ON THE UNBIASED CHARACTER OF LIKELIHOOD-RATIO TESTS FOR INDEPENDENCE IN NORMAL SYSTEMS

By Joseph F. Daly

1. **Introduction.** In the statistical interpretation of experimental data, the basic assumption is, of course, that we are dealing with a sample from a statistical population, the elements of which are characterized by the values of a number of random variables $x^1, \cdots, x^k$. But in many cases we are in a position to assume even more, namely, that the population has an elementary probability law $f(x^1, \cdots, x^k; \theta_1, \cdots, \theta_h)$, where the functional form of $f(x, \theta)$ is definitely specified, although the parameters $\theta_1, \cdots, \theta_h$ are to be left free for the moment to have values corresponding to any point of a set Ω in an h-dimensional space.

Under this assumption, the problem of obtaining from the data further information about the hypothetical distribution law $f(x, \theta)$ is considerably simplified. For it is then equivalent to that of deciding whether or not the data support the hypothesis that the population values of the θ's correspond to a point in a certain subset ω of Ω. For example, we may have reason to believe that the population K has a distribution law of the form

$$f(x^1, x^2; a^1, a^2, A_{11}, A_{12}, A_{22}) = \frac{|A_{ij}|^{\frac{1}{2}}}{2\pi} e^{-\frac{1}{2}\sum_{i,j} A_{ij}(x^i - a^i)(x^j - a^j)}$$

Here the set Ω is composed of all parameter points $(a^1, \cdots, A_{22})$ for which the matrix $\| A_{ij} \|$ $(i, j = 1, 2)$ is positive definite and for which $-\infty < a^i < \infty$. We may wish to decide, on the basis of N independent observations (x^1_α, x^2_α) drawn from K, whether A_{12} has the value zero for the population in question, without concerning ourselves at all about the values of the remaining parameters; in other words, we may wish to test the hypothesis H that the parameter point corresponding to K lies in that subset of Ω for which $A_{12} = 0$. One way to test this hypothesis is to select some (measurable) function $g(x)$ whose value can be determined from the data, say

$$g(x) = \frac{\sum_{\alpha=1}^{N} (x^1_\alpha - \bar{x}^1)(x^2_\alpha - \bar{x}^2)}{\left[\sum_{\alpha=1}^{N} (x^1_\alpha - \bar{x}^1)^2\right]^{\frac{1}{2}} \left[\sum_{\alpha=1}^{N} (x^2_\alpha - \bar{x}^2)^2\right]^{\frac{1}{2}}}$$

Now $g(x)$ is itself a random variable, so that it has a distribution law of its own when its constituent x's are drawn from any particular population K. Suppose then we choose a set of values of $g(x)$, say S, such that the probability is only .05 that $g(x)$ will lie in the set S when the x's are drawn independently from a population K for which the above hypothesis H is true. Ordinarily we would take S to be of the form $| g(x) | \geq g_0$, and the test would then reject H at the .05 probability level if the computed value of $g(x)$ came out too large. But for all that has been said so far, we are perfectly free to choose a different critical region S, and even a different function $g(x)$. The essential elements of this type of test are then a critical region S, a function of the data g, and a probability level ϵ, such that the probability is $\epsilon = .05$, say, that $g \subset S$ when H is true; in employing the test we reject H at the given probability level whenever the sample value of g falls in the critical region.

By the very nature of the problem, any inferences we make from a sample are subject to possible error. In the kind of test under consideration, the only error we can commit, strictly speaking, is that of rejecting H when it is true (an error of Type I in the terminology of Neyman and Pearson [9]). The risk of such an error is thus known in advance; for if we use the test consistently at, say, the .05 level, we know that the probability is .05 that we shall be led to reject a given hypothesis when it is true. On the other hand, it is quite conceivable that the test may be even less likely to reject H when it is false, or more precisely, when the true θ's correspond to a point of Ω which is not in ω. In this event the test is said to be biased. Let us make this term more definite by proposing the following definitions:

DEFINITION I. *A test is said to be completely unbiased if it has the property that for any probability level* ϵ $(0 < \epsilon < 1)$ *the probability of rejecting H is greater when the θ's correspond to a point of* $\Omega - \omega$ *than when they correspond to a point of* ω.

DEFINITION II. *A test is said to be locally unbiased if the set* Ω *contains a neighborhood U of* ω *such that for any probability level* ϵ $(0 < \epsilon < 1)$ *the probability of rejecting H is greater when the parameter values correspond to a point of* $U - \omega$ *than when they correspond to a point of* ω.

It is the purpose of this paper to consider the question of bias in connection with the Neyman-Pearson method of likelihood ratios [8] as applied to the testing of what may well be called hypotheses of independence in multivariate normal populations. The likelihood ratio method is undoubtedly a very familiar one, since the vast majority of tests in present statistical practice are based on this method. But for the sake of completeness we shall outline it briefly. Let the distribution law of the population K be of the form $f(x^1, \cdots, x^k; \theta_1, \cdots, \theta_h)$ where the θ's may correspond to any point in a set Ω, and let the hypothesis H to be tested be that the θ's actually belong to the subset ω of Ω. Form the likelihood function

$$P_N(x; \theta) = \prod_{\alpha=1}^{N} f(x_\alpha^1, \cdots, x_\alpha^k; \theta_1, \cdots, \theta_h)$$

i.e., the elementary probability law of a sample of N elements drawn independently from K. Denote by $P_N^\Omega(x)$ the maximum of P_N for fixed x where the θ's are allowed to range over Ω; and denote by $P_N^\omega(x)$ the corresponding maximum value when the θ's are restricted to ω. The test criterion is then

$$\lambda = \frac{P^\omega(x)}{P^\Omega(x)}.$$

Evidently λ depends only on the observable quantities x_α^i, and has the range $0 \le \lambda \le 1$, with a definite probability law depending on that of the basic population K. In this method the critical region S is taken to be $0 \le \lambda \le \lambda_\epsilon$, where λ_ϵ is so chosen that the probability $P\{\lambda \le \lambda_\epsilon\}$ is ϵ when the parameters of K correspond to a point in ω. (It may be noted here that in all the cases with which we shall have to deal the probability that λ lies in S when H is true is independent of the particular values of the θ's as long as they correspond to a point of ω.) The reason for taking the critical region to be of the form $0 \le \lambda \le \lambda_\epsilon$ and not, say, $\lambda_\epsilon' \le \lambda \le \lambda_\epsilon''$ or $\bar{\lambda}_\epsilon \le \lambda \le 1$ may become clearer when we examine the resulting tests for bias.

The recent work of Neyman and Pearson [10] has led them to lay considerable stress on the importance of unbiased tests. And though their attention has been directed mainly to the broader outlines of the theory of testing hypotheses, they have stimulated other writers to study particular tests of great practical importance. P. C. Tang [11] has obtained the general sampling distribution of $1 - \lambda^{2/N}$ for what we shall call the regression problem with one dependent variate, and has given tables for $P\{\lambda \le \lambda_\epsilon\}$—essentially proving the unbiased character of the test—which should be extremely useful. His article also contains an excellent discussion of the manner in which this test is related to the well known tests of linear hypotheses [7] and to the ordinary analysis of variance. P. L. Hsu [6] has shown that this same distribution is fundamental in the study of Hotelling's generalized T test [5] (a special but important case of what we shall call the general regression problem), and has proved that (locally) this test is not only unbiased but "most powerful" in a certain sense. On the other hand, it is not true that all likelihood ratio tests are unbiased [2]. Consequently, the knowledge that in a rather wide class of problems which arise in normal sampling theory the method of likelihood ratios furnishes tests which are either locally or completely unbiased would seem to be of some value, even when the exact sampling distribution of the criterion is too complicated to tabulate.

2. **The regression problem with one dependent variate.** Suppose that y is known to be normally distributed about a linear function of the fixed variables $x^1, \cdots, x^r$, so that the family of populations under consideration is characterized by a distribution function of the form

$$f(y \mid x, b, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\left(y - \sum_{i=1}^{r} b_i x^i\right)^2}, \tag{2.1}$$

where the set of admissible values of σ^2 and the b's is

$$\Omega:\ 0 < \sigma^2 < \infty, \qquad -\infty < b_i < \infty.$$

Let H be the hypothesis that the point $(\sigma^2, b_1, \cdots, b_r)$ lies in the subset of Ω defined by

$$\omega:\ b_{q+1} = b_{q+2} = \cdots = b_r = 0.$$

The likelihood ratio appropriate to testing the hypothesis H on the basis of N $(N > r)$ independent observations drawn from such a population is then

$$\lambda = \left\{ \frac{\sum_{\alpha=1}^{N} \left(y_\alpha - \sum_{i=1}^{r} \hat{b}_i x_\alpha^i \right)^2}{\sum_{\alpha=1}^{N} \left(y_\alpha - \sum_{k=1}^{q} \hat{b}_k^0 x_\alpha^k \right)^2} \right\}^{\frac{1}{2}N},$$

with the understanding that the values of the fixed variables $x_\alpha^1, \cdots, x_\alpha^r$ associated with the α-th observation have been so chosen that the matrix $\| a^{ij} \| = \left\| \sum_{\alpha=1}^{N} x_\alpha^i x_\alpha^j \right\|$ is positive definite. (The expression in the numerator is the minimum of $\sum_{\alpha=1}^{N} \left(y_\alpha - \sum_{i=1}^{r} b_i x_\alpha^i \right)^2$ for variations of the b's over Ω, while the denominator contains the corresponding minimum for variations of the b's over ω).

In order to show that the test is unbiased, we shall make use of the exact sampling distribution of the quantity

$$\xi = 1 - \lambda^{2/N},$$

first published by P. C. Tang [11]. Writing $\| A_{gh} \|$ for the inverse of the matrix $\| a^{gh} \|$ composed of the first q rows and columns of $\| a^{ij} \|$, let us put

$$G = \frac{1}{2\sigma^2} \sum_{k,l=q+1}^{r} \left(a^{kl} - \sum_{g,h=1}^{q} a^k A_{gh} a^{hl} \right) b_k b_l .$$

Since the critical region $0 \leq \lambda \leq \lambda_\epsilon$ corresponds to the region $1 - \lambda_\epsilon^{2/N} = \xi_\epsilon \leq \xi \leq 1$, it can then be shown that the probability of rejecting H when the population parameters have specified values $\sigma^2, b^1, \cdots, b^r$ is expressed by the series

$$I(G, \xi_\epsilon) = e^{-G} \sum_{\nu=0}^{\infty} \frac{G^\nu}{\nu!} \int_{\xi_\epsilon}^{1} \frac{\xi^{\frac{1}{2}(r-q)+\nu-1}(1-\xi)^{\frac{1}{2}(N-r)-1}}{B[\frac{1}{2}(r-q)+\nu, \frac{1}{2}(N-r)]} d\xi, \tag{2.2}$$

where

$$B(u, v) = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)} = \int_0^1 z^{u-1}(1-z)^{v-1} dz.$$

Now G is a positive definite quadratic form in the parameters $b^{q+1}, \cdots, b^r$, so that it vanishes if and only if the hypothesis is true. And if $0 < \epsilon < 1$, then $I(G, \xi_\epsilon)$ is a monotone increasing function of G. For by differentiating (2.2) we obtain

$$\frac{\partial}{\partial G} I(G, \xi_\epsilon) = e^{-G} \sum_{\nu=0}^{\infty} \frac{G^\nu}{\nu!} \int_{\xi_\epsilon}^{1} \left\{ \frac{\xi^{\frac{1}{2}(r-q)+\nu}(1-\xi)^{\frac{1}{2}(N-r)-1}}{B[\frac{1}{2}(r-q)+\nu+1, \frac{1}{2}(N-r)]} - \frac{\xi^{\frac{1}{2}(r-q)+\nu-1}(1-\xi)^{\frac{1}{2}(N-r)-1}}{B[\frac{1}{2}(r-q)+\nu, \frac{1}{2}(N-r)]} \right\} d\xi. \tag{2.3}$$

And from a property of incomplete Beta functions, which we shall demonstrate in the next section, it follows that each term in the series (2.3) is positive. Accordingly we have

THEOREM I. *The likelihood ratio test for the hypothesis that in a population of type* (2.1) *certain of the regression coefficients are zero, i.e., the hypothesis that* y *is independent of the fixed variables* $x^{q+1}, \cdots, x^r$, *is completely unbiased.*

Wilks [15] has noted that the ordinary analysis of variance and covariance amounts essentially to testing hypotheses of this nature by means of the function

$$\zeta = \frac{1 - \lambda^{2/N}}{\lambda^{2/N}}.$$

Consequently such tests are also completely unbiased, since the region of rejection is then taken to be of the form $\zeta \geq \zeta_\epsilon$.

3. An inequality relating to incomplete Beta functions.

Let us write

$$\mathrm{B}(u, v; t) = \int_t^1 z^{u-1}(1 - z)^{v-1}\, dz \qquad (0 \leq t \leq 1).$$

Now,

$$\int_t^1 z^{u-1}(1 - z)^v\, dz = \frac{z^u(1 - z)^v}{u}\Big]_t^1 + \frac{v}{u}\int_t^1 z^u(1 - z)^{v-1}\, dz.$$

The integrated term on the right is non-positive, so that

$$\mathrm{B}(u, v + 1; t) \leq \frac{v}{u}\mathrm{B}(u + 1, v; t) \tag{3.1}$$

in which the equality holds if and only if $t = 0$ or $t = 1$. Again, since

$$z^u(1 - z)^{v-1} + z^{u-1}(1 - z)^v \equiv z^{u-1}(1 - z)^{v-1},$$

we have

$$\mathrm{B}(u + 1, v; t) + \mathrm{B}(u, v + 1; t) \equiv \mathrm{B}(u, v; t). \tag{3.2}$$

Combining these results, we find that

$$\frac{u + v}{u}\mathrm{B}(u + 1, v; t) \geq \mathrm{B}(u, v; t) \tag{3.3}$$

with equality only when $t = 0$ or $t = 1$. Hence we have

LEMMA 1: *If* $0 < t < 1$, *then*

$$\frac{\mathrm{B}(u + 1, v; t)}{\mathrm{B}(u + 1, v)} > \frac{\mathrm{B}(u, v; t)}{\mathrm{B}(u, v)}.$$

4. The multiple correlation coefficient.

Suppose the distribution law of the underlying population is known to be of the form

$$f(x^1, \cdots, x^t \mid x^{t+1}, \cdots, x^m) = \frac{|B_{ij}|^{\frac{1}{2}}}{\pi^{\frac{1}{2}t}} e^{-B_{ij}(x^i - a^i - C_p^i x^p)(x^j - a^j - C_q^j x^q)}. \tag{4.1}$$

The indices appearing in this expression take the values $i, j = 1, \cdots, t$ and $p, q = t + 1, \cdots, m$. The summation convention of repeated indices will be used, for example, $\sum_{p=t+1}^{m} C^i_p x^p$ will be denoted by $C^i_p x^p$. We shall also have occasion to use indices r, s with the range $r, s = 1, \cdots, m$. The set of possible values of the a's, B's, and C's is

$$\Omega: \ \| B_{ij} \| \text{ positive definite}; \ -\infty < a^i < \infty; -\infty < C^i_p < \infty.$$

We shall consider the λ test for the hypothesis H that x^1 is independent of the remaining variables $x^2, \cdots, x^m$, i.e., that the parameters belong to that subset of Ω defined by

$$\omega: B_{1k} = 0, \qquad (k = 2, \cdots, t); \qquad C^1_p = 0.$$

Let us write $v^{rs} = \sum_{\alpha=1}^{N} (x^r_\alpha - \bar{x}^r)(x^s_\alpha - \bar{x}^s)$, and assume that the values of the fixed variables x^p_α have been so selected that the matrix $\| v^{pq} \|$ is positive definite. The likelihood ratio can then be expressed in the form

$$\lambda = \left(\frac{| v^{rs} |}{v^{11} \cdot \bar{v}_{11}} \right)^{\frac{1}{2}N} = (1 - R^2)^{\frac{1}{2}N},$$

where $\bar{v}_{11}$ is the complement of v^{11} in the determinant $| v^{rs} |$. If $N \geq m + 1$, the general sampling distribution of R^2 (the multiple correlation coefficient between x^1 and $m - 1$ other variates), for this case in which $x^2, \cdots, x^t$ are subject to sampling variation and the remainder are fixed, is

$$(4.2) \quad \begin{aligned} F(R^2)\, d(R^2) &= \frac{(1 - \rho^2)^{\frac{1}{2}(N-1)} e^{-\frac{1}{2}y^2} (1 - R^2)^{\frac{1}{2}(N-m)-1} (R^2)^{\frac{1}{2}(m-1)-1}}{\Gamma[\frac{1}{2}(N - m)]} \\ &\times \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\infty} \frac{\frac{1}{2}(y^2)^\mu (1 - \rho^2)^\mu (\rho^2)^\nu (R^2)^{\mu+\nu} \Gamma^2[\frac{1}{2}(N - 1) + \mu + \nu]}{\mu! \nu! \Gamma[\frac{1}{2}(N - 1) + \mu] \Gamma[\frac{1}{2}(m - 1) + \mu + \nu]} d(R^2), \end{aligned}$$

where

$$1 - \rho^2 = \frac{| B_{ij} |}{B_{11} \bar{B}^{11}}, \qquad \tfrac{1}{2}y^2 = \frac{v^{pq}}{\bar{B}^{11}} C^1_p C^1_q, \qquad \| B^{ij} \| = \| B_{ij} \|^{-1}.$$

This distribution was first obtained by Wilks [13], although Fisher [3] had previously treated the two extreme cases in which (1) all independent variables are subject to sampling fluctuation, and (2) all independent variables are fixed.

To simplify the presentation, let us put $\bar{\rho} = \rho^2$, $\bar{y} = \frac{1}{2}y^2$ and $\bar{R} = R^2$, and note that $\bar{y} = 0$ if and only if $C^1_p = 0$ $(p = t + 1, \cdots, m)$ while $\bar{\rho} = 0$ if and only if $B_{1k} = 0$ $(k = 2, \cdots, t)$, so that $\bar{y} = \bar{\rho} = 0$ means that the hypothesis H is true. On any alternative hypothesis, one or the other or both of these quantities will be positive. Let the region of rejection be taken to be

$$R_\epsilon \leq \bar{R} \leq 1,$$

which corresponds to

$$0 \leq \lambda \leq (1 - \bar{R}_\epsilon)^{\frac{1}{2}N}.$$

The probability of rejecting H is then

$$I(\bar{\rho}, \bar{y}, \bar{R}_\epsilon) = e^{-\bar{y}} \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\infty} \frac{\bar{y}^\mu}{\mu!} (1 - \bar{\rho})^{\frac{1}{2}(N-1)+\mu} \frac{\bar{\rho}^\nu}{\nu!} \frac{\Gamma[\frac{1}{2}(N-1)+\mu+\nu]}{\Gamma[\frac{1}{2}(N-1)+\mu]} \times \int_{\bar{R}_\epsilon}^{1} \frac{\bar{R}^{\frac{1}{2}(m-1)+\mu+\nu-1}(1-\bar{R})^{\frac{1}{2}(N-m)-1}}{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m)]} d\bar{R}. \tag{4.4}$$

We shall show that $I(\bar{\rho}, \bar{y}, \bar{R}_\epsilon)$ is a strictly monotone increasing function of $\bar{\rho}$ for each $\bar{y}$, and that $I(0, \bar{y}, R_\epsilon)$ is a strictly monotone increasing function of $\bar{y}$. First consider $\frac{\partial I}{\partial \bar{\rho}}$. We can write (4.4) in the form

$$I(\bar{\rho}, \bar{y}, R_\epsilon) = e^{-\bar{y}} \sum_{\mu=0}^{\infty} \frac{\bar{y}^\mu}{\mu!} \frac{1}{\Gamma[\frac{1}{2}(N-1)+\mu]} \cdot \sum_{\nu=0}^{\infty} \frac{\bar{\rho}^\nu}{\nu!} (1-\bar{\rho})^{\frac{1}{2}(N-1)+\mu} \varphi_{\mu,\nu},$$

where

$$\varphi_{\mu,\nu} = \Gamma[\tfrac{1}{2}(N-1)+\mu+\nu] \frac{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m); \bar{R}_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m)]}.$$

Then, formally,

$$\frac{\partial}{\partial \bar{\rho}} \left(\sum_{\nu=0}^{\infty} \frac{\bar{\rho}^\nu}{\nu!} (1-\bar{\rho})^{\frac{1}{2}(N-1)+\mu} \varphi_{\mu,\nu} \right)$$
$$= \sum_{\nu=0}^{\infty} \frac{\nu \bar{\rho}^{\nu-1}}{\nu!} (1-\bar{\rho})^{\frac{1}{2}(N-1)+\mu} - \sum_{\nu=0}^{\infty} \frac{\bar{\rho}^\nu}{\nu!} (1-\bar{\rho})^{\frac{1}{2}(N-1)+\mu-1} [\tfrac{1}{2}(N-1)+\mu] \varphi_{\mu,\nu}.$$

Taking out the factor $(1-\bar{\rho})^{\frac{1}{2}(N-1)+\mu-1}$, we have left

$$\sum_{\nu=0}^{\infty} \frac{\nu \rho^{\nu-1}}{\nu!} \varphi_{\mu,\nu} - \sum_{\nu=0}^{\infty} \frac{\nu \rho^\nu}{\nu!} \varphi_{\mu,\nu} - \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu!} [\tfrac{1}{2}(N-1)+\mu] \varphi_{\mu,\nu}$$
$$= \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu!} \{\varphi_{\mu,\nu+1} - [\tfrac{1}{2}(N-1)+\mu+\nu]\varphi_{\mu,\nu}\}.$$

And the expression $\varphi_{\mu,\nu+1} - [\frac{1}{2}(N-1)+\mu+\nu]\varphi_{\mu,\nu}$ is the same as

$$\Gamma[\tfrac{1}{2}(N-1)+\mu+\nu+1] \left\{ \frac{B[\frac{1}{2}(m-1)+\mu+\nu+1, \frac{1}{2}(N-m), \bar{R}_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+\nu+1, \frac{1}{2}(N-m)]} - \frac{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m), \bar{R}_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m)]} \right\}$$

and is therefore positive, by Lemma 1. Consequently

$$\frac{\partial}{\partial \bar{\rho}} I(\bar{\rho}, \bar{y}, \bar{R}_\epsilon) \geq 0,$$

with equality holding only if $\bar{\rho} = 1$, or if the critical region is taken as the whole interval or the null set.

We have yet to investigate $\frac{\partial}{\partial \bar{y}} I(0, \bar{y}, \bar{R}_\epsilon)$. In this case (4.4) becomes

$$I(0, \bar{y}, \bar{R}_\epsilon) = e^{-\bar{y}} \sum_{\mu=0}^{\infty} \frac{\bar{y}^\mu}{\mu!} \frac{\mathrm{B}[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m), \bar{R}_\epsilon]}{\mathrm{B}[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m)]}. \tag{4.5}$$

(Note that this agrees with (2.2) if we make use of the relations $r = m, q = 1$, and $B^{11} = 2\sigma^2$.) We then obtain

$$\frac{\partial}{\partial \bar{y}} I(0, \bar{y}, \bar{R}_\epsilon) = e^{-\bar{y}} \sum_{\mu=0}^{\infty} \frac{y^\mu}{\mu!} \left\{ \frac{\mathrm{B}[\frac{1}{2}(m-1)+\mu+1, \frac{1}{2}(N-m); \bar{R}_\epsilon]}{\mathrm{B}[\frac{1}{2}(m-1)+\mu+1, \frac{1}{2}(N-m)]} - \frac{\mathrm{B}[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m); \bar{R}_\epsilon]}{\mathrm{B}[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m)]} \right\}$$

which the lemma shows to be positive when $0 < R_\epsilon < 1$.

This concludes the proof of

THEOREM II. *If the underlying population has a distribution law of the form (4.1), then the likelihood ratio test for the hypothesis that x^1 is independent of $x^2, \cdots, x^m$, where $x^{t+1}, \cdots, x^m$ are fixed and $x^2, \cdots, x^t$ are subject to sampling variation, is completely unbiased.*

5. Mutual independence of several sets of random variables.[1]

Let the distribution law of the m-variate population be of the form

$$\frac{|B_{ij}|^{\frac{1}{2}}}{\pi^{\frac{1}{2}m}} e^{-B_{ij}(x^i - a^i)(x^j - a^j)}. \tag{5.1}$$

Here Ω is the set $\| B_{ij} \|$ positive definite; $-\infty < a^i < \infty$. Suppose we wish to test the hypothesis H_I that the variates $\{x^1, \cdots, x^{m_1}\}, \cdots, \{x^{m_{p-1}+1}, \cdots, x^{m_p}\}$ are mutually independent in sets [14], where $0 = m_0 < m_1 < \cdots < m_p = m$. Then the ω set is that defined by

$$\| B_{ij} \| = \| B_{i_1 j_1} \| \dotplus \cdots \dotplus \| B_{i_p j_p} \| = \| B_1 \| \dotplus \cdots \dotplus \| B_p \|,$$

that is, we have $B_{ij} = 0$ unless the indices i and j both relate to the same set of variates.

Associated with the population of random samples O_N ($N \geq m + 1$) drawn from a universe characterized by (5.1), we have the distribution function

$$P(x; B, a) = \frac{|B_{ij}|^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x_\alpha^i - a^i)(x_\alpha^j - a^j)}$$

The maximum of P with respect to variations of the parameters B_{ij}, a^i in Ω is

$$P_\Omega = |v^{ij}|^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nm} e^{-\frac{1}{2}N},$$

[1] In this and in subsequent sections an index occurring both above and below indicates summation in accordance with the usual convention.

where

$$v^{ij} = \sum_{\alpha=1}^{N} (x_\alpha^i - \bar{x}^i)(x_\alpha^j - \bar{x}^j).$$

And the maximum when the parameters are restricted to ω is

$$P_\omega = [v_1 \cdots v_p]^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nm} e^{-\frac{1}{2}N},$$

where v_μ stands for the determinant of the v's connected with the μ-th set of x's. Thus the appropriate likelihood-ratio is given by

$$\lambda_I^{2/N} = \frac{|v^{ij}|}{v_1 \cdots v_p}.$$

It is easy to see that the value of λ_I is unaltered if we replace $x^i - a^i$ by x^i, so that we can express the probability that λ_I will lie between 0 and λ_ϵ in the form

$$I(B, \lambda_\epsilon) = \frac{B^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda<\lambda_\epsilon} e^{-\sum_{\alpha=1}^{N} B_{ij}x_\alpha^i x_\alpha^j} dx_1^1 \cdots dx_N^m.$$

Furthermore, λ_I is invariant under the operation of replacing any x by a linear combination of x's belonging to the same set. And since the assumption that $\| B_{ij} \|$ is positive definite implies that the matrices $\| B_{i_\mu j_\mu} \|$ have the same property, we can transform the x's in each set among themselves by orthogonal transformations in such a way as to reduce each of the expressions

$$B_{i_\mu j_\mu} x^{i_\mu} x^{j_\mu}$$

to sums of squares. Thus we have

$$(5.2) \qquad I(B, \lambda_\epsilon) = \frac{B^{*\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda<\lambda_\epsilon} e^{-\sum_{\alpha=1}^{N} B^*_{ij}x_\alpha^i x_\alpha^j} dx_1^1 \cdots dx_N^m = I(B^*, \lambda_\epsilon),$$

where

$$(5.3) \qquad B^*_{i_\mu j_\nu} = \alpha_{i_\mu}^{h_\mu} B_{h_\mu k_\nu} \alpha_{j_\nu}^{k_\nu} \qquad (h_\mu, i_\mu, j_\mu, k_\mu = m_{\mu-1} + 1, \cdots, m_\mu),$$

$$(5.4) \qquad B^*_{i_\sigma j_\sigma} = 0 \qquad i_\sigma \neq j_\sigma,$$

and the subscripts on the indices indicate the sets of values over which they range; e.g., i_2 runs over the numbers corresponding to the columns of the matrix $\| B_2 \|$. From (5.3) and (5.4) it is clear that $\| B^*_{ij} \|$ reduces to a diagonal matrix when H is true.

In order to show that the test is locally unbiased, we may consider the derivatives

$$\left(\frac{\partial}{\partial B^*_{i_\mu j_\nu}} I(B^*, \lambda_\epsilon)\right)_0, \qquad \left(\frac{\partial^2}{\partial B^*_{i_\mu j_\nu} \partial B^*_{h_\sigma k_\tau}} I(B^*, \lambda_\epsilon)\right)_0, \qquad (\mu \neq \nu, \sigma \neq \tau)$$

for the B^*'s are linear functions of the B's; and the positive definiteness of one matrix of second partials implies that of the other. We have at once

$$\left(\frac{\partial B^*}{\partial B^*_{i_\mu j_\nu}}\right)_0 = 0, \qquad \left(\frac{\partial^2 B^*}{\partial B^*_{i_\mu j_\nu} \partial B_{h_\sigma k_\tau}}\right)_0 = 0, \qquad (\mu \neq \nu, \sigma \neq \tau)$$

unless the second derivative is taken twice with respect to the same B^*. Thus

$$\left(\frac{\partial I(B^*, \lambda_\epsilon)}{\partial B^*_{i_\mu j_\nu}}\right)_0 = -2 \frac{B_0^{*\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda < \lambda_\epsilon} \sum_{\alpha=1}^{N} x_\alpha^{i_\mu} x_\alpha^{j_\nu} e^{-\sum_{\alpha=1}^{N} B^*_{ij0} x_\alpha^i x_\alpha^j} dx,$$

where the B_0^* indicates that the B's have the diagonal form associated with H. And since whenever the point $x_1^1, \cdots ; x_1^i, \cdots, x_N^i ; \cdots, x_N^m$ is in the region $\lambda \leq \lambda_\epsilon$, so also is the point $x_1^1, \cdots ; -x_1^i, \cdots, -x_N^i ; \cdots, x_N^m$ it follows that

$$\frac{\partial}{\partial B^*_{i_\mu j_\nu}} I(B_0^*, \lambda_\epsilon) = 0, \qquad (\mu \neq \nu).$$

Similar considerations show that the non-repeated second derivatives

$$\frac{\partial^2}{\partial B^*_{i_\mu j_\nu} \partial B^*_{h_\sigma k_\tau}} I(B_0^*, \lambda_\epsilon) = 4 \frac{B_0^{*\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda < \lambda_\epsilon} \left(\sum_{\alpha=1}^{N} x_\alpha^{i_\mu} x_\alpha^{j_\nu}\right)\left(\sum_{\beta=1}^{N} x_\beta^{h_\sigma} x_\beta^{k_\tau}\right) e^{-\sum_{\alpha=1}^{N} B^*_{ij0} x_\alpha^i x_\alpha^j} dx$$

must vanish.

Finally, we must show that the repeated second derivatives are positive when evaluated at a point in ω, except of course in the trivial cases $\lambda_\epsilon = 0$, $\lambda_\epsilon = 1$, when they must be zero. In order to do this, we shall make use of the fact that the v's which go to make up λ have the Wishart distribution [17]

$$(5.5) \qquad \frac{B^{\frac{1}{2}(N-1)}}{\pi^{\frac{1}{4}(m(m-1))} \prod_{i=1}^{m} \Gamma[\frac{1}{2}(N-i)]} \cdot v^{\frac{1}{2}(N-m)-1} e^{-B_{ij}v^{ij}} dv^{11} \cdots dv^{mm}.$$

(Because of the relation $v^{ij} = v^{ji}$, only $\frac{1}{2}m(m+1)$ of the v's appear as differentials). It will be useful to have the notation

$$G(B, N-1, m) = \frac{B^{\frac{1}{2}(N-1)}}{\pi^{\frac{1}{4}(m(m-1))} \prod_{i=1}^{m} \Gamma[\frac{1}{2}(N-i)]},$$

$$V(B, N-1, m) = v^{\frac{1}{2}(N-m)-1} e^{-B_{ij}v^{ij}}$$

With the aid of (5.5) we shall now compute the moments

$$E[(\lambda^{2/N})^h], \qquad h = 0, 1, \cdots,$$

for the case in which the matrix $\| B_{ij} \|$ has the form

$$(5.6)\qquad \left\| \begin{matrix} B_{11} & \cdots & B_{1m_1} & 0 & \cdots & 0 B_{1m} \\ \cdot & & & & & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{m_1 1} & \cdots & B_{m_1 m_1} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & & & \\ \vdots & & \vdots & & \| \bar{B} \| & \\ 0 & & & & & \\ B_{m1} 0 & \cdots & 0 & & & \end{matrix} \right\|,$$

where $\| \bar{B} \|$ stands for $\| B_2 \| \dotplus \cdots \dotplus \| B_p \|$, and all other B's, except those indicated, are zero. Let us designate by $(\bar{v})$ the set of v^{ij} which correspond to the rows and columns of $\bar{B}$, and by $(v - \bar{v})$ the remaining v's. We then remark that the result of integrating (5.5) with respect to the v's in $(v - \bar{v})$ is to reduce it to the corresponding distribution for the variables in the set $\bar{v}$, thus:

$$(5.7)\qquad \begin{aligned} G(B, N-1, m) \int V(B, N-1, m)\, d(v - \bar{v}) \\ = G(\tilde{B}, N-1, m-m_1) V(\tilde{B}, N-1, m-m_1), \end{aligned}$$

where $\| \tilde{B}_{kl} \|$ is the inverse of the matrix obtained by inverting $\| B_{ij} \|$, and striking out the first m_1 rows and columns, that is

$$\tilde{B}^{kl} = B^{kl}, \qquad (k, l = m_1 + 1, \cdots, m).$$

Then,

$$G(B, N-1, m) \int \frac{v^h}{v_2^h \cdots v_p^h} V(B, N-1, m)\, d(v - \bar{v})$$

can be written as

$$(5.8)\qquad \begin{aligned} \frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot G(B, N-1+2h, m) \int v_2^{-h} \cdots v_p^{-h} \\ \times V(B, N-1+2h, m)\, d(v - \bar{v}) \\ = \frac{G(B, N-1, m)}{G(B, N-1+2h, m)} G(\tilde{B}, N-1+2h, m-m_1) \\ \times v_2^{-h} \cdots v_p^{-h} V(\tilde{B}, N-1+2h, m-m_1). \end{aligned}$$

It can be seen from (5.6) that

$$\| \tilde{B} \| = \| B_2 \| + \cdots \dotplus \| B_{p-1} \| \dotplus \| \tilde{B}_p \|$$

since of all the rows and columns of $\| B_{ij} \|$ which are involved in $\| \bar{B} \|$ it is only the last in which a non zero element appears outside of the blocks $\| B_2 \|$, $\cdots$, $\| B_p \|$. Consequently, the v's corresponding to the determinants $v_2, \cdots,$

v_p are independently distributed, so that if in (5.8) we integrate out all the remaining v's but these, we shall be left with a product of factors

$$\frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot \prod_{t=2}^{p-1} \frac{G(B_t, N-1+2h, k_t)}{G(B_t, N-1, k_t)} \times G(B_t, N-1, k_t) v_t^{-h} V(B_t, N-1+2h, k_t)$$

$$\times \frac{G(\tilde{B}_p, N-1+2h, k_p)}{G(\tilde{B}_p, N-1, k_p)} \cdot G(\tilde{B}_p, N-1, k_p) v_p^{-h} V(\tilde{B}_p, N-1+2h, k_p),$$

where k_μ stands for the order of $\| B_\mu \|$. And this, when integrated with respect to the v's in $v_2, \cdots, v_p$, yields

$$\frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot \prod_{t=2}^{p-1} \frac{G(B_t, N-1+2h, k_t)}{G(B_t, N-1, k_t)} \times \frac{G(\tilde{B}_p, N-1+2h, k_p)}{G(\tilde{B}_p, N-1, k_p)},$$

which, because of the definition of the G's, reduces to

$$\prod_{i=1}^{m} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot \prod_{t=2}^{p} \prod_{i=1}^{k_t} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]} \times B^{-h} B_2^h \cdots B_{p-1}^h \tilde{B}_p^h.$$

Denoting the product of ratios of Γ's by K_h, and recalling the form of $\| B_{ij} \|$, we therefore have

$$E\left[\frac{v^h}{v_2^h \cdots v_p^h}\right] = K_h \tilde{B}_p^h B'^{-h} \tag{5.9}$$

with

$$\| B' \| = \begin{Vmatrix} B_{11} & \cdots & B_{1m_1} & 0 & \cdots & 0 B_{1m} \\ & & & & & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{m_1 1} & \cdots & B_{m_1 m_1} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & & & \\ \vdots & & \vdots & & \| B_p \| & \\ 0 & & & & & \\ B_{m1} 0 & \cdots & 0 & & & \end{Vmatrix}$$

But it is not difficult to see that under the condition (5.6), the matrix $\| \tilde{B}_p \|$ is also the inverse of the matrix obtained by striking out the first m rows and columns in the inverse of $\| B' \|$. Making use of this relation, we can apply the Jacobi theorem to (5.9), and put that expression in the form

$$E\left[\frac{v^h}{v_2^h \cdots v_p^h}\right] = K_h B_1^{-h},$$

where $\| B_1 \|$ is the matrix in the upper left hand corner of $\| B' \|$, namely $\| B_{i_1 j_1} \|$.

Let the subscript β on a B stand for the result of replacing $B_{i_1 j_1}$ by $B_{i_1 j_1} + \beta_{i_1 j_1}$. For sufficiently small values of the β's the matrix $\| B_{ij\beta} \|$ will still be positive definite, so that we shall have

$$\frac{B_\beta^{\frac{1}{2}(N-1)}}{\pi^{\frac{1}{4}(m(m-1))} \prod_{i=1}^{m} \Gamma[\frac{1}{2}(N-i)]} \int \frac{v^h}{v_2^h \cdots v_p^h} v^{\frac{1}{2}(N-m)-1} e^{-B_{ij\beta} v^{ij}} \, dv = K_h B_{1\beta}^{-h},$$

which we can put in the form

$$K' \int \frac{v^h}{v_2^h \cdots v_p^h} v^{\frac{1}{2}(N-m)-1} e^{-B_{ij\beta} v^{ij}} \, dv = \frac{K_h}{B_{1\beta}^h B_\beta^{\frac{1}{2}(N-1)}}. \tag{5.10}$$

Wilks [13] has shown how to generate moments of determinants by the device of replacing $\beta_{i_1 j_1}$ by $\beta_{i_1 j_1} + \xi_{i_1} \xi_{j_1}$, and integrating with respect to the ξ's from $-\infty$ to ∞. Applying this process $2h$ times to the left hand side of (5.10) gives

$$\pi^{t_1 h} K' \int \left(\frac{v}{v_1 \cdots v_p} \right)^h V(B_\beta, N-1, m) \, dv,$$

which when multiplied by $\pi^{-t_1 h} B^{\frac{1}{2}(N-1)}$ yields

$$E[(\lambda^{2/N})^h]$$

when the β's are set equal to zero.

To obtain the value of this expression, we may perform the same operations on the right hand side of (5.10). But before so doing, we shall put B_β in a more convenient form. We have

$$B_\beta = B_{1\beta} \cdot \bar{B} - B_{1m}^2 \cdot \bar{B} \bar{B}^{mm} \cdot B_{1\beta}'^{11},$$

where $\bar{B}^{mm}$ is the inverse element of B_{mm} in $\| \bar{B} \|$, and $B_{1\beta}'^{11}$ is the cofactor of $B_{11\beta}$ in $B_{1\beta}$, the result being obtained by expanding B_β according to minors of the first row and first column. Similarly,

$$B = B_1 \cdot \bar{B} - B_{1m}^2 \cdot \bar{B} \bar{B}^{mm} \cdot B_1'^{11}. \tag{5.11}$$

From (5.11) we have

$$\frac{B}{B_1 \cdots B_p} = 1 - B_{1m}^2 \bar{B}^{mm} \cdot \frac{B_1'^{11}}{B_1},$$

so that if we put $B \cdot B_1^{-1} \cdots B_p^{-1} = \Lambda$, we find that

$$B_\beta = B_{1\beta} \cdot \bar{B} \left\{ 1 - B_{1m}^2 \bar{B}^{mm} \cdot \frac{B_1'^{11}}{B_1} \cdot \frac{B_1}{B_1'^{11}} \cdot \frac{B_{1\beta}'^{11}}{B_{1\beta}} \right\}$$

$$= B_{1\beta} \bar{B} \left\{ 1 - \frac{B_1}{B_1'^{11}} (1 - \Lambda) \frac{B_{1\beta}'^{11}}{B_{1\beta}} \right\}.$$

Thus the result of multiplying (5.10) through by $B^{\frac{1}{2}(N-1)}$ (where no β's are substituted in this determinant) can be put in the form

$$(5.12)\qquad \left(\frac{B}{B_2 \cdots B_p}\right)^{\frac{1}{2}(N-1)} \left\{1 - \frac{B_1}{B_1'^{11}}(1-\Lambda)\frac{B_{1\beta}'^{11}}{B_{1\beta}}\right\}^{-\frac{1}{2}(N-1)} B_{1\beta}^{-h}.$$

Expanding the expression in curled brackets, we get

$$\Lambda^{\frac{1}{2}(N-1)} B_1^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} B_1^{\nu} \left(\frac{B_{1\beta}'^{11}}{B_1'^{11}}\right)^{\nu} B_{1\beta}^{-[\frac{1}{2}(N-1)+h+\nu]} (1-\Lambda)^{\nu}.$$

If we let $B_{1\beta t}$ stand for the result of replacing B_{11} by $B_{11} - t$ in $B_{1\beta}$, we can write this as

$$(5.13)\qquad \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} (1-\Lambda)^{\nu} (B_1'^{11})^{-\nu} B_1^{\frac{1}{2}(N-1)+\nu} \times \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \frac{\partial^{\nu}}{\partial t^{\nu}} B_{1\beta t}^{-[\frac{1}{2}(N-1)+h]}$$

the derivatives being evaluated at $t = 0$.

Now Wilks' results show that the operation of introducing $\beta_{i_1 j_1} + \xi_{i_1}\xi_{j_1}$ into $B_{1\beta t}$ to replace $\beta_{i_1 j_1}$ and integrating with respect to the ξ's, when repeated $2h$ times on $B_{1\beta t}^{-[\frac{1}{2}(N-1)+h]}$, produces

$$\pi^{m_1 h} B_{1t}^{-\frac{1}{2}(N-1)} \prod_{i=1}^{m_1} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$

when the β's are finally set equal to zero. Reversing the order of summation, differentiation and integration in (5.13), we thus obtain

$$(5.14)\qquad \pi^{m_1 h} \prod_{i=1}^{m_1} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]} \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \times (1-\Lambda)^{\nu}(B_1'^{11})^{-\nu} B_1^{\frac{1}{2}(N-1)+\nu} \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \left(\frac{\partial^{\nu}}{\partial t^{\nu}} B_{1t}^{-\frac{1}{2}(N-1)}\right)_0.$$

Now

$$\left(\frac{\partial^{\nu}}{\partial t^{\nu}} B_{1t}^{-\frac{1}{2}(N-1)}\right)_0 = \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)]} \cdot (B_1'^{11})^{\nu} B_1^{-[\frac{1}{2}(N-1)+\nu]},$$

so that (5.14) becomes

$$\pi^{m_1 h} \prod_{i=1}^{m_1} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]} \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \times (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)]}.$$

From this it appears that the h-th moment of $\lambda_I^{2/N}$ is given by

$$E[(\lambda^{2/N})^h] = \prod_{i=1}^{m} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot \prod_{t=1}^{p} \prod_{i=1}^{k_t} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$
$$\times \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\Gamma[\frac{1}{2}(N-1)]} \tag{5.15}$$
$$\times \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)]}.$$

A considerable amount of cancellation will take place in (5.15), for m is greater than any k_t. Suppose the largest k_t is $k_{t'}$. Then we can cancel its product into the first one, with the assurance that there will be at least one factor

$$\frac{\Gamma[\frac{1}{2}(N-1)]}{\Gamma[\frac{1}{2}(N-1)+h]} \tag{5.16}$$

to cancel the corresponding factor under the summation sign. Hence we have

$$E[(\lambda^{2/N})^h] = \prod_{i=k_{t'}+1}^{m} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot \prod_{t=1}^{p}{}' \prod_{i=1}^{k_t}{}'' \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$
$$\times \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\Gamma[\frac{1}{2}(N-1)]} \cdot \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]}, \tag{5.17}$$

where Π' indicates that t' has been omitted, and Π'' indicates that one factor (5.16) has been cancelled. Then we can take out the factor $i = m$ in the first product, putting it under the summation sign, where, together with the final factor in each term of the sum, it gives rise to the combination

$$\frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-m)]\Gamma[\frac{1}{2}(m-1)+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(N-m)+h]\Gamma[\frac{1}{2}(m-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]}.$$

After making this reduction, we obtain

$$E[(\lambda^{2/N})^h] = \prod_{i=k_{t'}+1}^{m-1} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot \prod_{t=2}^{p}{}' \prod_{i=1}^{k_t}{}' \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$
$$\times \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\Gamma[\frac{1}{2}(N-1)]} \frac{B[\frac{1}{2}(N-m)+h, \frac{1}{2}(m-1)+\nu]}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]}. \tag{5.18}$$

The products of ratios in the first part of (5.18) are of the type discussed by Wilks in connection with integral equations of type B [12]. It follows from his results that $\lambda_I^{2/N}$ is distributed like the product

$$z \cdot \theta_1 \cdots \theta_{m'} \qquad (m' = m - k_{t'} - 1),$$

where z and the θ's are independently distributed, with the distribution of the θ's given by

$$f(\theta_1, \cdots, \theta_{m'}) = \prod_{i=1}^{m'} \frac{\Gamma(c_i)}{\Gamma(b_i)\Gamma(c_i-b_i)} \cdot \theta_i^{b_i-1}(1-\theta_i)^{c_i-b_i-1},$$

where the b_i and c_i are constants which depend on N, m, and the sizes of the blocks, but not on Λ, and the distribution of z is given by

$$F(z) = \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1 - \Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\Gamma[\frac{1}{2}(N-1)]} \cdot \frac{z^{\frac{1}{2}(N-m)-1}(1-z)^{\frac{1}{2}(m-1)+\nu-1}}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]}.$$

Consequently, the probability that λ lies between zero and λ_ϵ is

$$J(\Lambda, \lambda_\epsilon) = \Lambda^{\frac{1}{2}(N-1)} \int_S \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\Gamma[\frac{1}{2}(N-1)]} \times f(\theta) \frac{z^{\frac{1}{2}(N-m)-1}(1-z)^{\frac{1}{2}(m-1)+\nu-1}}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]} dz\, d\theta,$$

where the integral is to be extended over the region

$$S: 0 \leq z \cdot \theta_1 \cdots \theta_{m'} < \lambda_\epsilon^{2/N}, \qquad 0 \leq \theta_i \leq 1, \qquad 0 \leq z \leq 1.$$

Let us integrate first with respect to z and then with respect to the θ's; we have

$$(5.19)\quad J(\Lambda, \lambda_\epsilon) = \int_{S_\theta} \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\Gamma[\frac{1}{2}(N-1)]} \times \frac{B'[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu; \varphi]}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]} f(\theta)\, d\theta,$$

where S_θ is the set $\Pi\theta_i < \lambda_\epsilon^{2/N}$, $0 \leq \theta_i \leq 1$, and

$$(5.20)\quad \begin{aligned} B'(u, v, \varphi) &= \int_0^{\varphi} z^{u-1}(1-z)^{v-1} dz \\ &= \int_{1-\varphi}^{1} z^{v-1}(1-z)^{u-1} dz = B(v, u, 1-\varphi), \end{aligned}$$

$\varphi(\theta)$ being the upper limit for z for fixed θ. It is clear that the subset of s_θ for which $\varphi(\theta) < 1$ will not be of measure zero in the θ-space, since we assume that $0 < \lambda_\epsilon < 1$.

The relation between (5.19) and the corresponding expression for the multiple correlation coefficient without fixed variates—the case $\bar{y} = 0$ in (4.4)—may be clearer if we put

$$(5.21)\qquad \bar{\rho} = 1 - \Lambda = B_{1m}^2 \bar{B}^{mm} B_1^{11},$$

where $\bar{B}^{mm}$ is the inverse of B_{mm} in $\| \bar{B} \|$, and B_1^{11} is the inverse of B_{11} in $\| B_1 \|$. Then the required probability of rejection when $\bar{\rho}$ has any fixed value is

$$I(\bar{\rho}, 1 - \lambda_\epsilon^{2/N}) = \int_{S_\theta} \sum_{\nu=0}^{\infty} \frac{\bar{\rho}^{\nu}}{\nu!} (1-\bar{\rho})^{\frac{1}{2}(N-1)} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)]} \times \frac{B[\frac{1}{2}(m-1)+\nu, \frac{1}{2}(N-m), 1-\varphi]}{B[\frac{1}{2}(m-1)+\nu, \frac{1}{2}(N-m]} f(\theta)\, d\theta,$$

where we have used the relation (5.20) between the incomplete Beta functions. Differentiating with respect to $\bar{\rho}$ before performing the integration with respect to the θ's, we find by a computation similar to that in section 4 that each term in the series is positive except where $\varphi(\theta) = 1$; so that we have

$$\frac{\partial I}{\partial \bar{p}}(\bar{p}, 1 - \lambda_\epsilon^{2/N}) > 0 \qquad (\lambda_\epsilon \neq 1, 0).$$

And by (5.21), we then have

$$\frac{\partial^2 I}{\partial B_{1m}^2} > 0.$$

Since the argument is clearly independent of which $B_{i_\mu j_\nu}$ $(\mu \neq \nu)$ we take, it follows that the test is locally unbiased. We have therefore proved:

THEOREM III. *If $x^1, \cdots, x^m$ have the joint normal distribution* (5.1), *then the likelihood ratio test for the hypothesis that the x's are independent in sets is locally unbiased.*

In certain types of statistical material it may be important to consider, not the independence of the x's themselves, but of their deviations from regression functions. For example, in the case of several related time series, it may be desirable to eliminate the trend of each x^i by means of, say, a second degree polynomial in t. Consider then in general a population whose distribution function is of the form

$$\frac{B^{\frac{1}{2}}}{\pi^{\frac{1}{2}m}} e^{-B_{ij}(x^i - C_\mu^i x^\mu)(x^j - C_\nu^j x^\nu)} \qquad (\mu, \nu = m + 1, \cdots, m + q)$$

with unknown B_{ij} and C_μ^i. The likelihood ratio for testing the hypothesis H_I that the sets of deviations

$$x^1 - C_\mu^1 x^\mu, \cdots, x^{m_1} - C_\mu^{m_1} x^\mu; \cdots; x^{m_{p-1}+1} - C_\mu^{m_{p-1}+1} x^\mu, \cdots, x^m - C_\mu^m x^\mu$$

are independent is

$$\lambda_I = \left\{\frac{|d^{ij}|}{d_1 \cdots d_p}\right\}^{\frac{1}{2}N}$$

where

$$d^{ij} = \Sigma(x_\alpha^i - \hat{C}_\mu^i x_\alpha^\mu)(x_\alpha^j - \hat{C}_\nu^j x_\alpha^\nu)$$

and $\hat{C}_\mu^i$ is the usual least squares estimate of C_μ^i, given by

$$\hat{C}_\mu^i a^{\mu\nu} = a^{i\nu}$$

with

$$a^{rs} = \Sigma x_\alpha^r x_\alpha^s \qquad (r, s = 1, \cdots, m + q).$$

An examination of the characteristic function of the d^{ij} shows that their distribution law is the same as that of the v^{ij} of the preceding discussion, except for the fact that $N - 1$ is replaced by $N - q$. Consequently the above results on freedom from bias, and also those of the next section, apply equally well to the λ_I test for the independence of deviations from regression functions.

6. **On the moments of $\lambda_I^{2/N}$.** Although we have succeeded in proving the unbiased nature of the preceding test only in the local sense, we can show that the moments of the criterion $\lambda_I^{2/N}$ have a property which seems very closely related to that of furnishing a completely unbiased test. For it can be shown that each of the quantities

$$E[(\lambda^{2/N})^h] \qquad h = \tfrac{1}{2}, 1, 1\tfrac{1}{2}, \cdots$$

is greater, when H_I is true than when any alternative H' holds. It will perhaps be sufficient to prove this statement in detail for the case where $h = 1$ and where H_I is the hypothesis that the matrix $\| B_{ij} \|$ has the form $\| \breve{B}_0 \| \dotplus \|B_{i_3 j_3}\|$:

$$\left\| \begin{matrix} \begin{matrix} B_{11} B_{12} \\ B_{21} B_{22} \end{matrix} & 0 & 0 \\ 0 & \begin{matrix} B_{33} B_{34} \\ B_{43} B_{44} \end{matrix} & 0 \\ 0 & 0 & \| B_{i_3 j_3} \| \end{matrix} \right\| .$$

in the notation of the preceding section we then have

$$i_1, j_1 = 1, 2,; \qquad i_2, j_2 = 3, 4; \qquad i_3, j_3 = 5, \cdots, m.$$

Even when H is not true we find that

$$(6.1) \quad E[| v^{ij} |^h | v^{i_3 j_3} |^{-h}] = \frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot \frac{G(\tilde{B}, N-1+2h, m-4)}{G(\tilde{B}, N-1, m-4)},$$

where $\tilde{B}^{i_3 j_3} = B^{i_3 j_3}$ Using the definition of the G's in section 5 and the Jacobi theorem, we can write (6.1) in the form

$$E[| v^{ij} |^h | v^{i_3 j_3} |^{-h}] = K_h \breve{B}^{-h}$$

where $\breve{B}$ is the determinant of the matrix composed of the first four rows and columns of $\| B_{ij} \|$. In the general case we therefore have

$$\| \breve{B} \| = \left\| \begin{matrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{matrix} \right\| .$$

Thus if we set $h = 1$, and replace $B_{i_1 j_1}$ and $B_{i_2 j_2}$ by $B_{i_1 j_1} + \xi_{i_1}^{(1)} \xi_{j_1}^{(1)} + \xi_{i_1}^{(2)} \xi_{j_1}^{(2)}$ and $B_{i_2 j_2} + \xi_{i_2}^{(3)} \xi_{j_2}^{(3)} + \xi_{i_2}^{(4)} \xi_{j_2}^{(4)}$ respectively, indicating this replacement by a prime, we obtain

$$(6.2) \qquad E[(\lambda^{2/N})^1] = K_1 \int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1} \, d\xi.$$

Treating B' as a bordered determinant, we can reduce it to

$$\begin{aligned} B' &= B_{(123)}(1 + B_{(123)}^{i_2 j_2}\xi_{i_2}^{(4)}\xi_{j_2}^{(4)}) \\ &= B_{(12)}(1 + B_{(12)}^{i_2 j_2}\xi_{i_2}^{(3)}\xi_{j_2}^{(3)})(1 + B_{(123)}^{i_2 j_2}\xi_{i_2}^{(4)}\xi_{j_2}^{(4)}) \\ &= B_{(1)}(1 + B_{(1)}^{i_1 j_1}\xi_{i_1}^{(2)}\xi_{j_1}^{(2)})(1 + B_{(12)}^{i_2 j_2}\xi_{i_2}^{(3)}\xi_{j_2}^{(3)})(1 + B_{(123)}^{i_2 j_2}\xi_{i_2}^{(4)}\xi_{j_2}^{(4)}) \\ &= B(1 + B^{i_1 j_1}\xi_{i_1}^{(1)}\xi_{j_1}^{(1)})(1 + B_{(1)}^{i_1 j_1}\xi_{i_1}^{(2)}\xi_{j_1}^{(2)})(1 + B_{(12)}^{i_2 j_2}\xi_{i_2}^{(3)}\xi_{j_2}^{(3)})(1 + B_{(123)}^{i_2 j_2}\xi_{i_2}^{(4)}\xi_{j_2}^{(4)}), \end{aligned}$$

where the subscripts on the B's indicate the sets of ξ's still contained in the determinants, and $\| B^{ij} \| = \| B_{ij} \|^{-1}$. Similarly,

$$\begin{aligned} (6.4)\quad & \breve{B}' \\ &= \breve{B}(1 + \breve{B}^{i_1 j_1}\xi_{i_1}^{(1)}\xi_{j_1}^{(1)})(1 + \breve{B}_{(1)}^{i_1 j_1}\xi_{i_1}^{(2)}\xi_{j_1}^{(2)})(1 + \breve{B}_{(12)}^{i_2 j_2}\xi_{i_2}^{(3)}\xi_{j_2}^{(3)})(1 + \breve{B}_{(123)}^{i_2 j_2}\xi_{i_2}^{(4)}\xi_{j_2}^{(4)}), \end{aligned}$$

the inverse now being taken with respect to $\| \breve{B} \|$.

But between, say, $\breve{B}_{(12)}^{i_2 j_2}$ and $B_{(12)}^{i_2 j_2}$, there is the relation

$$\breve{B}_{(12)}^{i_2 j_2} = B_{(12)}^{i_2 j_2} - B_{(12)}^{i_2 i_3} B_{(12)\, i_3 j_3} B_{(12)}^{j_3 j_2}, \tag{6.5}$$

where $\| B_{(12)\, i_3 j_3} \| = \| B_{(12)}^{i_3 j_3} \|^{-1}$, that is, the inverse of the matrix obtained by deleting the first four rows and columns of $\| B_{(12)}^{ij} \|$. Consequently

$$\breve{B}_{(12)}^{i_2 j_2}\xi_{i_2}^{(3)}\xi_{j_2}^{(3)} \leq B_{(12)}^{i_2 j_2}\xi_{i_2}^{(3)}\xi_{j_2}^{(3)}$$

with equality holding only for those values of the ξ's for which

$$\xi_{i_2}^{(3)} B_{(12)}^{i_2 i_3} = 0 \qquad i_3 = 5, \cdots, m.$$

And this set of ξ's will not make up the entire ξ space unless $\| B_{ij} \| = \| \breve{B} \| \dotplus \| B_{i_3 j_3} \|$. Applying the same kind of reasoning to the other quadratic forms in (6.4), we can therefore show that

$$\begin{aligned} &\int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1} d\xi \\ &\qquad \leq \breve{B}^{-1} \int (1 + \breve{B}^{i_1 j_1}\xi_{i_1}^{(1)}\xi_{j_1}^{(1)})^{-\frac{1}{2}(N+1)} \cdots (1 + \breve{B}_{(123)}^{i_2 j_2}\xi_{i_2}^{(4)}\xi_{j_2}^{(4)})^{-\frac{1}{2}(N+1)} d\xi. \end{aligned}$$

The last form can be reduced to a sum of squares with unit coefficients by a linear transformation of the $\xi^{(4)}$'s; thus

$$\begin{aligned} &\int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1} d\xi \\ (6.6)\quad &\qquad \leq \breve{B}^{-1} \int | \breve{B}_{(123)}^{i_2 j_2} |^{-\frac{1}{2}} (1 + \breve{B}^{i_1 j_1}\xi_{i_1}^{(1)}\xi_{j_1}^{(1)})^{-\frac{1}{2}(N-1)} \cdots (1 + \Sigma\xi_{i_2}^{(4)}\xi_{i_2}^{(4)})^{-\frac{1}{2}(N+1)} d\xi. \end{aligned}$$

And by making use of the fact that

$$\breve{B}_{(123)}^{i_2 j_2} = \breve{B}_{(123)}^{-1} \cdot | B_{(12)\, i_1 j_2} |,$$

we can express the right-hand side of (6.6) as

$$\breve{B}^{-1}\int \breve{B}^{\frac{1}{2}}_{(123)}\cdot|B_{(12)i_1j_1}|^{-\frac{1}{2}}(1+\breve{B}^{i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}(N+1)}\cdots(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.$$

This in turn becomes [c. f. (6.4)]

$$\breve{B}^{-\frac{1}{2}}\int|B_{(12)i_1j_1}|^{-\frac{1}{2}}(1+\breve{B}^{i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}N}(1+\breve{B}^{i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}N}$$
$$\times(1+\breve{B}^{i_2j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{j_2})^{-\frac{1}{2}(N+1)}\,d\xi$$
$$=\int|B_{(12)i_1j_1}|^{-1}(1+\breve{B}^{i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}N}(1+\breve{B}^{i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}(N-1)}$$
$$\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.$$

At this stage we can write

$$|B_{(12)i_1j_1}|=|B_{i_1j_1}|(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})(1+B^{*i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1}),$$

where $\|B^{*i_1j_1}_{(1)}\|=\|B_{(1)i_1j_1}\|^{-1}$, and apply the relation

$$B^{*i_1j_1}_{(1)}=\breve{B}^{i_1j_1}_{(1)}-\breve{B}^{i_1i_2}_{(1)}\breve{B}_{(1)i_2j_2}\breve{B}^{j_2j_1}_{(1},\qquad\|\breve{B}_{(1)i_2j_2}\|=\|\breve{B}^{i_2j_2}_{(1)}\|^{-1}.$$

Therefore,

$$B^{*i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1}<\breve{B}^{i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1},$$

unless $\xi^{(2)}_{i_1}\breve{B}^{i_1i_2}_{(1)}=0$ $(i_2=3,4)$. We can thus continue as follows

$$\int B^{\frac{1}{2}(N-1)}B'^{-\frac{1}{2}(N-1)}\breve{B}'^{-1}\,d\xi$$
$$\leq|B_{i_1j_1}|^{-1}\int(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}(N+1)}(1+B^{*i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}(N+1)}$$
$$\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.$$

Transforming the $\xi^{(2)}$'s, we get

$$|B_{i_1j_1}|^{-1}\int|B^{*i_1j_1}_{(1)}|^{-\frac{1}{2}}(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}(N+1)}(1+\Sigma\xi^{(2)}_{i_1}\xi^{(2)}_{i_1})^{-\frac{1}{2}(N+1)}$$
$$\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.$$

Since $|B^{*i_1j_1}_{(1)}|^{-1}=|B_{(1)i_1j_1}|$, this becomes

$$|B_{i_1j_1}|^{-\frac{1}{2}}\int(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}N}(1+\Sigma\xi^{(2)}_{i_1}\xi^{(2)}_{i_1})^{-\frac{1}{2}(N+1)}$$
$$\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{j_2})^{-\frac{1}{2}(N+1)}\,d\xi$$
$$=\int(1+\Sigma\xi^{(1)}_{i_1}\xi^{(1)}_{i_1})^{-\frac{1}{2}N}(1+\Sigma\xi^{(2)}_{i_1}\xi^{(2)}_{i_1})^{-\frac{1}{2}(N+1)}$$
$$\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.$$

Collecting these results, we finally obtain

$$
\begin{aligned}
(6.7) \quad & K_1 \int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1} \, d\xi \\
& \leq K_1 \int (1 + \Sigma \xi_{i_1}^{(1)} \xi_{i_1}^{(1)})^{-\frac{1}{2}N} (1 + \Sigma\, \xi_{i_1}^{(2)} \xi_{i_1}^{(2)})^{-\frac{1}{2}(N+1)} \\
& \qquad \times (1 + \Sigma \xi_{i_2}^{(3)} \xi_{i_2}^{(3)})^{-\frac{1}{2}N} (1 + \Sigma \xi_{i_2}^{(4)} \xi_{i_2}^{(4)})^{-\frac{1}{2}(N+1)} \, d\xi
\end{aligned}
$$

with equality only in case H_I is true. But the right side of (6.7) is the first moment of $\lambda_I^{2/N}$ computed under the hypothesis H_I, while the left side gives the corresponding moment in the general case.

The possibility of carrying out this reduction for the case in which the matrix $\| \breve{B} \|$ has more than two blocks, or blocks of unequal size, seems sufficiently clear. And to obtain higher moments, we have only to introduce the proper number of ξ's into each set. We then have:

THEOREM IIIa. *Let λ_i be the likelihood ratio appropriate to testing the hypothesis H_I that the normally distributed variates $x^1, \cdots, x^m$ fall into the mutually independent sets $x^1, \cdots, x^{m_1}; \cdots; x^{m_{p-1}+1}, \cdots, x^m$. Then the expected value of $(\lambda_I^{2/N})^h$, $h = \frac{1}{2}, 1, 1\frac{1}{2}, \cdots$, is greater under the null hypothesis H_I than under any alternative hypothesis in Ω.*

7. **The general regression problem.** Let the variates $x^1, \cdots, x^t$ be distributed according to the law

$$
(7.1) \qquad \frac{| B_{ij} |^{\frac{1}{2}}}{\pi^{\frac{1}{2}t}} e^{-B_{ij}(x^i - C_\mu^i x^\mu - C_\sigma^i x^\sigma)(x^j - C_\nu^j x^\nu - C_\tau^j x^\tau)}.
$$

Throughout this section, let the ranges of the indices be

$$
\begin{aligned}
i, j &= 1, \cdots, t & p, q &= t + 1, \cdots, m \\
r, s &= 1, \cdots, m & r', s' &= 1, \cdots, t + q \\
\mu, \nu &= t + 1, \cdots, t + q & \sigma, \tau &= t + q + 1, \cdots, m.
\end{aligned}
$$

In (7.1) we therefore have t random variates, and $m - t$ fixed variates. Consider the hypothesis H that the x^i are independent of the last set of x's, namely x^σ. We have

$$
\Omega: \quad \| B_{ij} \| \text{ positive definite}, \ -\infty < C_p^i < \infty,
$$

while for ω we impose the additional requirement

$$
C_\sigma^i = 0.
$$

Thus in general we have for the distribution of random samples 0_N, $N \geq m$,

$$
(7.2) \qquad \frac{| B_{ij} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x_\alpha^i - C_p^i x_\alpha^p)(x_\alpha^j - C_q^j x_\alpha^q)}
$$

while when H is true, we have

$$(7.3) \qquad P = \frac{|B_{ij}|^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x_\alpha^i - C_\mu^i x_\alpha^\mu)(x_\alpha^j - C_\nu^j x_\alpha^\nu)}$$

Differentiating (7.2) with respect to the B's and C's and setting the derivatives equal to zero gives us the conditions

$$(7.4) \qquad \sum_{\alpha=1}^{N} C_p^i x_\alpha^p x_\alpha^q = \sum_{\alpha=1}^{N} x_\alpha^i x_\alpha^q,$$

$$(7.5) \qquad B^{ij} = \frac{2}{N} \sum_{\alpha=1}^{N} (x_\alpha^i - C_p^i x_\alpha^p)(x_\alpha^j - C_q^j x_\alpha^q).$$

As in section 2, we put

$$a^{rs} = \sum_{\alpha=1}^{N} x_\alpha^r x_\alpha^s.$$

and assume that the fixed values x_α^p have been so chosen that $\| a^{pq} \|$ is positive definite. Then (7.4) and (7.5) can be combined to give

$$B^{ij} = \frac{2}{N} (a^{ij} - a^{ip} a'_{pq} a^{qj}) = \frac{2}{N} \tilde{a}^{ij},$$

where $\| a'_{pq} \|^{-1} = \| a^{pq} \|$. It then follows that

$$P_\Omega = |\tilde{a}^{ij}|^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nt} e^{-\frac{1}{2}N}.$$

Similarly,

$$P_\omega = |\tilde{a}_0^{ij}|^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nt} e^{-\frac{1}{2}N},$$

where

$$\tilde{a}_0^{ij} = a^{ij} - a^{i\mu} a'_{\mu\nu} a^{\nu j}, \qquad \| a'_{\mu\nu} \|^{-1} = \| a^{\mu\nu} \|.$$

The matrix $\| a^{rs} \|$ will be positive definite except for a set of probability zero, so that we can consider $\| \tilde{a}^{ij} \|$ as the inverse of the matrix obtained by removing the last $m - t$ rows and columns of the inverse of $\| a^{rs} \|$, and $\| \tilde{a}_0^{ij} \|$ as the inverse of the matrix obtained by removing the last q rows and columns of $\| a^{r's'} \|^{-1}$. Then by the Jacobi theorem

$$|\tilde{a}^{ij}|^{-1} = \frac{|a^{pq}|}{|a^{rs}|}, \qquad |\tilde{a}_0^{ij}|^{-1} = \frac{|a^{\mu\nu}|}{|a^{r's'}|}$$

so that the appropriate likelihood ratio is given by

$$\lambda^{2/N} = \frac{|a^{rs}|}{|a^{r's'}|} \cdot \frac{|a^{\mu\nu}|}{|a^{pq}|}.$$

It will be advantageous to complete the matrix $\| B_{ij} \|$ in (7.1) by defining

$$B_{ip} = -B_{ij}C^j_p, \qquad B_{pq} = C^i_p B_{ij} C^j_q. \tag{7.6}$$

(Evidently $B_{ip} = 0$ for $i = 1, \cdots, t$ and fixed p, if and only if $C^j_p = 0$, $j = 1, \cdots, t$). We can now write (7.2) as

$$P(x, B) = \frac{| B_{ij} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x^i_\alpha + B^{ih}B_{hp}x^p_\alpha)(x^j_\alpha + B^{jk}B_{kq}x^q_\alpha)} \tag{7.7}$$

We next notice that λ is invariant under the transformations

$$x^i \rightarrow \alpha^i_j x^j, \qquad x^\sigma \rightarrow \beta^\sigma_\tau x^\tau,$$

so that if we put

$$I(B, \lambda_\epsilon) = \int_S P(x, B)\, dx^1_1 \cdots dx^t_N,$$

where the integral is extended over the region

$$S: \ 0 \leq \lambda \leq \lambda_\epsilon,$$

it turns out that

$$I(B, \lambda_\epsilon) \equiv I(B^*, \lambda_\epsilon),$$

provided

$$B^*_{ij} = \alpha^k_i B_{kl} \alpha^l_j, \qquad B^*_{i\mu} = \alpha^k_i B_{k\mu}, \qquad B^*_{i\sigma} = \alpha^k_i B_{k\tau} \beta^\tau_\sigma.$$

To prove the locally unbiased character of the test, we may therefore consider the derivatives

$$\frac{\partial}{\partial B^*_{i\sigma}} I(B^*_0, \lambda_\epsilon), \qquad \frac{\partial^2}{\partial B^*_{i\sigma} \partial B^*_{j\tau}} I(B^*_0, \lambda_\epsilon),$$

and assume that $\| B^*_{ij} \|$ and $\| a^{\mu\nu} \|$ are in diagonal form. We also observe that λ is unaltered by the transformation

$$x^i \rightarrow x^i + B^{ik} B_{k\mu} x^\mu.$$

We therefore have

$$I(B^*, \lambda_\epsilon) = \frac{| B^*_{ij} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} \int_S e^{-\sum_{\alpha=1}^{N} B^*_{ij}(x^i_\alpha + B^{*ik}B^*_{k\sigma}x^\sigma_\alpha)(x^j_\alpha + B^{*jl}B^*_{l\tau}x^\tau_\alpha)}\, dx.$$

Thus,

$$\frac{\partial}{\partial B^*_{k\sigma}} I(B^*_0, \lambda_\epsilon) = -2 \frac{| B^*_{ij0} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} \int_S \sum_{\alpha=1}^{N} x^k_\alpha x^\sigma_\alpha e^{-\sum_{\alpha=1}^{N} B^*_{ij0} x^i_\alpha x^j_\alpha}\, dx,$$

which is easily seen to be zero. Again, consider a non-repeated second partial derivative, say

$$\frac{\partial^2}{\partial B^*_{k\sigma}\partial B^*_{l\tau}} I(B^*_0, \lambda_\epsilon)$$

$$= -2 \frac{|B^*_{ij}|^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} \int_S \left(B^{*kl} a^{\sigma\tau} - 2 \sum_{\alpha=1}^N x^\sigma_\alpha x^k_\alpha \cdot \sum_{\beta=1}^N x^\tau_\beta x^l_\beta\right) e^{-\sum_{\alpha=1}^N B^*_{ij} x^i_\alpha x^j_\alpha} dx.$$

This plainly vanishes if $k \neq l$; but it is by no means easy to see what happens when $k = l$, even when $\sigma \neq \tau$. Let us therefore study the distribution law of $\lambda^{2/N}$ for the case,

$$B_{i\sigma} = 0, \qquad i \neq 1.$$

(We shall not, however, assume that the transformation $B \to B^*$ has been made on the B's.)

Define

$$\tilde{B}_{pq} = B_{pq} - B_{pi}B^{ij}B_{jq},$$

$$\tilde{a}^{\sigma\tau} = a^{\sigma\tau} - a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau},$$

where $\| a_{\mu\nu} \|$ now stands for the inverse of $\| a^{\mu\nu} \|$. These expressions will arise when we adapt Wilks' method of moment generating operators [13], based on the identity

$$\int e^{-B_{rs}a^{rs}} dx^1_1 \cdots dx^t_N = \pi^{\frac{1}{2}Nt} B^{-\frac{1}{2}N} \exp(-\tilde{B}_{pq}a^{pq}) \tag{7.8}$$

to the problem. We shall understand from now on that $B = |B_{ij}|$ and $\| B^{ij} \| = \| B_{ij} \|^{-1}$. Let us rearrange the form in the exponential on the right, thus:

$$\begin{aligned}\tilde{B}_{pq}a^{pq} &= (\tilde{B}_{\mu\nu}a^{\mu\nu} + 2B_{\mu\sigma}a^{\mu\sigma} + B_{\sigma\tau}a^{\sigma\tau} - 2B_{\mu i}B^{ij}B_{j\sigma}a^{\mu\sigma} \\ &\qquad - B_{\sigma i}B^{ij}B_{j\tau}a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau}) - B_{\sigma i}B^{ij}B_{j\tau}\tilde{a}^{\sigma\tau} \\ &= Q - B_{\sigma i}B^{ij}B_{j\tau}\tilde{a}^{\sigma\tau} \\ &= Q - B^{ij}y_{ij}.\end{aligned}$$

A subscript β will denote the result of replacing $B_{r's'}$ by $B_{r's'} + \beta_{r's'}$, and a prime will indicate that each $\beta_{r's'}$ has been replaced by $\beta_{r's'} + \xi_{r'}\xi_{s'}$. Consider now the result of integrating the right hand side of (7.8) after these replacements have been made:

$$\begin{aligned}&\pi^{\frac{1}{2}Nt} \int B'^{-\frac{1}{2}N}_\beta \exp(-\tilde{B}'_{pq\beta}a^{pq})\, d\xi_1 \cdots d\xi_t\, d\xi_{t+1} \cdots d\xi_{t+q} \\ &\qquad = \pi^{\frac{1}{2}Nt} \int B'^{-\frac{1}{2}N}_\beta e^{B'^{ij}_\beta y_{ij}} \left(\int e^{-Q'_\beta} d\xi_\mu\right) d\xi_i,\end{aligned} \tag{7.9}$$

Let us integrate first with respect to the ξ_μ. Wilks has shown how to write Q'_β in the form

$$Q'_\beta = -Q'_{1\beta} + B_{pq\beta}a^{pq} + \frac{B_\beta}{B'_\beta} a^{\mu\nu}\xi_\mu\xi_\nu - 2B_{pi\beta}B'^{ij}_\beta a^{p\nu}\xi_j\xi_\nu,$$

where

$$Q'_{1\beta} = B_{\mu i\beta}B'^{ij}_\beta B_{j\nu\beta}a^{\mu\nu} + 2B_{\mu i\beta}B'^{ij}_\beta B_{j\sigma}a^{\mu\sigma} + B_{\sigma i}B'^{ij}_\beta B_{j\tau}a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau}.$$

This latter expression is thus free of the ξ_μ. Consequently,

$$\int e^{-Q'_\beta}\,d\xi_\mu = \left(\frac{B'_\beta}{B_\beta}\right)^{\frac{1}{2}q} |\,a^{\mu\nu}\,|^{-\frac{1}{2}}\pi^{\frac{1}{2}q}e^{-B_{pq\beta}a^{pq}}e^{Q'_{1\beta}+Q'_{2\beta}},$$

where

$$Q'_{2\beta} = \frac{B'_\beta}{B_\beta} a_{\mu\nu}(a^{\mu p}B_{pi\beta}B'^{ij}_\beta\xi_j)(a^{\nu q}B_{qk\beta}B'^{kl}_\beta\xi_l),$$

which can be written

$$\frac{B'_\beta}{B_\beta}\{B_{\mu i\beta}B'^{ij}_\beta B_{\nu k\beta}B'^{kl}_\beta\xi_j\xi_l a^{\mu\nu} + 2B_{\mu i\beta}B'^{ij}_\beta B_{\sigma k}B'^{kl}_\beta\xi_j\xi_l a^{\mu\sigma}$$
$$+ B_{\sigma i}B'^{ij}_\beta B_{\tau k}B'^{kl}_\beta\xi_j\xi_l a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau}\}.$$

The method of reduction used by Wilks can now be applied to $Q'_{1\beta}$ and $Q'_{2\beta}$, and gives

$$Q'_{1\beta} + Q'_{2\beta} = B_{\mu i\beta}B^{ij}_\beta B_{j\nu\beta}a^{\mu\nu} + 2B_{\mu i\beta}B^{ij}_\beta B_{j\sigma}a^{\mu\sigma} + B_{\sigma i}B^{ij}_\beta B_{j\tau}a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau},$$

an expression which does not involve the ξ's. Thus

$$\int e^{-Q'_\beta}\,d\xi_\mu = \pi^{\frac{1}{2}q}|\,a^{\mu\nu}\,|^{-\frac{1}{2}}B_\beta^{-\frac{1}{2}q}e^{-Q_\beta}\cdot B'^{\frac{1}{2}q}_\beta. \tag{7.10}$$

Now the quantity

$$e^{B'^{ij}_\beta y_{ij}} = \sum_{\nu=0}^{\infty}\frac{(y_{ij}\bar{B}'^{ij}_\beta)^\nu}{\nu!}B'^{-\nu}_\beta,$$

where $\bar{B}^{ij}$ stands for the cofactor of B_{ij} in $\|\,B_{ij}\,\|$, can be expressed in terms of B'_β, provided we use our assumption that $B_{i\sigma} = 0$, $i \neq 1$, whereupon $y_{ij}B'^{ij}_\beta$ reduces to the single term yB'^{11}_β. In fact, we have

$$\begin{aligned} E[g_\beta|\,a^{rs}\,|^h] &= \bar{K}\prod_{i=1}^{t}\psi(N-m+t+1-i,2h)\,|\,a^{pq}\,|^h\pi^{\frac{1}{2}Nt}B_\beta^{-(\frac{1}{2}N+h)} \\ \times \exp(-\tilde{B}_{pq\beta}a^{pq}) &= \bar{K}\prod_{i=1}^{j}\psi\cdot\pi^{\frac{1}{2}Nt}|\,a^{pq}\,|^hB_\beta^{-\frac{1}{2}q}e^{-Q_\beta}\sum_{\nu=0}^{\infty}\frac{(y\bar{B}^{11}_\beta)^\nu}{\nu!}B_\beta^{-[\frac{1}{2}(N-q)+h+\nu]}, \end{aligned} \tag{7.11}$$

where, following the notation used by Wilks [13],

$$g_\beta = e^{-\beta_{rs}a^{rs}}, \qquad \bar{K} = \pi^{-\frac{1}{2}Nt}B^{\frac{1}{2}N}\exp(-\tilde{B}_{pq}a^{pq}),$$

$$\psi(a, b) = \frac{\Gamma[\frac{1}{2}(a+b)]}{\Gamma[\frac{1}{2}a]}.$$

And (7.11) can be written as

$$E[g_\beta \mid a^{rs}\mid^h] = \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi\cdot\mid a^{pq}\mid^h B_\beta^{-\frac{1}{2}q}e^{-Q_\beta} \times \sum_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}(B_{\beta u}^{-[\frac{1}{2}(N-q)+h]})_{u=0}, \tag{7.12}$$

where B_u stands for the result of replacing B_{11} by $B_{11} - u$. Changing $\beta_{r's'}$ into $\beta_{r's'} + \xi_{r'}\xi_{s'}$ and integrating, we then find that by virtue of (7.10)

$$E[g_\beta \mid a^{rs}\mid^h \mid a^{r's'}\mid^{-\frac{1}{2}}] = \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi\cdot\mid a^{pq}\mid^h\mid a^{\mu\nu}\mid^{-\frac{1}{2}}B_\beta^{-\frac{1}{2}q}e^{-Q_\beta} \times \pi^{-\frac{1}{2}t}\prod_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}\int B_{\beta u}^{\prime-[\frac{1}{2}(N-q)+h]}d\xi_i\Big]_{u=0}. \tag{7.13}$$

Now

$$\int B_{\beta u}^{\prime-[\frac{1}{2}(N-q)+h]}d\xi_i = B_{\beta u}^{-[\frac{1}{2}(N-1-q)+h]}\pi^{\frac{1}{2}t}\prod_{i=1}^{t}\psi(N-q+2h+1-i, -1),$$

so that (7.13) becomes

$$E[g_\beta \mid a^{rs}\mid^h\mid a^{r's'}\mid^{-\frac{1}{2}}] = \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi(N-m+t+1-i, 2h) \times \prod_{i=1}^{t}\psi(N-q+2h+1-i, -1)\mid a^{pq}\mid^h\mid a^{\mu\nu}\mid^{-\frac{1}{2}} \times B_\beta^{-\frac{1}{2}q}e^{-Q_\beta}\sum_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}(B_{\beta u}^{-[\frac{1}{2}(N-1-q)+h]})_{u=0}. \tag{7.14}$$

Comparing (7.14) with (7.12), and making use of the fact that

$$\psi(a, -1)\psi(1-1, -1)\cdots\psi(a-2h+1, -1) = \psi(a, -2h),$$

we thus have

$$E[g_\beta \mid a^{rs}\mid^h\mid a^{r's'}\mid^{-h}] = \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi(N-m+t+1-i, 2h) \times \prod_{i=1}^{t}\psi(N-q+2h+1-i, -2h)\mid a^{pq}\mid^h\mid a^{\mu\nu}\mid^{-h}B_\beta^{-\frac{1}{2}q}e^{-Q_\beta} \times \sum_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}[B_{\beta u}^{-\frac{1}{2}(N-q)}]_{u=0}.$$

Setting the β's equal to zero, performing the differentiation, and recalling the definitions of $\bar{K}$ and Q_β, we then find

$$E[(\lambda^{2/N})^h] = \prod_{i=1}^{t} \psi(N - m + t + 1 - i, 2h) \prod_{i=1}^{t} \psi(N - q + 2h + 1 - i, -2h) \tag{7.15}$$
$$\times e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]} \frac{\Gamma[\frac{1}{2}(N-q)+\nu]}{\Gamma[\frac{1}{2}(N-q)]}.$$

Taking the first factor from each product, we can convert (7.15) into

$$\prod_{i=2}^{t} \psi(N - m + t + 1 - i, 2h) \prod_{t=2}^{t} \psi(N - q + 2h + 1 - i, -2h)$$
$$\times e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{\Gamma[\frac{1}{2}(N-m+t)+h]}{\Gamma[\frac{1}{2}(N-m+t)]} \cdot \frac{\Gamma[\frac{1}{2}(N-q)+\nu]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}.$$

This last product of ratios of Γ's is equivalent to

$$\frac{\Gamma[\frac{1}{2}(N-q)+\nu]}{\Gamma[\frac{1}{2}(N-m+t)]\Gamma[\frac{1}{2}(m-t-q)+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(m-t-q)+\nu]\Gamma[\frac{1}{2}(N-m+t)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}.$$

Thus the moments of $\lambda^{2/N}$ are connected with an integral equation of type B [12] and $\lambda^{2/N}$ is distributed like the product

$$z \cdot \theta_2 \cdots \theta_t \qquad 0 \le z \le 1, 0 \le \theta_i \le 1,$$

where the joint distribution of the θ's is

$$f(\theta) = \prod_{i=2}^{t} \frac{\Gamma[\frac{1}{2}(N-q+1-i)]}{\Gamma[\frac{1}{2}(N-m+t+1-i)]\Gamma[\frac{1}{2}(m-t-q)]} \cdot \theta_i^{\frac{1}{2}(N-m+t+1-i)-1}(1-\theta_i)^{\frac{1}{2}(m-t-q)-1},$$

and z is distributed independently of the θ's with the distribution

$$F(z) = e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{z^{\frac{1}{2}(N-m+t)-1}(1-z)^{\frac{1}{2}(m-t-q)+\nu-1}}{B[\frac{1}{2}(N-m+t), \frac{1}{2}(m-t-q)+\nu]}. \tag{7.16}$$

The probability that $0 \le \lambda \le \lambda_\epsilon$ is therefore

$$I(y, \lambda_\epsilon) = \int_S f(\theta)F(z)\, dz\, d\theta_2 \cdots d\theta_t,$$

where S is the region $0 \le \theta_2 \cdots \theta_t, z < \lambda_\epsilon^{2/N}$. Putting $\varphi(\theta)$ for the upper limit of z in S for fixed θ, and S_θ for the projection of S into the θ space, we then have

$$I(y, \lambda_0) = \int_{S_\theta} f(\theta) \left\{ e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \int_0^\varphi \frac{z^{\frac{1}{2}(N-m+t)-1}(1-z)^{\frac{1}{2}(m-t-q)-1+\nu}}{B[\frac{1}{2}(N-m+t), \frac{1}{2}(m-t-q)+\nu]} dz \right\} d\theta.$$

If we replace z by $(1 - z)$ we then find

$$I(y, \lambda_0) = \int_{S_\theta} f(\theta) \tag{7.17}$$
$$\times \left\{ e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{B[\frac{1}{2}(m-t-q)+\nu, \frac{1}{2}(N-m+t); 1-\varphi]}{B[\frac{1}{2}(m-t-q)+\nu, \frac{1}{2}(N-m+t)]} \right\} d\theta.$$

As far as y is concerned, (7.17) is essentially the same as (2.8). The computation which was made there, together with the type of reasoning employed in the latter part of section 5 in connection with the independence test for several blocks, then shows that

$$\frac{\partial}{\partial y} I(y, \lambda_\epsilon) > 0 \qquad (0 < \epsilon < 1).$$

Remembering that

$$y = \tilde{a}^{\sigma\tau} B_{\sigma 1} B_{\tau 1},$$

we see that

$$\left(\frac{\partial y}{\partial B_{\sigma 1}}\right)_0 = 0, \qquad \frac{\partial^2 y}{\partial B_{\sigma 1}\, \partial B_{\tau 1}} = 2\tilde{a}^{\sigma\tau},$$

and we remark that the assumed positive definiteness of $\| a^{pq} \|$ implies that of $\| \tilde{a}^{\sigma\tau} \|$. Hence the relation

$$\left(\frac{\partial^2}{\partial B_{1\sigma}\, \partial B_{1\tau}} I(y, \lambda_\epsilon)\right)_0 = \left(\frac{\partial I}{\partial y}\right)_0 \tilde{a}^{\sigma\tau}$$

together with the fact that we could have obtained the analogue of (7.17) under the assumption

$$B_{i\sigma} = 0 \qquad i \neq i_0,$$

where i_0 is any fixed number in the set $1, \cdots, t$, shows that the matrix of second partial derivatives is positive definite when H is true.

Thus we have

THEOREM IV. *Let $x^1, \cdots, x^t$ be normally distributed about means which are linear functions of certain fixed variates $x^{t+1}, \cdots, x^m$. Then the likelihood ratio test for the hypothesis that the distribution of $x^1, \cdots, x^t$ depends only on a selected subset $x^{t+1}, \cdots, x^{t+q}$ of the fixed variates is locally unbiased.*

The result of this section has its most immediate application to those problems in the analysis of variance which require simultaneous consideration of several interrelated dependent variables $x^1, \cdots, x^t$ in conjunction with a given set of independent variables $x^{t+1}, \cdots, x^m$ [15]. For the usual hypothesis to be tested in this case is that $x^1, \cdots, x^t$ are jointly independent of, say, $x^{t+q+1}, \cdots, x^m$.

To return to the general case of (7.1), the method of this section can also be used to test the hypothesis that the regression coefficients referring to the x^σ have particular values, say

$$C^i_\sigma = C^i_{\sigma 0} \qquad i = 1, \cdots, t; \sigma = t + q + 1, \cdots, m,$$

the remaining C's and the B's being left unspecified. Since we have

$$x^i - C^i_\mu x^\mu - C^i_\sigma x^\sigma = x^i - C^i_\mu x^\mu - (C^i_\sigma - C^i_{\sigma 0})x^\sigma - C^i_{\sigma 0} x^\sigma,$$

by the device of replacing x^i_α by $x^i_\alpha - C^i_{\sigma 0}x^\sigma_\alpha$, we can reduce this problem to that of testing the hypothesis that

$$C'^i_\sigma = C^i_\sigma - C^i_{\sigma 0} = 0.$$

Similarly, the problem of testing whether the linear functions $u^i_\sigma = \alpha^\tau_\sigma C^i_\tau$ have specified values $u^i_{\sigma 0}$ comes under the same heading [7].

A particularly interesting case of the general regression problem is that in which $m = t + q + 1$, so that the null hypothesis H states that the chance variables x^i are independent of the fixed variate x^m, though they may depend upon $x^{t+1}, \cdots, x^{m-1}$. In this case we are able to find the exact distribution law of $\lambda^{2/N}$ without assuming that any of the regression coefficients C^i are zero. For the quantity

$$\sum_{\nu=0}^{\infty} \frac{(y_{ij}\bar{B}^{ij})^\nu}{\nu!} B_\beta^{-[\frac{1}{2}(N-q)+h+\nu]}, \tag{7.18}$$

which would have occurred in (7.11) had it not been for the restriction $B_{i\sigma} = 0$ $(i \neq 1)$, can now be expressed in terms of B_β even without this restriction. By definition

$$y_{ij}\bar{B}^{ij} = \bar{a}^{mm}\bar{B}^{ij}B_{mi}B_{mj}$$

and the vanishing of the B_{mi} is equivalent to the vanishing of the regression coefficients C^i_m associated with x^m. And since

$$| B_{ij} - ua^{mm}B_{mi}B_{mj} | = B - ua^{mm}\bar{B}^{ij}B_{mi}B_{mj},$$

we can write (7.18) in the form

$$\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]} \cdot \frac{\partial^\nu}{\partial u^\nu} [B_{\beta u}^{-[\frac{1}{2}(N-q)+h]}]_{u=0},$$

where

$$|| B_{\beta u} || = || B_{ij\beta} - ua^{mm}B_{mi}B_{mj} ||$$

is positive definite provided u is sufficiently small. Thus the moments of $\lambda^{2/N}$ can be found from (7.15) if we put $a^{mm}B^{ij}B_{mi}B_{mj} = y_{ij}B^{ij}$ in place of yB^{11}. Moreover, it can be seen that when the value $m = t + q + 1$ is substituted into (7.15), that expression reduces to

$$E[(\lambda^{2/N})^h] = e^{-y_{ij}B^{ij}} \sum_{\nu=0}^{\infty} \frac{(y_{ij}B^{ij})^\nu}{\nu!} \frac{B[\frac{1}{2}(N-m+1)+h, \frac{1}{2}(m-q-1)+\nu]}{B[\frac{1}{2}(N-m+1), \frac{1}{2}(m-q-1)+\nu]}$$

so that $\lambda^{2/N}$ is distributed like w, where

$$f(w) = e^{-y_{ij}B^{ij}} \sum_{\nu=0}^{\infty} \frac{(y_{ij}B^{ij})^\nu}{\nu!} \frac{w^{\frac{1}{2}(N-m+1)-1}(1-w)^{\frac{1}{2}(m-q-1)-1+\nu}}{B[\frac{1}{2}(N-m+1), \frac{1}{2}(m-q-1)+\nu]}. \tag{7.19}$$

The distribution law of $\lambda^{2/N}$ for this case is thus closely related to that obtained in the treatment of the regression problem with one dependent variate in section 2. Applying the argument used there, we can obtain:

THEOREM IVa. *The likelihood ratio test for the hypothesis that in a population of the type* (7.1) *the variates x^i are independent of x^m—the case $m = t + q + 1$ of Theorem IV—is completely unbiased.*

If we specialize the problem somewhat further, considering the case $q = 0$, $x_\alpha^m = 1$ (so that $m = t + 1$), we find that the likelihood ratio takes the form

$$\lambda^{2/N} = \frac{1}{1 + Nv_{ij}\bar{x}^i\bar{x}^j} = \frac{1}{1 + T},$$

where $v^{ij} = \sum_{\alpha=1}^{N} (x_\alpha^i - \bar{x}^i)(x_\alpha^j - \bar{x}^j)$, and T is Hotelling's generalization [5] of Student's ratio. In this case we are testing the hypothesis that the x^i are distributed with zero means. The exact distribution law of

$$T = \frac{1 - \lambda^{2/N}}{\lambda^{2/N}}$$

was recently published by P. L. Hsu [6], who obtained it in a very elegant fashion by means of the Laplace transform. He has also shown that the resulting test is *most powerful* in the sense that, of all critical regions S for which

$$P\{x \subset S\} = \epsilon + \tfrac{1}{2}\alpha B^{ij}b_i b_j + R(b)$$

(where ϵ and α are independent of the B^{ij} and of the means b_i, and R is an infinitesimal of at least the third order as all b_i tend to zero), the critical region defined by

$$S: \quad T \geq T_\epsilon$$

has the largest possible value of α. Tang's tables [11] make it evident that this largest possible value of α is actually positive and that the test is in fact unbiased for all values of the b's when $\epsilon = .05$ or $\epsilon = .01$. The results of this section may be used to show that this property extends to all probability levels other than $\epsilon = 0$ and $\epsilon = 1$.

The application of Hotelling's T is by no means confined to the above case. Other hypotheses which can be tested by means of this statistic are discussed by Hsu [6]. In addition it is now known that the Studentized D^2, devised by Mahalanobis for measuring the "distance" between two normal multivariate populations, is proportional to Hotelling's T. This fact is pointed out by R. C. Bose and N. Roy [1], who have obtained the exact distribution of D^2 for the case in which the two populations from which the samples are drawn are assumed to have the same matrix of variances and covariances, but are allowed to have different sets of means; their work, however, is quite independent of Hsu's. They also note that D^2 is proportional to the ratio which arises in Fisher's method of multiple measurements [4].

8. Summary. The method of likelihood-ratios is of practical as well as theoretical importance, because it provides a unified approach to the problem of testing statistical hypotheses. In this paper we have investigated many of the tests which this method yields when applied to hypotheses about sets of regression coefficients and covariances in normal populations. By studying the probability functions of the corresponding λ-criteria we are able to show that these tests are "good," in the sense that they are unbiased even for small samples.

Among the completely unbiased tests which can be based on the likelihood-ratio method, our discussion includes: the multiple correlation coefficient, with or without fixed variates [13]; Hotelling's generalized T test [6] and the statistically equivalent "Studentized D^2" [1]; the ordinary analysis of variance and covariance for orthogonal or non-orthogonal data [11, 16], as well as related tests of linear hypotheses in the case of one chance variable.

With respect to the analysis of variance for two or more variables [15] and certain other hypotheses regarding regression coefficients in multivariate populations, though there are indications that the tests are completely unbiased, we have succeeded in demonstrating this property only in the local sense.

Finally, the likelihood-ratio test for the hypothesis that the variates fall into certain specified mutually independent sets [14] is shown to be unbiased, at least locally, and has the additional property described in Theorem IIIa.

In conclusion, much more than a word of acknowledgment is due to Professor S. S. Wilks of Princeton University, to whom the writer is greatly indebted for advice and encouragement.

REFERENCES

[1] R. C. Bose and N. Roy, "The Distribution of the Studentized D^2-statistic," *Sankhyā*, Vol. 4 (1938–39), pp. 19–38.

[2] G. W. Brown, "On the Power of the L_1 Test for Equality of Several Variances," *Annals of Math. Stat.*, Vol. 10 (1939), p. 124.

[3] R. A. Fisher, "The General Sampling Distribution of the Multiple Correlation Coefficient," *Proc. Roy. Soc. Lond.*, series A, Vol. 121 (1928), pp. 196–203.

[4] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, Vol. 7 (1936), pp. 179–188.

[5] H. Hotelling, "The Generalization of Student's Ratio," *Annals of Math. Stat.*, Vol. 2 (1931), pp. 359–378.

[6] P. L. Hsu, "Notes on Hotelling's Generalized *T*," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 231–243.

[7] S. Kolodziecczyk, "On an Important Class of Statistical Hypotheses," *Biometrika*, Vol. 27 (1935), pp. 161–190.

[8] J. Neyman and E. S. Pearson, "On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference," *Biometrika*, Vol. 20A (1928), pp. 175–240.

[9] J. Neyman and E. S. Pearson, "The Testing of Statistical Hypotheses in Relation to Probabilities a Priori," *Proc. Camb. Phil. Soc.*, Vol. 29 (1933), p. 493.

[10] J. Neyman and E. S. Pearson, "Contributions to the Theory of Testing Statistical Hypotheses," *Stat. Res. Mem.*, Vol. 1 (1936), pp. 1–37, Vol. 2 (1938), pp. 25–57.

[11] P. C. Tang, "The Power Function of the Analysis of Variance Tests with Tables and Illustrations of Their Use," *Stat. Res. Mem.*, Vol. 2 (1938), pp. 126–149 + viii.

[12] S. S. Wilks, "Certain Generalizations in the Analysis of Variance," *Biometrika*, Vol. 24 (1932), pp. 471–494.
[13] S. S. Wilks, "Moment-generating Operators for Determinants of Product Moments in Samples from a Norm: ʌ System," *Ann. of Math.*, Vol. 35 (1934), pp. 312–340.
[14] S. S. Wilks, "On the Independence of k Sets of Normally Distributed Statistical Variables," *Econometrica*, Vol. 3 (1935), pp. 309–326.
[15] S. S. Wilks, "The Analysis of Variance for Two or More Variables," *Report of Third Annual Research Conference on Economics and Statistics*, (1937), pp. 82–85.
[16] S. S. Wilks, "The Analysis of Variance and Covariance in Non-Orthogonal Data," *Metron*, Vol. XIII (1938), pp. 141–154.
[17] J. Wishart, "The Generalized Product Moment Distribution in a Normal Distribution," *Biometrika*, Vol. 20A (1928), pp. 32–52.

Princeton University,
Princeton, N. J.