators parallel to the intersection of $T_{i+1} = x_{i.} =$ constant, for $i = 1, 2, \cdots, N$. In other words and more precisely, a region w of the family $F(\epsilon)$ belongs to $F_2(\epsilon)$ only if the question of its including a given point E depends on $Nn - N$ of its coordinates, namely on $T_1, u_1, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2}$ and not on $T_2, T_3, \cdots, T_{N+1}$.

We easily show that any region w belonging to $F_2(\epsilon)$ possesses the property that its power function is independent of the ξ_i's. Denote by w' the set of systems of values of $T_1, u_1, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2}$ corresponding to points included in any given region w of the family $F_2(\epsilon)$. We see that the power function $\beta(\nu \mid w)$, equal to the integral of (68) over w, can be calculated by using the transformations (47), (49), and (51). Then the region of integration for $T_1, u_1, \cdots, u_{N-1}, z_{1,1}, \cdots, z_{N,n-2}$ is what we have just denoted by w' and the integrations for $T_{i+1} = x_{i.}$ extend from $-\infty$ to $+\infty$ irrespective of the fixed values of the other variables. These integrations are easily carried out by substituting

$$\tfrac{1}{2}nh_0\nu(x_{i.} - \xi_i)^2 = (1 + \tfrac{1}{2}nh_0\nu S_i^2)t_i^2 . \tag{81}$$

The final result is

$$\beta(\nu \mid w) = \int \cdots \int_{w'} p(T_1, u_1, \cdots u_{N-1}, z_{1,1}, \cdots z_{N,n-2})\, dT_1 \cdots dz_{N,n-2} \tag{82}$$

Here

$$\begin{aligned} p(T_1, u_1, &\cdots u_{N-1}, z_{1,1}, \cdots z_{N,n-2}) \\ &= c(\nu)\Phi(T_1, u, z) \Big/ \prod_{i=1}^{N} (1 + \tfrac{1}{2}nh_0\nu S_i^2)^{\frac{1}{2}(n-1)+1/\nu}, \end{aligned} \tag{83}$$

where $c(\nu)$ denotes a constant depending on ν, $\Phi(T_1, u, z)$ denotes a function of all the $N(n-1)$ variables involved, independent of ν, and S_i^2 denotes expressions (47) for short. We see that (82) is independent of the ξ_i's.

Since the region w belongs to $F(\epsilon)$, it is composed of sections $w(\varphi)$ selected separately on each hypersurface $T_1 =$ constant and $T_{i+1} =$ constant, $i = 1, 2, \cdots, N$. Because of the definition of the family $F_2(\epsilon)$, the sections $w(\varphi)$ are independent of $T_2, T_3, \cdots, T_{N+1}$ so that each of them can be selected only in accordance with the value of T_1. Therefore, we may denote them by $w(T_1)$. As far as property (80) is concerned, the choice is arbitrary. But the property of similarity requires the fulfillment of condition (60) which, in the present case, reduces to

$$\int \cdots \int_{w(T_1)} p(u_1, \cdots u_{N-1}, z_{1,1}, \cdots z_{N,n-2} \mid T_1, \cdots T_{N+1})\, du_1 \cdots dz_{N,n-2} = \epsilon \tag{84}$$

[2] Regions with the property (80) and belonging to $F(\epsilon)$ but not to $F_2(\epsilon)$ exist. Probably however, each of them differs from one of the regions of $F_2(\epsilon)$ by a set of measure zero only.

Applying the method already used, we find that sections $\bar{w}(T_1)$ of the region $\bar{w}$ belonging to $F_2(\epsilon)$ and maximizing the derivative $\partial\beta(\nu \mid w)/\partial\nu]_{\nu=0}$ are determined, separately for each value of T_1, by the inequality

$$\left.\frac{\partial \log p(T_1, u_1, \cdots u_{N-1}, z_{1,1}, \cdots z_{N,n-1})}{\partial \nu}\right]_{\nu=0} \geq k_2(T_1) \tag{85}$$

where $k_2(T_1)$ denotes a function of T_1 determined to satisfy (84).

Substituting (83) in (85) we easily find that this condition is equivalent to

$$\zeta \equiv \sum_{i=1}^{N-1} u_i^2 + \left(1 - \sum_{i=1}^{N-1} u_i\right)^2 \geq k_3(T_1) \tag{86}$$

where, again, $k_3(T_1)$ is determined for each particular value of T_1 to satisfy (84). As (86) does not imply any restrictions on the variability of $z_{1,1}, z_{1,2}, \cdots, z_{N,n-2}$, the integrations for the z's while calculating (84) must be carried out over the extreme limits (52). This will reduce the integrand to the relative probability law of $u_1, u_2, \cdots, u_{N-1}$ given all the T's. This law is easily calculated from (58) and is

$$\begin{aligned} p(u_1, u_2, \cdots u_{N-1} \mid T_1, T_2, \cdots T_{N+1}) \\ &= \frac{\Gamma(\frac{1}{2}N(n-1))}{\Gamma^N(\frac{1}{2}(n-1))} \left(\left(1 - \sum_{i=1}^{N-1} u_i\right) \prod_{i=1}^{N-1} u_i\right)^{\frac{1}{2}(n-3)} \\ &= p(u_1, u_2, \cdots u_{N-1}) \end{aligned} \tag{87}$$

As (87) is independent of $T_1, T_2, \cdots, T_{N+1}$, it is also the absolute probability law of the u's and hence $k_3(T_1)$ is independent of T_1. In accordance with the notation adopted for the left side of (86), namely ζ, and since the choice of $k_3(T_1)$ depends on ϵ, n, and N, we may use ζ_ϵ instead of $k_3(T_1)$. Then the region $\bar{w}$ is determined by the inequality

$$\zeta = \sum_{i=1}^{N-1} u_i^2 + \left(1 - \sum_{i=1}^{N-1} u_i\right)^2 \geq \zeta_\epsilon \tag{88}$$

or, returning to the original variables, by the inequality

$$\zeta \equiv \sum_{i=1}^{N} S_i^4 \Big/ \left(\sum_{i=1}^{N} S_i^2\right)^2 \geq \zeta_\epsilon \tag{89}$$

where ζ_ϵ is the root of the equation

$$\frac{\Gamma(\frac{1}{2}N(n-1))}{\Gamma^N(\frac{1}{2}(n-1))} \int \cdots \int_{\zeta \geq \zeta_\epsilon} \left(\left(1 - \sum_{i=1}^{N-1} u_i\right) \prod_{i=1}^{N-1} u_i\right)^{\frac{1}{2}(n-3)} du_1 \cdots du_{N-1} = \epsilon \tag{90}$$

This region $\bar{w}$ has the following property: of all the regions belonging to the family $F_2(\epsilon)$, the derivative of the power function of $\bar{w}$ at the point $\nu = 0$ is the greatest. Thus, as far as the values of ν close to zero are concerned, we may say that, for testing H, $\bar{w}$ is the most powerful critical region in the family $F_2(\epsilon)$.

8. **Methods of determining ζ_ϵ.** To calculate ζ_ϵ accurately we must calculate the integral probability law of ζ, that is to say,

$$P\{\zeta < z\} = \int_{\zeta<z} \cdots \int p(u_1, \cdots u_{N-1})\, du_1 \cdots du_{N-1} \tag{91}$$

for any z. The author was not able to achieve this. Therefore some methods of approximation had to be looked for. This task becomes somewhat simplified by noting that in most practical problems N will be very large, in the hundreds or thousands, while n will probably not exceed 5.

To start, we notice that the range of ζ is limited by

$$1/N \leq \zeta \leq 1. \tag{92}$$

The easiest way to see this is to look for maxima and minima of the sum

$$X = \sum_{i=1}^{N} S_i^4 \tag{93}$$

subject to the restriction that

$$\sum_{i=1}^{N} S_i^2 = T_1 \tag{94}$$

We then easily find that

$$T_1^2/N \leq X \leq T_1^2 \tag{95}$$

and (92) follows directly.

Since ζ is a polynomial of the second order in the u's, we may consider its moments. These will be functions of the expectations of the products $\prod_{i=1}^{N} u_i^{k_i}$ where, for short, $u_N = 1 - \sum_{i=1}^{N-1} u_i$. Using (87) we easily find that

$$E\left(\prod_{i=1}^{N} u_i^{k_i}\right) = \frac{\Gamma(\frac{1}{2}N(n-1))}{\Gamma\left(\frac{1}{2}N(n-1) + \sum_{i=1}^{N} k_i\right)} \prod_{i=1}^{N} \frac{\Gamma(\frac{1}{2}(n-1) + k_i)}{\Gamma(\frac{1}{2}(n-1))}. \tag{96}$$

In particular, if we let $(n-1)/2 = a$

$$E(u_i^2) = \frac{a(a+1)}{Na(Na+1)} \tag{97}$$

$$E(u_i^4) = \frac{a(a+1)(a+2)(a+3)}{Na(Na+1)(Na+2)(Na+3)} \tag{98}$$

$$E(u_i^2 u_j^2) = \frac{a^2(a+1)^2}{Na(Na+1)(Na+2)(Na+3)}. \tag{99}$$

Consequently and because $\zeta = \sum_{i=1}^{N} u_i^2$, we have

$$E(\zeta) = \mu_1' = (a + 1)/(Na + 1) \tag{100}$$

$$\begin{aligned} E(\zeta^2) = \mu_2' &= \sum_{i=1}^{N} E(u_i^4) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} E(u_i^2 u_j^2) \\ &= \frac{(a+1)(a+2)(a+3)}{(Na+1)(Na+2)(Na+3)} + \frac{(N-1)a(a+1)^2}{(Na+1)(Na+2)(Na+3)}. \end{aligned} \tag{101}$$

The variance σ_ζ^2 of ζ is therefore

$$\sigma_\zeta^2 = \frac{2a(a+1)(N-1)}{(Na+1)^2(Na+2)(Na+3)}. \tag{102}$$

By a similar procedure we find that

$$E(\zeta^3) = \mu_3' = \frac{\begin{aligned}(a+1)(a+2)(a+3)(a+4)(a+5) \\ + 3(N-1)a(a+1)^2(a+2)(a+3) \\ + (N-1)(N-2)a^2(a+1)^3\end{aligned}}{(Na+1)(Na+2)(Na+3)(Na+4)(Na+5)} \tag{103}$$

$$E(\zeta^4) = \mu_4' = \frac{\begin{aligned}\prod_{j=1}^{7} (a+j) + 4(N-1)a(a+1) \prod_{j=1}^{5} (a+j) \\ + 3(N-1)a \prod_{j=1}^{3} (a+j)^2 \\ + 6(N-1)(N-2)a^2(a+1)^3(a+2)(a+3) \\ + (N-1)(N-2)(N-3)a^3(a+1)^4\end{aligned}}{\prod_{j=1}^{7} (Na+j)} \tag{104}$$

One possible method of approximating ζ_ϵ is to use the formulae above, together with the higher moments whose formulae are easy to deduce. Some convenient known distribution, say $p_0(\zeta)$, could be fitted to have its first two or three moments coincide with those of the unknown true distribution of ζ. We would then look for better approximations by means of the functions

$$p_m(\zeta) = p_0(\zeta) \sum_{j=1}^{m} A_j \pi_j \tag{105}$$

where the π_j's denote polynomials which are orthogonal and normal with respect to $p_0(\zeta)$ so that

$$\int \pi_j \pi_k p_0(\zeta)\, d\zeta = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k. \end{cases} \tag{106}$$

The constant coefficients A_j are formed to minimize the integral

$$\int \left(p(\zeta) - p_0(\zeta) \sum_{j=1}^{m} A_j \pi_j \right)^2 p_0^{-1}(\zeta)\, d\zeta. \tag{107}$$

They are expressible in terms of the known moments of $p(\zeta)$.

This is one possible way to approximate $p(\zeta)$ which would eventually lead to the computation of ζ_ϵ even for small values of N.

Remembering that we are concerned with large N's, we can prove that the normalized distribution of ζ, that is, the distribution of

$$\frac{\zeta - \mu_1'}{\sigma_\zeta} \tag{108}$$

tends to be normal as $N \rightarrow \infty$. However, the process of tending to the limit is rather slow as may be seen from the following table of K. Pearson's β_1 and β_2.

TABLE II

Frequency constants of the distribution of ζ

n	N	μ_1'	σ_ζ	β_1	β_2
3	100	.0198	.001922	.8652	5.042
3	200	.0099	.000693	.4618	4.244
3	400	.0050	.000248	.2410	3.587

Because of this and also because the proof that the distribution of (108) tends to normality is not very straightforward, we shall not reproduce it. But it may be well to point out that the cause of this slowness in tending to the limit lies in the skewness of the distribution of each particular u_i and in the mutual dependency of all the u_i's.

The most promising method seems to be the following. First consider the two sums

$$T_1 = \sum_{i=1}^{N} S_i^2 \quad \text{and} \quad T_0 = \sum_{i=1}^{N} S_i^4. \tag{109}$$

Obviously, these two sums satisfy the conditions of the limiting theorem of S. Bernstein [29], [30] and, therefore, as $N \rightarrow \infty$, their joint normalized distribution tends to a normal surface. Also, we may expect the process of tending to the limit to be rapid in this case. If $p(T_0, T_1)$ denotes the limiting normal distribution, the probability that $\zeta \geq z$ can be approximately calculated by the integral

$$P\{\zeta \geq z\} = P\{T_0 \geq zT_1^2\} = \int_{-\infty}^{+\infty} dT_1 \int_{zT_1^2}^{\infty} p(T_0, T_1)\, dT_0. \tag{110}$$

To calculate the limiting distribution $p(T_0, T_1)$ we need only the expectations, say A and B, of T_1 and T_0 respectively, their standard errors, say σ_1 and σ_2, and their correlation coefficient R. These may be obtained from the moments of the S_i^2's.

Formula (110) can be used not only for tabulating the integral probability law of ζ and for determining ζ_ϵ, but also for an approximate calculation of the power function of the test. For, if the limiting probability law $p(T_0, T_1)$ is calculated using the moments of S_i^2 calculated from (70) with some $\nu > 0$, then the integral (110) calculated with $z = \zeta_\epsilon$ gives us the probability $P\{\zeta \geq \zeta_\epsilon \mid \nu\}$ of the test detecting the falsehood of the hypothesis tested, that is, the power function.

To save space, we shall now calculate the constants A, B, σ_1, σ_2, and R as functions of $\nu \geq 0$. The values appropriate to the case when the hypothesis tested is true will then be obtained from the general formulae by the mere substitution of $\nu = 0$.

Since all the constants above depend on the expectations of S_i^{2k}, we use formula (70) to calculate them. Denoting the expectation of S_i^{2k} by μ_k, we have

$$\mu_k = \frac{2(nh_0\nu/2)^{\frac{1}{2}(n-1)}}{\mathrm{B}(1/\nu, \frac{1}{2}(n-1))} \int_0^\infty \frac{S^{2k+n-2}}{(1 + \frac{1}{2}nh_0\nu S^2)^{\frac{1}{2}(n-1)+1/\nu}} dS. \tag{111}$$

Introducing the new variable

$$1 + \tfrac{1}{2}nh_0\nu S^2 = t^{-1} \tag{112}$$

makes the integration straightforward and gives

$$\mu_k = \left(\frac{2}{nh_0\nu}\right)^k \frac{\Gamma((1/\nu) - k)\Gamma(\frac{1}{2}(n-1) + k)}{\Gamma(1/\nu)\Gamma(\frac{1}{2}(n-1))}. \tag{113}$$

This formula holds good if $1/\nu > k$. Otherwise the kth moment μ_k is divergent. So this approximate method of calculating the power function of the test is applicable only for $\nu < .25$.

Substituting $k = 1, 2, 3, 4$ in (113), we have

$$\begin{aligned} \mu_1 &= \frac{1}{nh_0}\frac{n-1}{1-\nu} \\ \mu_2 &= \left(\frac{1}{nh_0}\right)^2 \frac{n^2-1}{(1-\nu)(1-2\nu)} \\ \mu_3 &= \left(\frac{1}{nh_0}\right)^3 \frac{(n^2-1)(n+3)}{(1-\nu)(1-2\nu)(1-3\nu)} \\ \mu_4 &= \left(\frac{1}{nh_0}\right)^4 \frac{(n^2-1)(n+3)(n+5)}{(1-\nu)(1-2\nu)(1-3\nu)(1-4\nu)}, \end{aligned} \tag{114}$$

and now we have

$$A = \frac{N}{nh_0}\frac{n-1}{1-\nu}, \qquad B = \frac{N}{(nh_0)^2}\frac{n^2-1}{(1-\nu)(1-2\nu)}, \tag{115}$$

$$\sigma_1^2 = \frac{N}{(nh_0)^2}\frac{(n-1)(2+\nu(n-3))}{(1-\nu)^2(1-2\nu)}, \tag{116}$$

$$\sigma_2^2 = \frac{2N}{(nh_0)^4}\frac{(n^2-1)(2+\nu(n-3))(2(n+2)-\nu(5n+7))}{(1-\nu)^2(1-2\nu)^2(1-3\nu)(1-4\nu)}, \tag{117}$$

$$R^2 = \frac{2(n+1)(1-2\nu)(1-4\nu)}{(2(n+2)-\nu(5n+7))(1-3\nu)}. \tag{118}$$

Inspecting formulae (115) to (118) makes us see that there is an advantage in substituting two new variables

$$t_1 = \frac{nh_0}{N(n-1)} T_1, \qquad t_2 = \frac{(nh_0)^2}{N(n^2-1)} T_0, \tag{119}$$

for T_1 and T_0. Their expectations, say ϑ_1 and ϑ_2, are

$$\vartheta_1 = \frac{1}{1-\nu}, \qquad \vartheta_2 = \frac{1}{(1-\nu)(1-2\nu)}. \tag{120}$$

Probably without any danger of confusion, the S.E.'s of t_1 and t_2 may be denoted by σ_1 and σ_2 also and we shall have

$$\begin{aligned} \sigma_1^2 &= \frac{2+\nu(n-3)}{N(n-1)(1-\nu)^2(1-2\nu)}, \\ \sigma_2^2 &= \frac{2(2+\nu(n-3))(2(n+2)-\nu(5n+7))}{N(n^2-1)(1-2\nu)^2(1-3\nu)(1-4\nu)}. \end{aligned} \tag{121}$$

Of course, the correlation coefficient of t_1 and t_2 is the same as that of T_1 and T_0, namely R. Obviously, the inequality $T_0 \geq zT_1^2$ is equivalent to $t_2 \geq z_1 t_1^2$ provided that

$$z = z_1 \frac{n+1}{N(n-1)}. \tag{122}$$

Now the problem of calculating (110) is reduced to finding

$$\begin{aligned} P\{\zeta \geq z\} &= P\{t_2 \geq z_1 t_1^2\} \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-R^2}} \iint_{t_2 \geq z_1 t_1^2} \exp\left[-\frac{1}{2(1-R^2)}\left\{\frac{(t_1-\vartheta_1)^2}{\sigma_1^2}\right.\right. \\ &\qquad \left.\left. - 2R\frac{(t_1-\vartheta_1)(t_2-\vartheta_2)}{\sigma_1\sigma_2} + \frac{(t_2-\vartheta_2)^2}{\sigma_2^2}\right\}\right] dt_1\,dt_2. \end{aligned} \tag{123}$$

We may conveniently see the workings of the test proposed by considering formula (123). First consider the case when the hypothesis tested is true. Both ϑ_1 and ϑ_2 reduce to unity. The region of highest frequency is around the point $t_1 = t_2 = 1$. If N is large then both σ_1 and σ_2 are small so that the region of significant frequency is rather small. The integral (123) is to be taken over the region above the parabola $t_2 = z_1 t_1^2$ passing through the origin of coordinates. When z_1 is small and the parabola passes far below the point $t_1 = t_2 = 1$, the probability $P\{\zeta \geq z\}$ will be close to unity. When $z_1 = 1$ this probability will be less than $\frac{1}{2}$ and it will diminish rapidly with further increases of z_1. Now suppose that we have found the value ζ_ϵ for which $P\{\zeta \geq \zeta_\epsilon \mid \nu = 0\} = \epsilon$ and consider what will happen to (123) when $z = \zeta_\epsilon$ if ν is increased. Clearly, neither of σ_1 and σ_2 nor R are very sensitive to slight changes in ν. Also θ_1 will not change very much. On the other hand, ϑ_2 will increase rather fast. The final conclusion is that the whole frequency surface corresponding to the integrand in (123) will not change shape much but will shift to bring a greater amount of frequency into the region of integration.

To facilitate numerical calculations introduce

$$(124) \qquad x = \frac{t_1 - \vartheta_1}{\sigma_1}, \qquad y = \frac{t_2 - \vartheta_2 - R\sigma_2(t_1 - \vartheta_1)/\sigma_1}{\sigma_2\sqrt{1 - R^2}}.$$

Now (123) may be rewritten as

$$(125) \qquad P\{\zeta \geqq z\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left\{ e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} \int_{y(x,z_1)}^{\infty} e^{-\frac{1}{2}y^2}\, dy \right\} dx$$

where

$$(126) \qquad y(x, z_1) = \frac{z_1(\vartheta_1 + \sigma_1 x)^2 - \vartheta_2 - R\sigma_2 x}{\sigma_2\sqrt{1 - R^2}}.$$

Using formulae (125), (126) and (119) to (122), the following numerical values were obtained.

TABLE III

$n = 3, \quad N = 100, \quad \nu = 0.$

z_1	$P\{\zeta \geq z \mid \nu = 0\}$
.8	.9126
.9	.7305
1.0	.4905
1.1	.2847
1.2	.1495
1.3	.0730
1.4	.0335
1.5	.0148
1.6	.00644
1.7	.00288
1.34450	.05000
1.54563	.01000

TABLE IV

Power of the test for $n = 3$ *and* $N = 100$.

ϵ	ζ_ϵ	$\nu = .01$	$\nu = .16$
.05	.02689	.05823	.37482
.01	.03091	.01234	.10699

The figures above are only approximate and we realize that the greater the value of ν the less satisfactory is the approximation of the power function. A check of the goodness of the approximation and, if it proves satisfactory, a few numerical tables for practical applications of the test must be postponed to another publication.

It is a pleasure to record the author's indebtedness to Miss Elizabeth Scott and also to Miss Julia Bowman for carrying out all the numerical work connected with the present paper.

REFERENCES

[1] J. Neyman, *Jour. Roy. Stat. Soc.*, Vol. 97 (1934), p. 558.
[2] J. Neyman, *Phil. Trans. Roy. Soc. London*, Vol. 236-A (1937), p. 333.
[3] "Student", *Biometrika*, Vol. 19, (1927), p. 151.
[4] J. Przyborowski, *Roczniki Nauk Rolniczych i Leśnych*, Vol. 30; (1933), p. 1.
[5] J. Neyman, *Lectures and conferences on mathematical statistics*, Washington, D. C. (1937).
[6] J. Berkson, *Jour. Am. Stat. Assn.*, Vol. 33, (1938), p. 526.
[7] J. Berkson, *Jour. Am. Stat. Assn.*, Vol. 35, (1940), p. 362.
[8] J. Neyman, *Annals of Math. Stat.*, Vol. 11 (1940), p. 478.
[9] J. Neyman and E. S. Pearson, *Bull. Int. Acad. Polon. Sci. Cracovie*, (A) Vol. 6 (1931), p. 460.
[10] B. L. Welch, *Biometrika*, Vol. 27 (1935), p. 145.
[11] B. L. Welch, *Stat. Res. Memoirs*, Vol. 1 (1936), p. 52.
[12] U. S. Nair, *Stat. Res. Memoirs*, Vol. 1 (1936), p. 38.
[13] M. S. Bartlett, *Proc. Roy. Soc.*, Vol. 160-A (1937), p. 268.
[14] S. S. Wilks and C. M. Thompson, *Biometrika*, Vol. 29 (1937), p. 124.
[15] D. T. Bishop and U. S. Nair, *Jour. Roy. Stat. Soc. Suppl.*, Vol. 6 (1939), p. 89.
[16] E. J. G. Pitman, *Biometrika*, Vol. 31, (1939), p. 200.
[17] H. O. Hartley, *Biometrika*, 31 (1940), p. 249.
[18] J. Neyman and E. S. Pearson, *Phil. Trans. Roy Soc. London*, 231-A (1933), p. 289.
[19] J. Neyman and E. S. Pearson, *Stat. Res. Memoirs*, Vol. 1 (1936), p. 1.
[20] J. Neyman, *Bull. Soc. Math. France*, Vol. 63 (1935), p. 246.
[21] H. Hotelling, *Ann. Math. Stat.*, Vol. 11 (1940), p. 271.
[22] W. Feller, *Stat. Res. Memoirs*, Vol. 2 (1938), p. 117.
[23] H. Cramér and H. Wold, *Jour. London Math. Soc.*, Vol. 11 (1936), p. 291.
[24] K. Pearson, *Tables of the Incomplete Beta Function*, Biometrika Office, University College, London, 1934.
[25] R. T. Birge, *Astrophysical Journal*, Vol. 39 (1914), p. 50.
[26] R. T. Birge, *Physical Review*, Vol. 40 (1932), p. 207.
[27] A. A. Michelson, F. G. Pease, and F. Pearson, *Astrophysical Journal*, Vol. 82 (1935), p. 26.
[28] J. Neyman and E. S. Pearson, *Stat. Res. Memoirs*, Vol. 2 (1938), p. 25.
[29] S. Bernstein, *Math. Ann.*, Vol. 97 (1926), p. 44.
[30] W. Kozakiewicz, *Ann. Soc. Polon. Math.*, Vol. 13 (1934), p. 24.