ON CLASSIFICATION BY THE STATISTICS R AND Z

By S. John

(Received May 28, 1962)

1. Introduction

$$W=(\bar{x}^{(2)}-\bar{x}^{(1)})S^{-1}x'-\tfrac{1}{2}(\bar{x}^{(2)}-\bar{x}^{(1)})S^{-1}(\bar{x}^{(1)}+\bar{x}^{(2)})',$$

$$R=\left(x-\frac{N_1}{N_1+N_2}\bar{x}^{(1)}-\frac{N_2}{N_1+N_2}\bar{x}^{(2)}\right)S^{-1}\left(x-\frac{N_1}{N_1+N_2}\bar{x}^{(1)}-\frac{N_2}{N_1+N_2}\bar{x}^{(2)}\right)'$$
$$-d\left[\frac{1+(N_1+N_2)^{-1}}{N_1^{-1}+N_2^{-1}}\right]^{\frac{1}{2}}(\bar{x}^{(2)}-\bar{x}^{(1)})S^{-1}\left(x-\frac{N_1}{N_1+N_2}\bar{x}^{(1)}-\frac{N_2}{N_1+N_2}\bar{x}^{(2)}\right)',$$

$$Z=\frac{N_1}{N_1+1}(x-\bar{x}^{(1)})S^{-1}(x-\bar{x}^{(1)})'-\eta\frac{N_2}{N_2+1}(x-\bar{x}^{(2)})S^{-1}(x-\bar{x}^{(2)})',$$

where $\bar{x}^{(k)}(k=1,2)$ is the mean of a sample of size N_k from population $P^{(k)}$, S an estimate of the variance-covariance matrix Σ of $P^{(1)}$ and $P^{(2)}$, η a constant and

$$d=2\left[\frac{1+(N_1+N_2)^{-1}}{N_1^{-1}+N_2^{-1}}\right]^{\frac{1}{2}}\Big/\frac{N_2-N_1}{N_1+N_2},$$

are three criteria that have been proposed for deciding the population to which an individual, with measurement x, known to belong to $P^{(1)}$ or $P^{(2)}$, really belongs. If $P^{(1)}$ and $P^{(2)}$ are normal, W is the statistic obtained by replacing the parameters in the logarithm of the likelihood-ratio by estimates of them from random samples; see Anderson (1951). The statistic

$$\left(x-\frac{N_1}{N_1+N_2}\bar{x}^{(1)}-\frac{N_2}{N_1+N_2}\bar{x}^{(2)}\right)\Sigma^{-1}\left(x-\frac{N_1}{N_1+N_2}\bar{x}^{(1)}-\frac{N_2}{N_1+N_2}\bar{x}^{(2)}\right)'$$
$$-d\left[\frac{1+(N_1+N_2)^{-1}}{N_1^{-1}+N_2^{-1}}\right]^{\frac{1}{2}}(\bar{x}^{(2)}-\bar{x}^{(1)})\Sigma^{-1}\left(x-\frac{N_1}{N_1+N_2}\bar{x}^{(1)}-\frac{N_2}{N_1+N_2}\bar{x}^{(2)}\right)'$$

was proposed by Rao (1954). He showed that it is best for discriminating between alternatives that are close to each other. The statistic Z is equivalent to a statistic derived by Anderson (1958, p. 142).

[Z with $\eta=1$ was considered by John (1960). It was proposed there because it appeared reasonable to give the individual to that population which, on testing, rejects it at a higher level*. Note that if $\eta=1$ the two terms, whose difference Z is, are the criteria used for the tests. We

* Rao (1954, p. 655) had previously stated this principle.

wish also to mention that A. Kudo (1959) has shown that the procedure of assigning the individual to $P^{(1)}$ or $P^{(2)}$ according as

$$\frac{N_1}{N_1+1}(x-\bar{x}^{(1)})\Sigma^{-1}(x-\bar{x}^{(1)})' - \frac{N_2}{N_2+1}(x-\bar{x}^{(2)})\Sigma^{-1}(x-\bar{x}^{(2)})' \gtreqless 0$$

is the best of all two-decision rules invariant under translation and rotation of axes.

The three statistics W, R and Z (with $\eta=1$) are asymptotically equivalent. W and R are equivalent if $N_1=N_2$. Z is equivalent to W, if $N_1/(N_1+1)=\eta\ N_2/(N_2+1)$.]

If W, R or Z is used to classify individuals, the probability (given $\bar{x}^{(1)}$, $\bar{x}^{(2)}$ and S) of assigning an individual to $P^{(1)}$ or $P^{(2)}$ depends on the realised S, $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$. John (1961a, 1962) gave the distributions and expected values of these probabilities in the case of W. With R and Z the problems are more difficult. The present paper will indicate what results have been obtained.

2. Notation

A glossary of the symbols used is given below. Notations introduced in the previous section are repeated for the sake of completeness.

$P^{(k)}(k=1, 2)$: the two parent populations (assumed p-variate normal) claiming the individual to be classified.

P: a third population (also assumed p-variate normal).

p: the number of characters used.

$x=(x_1, x_2, \cdots, x_p)$: the vector of measurements on the individual to be classified.

$x \in P$ means that x is the vector of measurements on an individual from P (etc).

$\mu^{(k)}=(\mu_1^{(k)}, \mu_2^{(k)}, \cdots, \mu_p^{(k)})$: the mean of $P^{(k)}$ $(k=1, 2)$.

$\mu=(\mu_1, \mu_2, \cdots, \mu_p)$: the mean of P.

$\Sigma=(\sigma_{ij})$: the variance-covariance matrix of $P^{(1)}$, $P^{(2)}$ and P.

$\delta=[(\mu^{(2)}-\mu^{(1)})\Sigma^{-1}(\mu^{(2)}-\mu^{(1)})']^{1/2}$.

$\bar{x}^{(k)}$: the arithmetic mean of the observations in a random sample of size N_k from $P^{(k)}$ $(k=1, 2)$.

S: an unbiased estimate of Σ, distributed independently of $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, and following the Wishart law with n degrees of freedom.

N_k: size of the sample from $P^{(k)}$ $(k=1, 2)$.

n: the degree of freedom of S.

$a_1 = N_1/(N_1+N_2)$; $a_2 = N_2/(N_1+N_2)$.

$a_3 = N_2^{-1}+N_1^{-1}$; $a_4 = 1+(N_1+N_2)^{-1}$.

$a_5 = N_1/(N_1+1)$; $a_6 = N_2/(N_2+1)$.

$d = 2(a_4/a_3)^{1/2}/(a_2-a_1)$.

$W = (\bar{x}^{(2)}-\bar{x}^{(1)})S^{-1}x' - \frac{1}{2}(\bar{x}^{(2)}-\bar{x}^{(1)})S^{-1}(\bar{x}^{(1)}+\bar{x}^{(2)})'$

$R = (x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})S^{-1}(x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})'$
$\quad - d(a_4/a_3)^{1/2}(\bar{x}^{(2)}-\bar{x}^{(1)})S^{-1}(x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})'$.

$Z = a_5(x-\bar{x}^{(1)})S^{-1}(x-\bar{x}^{(1)})' - \gamma a_6(x-\bar{x}^{(2)})S^{-1}(x-\bar{x}^{(2)})'$.

W_0, R_0, Z_0: respectively the same as W, R and Z with Σ substituted for S.

$Q = (\bar{x}^{(2)}-\bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(2)}-\bar{x}^{(1)})'$.

$T = (a_1\mu^{(1)}+a_2\mu^{(2)}-\mu)\Sigma^{-1}(\bar{x}^{(2)}-\bar{x}^{(1)})'/C_1^{1/2}$.

$C_1 = (a_1\mu^{(1)}+a_2\mu^{(2)}-\mu)\Sigma^{-1}(a_1\mu^{(1)}+a_2\mu^{(2)}-\mu)'$.

$C_2 = (a_1\mu^{(1)}+a_2\mu^{(2)}-\mu)\Sigma^{-1}(\mu^{(2)}-\mu^{(1)})'$.

$C_3 = \delta^2 - C_2^2/C_1$; $C_4 = C_2/C_1^{1/2}$; $C_5 = C_3^{1/2}$.

$F(Q, T)$: the joint density of Q and T.

$$I_x(r, s) = \frac{\Gamma(r+s)}{\Gamma(r)\,\Gamma(s)}\int_0^x z^{r-1}(1-z)^{s-1}\,dz.$$

$$\Delta_p(\chi^2 ; \lambda) = \frac{e^{-\lambda}}{2^{(1/2)p}}\sum_{r=0}^{\infty}\frac{(\frac{1}{2}\lambda)^r}{\Gamma(\frac{1}{2}p+r)r\,!}(\chi^2)^{(1/2)p+r-1}e^{-(1/2)\chi^2}.$$

A random variable having $\Delta_p(\chi^2 ; \lambda)$ as the density function of its distribution will be spoken of as a non-central chi-square with p degrees of freedom and non-centrality λ.

$L_p(\alpha, z)$: the value of λ satisfying the equation $\int_\alpha^\infty \Delta_p(\chi^2 ; \lambda)d\chi^2 = z$.

$e_k'(\bar{x}^{(1)}, \bar{x}^{(2)} ; S)$: the probability, given $\bar{x}^{(1)}$, $\bar{x}^{(2)}$ and S, of assigning an individual from P to $P^{(k)}(k=1, 2)$, if individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according as $R \gtreqless c$.

$e_k''(\bar{x}^{(1)}, \bar{x}^{(2)} ; S)$: the probability, given $\bar{x}^{(1)}$, $\bar{x}^{(2)}$ and S, of assigning an individual from P, to $P^{(k)}(k=1, 2)$, if individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according as $Z \gtreqless c$.

$e_k'(\bar{x}^{(1)}, \bar{x}^{(2)}) = e_k'(\bar{x}^{(1)}, \bar{x}^{(2)} ; \Sigma)$ $\qquad (k=1, 2)$.

$e_k''(\bar{x}^{(1)}, \bar{x}^{(2)}) = e_k''(\bar{x}^{(1)}, \bar{x}^{(2)} ; \Sigma)$ $\qquad (k=1, 2)$.

The symbol for a random variable preceded by E denotes its expected value.

Let $Y=(y_{ij})$ be a $m\times\nu$ random matrix whose elements are independent normal variables with unit variance. Let $Ey_{ij} = \theta_{ij}$

$(i=1, 2, \cdots, m; j=1, 2, \cdots, \nu)$. Let $\Lambda = \left(\sum_{r=1}^{\nu} \theta_{ir} \theta_{jr}\right)$. We shall call any random matrix having the same distribution as $Y\ Y'$ a non-central Wishart matrix of order m, degree of freedom ν and non-centrality $\frac{1}{2}\Lambda$. In case Λ is the null matrix, we shall call the random matrix simply a Wishart matrix of order m and degree of freedom ν.

3. Distribution of the probability when the statistic is R

If $N_1 = N_2$, R is equivalent to W. In what follows we shall, therefore, assume that $N_1 \neq N_2$. Whether individuals assigned to $P^{(1)}$ should be those for whom $R_0 \gtreqless c$ depends on whether $N_1 \gtreqless N_2$. For the sake of definiteness, we assume that $N_1 > N_2$. Suppose that the individuals assigned to $P^{(1)}$ are those with $R_0 \geqq c$. We denote the probability, given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, of assigning an individual from P to $P^{(i)}$ by $e_i'(\bar{x}^{(1)}, \bar{x}^{(2)})$. Since $e_1'(\bar{x}^{(1)}, \bar{x}^{(2)}) = 1 - e_2'(\bar{x}^{(1)}, \bar{x}^{(2)})$, we consider only $e_2'(\bar{x}^{(1)}, \bar{x}^{(2)})$.

The inequality $R_0 \geqq c$ is equivalent to the inequality

$$V \geqq c + \tfrac{1}{4} a_1 d^2 Q / a_2, \tag{3.1}$$

where

$$V = [x - a_1 \bar{x}^{(1)} - a_2 \bar{x}^{(2)} - \tfrac{1}{2} d (a_1/a_2)^{1/2} (\bar{x}^{(2)} - \bar{x}^{(1)})] \Sigma^{-1} \\ [x - a_1 \bar{x}^{(1)} - a_2 \bar{x}^{(2)} - \tfrac{1}{2} d (a_1/a_2)^{1/2} (\bar{x}^{(2)} - \bar{x}^{(1)})]'. \tag{3.2}$$

Given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, V is distributed as a non-central chi-square with p degrees of freedom and non-centrality

$$\tfrac{1}{2}[\mu - a_1 \bar{x}^{(1)} - a_2 \bar{x}^{(2)} - \tfrac{1}{2} d (a_1/a_2)^{1/2} (\bar{x}^{(2)} - \bar{x}^{(1)})] \Sigma^{-1} \\ [\mu - a_1 \bar{x}^{(1)} - a_2 \bar{x}^{(2)} - \tfrac{1}{2} d (a_1/a_2)^{1/2} (\bar{x}^{(2)} - \bar{x}^{(1)})]' = V' \text{ (say)}. \tag{3.3}$$

Therefore

$$e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}) = \int_{c+(1/4)a_1 d^2 Q/a_2}^{\infty} J_p(\chi^2; V') d\chi^2. \tag{3.4}$$

Hence,

$$\Pr[e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = \Pr[V' < L_p(c + \tfrac{1}{4} a_1 d^2 Q / a_2, z)]. \tag{3.5}$$

To find $\Pr[V' < L_p(c + \frac{1}{4} a_1 d^2 Q / a_2, z)]$, we first find

$$\Pr[V' < L_p(c + \tfrac{1}{4} a_1 d^2 Q / a_2, z) | \bar{x}^{(2)} - \bar{x}^{(1)}].$$

Given $\bar{x}^{(2)} - \bar{x}^{(1)}$, $2(N_1 + N_2) V'$ is distributed as a non-central chi-square with p degrees of freedom and non-centrality

$$\tfrac{1}{2}(N_1 + N_2)[\tfrac{1}{4}(a_2/a_1) d^2 Q + d (C_1 a_1/a_2)^{1/2} T + C_1] = V_0'' \text{ (say)}. \tag{3.6}$$

Therefore,

$$\Pr[V' < L_p(c+\tfrac{1}{4}a_4 d^2 Q/a_3, z) | \bar{x}^{(2)} - \bar{x}^{(1)}] = \int_{0 \le \chi^2 \le 2(N_1+N_2) L_p(c+(1/4)a_4 d^2 Q/a_3, z)} \Delta_p(\chi^2 ; V'') d\chi^2, \tag{3.7}$$

a function of z, Q and T, which we shall denote by the symbol Ω (z; Q, T). Let $F(Q, T)$ denote the joint density of Q and T. Then

$$\Pr[V' < L_p(c+\tfrac{1}{4}a_4 d^2 Q/a_3, z)] = \iint \Omega(z ; Q, T) F(Q, T) dQ dT, \tag{3.8}$$

where the domain of integration is the entire domain of variation of Q and T. That is,

$$\Pr[e_1'(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = \iint \Omega(z ; Q, T) F(Q, T) dQ dT. \tag{3.9}$$

The joint density of Q and T was required in the solution of problems connected with W, and will again be required when we consider the statistic Z. John (1962) shows that

$$F(Q, T) = \frac{(Q-T^2)^{(1/2)(p-3)}}{(2a_3)^{(1/2)p}\sqrt{\pi}} \exp\left[-\frac{1}{2a_3}(Q - 2C_1 T + \delta^2)\right]$$
$$\sum_{k=0}^{\infty} \frac{(\frac{1}{2}C_1/a_3)^{2k}(Q-T^2)^k}{k!\,\Gamma(\frac{1}{2}p+k-\frac{1}{2})} \, (-Q^{1/2} \leqq T \leqq Q^{1/2}), \tag{3.10}$$

if $C_1 \neq 0$; if $C_1 = 0$, Ω (z; Q, T) does not involve T, and, therefore, in equation (3.9), $F(Q, T)$ may be replaced by the density function of Q; the density function of Q is $a_3^{-1}\Delta_p(Q/a_3 ; \frac{1}{2}\delta^2/a_3)$.

4. Distribution of the probability when the statistic is Z

Suppose individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according as $Z_0 \gtreqless c$. We denote the probability, given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, of assigning the individuals to $P^{(k)} (k=1, 2)$ by $e_k''(\bar{x}^{(1)}, \bar{x}^{(2)})$. Since $e_1''(\bar{x}^{(1)}, \bar{x}^{(2)}) = 1 - e_2''(\bar{x}^{(1)}, \bar{x}^{(2)})$, we shall consider only $e_2''(\bar{x}^{(1)}, \bar{x}^{(2)})$.

The inequality $Z_0 \geqq c$ is equivalent to

$$[x - \bar{x}^{(1)} + \alpha(\bar{x}^{(2)} - \bar{x}^{(1)})] \Sigma^{-1} [x - \bar{x}^{(1)} + \alpha(\bar{x}^{(2)} - \bar{x}^{(1)})]'$$
$$\gtreqless \alpha(\alpha+1)Q + c' \tag{4.1}$$

according as $a_2 \gtrless \eta a_1$;

$$\alpha = \frac{N_2(N_1+1)\eta}{N_1(N_2+1) - N_2(N_1+1)\eta} ; \tag{4.2}$$

$$c' = \frac{c(N_1+1)(N_2+1)}{N_1(N_2+1) - N_2(N_1+1)\eta} . \tag{4.3}$$

For the sake of definiteness, we shall assume that $a_2 > \eta a_4$. (If $a_2 = \eta a_4$, Z_4 is equivalent to W_4.) Then $e_2''(\bar{x}^{(1)}, \bar{x}^{(2)})$ is the probability, given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, of the inequality (4.1) with the upper inequality sign being satisfied. It is easily seen to be equal to

$$\int_{\alpha(\alpha+1)Q+c'}^{\infty} \Delta_p(\chi^2 ; v) d\chi^2, \tag{4.4}$$

where

$$v = \tfrac{1}{2}[\mu - \bar{x}^{(1)} + \alpha(\bar{x}^{(2)} - \bar{x}^{(1)})] \Sigma^{-1} [\mu - \bar{x}^{(1)} + \alpha(\bar{x}^{(2)} - \bar{x}^{(1)})]'. \tag{4.5}$$

By methods similar to those of the previous section we can show that

$$\Pr[e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = \iint \Omega'(z ;\ Q, T) F(Q, T) dQ dT, \tag{4.6}$$

where

$$\Omega'(z ; Q, T) = \int_{0 \leq \chi^2 \leq 2(N_1+N_2) L_p(\alpha(\alpha+1)Q+c',\ z)} \Delta_p(\chi^2 ; v') d\chi^2 ; \tag{4.7}$$

$$v' = \tfrac{1}{2}(N_1 + N_2)[\beta^2 Q - 2\beta C_1^{1/2} T + C_1] ; \tag{4.8}$$

$$\beta = \alpha + a_2 . \tag{4.9}$$

If $C_1 = 0$, Ω' does not involve T, and, therefore, in equation (4.6), $F(Q, T)$ may be replaced by the density function of Q; i.e., by $a_3^{-1} \Delta_p(Q/a_3 ; \tfrac{1}{2}\delta^2/a_3)$.

5. Expected values

If individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according as $R_4 \lesseqgtr c (Z_4 \lesseqgtr c)$, the expected probability of assigning the individual to $P^{(2)}$ is clearly equal to the integral of the density function of $R_4(Z_4)$ from c to ∞. The density function of R_4 for $x \in P^{(1)}$, together with that of W_4 and another statistic, was given by John (1960). They all have the form

$$\exp[-\lambda_1 - \lambda_2 - \tfrac{1}{4}(b_1 - b_2)\theta] \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\lambda_1^r \lambda_2^s}{r!\, s!} \frac{b_1^{(1/2)p+r} b_2^{(1/2)p+s}}{\Gamma(\frac{1}{2}p + r)} \left(\frac{\theta}{2b_1 + 2b_2} \right)^{1/2(p+r+s)} \theta^{-1} W_{l,m}(\tfrac{1}{2}[b_1 + b_2]\theta) \tag{5.1}$$

for $\theta \geq 0$ and the form

$$\exp[-\lambda_1 - \lambda_2 - \tfrac{1}{4}(b_1 - b_2)\theta] \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\lambda_1^r \lambda_2^s}{r!\, s!} \frac{b_1^{(1/2)p+r} b_2^{(1/2)p+s}}{\Gamma(\frac{1}{2}p + s)} \left(\frac{-\theta}{2b_1 + 2b_2} \right)^{1/2(p+r+s)} \theta^{-1} W_{-l,m}(-\tfrac{1}{2}[b_1 + b_2]\theta) \tag{5.2}$$

for $\theta < 0$. Here

$$l=\tfrac{1}{2}(r-s),\quad m=\tfrac{1}{2}(r+s+p-1)\,; \tag{5.3}$$

$W_{l,m}$ is Whittaker's confluent hypergeometric function defined by the equation

$$W_{l,m}(\theta)=\frac{\theta^{m+(1/2)}}{\Gamma(m-l+\frac{1}{2})}e^{-(1/2)\theta}\int_0^\infty e^{-\alpha\theta}\alpha^{m-l-(1/2)}(1+\alpha)^{m+l-(1/2)}d\alpha.$$

The density functions of W_0, R_0 and Z_0 retain the form given above even if $x\in P$. If in (5.1) and (5.2) we take $\theta=R_0$,

$$\lambda_1=\frac{1}{2}\Big[\frac{\{\sqrt{(1+d^2)}+1\}^{1/2}}{(1+d^2)^{1/4}(2a_1)^{1/2}}(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})$$
$$-\frac{d\{\sqrt{(1+d^2)}+1\}^{-1/2}}{(1+d^2)^{1/4}(2a_2)^{1/2}}(\mu^{(2)}-\mu^{(1)})\Big]\Sigma^{-1}$$
$$\Big[\frac{\{\sqrt{(1+d^2)}+1\}^{1/2}}{(1+d^2)^{1/4}(2a_1)^{1/2}}(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})$$
$$-\frac{d\{\sqrt{(1+d^2)}+1\}^{-1/2}}{(1+d^2)^{1/4}(2a_2)^{1/2}}(\mu^{(2)}-\mu^{(1)})\Big]'. \tag{5.4}$$

$$\lambda_2=\frac{1}{2}\Big[\frac{\{\sqrt{(1+d^2)}-1\}^{1/2}}{(1+d^2)^{1/4}(2a_1)^{1/2}}(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})$$
$$+\frac{d\{\sqrt{(1+d^2)}-1\}^{-1/2}}{(1+d^2)^{1/4}(2a_2)^{1/2}}(\mu^{(2)}-\mu^{(1)})\Big]\Sigma^{-1}$$
$$\Big[\frac{\{\sqrt{(1+d^2)}-1\}^{1/2}}{(1+d^2)^{1/4}(2a_1)^{1/2}}(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})$$
$$+\frac{d\{\sqrt{(1+d^2)}-1\}^{-1/2}}{(1+d^2)^{1/4}(2a_2)^{1/2}}(\mu^{(2)}-\mu^{(1)})\Big]'. \tag{5.5}$$

$$b_1=(2/a_1)/[\sqrt{(1+d^2)}+1],\quad b_2=(2/a_1)/[\sqrt{(1+d^2)}-1], \tag{5.6}$$

we shall get the density function of R_0 (for $x\in P$). To get the density function of Z_0 (for $x\in P$) we should take, in (5.1) and (5.2), $\theta=Z_0$

$$\lambda_1=\tfrac{1}{2}[\{a_0^{1/2}(\gamma_1+\gamma_2)(\mu-\mu^{(1)})+(\eta a_0)^{1/2}(\gamma_1-\gamma_2)(\mu-\mu^{(2)})\}\Sigma^{-1}$$
$$\{a_0^{1/2}(\gamma_1+\gamma_2)(\mu-\mu^{(1)})+(\eta a_0)^{1/2}(\gamma_1-\gamma_2)(\mu-\mu^{(2)})\}']\div(2\gamma_1\gamma_2-2\eta+2), \tag{5.7}$$
$$\lambda_2=\tfrac{1}{2}[\{a_0^{1/2}(\gamma_1-\gamma_2)(\mu-\mu^{(1)})+(\eta a_0)^{1/2}(\gamma_1+\gamma_2)(\mu-\mu^{(2)})\}\Sigma^{-1}$$
$$\{a_0^{1/2}(\gamma_1-\gamma_2)(\mu-\mu^{(1)})+(\eta a_0)^{1/2}(\gamma_1+\gamma_2)(\mu-\mu^{(2)})\}']\div(2\gamma_1\gamma_2+2\eta-2), \tag{5.8}$$
$$b_1=2\{[(1+\eta)^2-4a_1a_2\eta]^{1/2}+1-\eta\}^{-1}, \tag{5.9}$$

and

$$b_2=2\{[(1+\eta)^2-4a_1a_2\eta]^{1/2}+\eta-1\}^{-1}; \tag{5.10}$$
$$\gamma_1=[1+\eta-2(\eta a_1a_2)^{1/2}]^{1/2};\quad \gamma_2=[1+\eta+2(\eta a_1a_2)^{1/2}]^{1/2}. \tag{5.11}$$

John (1961a, 1962) has given an expression of the form

$$e^{-(\lambda_1+\lambda_2)}\Big[\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\lambda_1^r\lambda_2^s}{r!\,s!}\Big\{1-I_{b_1/(b_1+b_2)}\Big(\frac{1}{2}p+r,\ \frac{1}{2}p+s\Big)\Big\}$$
$$+\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\lambda_1^r\lambda_2^s}{r!\,s!}\,I_{b_2/(b_1+b_2)}\Big(\frac{1}{2}p+s,\ \frac{1}{2}p+r\Big)\Big] \tag{5.12}$$

for $\Pr(W_0\geq 0)$ and has shown that it is approximately equal to

$$\int_{-\infty}^{a}\frac{1}{\sqrt{2\pi}}e^{-(1/2)t^2}dt, \tag{5.13}$$

where

$$a=\frac{2[b_1^{1/3}r_2^{-2/3}(1+\theta_2)-b_2^{1/3}r_1^{-2/3}(1+\theta_1)]+9[(r_1b_2)^{1/3}-(r_2b_1)^{1/3}]}{(18)^{1/2}[b_1^{2/3}r_2^{-4/3}(1+\theta_2)+b_2^{2/3}r_1^{-4/3}(1+\theta_1)]^{1/2}}\ ; \tag{5.14}$$

$$r_i=p+2\lambda_i\ \ (i=1,2)\ ; \tag{5.15}$$

$$\theta_i=2\lambda_i/(p+2\lambda_i)\ \ (i=1,2). \tag{5.16}$$

Since the density functions of R_0 and Z_0 are the same as that of W_0 in form, these same expressions will be equal to $\Pr(R_0>0)$ or $\Pr(Z_0>0)$ according as λ_1, λ_2, b_1 and b_2 are as in equations (5.4) to (5.6) or as in equations (5.7) to (5.10)*.

We shall now give some results regarding R and Z. John (1961b) shows that R is distributed as

$$(na_1/\chi^2)[B_{11}-d\{B_{12}-(n-p+2)^{-1/2}t|B|^{1/2}\}], \tag{5.17}$$

where

$$B=\begin{pmatrix}B_{11} & B_{12}\\ B_{12} & B_{22}\end{pmatrix},$$

t and χ^2 are independent random variables distributed as follows: B has the non-central Wishart distribution with p degrees of freedom and, if we let

$$\frac{1}{2}\Lambda=\frac{1}{2}\begin{pmatrix}\lambda_{11} & \lambda_{12}\\ \lambda_{12} & \lambda_{22}\end{pmatrix}$$

denote its non-centrality matrix, with

$$\lambda_{11}=(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})\Sigma^{-1}(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})'/a_1, \tag{5.18}$$
$$\lambda_{12}=(\mu-a_1\mu^{(1)}-a_2\mu^{(2)})\Sigma^{-1}(\mu^{(2)}-\mu^{(1)})'/(a_2a_1)^{1/2}\ \text{and}\ \lambda_{22}=\delta^2/a_2\ ;$$

t follows Student's law with $n-p+2$ degrees of freedom; χ^2 has the

* In my 1960 paper (1960a), another expression is given for $\Pr[Z_0>0]$ for $x\in P^{(1)}$. A similar expression can be given for $x\in P$ also, but will be less simple than what is given here.

chi-square distribution with $n-p+1$ degrees of freedom*. From this it follows that

$$\begin{aligned} E\, e_1'(\bar{x}^{(1)}, \bar{x}^{(2)}; S) \\ = E \Pr[\chi^2 \lesseqgtr (na_4/c)\{B_{11} - dB_{12} + d(n-p+2)^{-1/2}|B|^{1/2}t\}] \end{aligned} \tag{5.19}$$

according as $c \gtrless 0$. If $c=0$, we can write

$$E\, e_1'(\bar{x}^{(1)}, \bar{x}^{(2)}; S) = E \Pr[t > (n-p+2)^{1/2}|B|^{-1/2}\{B_{12} - B_{11}/d\}]. \tag{5.20}$$

These results will facilitate the evaluation of $E\, e_1'(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ by Monte Carlo methods. Similarly, it has been shown there that Z has the distribution of

$$(n/\chi^2)[(1-\eta)B_{11} + 2(\eta - a_3a_4\eta)^{1/2}\{B_{12} - (n-p+2)^{-1/2}|B|^{1/2}t\}], \tag{5.21}$$

where B_{11}, B_{12}, B_{22}, t and χ^2 have the same joint distribution as in the case of R except that

$$\begin{aligned} \lambda_{11} = \{[a_5^{1/2}(\mu-\mu^{(1)}) - (\eta a_6)^{1/2}(\mu-\mu^{(2)})]\Sigma^{-1} \\ [a_5^{1/2}(\mu-\mu^{(1)}) - (\eta a_6)^{1/2}(\mu-\mu^{(2)})]'\} \div [1+\eta-2(\eta a_5 a_6)^{1/2}], \end{aligned} \tag{5.22}$$

$$\begin{aligned} \lambda_{12} = \{[a_5^{1/2}(\mu-\mu^{(1)}) - (\eta a_6)^{1/2}(\mu-\mu^{(2)})]\Sigma^{-1} \\ [\beta_1(\mu-\mu^{(1)}) + \beta_2(\mu-\mu^{(2)})]'\} \div [1+\eta-2(\eta a_5 a_6)^{1/2}]^{1/2}, \end{aligned} \tag{5.23}$$

$$\lambda_{22} = [\beta_1(\mu-\mu^{(1)}) + \beta_2(\mu-\mu^{(2)})]\Sigma^{-1}[\beta_1(\mu-\mu^{(1)}) + \beta_2(\mu-\mu^{(2)})]'; \tag{5.24}$$

$$\beta_1 = \frac{a_5^{1/2}[2\eta - 2(a_3a_4\eta)^{1/2}]}{2(\eta - a_3a_4\eta)^{1/2}[1+\eta-2(a_3a_4\eta)^{1/2}]^{1/2}}; \tag{5.25}$$

$$\beta_2 = \frac{(\eta a_6)^{1/2}[2 - 2(\eta a_3 a_4)^{1/2}]}{2(\eta - a_3a_4\eta)^{1/2}[1+\eta-2(a_4a_4\eta)^{1/2}]^{1/2}}. \tag{5.26}$$

Hence it follows that

$$\begin{aligned} E\, e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}; S) \\ = E \Pr[\chi^2 \lesseqgtr (n/c)(1-\eta)B_{11} + 2(n/c)(\eta - a_3a_4\eta)^{1/2} \\ \{B_{12} - (n-p+2)^{-1/2}|B|^{1/2}t\}] \end{aligned} \tag{5.27}$$

according as $c \gtrless 0$. If $c=0$, we can write

$$\begin{aligned} E\, e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}; S) \\ = E \Pr[t < (n-p+2)^{1/2}|B|^{-1/2}\{B_{12} + \tfrac{1}{2}(\eta - a_3a_4\eta)^{-1/2}(1-\eta)B_{11}\}]. \end{aligned} \tag{5.28}$$

INDIAN STATISTICAL INSTITUTE, CALCUTTA

---

* This result is similar to that of Bowker (1960) for *W*. Some implications of this representation of *R* and that of *Z*, given below, will be developed elsewhere.

REFERENCES

[1] T. W. Anderson, "Classification by multivariate analysis" *Psychometrika*, Vol. 16 (1951), pp. 31-50.

[2] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley and Sons, (1958).

[3] Albert H. Bowker, "A representation of Hotelling's T^2 and Anderson's classification statistic W in terms of simple statistics", *Contributions to Probability and Statistics* 142-150. Stanford: Stanford University Press (1960).

[4] S. John, "On some classification problems—I, *Sankhyā*, Vol. 22 (1960a), pp. 301-308.

[5] S. John, "On some classification statistics," *Sankhyā*, Vol. 22 (1960b), pp. 309-316

[6] S. John, "Errors in discrimination," *Ann. Math. Stat.*, Vol. 32 (1961a), pp. 1125-1144.

[7] S. John, *Distribution Theory of Some Classification Statistics* (Unpublished) (1961b).

[8] S. John, "Further results on classification by W" (Submitted for publication in the *Sankhyā*), (1962).

[9] Akio Kudo, "The classificatory problem viewed as a two-decision problem," *Memoirs of the Faculty of Science, Kyushu University*, Series A, Vol. 13 (1959), pp. 96-125.

[10] C. Radhikrishna Rao, "A general theory of discrimination when the information on alternative hypotheses is based on samples," *Ann. Math. Stat.*, Vol. 25, (1954), pp. 651-670.