DISTRIBUTION THEORY OF SOME CLASSIFICATION STATISTICS

by

S. JOHN

Indian Statistical Institute

Calcutta

[Thesis submitted for the award of the Ph. D. degree of the Indian Statistical Institute]

ACKNOWLEDGEMENT

The research leading to this thesis was done at the Indian Statistical Institute, under the supervision of Prof. C. R. Rao, when the author was one of its research scholars. He takes this opportunity to thank Prof. Rao and the Institute.

C O N T E N T S

-- --------------------

P R E F A C E

This thesis gives the distribution theory of four classification statistics. Some parts of it were published a few years ago. The relevant publications are listed at the end.

What the author knows of related work by other authors is set out in the Introduction; little material of this nature is given elsewhere in the thesis. If a result is not explicitly credited to another author, it may be taken to mean that the author believes it to be his work.

It will be noticed that three (U, R and Z) of the four statistics receive less space than the remaining one (W). One reason for this is that the methods required in the treatment of U, R and Z are more or less similar to those employed in the case of W. Another is that the various statistics are equivalent under certain conditions.

Simplifying situations receive special attention. Generally, they also correspond to conditions explicitly excluded while deriving the general results.

December, 1961.

GLOSSARY OF SYMBOLS

$P^{(k)}$ $(k = 1, 2)$: 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 characteristics used.

$x = (x_1, x_2, \ldots, x_p)$: the vector of measurements on the individual to be classified, also used for x_1 when $p=1$.

$x \varepsilon P$ means that x is the vector of measurements of an individual from P (e.t.c).

$\mu^{(k)} = (\mu_1^{(k)}, \mu_2^{(k)}, \ldots, \mu_p^{(k)})$: the mean of $P^{(k)}$ $(k = 1, 2)$; also used for $\mu_1^{(k)}$ when $p=1$ $(k= 1,2)$.

$\mu = (\mu_1, \mu_2, \ldots, \mu_p)$: the mean of P; also used for μ_1 when $p=1$.

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

σ : the symbol for $\sigma_{11}^{\frac{1}{2}}$ when $p = 1$.

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

$\bar{x}^{(k)}$: the arithmetic mean of the observations in a random sample of size N_k from $P^{(k)}$ $(k=1, 2)$; also used for the element of this row-vector when $p=1$.

S : an unbiased estimate of Σ, distributed independently of $\bar{x}^{(k)}$ $(k=1,2)$ and following the Wishart law of n degrees of freedom—conditions which are satisfied by the matrix of mean pooled corrected sum of products, with $n = N_1 + N_2 - 2$.

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

n : the degree of freedom of S.

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

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

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

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

$$R = y S^{-1} y' - d(a_4 / a_3)^{\frac{1}{2}} (\bar{x}^{(2)} - \bar{x}^{(1)}) S^{-1} y', \text{ where}$$

$$y = x - a_1 \bar{x}^{(1)} - a_2 \bar{x}^{(2)} \quad \text{and} \quad d = 2(a_4 / a_3)^{\frac{1}{2}} / (a_2 - a_1).$$

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

$$U = (\bar{x}^{(2)} - \bar{x}^{(1)}) S^{-1} x'.$$

W_o, R_o, Z_o, U_o : respectively the same as W, R, Z and U with Σ substituted for S.

$$Q = (\bar{x}^{(2)} - \bar{x}^{(1)}) \Sigma^{-1} (\bar{x}^{(2)} - \bar{x}^{(1)})' ; \quad Q_1 = (\bar{x}^{(2)} - \mu^{(1)}) \Sigma^{-1} (\bar{x}^{(2)} - \mu^{(1)})'$$

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

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

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

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

$g(\alpha)$: the standard normal density.

$G(\alpha)$: the integral of $g(\alpha)$ from $-\infty$ to α.

$G^{-1}(\alpha)$: the function inverse to $G(\alpha)$.

$$G(\alpha_1, \alpha_2 ; \varrho) = \frac{(1-\varrho^2)^{-\frac{1}{2}}}{2\pi} \int_{-\infty}^{\alpha_1} dz_1 \int_{-\infty}^{\alpha_2} \exp\left[-\frac{1}{2(1-\varrho^2)}(z_1^2 - 2\varrho z_1 z_2 + z_2^2)\right] dz_2 .$$

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

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

$L_p(\alpha, z)$: the value of λ satisfying the equation

$$\int_{\alpha}^{\infty} \Delta_p(\chi^2; \lambda) d\chi^2 = z .$$

I : the identity matrix of order p.

If $u = (u_1, u_2, \ldots, u_p)$,

$du = du_1 \, du_2, \ldots, du_p$.

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 $E(y_{ij}) = \theta_{ij}$ $(i = 1,2,\ldots,m;\ j=1,2,\ldots,\nu)$. Let $\Lambda = (\sum_{r=1}^{\nu} \theta_{ir}\,\theta_{jr})$. We shall call any random matrix having the same distribution as YY' 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.

Chapter One

INTRODUCTION

1. THE CLASSIFICATION PROBLEM

About an individual it may be known that it belongs either to population $P^{(1)}$ or to population $P^{(2)}$, but which one of them is the real parent population may not be known. Statistical theory was first harnessed to the solution of such problems about twentyfive years ago, when M. M. Barnard, [4], following a suggestion of R. A. Fisher, [6], used the discriminant function to classify skeletal remains. Fisher saw the analogy of the classification problem to the problem of prediction. The discriminant function is, indeed, the best linear function for predicting the value of a variable which takes one value with individuals of one population and a different value with individuals of the other. Four years later Welch, [21], pointed out that, if one wishes to minimize the probabilities of misclassification, the likelihood-ratio should be the criterion of classification. The discriminant function and the likelihood-ratio are equivalent, if the alternative populations are multivariate normal and have their dispersion matrices equal. Throughout this thesis we shall assume that the alternative populations are of this type.

2. THE PROBLEMS CONSIDERED

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

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

$$R = (x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})S^{-1}(x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})'$$

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

and

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

(η a suitable constant; $d = 2(a_4/a_3)^{\frac{1}{2}}/(a_2-a_1)$) are four statistics which have been proposed for use in classifying individuals to $P^{(1)}$ or $P^{(2)}$ when the parameters of $P^{(1)}$ and $P^{(2)}$ require to be estimated. W is the statistic obtained by replacing the parameters in the logarithm of the likelihood-ratio by their estimates from a random sample. The logarithm of the likelihood-ratio is

$$(\mu^{(2)} - \mu^{(1)})\,\Sigma^{-1}x' - \tfrac{1}{2}(\mu^{(2)} - \mu^{(1)})\,\Sigma^{-1}(\mu^{(1)}+\mu^{(2)})'.$$

A statistic equivalent to this is

$$(\mu^{(2)} - \mu^{(1)})\Sigma^{-1}x' .$$

U is the statistic obtained from this by replacing $\mu^{(1)}$, $\mu^{(2)}$ and Σ by their estimates. Wald, [20], proposed it, because he thought the distribution theory of U would be simpler than that of W. The statistic

$$(x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})\,\Sigma^{-1}(x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})'$$

$$-\,d(a_4/a_3)^{\frac{1}{2}}(\bar{x}^{(2)}-\bar{x}^{(1)})\,\Sigma^{-1}(x-a_1\bar{x}^{(1)}-a_2\bar{x}^{(2)})'$$

was proposed by Rao, [15]. He showed that it is best for discriminating between alternatives which are close to each other. The fourth statistic Z is equivalent to a statistic derived by Anderson, [3],(p.142).

[Z with $\eta = 1$ was considered by the author in [9]. It was proposed there, because it appeared to be 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 two tests. We wish to mention that A. Kudo, [12], has shown that the procedure of assigning the individual to $P^{(1)}$ or $P^{(2)}$ according as

$$a_5(x-\bar{x}^{(1)})\Sigma^{-1}(x-\bar{x}^{(1)})' - a_6(x-\bar{x}^{(2)})\Sigma^{-1}(x-\bar{x}^{(2)})' \lesseqgtr 0$$

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

The four statistics W, U, 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 $a_5 = \eta a_6$.]

If W, U, 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 $\bar{x}^{(1)}$, $\bar{x}^{(2)}$ and S. In this thesis we seek to find out the

* Rao, [15], had previously stated this principle. See p.655.

distributions of these probabilities as well as their expected values*. Since the expected value can be calculated using the distribution of the statistic used, we shall obtain the distributions of the classification statistics also.

3. PREVIOUS WORK

Among the problems mentioned in the previous section, the distributions of the classification statistics W and U have received some attention from earlier authors. An account of their work will now be given.

In [20], Wald attempted to derive the distribution of U. Realizing that, when the alternative populations are normal, the distribution of U is a special case of that of $y^{(1)} S^{-1} y^{(2)'}$, where $y^{(1)}$, $y^{(2)}$ and S are independent random variables, $y^{(1)}$ and $y^{(2)}$ following the p-variate normal law, each with dispersion matrix Σ (say), and S the Wishart law with n degrees of freedom (say), Wald devoted his efforts to finding the distribution of $y^{(1)} S^{-1} y^{(2)'}$. Wald showed that $y^{(1)} S^{-1} y^{(2)'}$ could be expressed as a function of three statistics m_1, m_2 and m_3 and gave the joint distribution of m_1, m_2 and m_3 in the form of a product of three rather complicated factors. One of these factors was the unevaluated expected value of a power of a random

---

* Since writing this thesis, I have found that this problem was raised by Tocher during the discussion which followed the reading of [14] (p.198). [It is not clear which statistic Tocher had in mind. It could not have been R or Z, for, they were proposed only later]. From the speaker's reply (p.203 of [14]) we should gather that in 1948 no results on this problem existed. The author knows of no later work.

determinant. In 1951 Anderson, [2], evaluated this expectation in the case of linearly dependent $\xi^{(1)}$ and $\xi^{(2)}$. [$\xi^{(k)} (k = 1, 2)$ denotes the expectation of $y^{(k)}$.] Harter, [7], obtained the distribution of $y^{(1)} S^{-1} y^{(2)'}$ in the univariate case under the assumption $\xi^{(1)} = 0$. In [20], Wald had shown that, if n is large, $y^{(1)} S^{-1} y^{(2)'}$ has approximately the same distribution as nm_3. Harter, [7], obtains the marginal distribution of m_3, when both $\xi^{(1)}$ and $\xi^{(2)}$ are null vectors.

Sitgreaves, [16], showed that the statistic W also could be expressed in terms of three functions of $y^{(1)}$, $y^{(2)}$ and S having the same joint distribution as m_1, m_2 and m_3 and gave a derivation of the joint distribution of m_1, m_2 and m_3 different from that of Wald. Sitgreaves also requires that $\xi^{(1)}$ should be a scalar multiple of $\xi^{(2)}$.

Bowker, [5], expresses the statistic W as a function of the elements of two independent 2 x 2 Wishart matrices, one central and the other non-central. An advantage of Bowker's expression over that of Sitgreaves' in terms of m_1, m_2 and m_3 seems to be that the joint distribution of the variables appearing in Bowker's expression is explicitly known, whereas no explicit expression for the density function of m_1, m_2 and m_3 is available, except when $\xi^{(1)}$ and $\xi^{(2)}$ are linearly dependent. In principle, to get the distribution of W we have only to integrate the joint density of the elements of the two Wishart matrices over a region of the six-dimensional space, but attempts to further simplify the integral have not succeeded.

4. SYNOPSIS OF THE THESIS

In the case of unknown $\mu^{(1)}$, $\mu^{(2)}$ and Σ we exhibit U, W, R and Z as functions of a few random variables with a fairly simple joint distribution (Chapter two, equations (2.3.2),(2.4.2),(2.5.2) and (2.6.5)); using these we derive certain results for the expected probabilities, which may facilitate Monte Carlo methods (Chapter eight); thirdly, the limiting distribution of the probability of assigning the individual to $P^{(1)}$ or $P^{(2)}$ is obtained (Chapter four, equation (4.4.4) and Chapter five, equation (5.2.4))*

In the case of known Σ and unknown $\mu^{(1)}$ and $\mu^{(2)}$, explicit expressions for the density functions of U, W, R and Z are given (Chapter two); exact distributions of the (conditional) probabilities are obtained (Chapters three to seven); their expected values are evaluated (Chapter eight); for the expected values, approximations, which, unlike exact expressions, are easy to compute, are worked out (Chapter eight, equation (8.3.2)); special attention is paid to the important case where the origin of the individual to be classified is one of the two alternative populations (Chapters four and five); situations where we can have simpler results are explored (Chapters three and four).

---

* The asymptotic distribution of the probability is given only in the case of W; for, the four statistics U, W, R and Z (with $\eta = 1$) are asymptotically equivalent.

Chapter Two

DISTRIBUTION OF THE CLASSIFICATION STATISTICS

1. INTRODUCTION

In the introduction to this thesis we wrote about four classification statistics U, W, R and Z. Suppose that f_k $(k = 1, 2)$ denotes the density function of the classification statistic when $x \varepsilon P^{(k)}$. Then the probability of misclassifying an individual from $P^{(i)}$ $(i = 1, 2)$ is the integral of f_i over the set of values of the statistic which lead to the assignment of the individual to $P^{(j)}$ $(j = 1, 2;\ j \neq i)$. Since it is interesting to know the probability of assignment to $P^{(k)}$ $(k = 1, 2)$ for individuals from P also, we shall derive the distributions with $x \varepsilon P$; we can, of course, get the distribution when $x \varepsilon P^{(k)}$ $(k = 1, 2)$ by setting $\mu = \mu^{(k)}$.

The statistics U, W, R and Z are particular cases of a general statistic introduced in the next section. We shall first derive the distribution of this statistic and from it deduce the distributions of U, W, R and Z.

2. TWO GENERAL RESULTS

Let $y^{(1)}$ and $y^{(2)}$ be two independent p-dimensional random normal vectors. Let $\theta^{(k)}$ $(k = 1, 2)$ denote $E\, y^{(k)}$, and let Σ denote the dispersion matrix of each of them. Let

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

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

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

denote its noncentrality matrix, with

$$(2.2.1) \qquad \lambda_{11} = \theta^{(1)} \Sigma^{-1} \theta^{(1)'}, \ \lambda_{12} = \theta^{(1)} \Sigma^{-1} \theta^{(2)'},$$
$$\lambda_{22} = \theta^{(2)} \Sigma^{-1} \theta^{(2)'} ;$$

t follows Students' law with $n-p+2$ degrees of freedom; χ^2 is a chi-square variable with $n-p+1$ degrees of freedom. Then, any function of $y^{(1)} S^{-1} y^{(1)'}$ and $y^{(1)} S^{-1} y^{(2)'}$ has the same distribution as the function of B_{11}, B_{12}, B_{22}, t and χ^2 obtained by substituting nB_{11}/χ^2 for $y^{(1)} S^{-1} y^{(1)'}$ and $n[B_{12} - (n-p+2)^{-\frac{1}{2}} t \,|B|^{\frac{1}{2}}]/\chi^2$ for $y^{(1)} S^{-1} y^{(2)'}$.

We shall later exhibit the statistics Z, U, W and R as special cases of the statistic introduced above. In [5], Bowker has expressed the statistic W as a function of B_{11}, B_{12}, B_{22} and the elements of another Wishart matrix distributed independently of B. We omit the proof of the result given in the previous paragraph; for, the reasoning that led us to it happens to be similar to that of [5].

If Σ is known, we use Σ, rather than S, to construct the classification statistic. Each of the ~~three~~ classification statistics has then the same distribution as $y^{(1)} \Sigma^{-1} y^{(1)\prime}/b_1 - y^{(2)} \Sigma^{-1} y^{(2)\prime}/b_2$, with appropriate b_1, b_2, $\theta^{(1)}$ and $\theta^{(2)}$; i.e., the same distribution as $(w_1/b_1) - (w_2/b_2)$, where w_1 and w_2 are independent noncentral chi-squares. If we denote the non-centrality of w_k $(k = 1, 2)$ by λ_k, the joint density of w_1 and w_2 is

$$\frac{e^{-(\lambda_1+\lambda_2)}}{2^p} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \left(\tfrac{1}{2}\right)^{r+s} \frac{\lambda_1^r \lambda_2^s}{r!\, s!} \frac{w_1^{\frac{1}{2}p+r-1} w_2^{\frac{1}{2}p+s-1}}{\Gamma(\frac{1}{2}p+r)\,\Gamma(\frac{1}{2}p+s)} e^{-(w_1+w_2)/2}$$

Therefore, the

$$\Pr(z < w_1/b_1 - w_2/b_2 < z + dz)$$

$$(2.2.1) \quad = e^{-(\lambda_1+\lambda_2)} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\lambda_1^r \lambda_2^s}{r!\, s!} \int\int 2^{-(p+r+s)} \frac{w_1^{\frac{1}{2}p+r-1} w_2^{\frac{1}{2}p+s-1}}{\Gamma(\frac{1}{2}p+r)\,\Gamma(\frac{1}{2}p+s)} \times e^{-\frac{1}{2}(w_1+w_2)} \, dw_1 \, dw_2,$$

where the integration with respect to w_1 and w_2 is over the set of pairs of w_1 and w_2 satisfying the condition $z < w_1/b_1 - w_2/b_2 < z + dz$. That is, if

$$(2.2.3) \qquad \Psi_{r,s}(z)\, dz = \int_{z < \frac{w_1}{b_1} - \frac{w_2}{b_2} < z+dz} \int 2^{-(p+r+s)} \frac{w_1^{\frac{1}{2}p+r-1} w_2^{\frac{1}{2}p+s-1}}{\Gamma(\frac{1}{2}p+r)\,\Gamma(\frac{1}{2}p+s)} \times e^{-\frac{1}{2}(w_1+w_2)}\, dw_1\, dw_2,$$

the density function of $z \; (= w_1/b_1 - w_2/b_2)$ is

$$(2.2.4) \qquad e^{-(\lambda_1+\lambda_2)} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\lambda_1^r \lambda_2^s}{r!\, s!} \Psi_{r,s}(z) .$$

In terms of $\theta^{(1)}$ and $\theta^{(2)}$,

$$(2.2.5) \qquad \lambda_1 = \frac{1}{2}\theta^{(1)} \Sigma^{-1} \theta^{(1)\prime}, \quad \lambda_2 = \frac{1}{2}\theta^{(2)} \Sigma^{-1} \theta^{(2)\prime}.$$

In [13], Karl Pearson gives the equation

$$(2.2.6) \qquad \Psi_{r,s}(z) = \frac{2e^{-\frac{1}{4}(b_1-b_2)z}}{4^{p+r+s}} \frac{b_1^{\frac{1}{2}p+r}\, b_2^{\frac{1}{2}p+s}}{\Gamma(\frac{1}{2}p+r)\,\Gamma(\frac{1}{2}p+s)} |z|^{p+r+s-1} \int_1^{\infty} e^{-\frac{1}{4}(b_1+b_2)|z|\alpha} (\alpha+1)^{\frac{1}{2}p+r-1} (\alpha-1)^{\frac{1}{2}p+s-1}\, d\alpha \qquad (z \geq 0).$$

If $z < 0$,

$$(2.2.7)\quad \Psi_{r,s}(z) = \frac{2e^{-\frac{1}{4}(b_1-b_2)z}}{4^{p+r+s}} \frac{b_1^{\frac{1}{2}p+r}\, b_2^{\frac{1}{2}p+s}}{\Gamma(\frac{1}{2}p+r)\Gamma(\frac{1}{2}p+s)} |z|^{p+r+s-1}$$

$$\int_1^\infty e^{-\frac{1}{4}(b_1+b_2)|z|\alpha} (\alpha-1)^{\frac{1}{2}p+r-1} (\alpha+1)^{\frac{1}{2}p+s-1} d\alpha.$$

Pearson puts $\Psi_{r,s}(z)$ in terms of the double Bessel function and studies it in detail. It can also be expressed in terms of Whittaker's confluent hypergeometric function $W_{k,m}(z)$; for

$$(2.2.8)\quad \frac{1}{\Gamma(m-k+\frac{1}{2})} \int_1^\infty e^{-\frac{1}{2}z\alpha} (\alpha-1)^{m-k-\frac{1}{2}} (\alpha+1)^{m+k-\frac{1}{2}} d\alpha$$

$$= 2^{2m} z^{-(m+\frac{1}{2})} W_{k,m}(z)\text{([17], p.51)}.$$

3. THE DISTRIBUTION OF U

The first result of section two may be used to express U in terms of B_{11}, B_{12}, B_{22}, t and χ^2. Take

$$(2.3.1)\quad y^{(1)} = (\bar{x}^{(2)} - \bar{x}^{(1)})/\sqrt{a_3}\,, \quad y^{(2)} = x.$$

$$(2.3.2)\quad U = n\sqrt{a_3}\,\frac{B_{12} - (n-p+2)^{-\frac{1}{2}}\, t\,|B|^{\frac{1}{2}}}{\chi^2}.$$

$$(2.3.3)\quad \lambda_{11} = \delta^2/a_3,\ \lambda_{12} = \left\{(\mu^{(2)}-\mu^{(1)})\,\Sigma^{-1}\,\mu'\right\}/\sqrt{a_3},\ \lambda_{22} = \mu\,\Sigma^{-1}\mu'.$$

When Σ is known, we use $U_o (= (\bar{x}^{(2)} - \bar{x}^{(1)}) \Sigma^{-1} x')$ rather than U. The second result of section two may be used to write down the distribution of U_o. Take

(2.3.4) $y^{(1)} = (\bar{x}^{(2)} - \bar{x}^{(1)} + \sqrt{a_3}\,x)/\sqrt{(2a_3)}$, $y^{(2)} = (\bar{x}^{(2)} - \bar{x}^{(1)} - \sqrt{a_3}\,x)/\sqrt{(2a_3)}$,

(2.3.5) $b_1 = 2/\sqrt{a_3}$, $b_2 = 2/\sqrt{a_3}$.

(2.3.6) $\lambda_1 = \frac{1}{4a_3}(\mu^{(2)} - \mu^{(1)} + \sqrt{a_3}\,\mu)\,\Sigma^{-1}(\mu^{(2)} - \mu^{(1)} + \sqrt{a_3}\,\mu)'$;

$\lambda_2 = \frac{1}{4a_3}(\mu^{(2)} - \mu^{(1)} - \sqrt{a_3}\,\mu)\,\Sigma^{-1}(\mu^{(2)} - \mu^{(1)} - \sqrt{a_3}\,\mu)'$.

4. THE DISTRIBUTION OF W

The first result of section two may be used to express W in terms of B_{11}, B_{12}, B_{22}, t and χ^2. Take

(2.4.1) $y^{(1)} = (\bar{x}^{(2)} - \bar{x}^{(1)})/\sqrt{a_3}$, $y^{(2)} = (x - a_1\bar{x}^{(1)} - a_2\bar{x}^{(2)})/\sqrt{a_4}$.

(2.4.2) $W = \frac{n}{\chi^2}[\frac{1}{2}(N_1^{-1} - N_2^{-1})B_{11} + (a_3a_4)^{\frac{1}{2}}\{B_{12} - (n-p+2)^{-\frac{1}{2}}\, t\, |B|^{\frac{1}{2}}\}]$.

(2.4.3) $\lambda_{11} = \delta^2/a_3$; $\lambda_{12} = (\mu^{(2)} - \mu^{(1)})\Sigma^{-1}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)})'/\sqrt{(a_3a_4)}$;

$\lambda_{22} = (\mu - a_1\mu^{(1)} - a_2\mu^{(2)})\Sigma^{-1}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)})'/a_4$.

When Σ is known, we use

$$W_o(\blacksquare(\bar{x}^{(2)}-\bar{x}^{(1)})\Sigma^{-1} x - \tfrac{1}{2}(\bar{x}^{(2)}-\bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(1)}+\bar{x}^{(2)})')$$

rather than W. The second result of section two may be used to obtain the distribution of W_o. Take

$$(2.4.4)\quad y^{(1)} = \frac{(1+4/a_3)^{\frac{1}{4}}}{\sqrt{[\sqrt{(1+4/a_3)}+a_2-a_1]}}\left[\frac{\bar{x}^{(2)}-\bar{x}^{(1)}}{\sqrt{(2a_3)}} + \frac{x-\frac{1}{2}\bar{x}^{(1)}-\frac{1}{2}\bar{x}^{(2)}}{\sqrt{(2+a_3/2)}}\right],$$

$$y^{(2)} = \frac{(1+4/a_3)^{\frac{1}{4}}}{\sqrt{[\sqrt{(1+4/a_3)}-a_2+a_1]}}\left[\frac{\bar{x}^{(2)}-\bar{x}^{(1)}}{\sqrt{(2a_3)}} - \frac{x-\frac{1}{2}\bar{x}^{(1)}-\frac{1}{2}\bar{x}^{(2)}}{\sqrt{(2+a_3/2)}}\right],$$

$$(2.4.5)\quad b_1 = \frac{4}{a_3[\sqrt{(1+4/a_3)}+a_2-a_1]}\ ;\quad b_2 = \frac{4}{a_3[\sqrt{(1+4/a_3)}-a_2+a_1]}.$$

$$\lambda_1 = \frac{\sqrt{(1+4/a_3)}}{2[\sqrt{(1+4/a_3)}+a_2-a_1]}\left(\frac{\mu^{(2)}-\mu^{(1)}}{\sqrt{(2a_3)}}+\frac{\mu-\frac{1}{2}\mu^{(1)}-\frac{1}{2}\mu^{(2)}}{\sqrt{(2+a_3/2)}}\right)\Sigma^{-1}\left(\frac{\mu^{(2)}-\mu^{(1)}}{\sqrt{(2a_3)}}+\frac{\mu-\frac{1}{2}\mu^{(1)}-\frac{1}{2}\mu^{(2)}}{\sqrt{(2+a_3/2)}}\right)';$$

(2.4.6)

$$\lambda_2 = \frac{\sqrt{(1+4/a_3)}}{2[\sqrt{(1+4/a_3)}-a_2+a_1]}\left(\frac{\mu^{(2)}-\mu^{(1)}}{\sqrt{(2a_3)}}-\frac{\mu-\frac{1}{2}\mu^{(1)}-\frac{1}{2}\mu^{(2)}}{\sqrt{(2+a_3/2)}}\right)\Sigma^{-1}\left(\frac{\mu^{(2)}-\mu^{(1)}}{\sqrt{(2a_3)}}-\frac{\mu-\frac{1}{2}\mu^{(1)}-\frac{1}{2}\mu^{(2)}}{\sqrt{(2+a_3/2)}}\right)'$$

5. THE DISTRIBUTION OF R

Taking

$$(2.5.1) \qquad y^{(1)} = \frac{x - a_1\bar{x}^{(1)} - a_2\bar{x}^{(2)}}{\sqrt{a_4}}, \quad y^{(2)} = \frac{\bar{x}^{(2)} - \bar{x}^{(1)}}{\sqrt{a_3}},$$

we apply the first result of section two and get

$$(2.5.2) \qquad R = \frac{na_4}{\chi^2}[B_{11} - d\left\{B_{12} - (n-p+2)^{-\frac{1}{2}}\, t\,|B|^{\frac{1}{2}}\right\}].$$

$$(2.5.3) \qquad \lambda_{11} = (\mu - a_1\mu^{(1)} - a_2\mu^{(2)})\Sigma^{-1}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)})'/a_4;$$

$$\lambda_{12} = (\mu - a_1\mu^{(1)} - a_2\mu^{(2)})\Sigma^{-1}(\mu^{(2)} - \mu^{(1)})'/\sqrt{(a_3a_4)}; \quad \lambda_{22} = \delta^2/a_3.$$

When Σ is known, we use

$$R_o (= y\,\Sigma^{-1}\,y' - d[a_4/a_3]^{\frac{1}{2}}\,[\bar{x}^{(2)} - \bar{x}^{(1)}]\,\Sigma^{-1}\,y')$$

rather than R. Verify that R_o is a special case of the statistic $y^{(1)}\,\Sigma^{-1}y^{(1)}/\,b_1 - y^{(2)}\,\Sigma^{-1}\,y^{(2)}{}'/b_2$

considered in section two, taking

$$(2.5.4) \qquad y^{(1)} = \frac{[\sqrt{(1+d^2)}+1]^{\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}\,(2a_4)^{\frac{1}{2}}}(x - a_1\bar{x}^{(1)} - a_2\bar{x}^{(2)}) - \frac{d[\sqrt{(1+d^2)}+1]^{-\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}\,(2a_3)^{\frac{1}{2}}}(\bar{x}^{(2)} - \bar{x}^{(1)}),$$

$$y^{(2)} = \frac{[\sqrt{(1+d^2)}-1]^{\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}\,(2a_4)^{\frac{1}{2}}}(x - a_1\bar{x}^{(1)} - a_2\bar{x}^{(2)}) + \frac{d[\sqrt{(1+d^2)}-1]^{-\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}\,(2a_3)^{\frac{1}{2}}}(\bar{x}^{(2)} - \bar{x}^{(1)}),$$

$$(2.5.5)\qquad b_1 = (2/a_4)/[\sqrt{(1+d^2)}+1], \quad b_2 = (2/a_4)/[\sqrt{(1+d^2)}-1].$$

The density function of W_o is, therefore, the function (2.2.4) with

$$(2.5.6)\qquad \lambda_1 = \frac{1}{2}\left[\frac{\{\sqrt{(1+d^2)}+1\}^{\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_4)^{\frac{1}{2}}}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)}) - \frac{d\{\sqrt{(1+d^2)}+1\}^{-\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_3)^{\frac{1}{2}}}(\mu^{(2)} - \mu^{(1)})\right]\Sigma^{-1}\left[\frac{\{\sqrt{(1+d^2)}+1\}^{\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_4)^{\frac{1}{2}}}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)}) - \frac{d\{\sqrt{(1+d^2)}+1\}^{-\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_3)^{\frac{1}{2}}}(\mu^{(2)} - \mu^{(1)})\right]',$$

$$\lambda_2 = \frac{1}{2}\left[\frac{\{\sqrt{(1+d^2)}-1\}^{\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_4)^{\frac{1}{2}}}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)}) + \frac{d\{\sqrt{(1+d^2)}-1\}^{-\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_3)^{\frac{1}{2}}}(\mu^{(2)} - \mu^{(1)})\right]\Sigma^{-1}\left[\frac{\{\sqrt{(1+d^2)}-1\}^{\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_4)^{\frac{1}{2}}}(\mu - a_1\mu^{(1)} - a_2\mu^{(2)}) + \frac{d\{\sqrt{(1+d^2)}-1\}^{-\frac{1}{2}}}{(1+d^2)^{\frac{1}{4}}(2a_3)^{\frac{1}{2}}}(\mu^{(2)} - \mu^{(1)})\right]'.$$

6. THE DISTRIBUTION OF Z

With

$$\beta_1 = \frac{a_5^{\frac{1}{2}}[2\eta - 2(a_5 a_6 \eta)^{\frac{1}{2}}]}{2(\eta - a_5 a_6 \eta)^{\frac{1}{2}}[1 + \eta - 2(a_5 a_6 \eta)^{\frac{1}{2}}]^{\frac{1}{2}}}, \tag{2.6.1}$$

$$\beta_2 = \frac{(\eta a_6)^{\frac{1}{2}}[2 - 2(\eta a_5 a_6)^{\frac{1}{2}}]}{2(\eta - a_5 a_6 \eta)^{\frac{1}{2}}[1 + \eta - 2(a_5 a_6 \eta)^{\frac{1}{2}}]^{\frac{1}{2}}}, \tag{2.6.2}$$

$$y^{(1)} = \frac{a_5^{\frac{1}{2}}(x - \bar{x}^{(1)}) - (\eta a_6)^{\frac{1}{2}}(x - \bar{x}^{(2)})}{[1 + \eta - 2(\eta a_5 a_6)^{\frac{1}{2}}]^{\frac{1}{2}}}, \tag{2.6.3}$$

and

$$y^{(2)} = \beta_1(x - \bar{x}^{(1)}) + \beta_2(x - \bar{x}^{(2)}), \tag{2.6.4}$$

$$Z = (1 - \eta) y^{(1)} S^{-1} y^{(1)'} + 2(\eta - a_5 a_6 \eta)^{\frac{1}{2}} y^{(1)} S^{-1} y^{(2)'}. \tag{2.6.5}$$

Applying the first result of section two, we conclude that Z has the distribution of

$$(n/\chi^2)[(1-\eta)B_{11} + 2(\eta - a_5 a_6 \eta)^{\frac{1}{2}} \{B_{12} - (n-p+2)^{-\frac{1}{2}} |B|^{\frac{1}{2}} t\}], \tag{2.6.6}$$

with

$$\lambda_{11} = \{ [a_5^{\frac{1}{2}}(\mu-\mu^{(1)}) - (\eta a_6)^{\frac{1}{2}}(\mu-\mu^{(2)})] \Sigma^{-1}$$

$$[a_5^{\frac{1}{2}}(\mu-\mu^{(1)})-(\eta a_6)^{\frac{1}{2}}(\mu-\mu^{(2)})]'\} \div$$

$$[1+\eta - 2(\eta a_5 a_6)^{\frac{1}{2}}],$$

(2.6.7)
$$\lambda_{12} = \{[a_5^{\frac{1}{2}}(\mu-\mu^{(1)})- (\eta a_6)^{\frac{1}{2}}(\mu-\mu^{(2)})] \Sigma^{-1}$$

$$[\beta_1(\mu-\mu^{(1)})+\beta_2(\mu-\mu^{(2)})]'\} \div [1+\eta - 2(\eta a_5 a_6)^{\frac{1}{2}}]^{\frac{1}{2}},$$

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

$$[\beta_1(\mu - \mu^{(1)}) + \beta_2(\mu - \mu^{(2)})']'.$$

(2.6.8)
$$Z_o = \frac{1}{b_1} y^{(1)} \Sigma^{-1} y^{(1)'} - \frac{1}{b_2} y^{(2)} \Sigma^{-1} y^{(2)'},$$

if

(2.6.9)
$$\gamma_1 = [1+\eta - 2(\eta a_5 a_6)^{\frac{1}{2}}]^{\frac{1}{2}},$$

(2.6.10)
$$\gamma_2 = [1+\eta + 2(\eta a_5 a_6)^{\frac{1}{2}}]^{\frac{1}{2}},$$

$$(2.6.11) \qquad y^{(1)} = [a_5^{\frac{1}{2}}(\gamma_1 + \gamma_2)(x - \bar{x}^{(1)}) + (\eta a_6)^{\frac{1}{2}}(\gamma_1 - \gamma_2)(x - \bar{x}^{(2)})] \div$$

$$(2\gamma_1 \gamma_2 - 2\eta + 2)^{\frac{1}{2}},$$

$$(2.6.12) \qquad y^{(2)} = [a_5^{\frac{1}{2}}(\gamma_1 - \gamma_2)(x - \bar{x}^{(1)}) + (\eta a_6)^{\frac{1}{2}}(\gamma_1 + \gamma_2)(x - \bar{x}^{(2)})] \div$$

$$(2\gamma_1 \gamma_2 + 2\eta - 2)^{\frac{1}{2}},$$

$$(2.6.13) \qquad b_1 = 2\left\{ [(1+\eta)^2 - 4a_5 a_6 \eta]^{\frac{1}{2}} + 1 - \eta \right\}^{-1},$$

$$(2.6.14) \qquad b_2 = 2\left\{ [(1+\eta)^2 - 4a_5 a_6 \eta]^{\frac{1}{2}} + \eta - 1 \right\}^{-1}.$$

Applying the second result of section two, we conclude that Z_0 has the density function (2.2.4), with b_1 and b_2 as given by equations (2.6.13) and (2.6.14) and with

$$(2.6.15) \qquad \lambda_1 = \tfrac{1}{2}[\{a_5^{\frac{1}{2}}(\gamma_1 + \gamma_2)(\mu - \mu^{(1)}) + (\eta a_6)^{\frac{1}{2}}(\gamma_1 - \gamma_2)(\mu - \mu^{(2)})$$

$$\Sigma^{-1}\{a_5^{\frac{1}{2}}(\gamma_1 + \gamma_2)(\mu - \mu^{(1)}) + (\eta a_6)^{\frac{1}{2}}(\gamma_1 - \gamma_2)(\mu - \mu^{(2)})\}'] \div$$

$$(2\gamma_1 \gamma_2 - 2\eta + 2),$$

and

$$(2.6.16) \qquad \lambda_2 = \tfrac{1}{2}[\{a_5^{\frac{1}{2}}(\gamma_1 - \gamma_2)(\mu - \mu^{(1)}) + (\eta a_6)^{\frac{1}{2}}(\gamma_1 + \gamma_2)(\mu - \mu^{(2)})$$

$$\Sigma^{-1}\{a_5^{\frac{1}{2}}(\gamma_1 - \gamma_2)(\mu - \mu^{(1)}) + (\eta a_6)^{\frac{1}{2}}(\gamma_1 + \gamma_2)(\mu - \mu^{(2)})\}'] \div$$

$$(2\gamma_{1\,2} + 2\eta - 2).$$

Chapter Three

W: THE UNIVARIATE CASE

1. INTRODUCTION

Let $\pi^{(k)}$ $(k = 1, 2)$ be the a priori probability that the individual to be classified belongs to $P^{(k)}$. Let c_{12} be the loss in assigning an individual from $P^{(1)}$ to $P^{(2)}$, and let c_{21} be the loss in assigning an individual from $P^{(2)}$ to $P^{(1)}$. Then it is known that, if the individual is assigned to $P^{(1)}$ or $P^{(2)}$ according as

$$(\mu^{(2)} - \mu^{(1)}) \Sigma^{-1} x' - \frac{1}{2}(\mu^{(2)} - \mu^{(1)}) \Sigma^{-1} (\mu^{(1)} + \mu^{(2)})' \lesseqgtr c,$$

$$c = \log_e \frac{\pi^{(1)} c_{12}}{\pi^{(2)} c_{21}},$$

the loss due to misclassification can be expected to be the least possible.* In case $\mu^{(k)}$ $(k = 1, 2)$ and Σ are unknown, we may decide to assign the individual to $P^{(1)}$ or $P^{(2)}$ according as

$$(\bar{x}^{(2)} - \bar{x}^{(1)}) S^{-1} x' - \frac{1}{2}(\bar{x}^{(2)} - \bar{x}^{(1)}) S^{-1} (\bar{x}^{(1)} + \bar{x}^{(2)})' \lesseqgtr c.$$

As we have already explained in the introduction to this thesis, the probabilities of misclassification are then functions of $\bar{x}^{(1)}$, $\bar{x}^{(2)}$ and S and, therefore, random variables. It is of interest to study their distributions. This chapter discusses the univariate case, and chapters four to six

* See, for instance, p. 134 of [3].

discuss the multivariate case. The results are first worked out for the case $c = 0$ and later extended to the more general case. The case $c = 0$ obtains if $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$ and $c_{12} = c_{21}$, which is the case usually.

In the univariate case, we shall for convenience write μ for μ_1, $\mu^{(k)}$ for $\mu_1^{(k)}$ $(k = 1, 2)$, σ for $\sqrt{\sigma_{11}}$, $\bar{x}^{(k)}$ for $\bar{x}_1^{(k)}$ $(k = 1, 2)$ and x for x_1.

It is easy to see that the classification procedure described earlier reduces in this case to the following: if $\bar{x}^{(2)} > \bar{x}^{(1)}$, assign the individual to $P^{(1)}$ or $P^{(2)}$ according as $x \lessgtr (\bar{x}^{(1)} + \bar{x}^{(2)})/2$; if $\bar{x}^{(2)} \leq \bar{x}^{(1)}$, assign the individual to $P^{(1)}$ or $P^{(2)}$ according as $x \gtrless (\bar{x}^{(1)} + \bar{x}^{(2)})/2$.

We shall denote by $e_1(\bar{x}^{(1)}, \bar{x}^{(2)})$ the probability of assigning an individual from P to $P^{(1)}$ given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, and by $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$ the probability of assigning the individual to $P^{(2)}$. Since $e_1(\bar{x}^{(1)}, \bar{x}^{(2)}) = 1 - e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$, it is enough to consider either $e_1(\bar{x}^{(1)}, \bar{x}^{(2)})$ or $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$; we consider $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$.

2. THE DISTRIBUTION OF $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$

It is easy to see that

$$(3.2.1) \quad e_2(\bar{\bar{x}}^{(1)},\bar{\bar{x}}^{(2)}) = \begin{cases} 1-G([\frac{1}{2}\{\bar{\bar{x}}^{(1)}+\bar{\bar{x}}^{(2)}\}-\mu]/\sigma), & \text{if } \bar{\bar{x}}^{(1)} < \bar{\bar{x}}^{(2)}, \\ G([\frac{1}{2}\{\bar{\bar{x}}^{(1)}+\bar{\bar{x}}^{(2)}\}-\mu]/\sigma), & \text{if } \bar{\bar{x}}^{(1)} \geq \bar{\bar{x}}^{(2)}. \end{cases}$$

Therefore, $e_2(\bar{\bar{x}}^{(1)}, \bar{\bar{x}}^{(2)}) < z$, if and only if either of the following two events happen:

$$\bar{\bar{x}}^{(1)} < \bar{\bar{x}}^{(2)} \quad \text{and} \quad \tfrac{1}{2}(\bar{\bar{x}}^{(1)} + \bar{\bar{x}}^{(2)}) - \mu > -\sigma G^{-1}(z),$$

or

$$\bar{\bar{x}}^{(1)} \geq \bar{\bar{x}}^{(2)} \quad \text{and} \quad \tfrac{1}{2}(\bar{\bar{x}}^{(1)} + \bar{\bar{x}}^{(2)}) - \mu < \sigma G^{-1}(z).$$

The distribution function of $e_2(\bar{\bar{x}}^{(1)}, \bar{\bar{x}}^{(2)})$ is, therefore, given by the equation

$$(3.2.2) \quad \Pr[e_2(\bar{\bar{x}}^{(1)}, \bar{\bar{x}}^{(2)}) < z] = G(h_{11}, h_{12}; \varrho) + G(h_{21}, h_{22}; \varrho),$$

where

$$h_{11} = \frac{\mu^{(2)} - \mu^{(1)}}{\sigma\sqrt{a_3}}, \quad h_{12} = \frac{\mu^{(1)} + \mu^{(2)} - 2\mu + 2\sigma G^{-1}(z)}{\sigma\sqrt{a_3}},$$

$$(3.2.3) \quad h_{21} = -h_{11}, \quad h_{22} = \frac{2\mu - \mu^{(1)} - \mu^{(2)} + 2\sigma G^{-1}(z)}{\sigma\sqrt{a_3}},$$

$$\varrho = a_1 - a_2.$$

We note that, if $N_1 = N_2$, we can write

$$(3.2.4) \quad \Pr[e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = G(h_{11})G(h_{12}) + G(h_{21})G(h_{22}).$$

3. EXPECTED VALUE OF $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$

The expected value of $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$ can be calculated from the equation

$$(3.3.1) \quad E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) = G(a_{11}, a_{12}; \varrho) + G(a_{21}, a_{22}; \varrho),$$

where

$$a_{11} = \frac{\mu^{(2)} - \mu^{(1)}}{\sigma\sqrt{a_3}}, \quad a_{12} = \frac{2\mu - \mu^{(1)} - \mu^{(2)}}{\sigma\sqrt{(4+ a_3)}},$$

$$(3.3.2) \quad a_{21} = -a_{11}, \qquad a_{22} = -a_{12}, \quad \varrho = \frac{a_2 - a_1}{\sqrt{(1+4/a_3)}}.$$

Equation (3.3.1) is easily established, if we observe that the assignment of the individual to $P^{(2)}$ corresponds to the occurrence of either of the following two events:

$$\bar{x}^{(1)} < \bar{x}^{(2)} \qquad \text{and} \qquad x \geq \tfrac{1}{2}(\bar{x}^{(1)} + \bar{x}^{(2)}),$$

or

$$\bar{x}^{(1)} > \bar{x}^{(2)} \qquad \text{and} \qquad x < \tfrac{1}{2}(\bar{x}^{(1)} + \bar{x}^{(2)}).$$

We note that, if $N_1 = N_2$, we can write

$$(3.3.3) \quad E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) = G(a_{11})G(a_{12}) + G(a_{21})G(a_{22}).$$

Below we give, for some values of $N(= N_1 = N_2)$ and δ, a table of the probability of assigning an individual from $P^{(1)}$ to $P^{(2)}$, computed using equation (3.3.3).

Table I

Probability of misclassification.

δ \ N	8	18	32	50	∞
0.1	0.4969	0.4954	0.4950	0.4924	0.4801
0.3	0.4739	0.4628	0.4545	0.4486	0.4404
0.5	0.4346	0.4156	0.4065	0.4030	0.4013
1.0	0.3223	0.3114	0.3099	0.3094	0.3085
2.0	0.1660	0.1620	0.1606	0.1599	0.1587

We note that the probability of misclassification decreases monotonically as N or δ increases. When δ is very small, the probability of misclassification is very nearly equal to 0.5 whatever be the value of N, and, therefore, it does not differ much from its value when the parameters are known. Similar is the case when δ is large, the probability of misclassification being then nearly equal to zero for all values of N. But when δ is neither too large nor too small, the difference is not negligible. If the sampling behaviour of other multivariate statistics is any guide, we should expect to find a more pronounced difference for larger values of p.

Chapter Four

W: THE MULTIVARIATE CASE

1. INTRODUCTION

In this chapter we take up for consideration the multivariate case. The procedure discussed is an adaptation of the standard discriminant function analysis to situations where the parameters required for the construction of the discriminant function are unknown.

The discussion will proceed in three stages. At stage one, we shall assume that only $\mu^{(2)}$ is unknown. The case where only $\mu^{(1)}$ is unknown is completely analogous and does not require separate consideration. At stage two, we shall only assume that the dispersion matrix Σ is known. In the third stage we do not assume that either $\mu^{(1)}$, $\mu^{(2)}$ or Σ are known.

2. CASE ONE: ONLY $\mu^{(2)}$ IS UNKNOWN

Before starting discussion of this case, let us note that we shall not be erring seriously if we take $\bar{x}^{(1)}$ to be $\mu^{(1)}$ and S to be Σ, provided N_1 is large.

For constructing the discriminant function, $\mu^{(2)}$ has to be estimated. If in the discriminant function we substitute $\bar{x}^{(2)}$ for $\mu^{(2)}$, we have $(\bar{x}^{(2)} - \mu^{(1)}) \Sigma^{-1} x'$. Individuals will be assigned to $P^{(1)}$ or $P^{(2)}$ according as

$$(\bar{x}^{(2)} - \mu^{(1)}) \Sigma^{-1} x' \lessgtr \tfrac{1}{2} (\bar{x}^{(2)} - \mu^{(1)}) \Sigma^{-1} (\bar{x}^{(2)} + \mu^{(1)})'.$$

The distribution of the probability of misclassification. Given $\bar{x}^{(2)}$, the probability of misclassifying an individual from $P^{(1)}$ is $1 - G(\tfrac{1}{2}\sqrt{Q_1})$, where

$$Q_1 = (\bar{x}^{(2)} - \mu^{(1)}) \Sigma^{-1} (\bar{x}^{(2)} - \mu^{(1)})'.$$

We shall denote this probability by $e_{12}(\bar{x}^{(2)})$. Clearly, $e_{12}(\bar{x}^{(2)})$ is a random variable. The distribution function of $e_{12}(\bar{x}^{(2)})$ is given by the equation

$$\text{(4.2.1)} \qquad \Pr\,[e_{12}(\bar{x}^{(2)}) < z] = \Pr\left\{ N_2 Q_1 > 4N_2 [G^{-1}(z)]^2 \right\} (\tfrac{1}{2} \geq z \geq 0).$$

Now, $N_2 Q_1$ is a non-central chi-square with p degrees of freedom and non-centrality $\tfrac{1}{2} N_2 \delta^2$. Therefore, $\Pr[e_{12}(\bar{x}^{(2)}) < z]$ can be determined from tables of the non-central chi-square distributions.

It is interesting to note that $\mu^{(1)}$, $\mu^{(2)}$ and Σ enter the distribution function of $e_{12}(\bar{x}^{(2)})$ only in the form of δ.

<u>Expected value of</u> $e_{12}(\bar{x}^{(2)})$. In the previous section we saw that

$$(4.2.2) \qquad e_{12}(\bar{x}^{(2)}) = \int_{\frac{1}{2}\sqrt{Q}}^{\infty} g(\alpha)d\alpha .$$

The random variable $N_2 Q_1$ has the distribution of the non-central chi-square with p degrees of freedom and non-centrality $\frac{1}{2} N_2 \delta^2$. Therefore,

$$(4.2.3) \qquad \begin{aligned} E\, e_{12}(\bar{x}^{(2)}) &= N_2 \int_0^{\infty} \Delta_p(N_2 Q_1 ; \tfrac{1}{2} N_2 \delta^2)\left[\int_{\frac{1}{2}\sqrt{Q}}^{\infty} g(\alpha)d\alpha\right] dQ_1 , \\ &= \tfrac{1}{2} e^{-\frac{1}{2}N_2\delta^2} \sum_{r=0}^{\infty} \frac{(\frac{1}{2} N_2 \delta^2)^r}{r!} I_a(\tfrac{1}{2} p+r, \tfrac{1}{2}), \text{ where} \end{aligned}$$

$$(4.2.4) \qquad a = 4 N_2 /(4 N_2 + 1).$$

That is, $E\, e_{12}(\bar{x}^{(2)}) = \frac{1}{2}$(the probability that a non-central F with degrees of freedom p and one and non-centrality $\frac{1}{2} N_2 \delta^2$ is $\leq 4 N_2/p$.).

<u>The distribution of the probability of misclassifying an individual from</u> $P^{(2)}$. Till now we were concerned with the probability of assigning an individual from $P^{(1)}$ to $P^{(2)}$. We shall now consider the probability of wrongly assigning an individual from $P^{(2)}$ to $P^{(1)}$.

Given $\bar{x}^{(2)}$, the probability of misclassifying an individual from $P^{(2)}$ is $G(w)$, where

(4.2.5) $$w = \tfrac{1}{2}\sqrt{Q_1} - [(\mu^{(2)} - \mu^{(1)})\, \Sigma^{-1} (\bar{x}^{(2)} - \mu^{(1)})'] / \sqrt{Q_1}.$$

This probability we denote by $e_{21}(\bar{x}^{(2)})$. Obviously, $e_{21}(\bar{x}^{(2)})$ is a random variable. We shall derive its distribution.

The distribution function of $e_{21}(\bar{x}^{(2)})$ is given by the equation

(4.2.6) $$\Pr[e_{21}(\bar{x}^{(2)}) < z] = \Pr[w < G^{-1}(z)].$$

This equation shows that it is enough to find the distribution of w.

If we set

(4.2.7) $$t_1 = (\mu^{(2)} - \mu^{(1)})\, \Sigma^{-1} (\bar{x}^{(2)} - \mu^{(1)})'/\delta,$$

We have

(4.2.8) $$w = \tfrac{1}{2}\sqrt{Q_1} - (\delta t_1) / \sqrt{Q_1}.$$

As a first step in finding the distribution of w we find the distribution of Q_1 and t_1. In finding the joint distribution of Q_1, t_1 we may, without loss of generality, assume that $\mu^{(1)} = 0$ and $\Sigma = I$. The density function of $\bar{x}^{(2)}$ is then

(4.2.9) $$(\tfrac{1}{2} N_2 / \pi)^{\frac{1}{2}p} \exp [-\tfrac{1}{2} N_2 (\bar{x}^{(2)} - \mu^{(2)})(\bar{x}^{(2)} - \mu^{(2)})']$$
$$= (\tfrac{1}{2} N_2 / \pi)^{\frac{1}{2}} \exp [-\tfrac{1}{2} N_2 (Q_1 - 2\delta t_1 + \delta^2)].$$

Therefore, if we denote by $f(Q_1, t_1)$ the joint density function of Q_1 and t_1,

(4.2.10)

$$f(Q_1,t_1)dQ_1dt_1 = \underset{\substack{Q_1 < \bar{x}^{(2)}\bar{x}^{(2)'} < Q_1+dQ_1 \\ t_1 < (\mu^{(2)}\bar{x}^{(2)'})/\delta < t_1+dt_1}}{\int \cdots \int} (\tfrac{1}{2}N_2/\pi)^{\frac{1}{2}p}\exp[-\tfrac{1}{2}N_2(Q_1-2\delta t_1+\delta^2]d\bar{x}^{(2)},$$

$$= (\tfrac{1}{2} N_2/\pi)^{\frac{1}{2}p} \exp[-\tfrac{1}{2} N_2(Q_1-2\delta t_1 + \delta^2)] \underset{\substack{Q_1 < \bar{x}^{(2)} \bar{x}^{(2)'} < Q_1+ dQ_1 \\ t_1 < (\mu^{(2)} \bar{x}^{(2)'})/\delta < t_1+dt_1}}{\int \cdots\cdots \int} d\bar{x}^{(2)},$$

$$= \frac{(\tfrac{1}{2} N_2)^{\frac{1}{2}p}}{\sqrt{\pi}} \frac{(Q_1-t_1^2)^{\frac{1}{2}(p-3)}}{\Gamma(\tfrac{1}{2}[p-1])} \exp[-\tfrac{1}{2} N_2(Q_1-2\delta t_1 + \delta^2)]dQ_1\, dt_1 ,$$

since we have, if we use a result given in [18],

(4.2.11)

$$\underset{\substack{Q_1 < \bar{x}^{(2)} \bar{x}^{(2)'} < Q_1+dQ_1 \\ t_1 < (\mu^{(2)} \bar{x}^{(2)'})/\delta < t_1+dt_1}}{\int \cdots\cdots \int} d\bar{x}^{(2)} = \pi^{\frac{1}{2}(p-1)} \frac{(Q_1-t_1^2)^{\frac{1}{2}(p-3)}}{\Gamma(\tfrac{1}{2}[p-1])} dQ_1\, dt_1 .$$

Let $u = \sqrt{Q_1}$. The joint density of Q_1 and t_1 yields the joint density of u and w:

(4.2.12) $$\frac{u^{p-1}}{K}\left[1 - (w - \tfrac{1}{2}u)^2/\delta^2\right]^{\frac{1}{2}(p-3)} \exp\left[-\tfrac{1}{2}N_2(2uw + \delta^2)\right],$$

where

(4.2.13) $$K = \sqrt{\pi}\,\delta\, 2^{\frac{1}{2}p-1} N_2^{-\frac{1}{2}p}\, \Gamma(\tfrac{1}{2}[p-1]).$$

Integrating out u from (4.2.12), we obtain the density of w, $h(w)$ (say).

(4.2.14)
$$h(w) = \frac{e^{-\frac{1}{2}N_2\delta^2}}{K}\int_{2(w-\delta)}^{2(w+\delta)} \left[1-(w-\tfrac{1}{2}u)^2/\delta^2\right]^{\frac{1}{2}(p-3)} u^{p-1} e^{-N_2 uw}\, du, \quad \text{if } w \geq \delta,$$

$$= \frac{e^{-\frac{1}{2}N_2\delta^2}}{K}\int_{0}^{2(w+\delta)} \left[1-(w-\tfrac{1}{2}u)^2/\delta^2\right]^{\frac{1}{2}(p-3)} u^{p-1} e^{-N_2 uw}\, du, \quad \text{if } -\delta \leq w \leq \delta, \text{ and}$$

$$= 0, \quad \text{if } w < -\delta.$$

If p is odd, we may expand the expression within square brackets in the integrand and integrate each term by parts. For example, if $p = 3$,

$$h(w) = [(2N_2/\pi)^{\frac{1}{2}}/(\delta w)]e^{-\frac{1}{2}N_2\delta^2} \cdot$$

$$[\{2(w-\delta)^2 + 2(w-\delta)/(N_2 w) + (N_2 w)^{-2}\} \exp\{-2N_2 w(w-\delta)\}$$

$$- \{2(w+\delta)^2 + 2(w+\delta)/(N_2 w) + (N_2 w)^{-2}\} \exp\{-2N_2 w(w+\delta)\}]$$

if $w \geq \delta$,

(4.2.15) $$= [(2N_2/\pi)^{\frac{1}{2}}/(\delta w)]e^{-\frac{1}{2}N_2\delta^2} \cdot$$

$$[(N_2 w)^{-2} - \{2(w+\delta)^2 + 2(w+\delta)/(N_2 w) + (N_2 w)^{-2}\} e^{-2N_2 w(w+\delta)}]$$

if $-\delta \leq w \leq \delta$ but $w \neq 0$,

$$= \frac{4}{3}(2N_2^3/\pi)^{\frac{1}{2}}\delta^2 e^{-\frac{1}{2}N_2\delta^2}, \text{ if } w = 0, \text{ and}$$

$$= 0, \text{ if } w < -\delta.$$

For even values of p recourse must be had either to interpolation or to numerical integration.

Let us observe here that with probability one $w > -\delta$. Therefore, with probability one

(4.2.16) $$e_{21}(\bar{x}^{(2)}) > G(-\delta).$$

On the other hand, $e_{12}(\bar{x}^{(2)})$ can go down even to zero.

The limiting distribution of $e_{21}(\bar{x}^{(2)})$. Since the exact distribution of $e_{21}(\bar{x}^{(2)})$ is somewhat complicated, it may be useful to note that, as $N_2 \to \infty$, the distribution of $2N_2^{\frac{1}{2}}[e_{21}(\bar{x}^{(2)}) - G(-\frac{1}{2}\delta)]/g(\frac{1}{2}\delta)$ tends (weakly) to the normal distribution with mean zero and variance unity*.

Perhaps it is better to use the limiting distribution of w together with equation (4.2.6). The limiting distribution of $2N_2^{\frac{1}{2}}(w+\frac{1}{2}\delta)$ is normal with mean zero and variance unity. From this and equation (4.2.6) we deduce that

$$\Pr\left[e_{21}(\bar{x}^{(2)}) < z\right] \approx G(2N_2^{\frac{1}{2}}\left[G^{-1}(z) + \tfrac{1}{2}\delta\right]). \tag{4.2.17}$$

3. CASE TWO : ONLY Σ IS UNKNOWN

Since $\mu^{(1)}$ and $\mu^{(2)}$ are unknown, we construct the discriminant function using $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$. The classification procedure consists in assigning individuals to $P^{(1)}$ or $P^{(2)}$ according as

$$(\bar{x}^{(2)}-x^{(1)})\Sigma^{-1}x' \lesseqgtr \tfrac{1}{2}(\bar{x}^{(2)}-\bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(1)}+\bar{x}^{(2)})'. \tag{4.3.1}$$

Given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, the probability of misclassifying an individual from $P^{(1)}$ is $1 - G(u_1)$, where

$$u_1 = \tfrac{1}{2}\sqrt{Q} + \left[(\bar{x}^{(2)}-\bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(1)}-\mu^{(1)})'\right]/\sqrt{Q}. \tag{4.3.2}$$

We shall denote this probability by $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$. Similarly, let

---

* All the asymptotic results in this thesis can be derived by standard methods. We omit the intricate algebra.

$e_{21}(\bar{x}^{(1)}, \bar{x}^{(2)})$ denote the probability of misclassifying an individual from $P^{(2)}$. Being functions of random variables, $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$ and $e_{21}(\bar{x}^{(1)}, \bar{x}^{(2)})$ are random variables. We shall obtain the distribution of $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$. Since we are free to regard either of the two populations as $P^{(1)}$, it is not necessary to consider $e_{21}(\bar{x}^{(1)}, \bar{x}^{(2)})$ separately.

<u>The distribution of</u> $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$. Since

(4.3.3) $$\Pr\left[e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}) < z\right] = \Pr\left[u_1 > -G^{-1}(z)\right],$$

the distribution function of $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$ will be determined when the distribution of u_1 is obtained.

Given $(\bar{x}^{(2)} - \bar{x}^{(1)})$, the distribution of u_1 is normal with variance $(N_1 + N_2)^{-1}$ and mean $\frac{1}{2}(a_1-a_2)\sqrt{Q}+a_2\delta t_2/\sqrt{Q}$ ($= u_1'$, say), where

(4.3.4) $$t_2 = [(\mu^{(2)} - \mu^{(1)}) \Sigma^{-1} (\bar{x}^{(2)} - \bar{x}^{(1)})']/\delta.$$

Hence, if $h(u_1)$ denotes the density function of u_1, and $F(u_1')$ the density function of u_1',

(4.3.5) $$h(u_1) = \left(\frac{N_1+N_2}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} F(u_1')\exp\left[-\frac{N_1+N_2}{2}(u_1-u_1')^2\right]du_1'.$$

~~where the limits of integration are $-a_2\delta$ and infinity if $N_1 < N_2$ and $a_2\delta$ if $N_1 = N_2$ and $\frac{1}{2}\delta$ if $N_1 > N_2$.~~

The distribution of u_1' can be found, if we have the joint density of Q and t_2. The joint density of Q and t_2 can be found by the method by which we found the joint density of Q_1 and t_1 in section two and is

$$(4.3.6) \qquad \frac{(Q - t_2^2)^{\frac{1}{2}(p-3)}}{\sqrt{\pi}(2a_3)^{\frac{1}{2}p}\,\Gamma(\frac{1}{2}[p-1])} \exp\left[-\frac{1}{2a_3}(Q - 2\delta t_2 + \delta^2)\right].$$

From the joint density of Q and t_2 we find the joint density of Q and u_1'. Integrating out Q from the joint density of Q and u_1', we get the density function of u_1':

$$(4.3.7) \qquad \int \frac{(2a_3)^{-\frac{1}{2}p}\, e^{-\frac{1}{2}\delta^2/a_3}}{a_2\delta\sqrt{\pi}\,\Gamma(\frac{1}{2}[p-1])} \left[1 - \left\{ u_1' - \tfrac{1}{2}(a_1 - a_2)\sqrt{Q} \right\}^2/(a_2\delta)^2\right]^{\frac{1}{2}(p-3)} Q^{\frac{1}{2}p-1}$$

$$\exp\left[-\frac{1}{2a_2a_3}(a_1Q - 2u_1'\sqrt{Q})\right]dQ,$$

where the limits of integration are

$\left[\max\left\{0,\ 2(u_1' - a_2\delta)/(a_1 - a_2)\right\}\right]^2$ and $\left[\max\left\{0,\ 2(u_1' + a_2\delta)/(a_1 - a_2)\right\}\right]^2$

if $N_1 > N_2$,

$[\max \{0, 2(u_1'+a_2\delta)/(a_1-a_2)\}]^2$ and $[\max\{0, 2(u_1'-a_2\delta)/(a_1-a_2)\}]^2$

if $N_1 < N_2$, and zero and infinity if $N_1 = N_2$.

<u>Another expression for the density function of</u> u_1 <u>when</u> $N_1 = N_2$. By a method different from that of the previous paragraph, we shall derive here another expression for the density function of u_1. Given $(\bar{x}^{(2)} - \bar{x}^{(1)})$, u_1 has the normal distribution with variance $(2N)^{-1}$ and mean $\frac{1}{2}\delta t_2 / \sqrt{Q} = \frac{1}{2}\delta t$ (say). Therefore, the conditional characteristic function of u_1 is $\exp[\frac{1}{2} i\delta t\theta - \theta^2/(4N)]$. Hence, if $f(t)$ is the density function of t, and $\phi_{u_1}(\theta)$ the characteristic function of u_1,

$$(4.3.8) \qquad \phi_{u_1}(\theta) = \int_{-1}^{1} f(t) \exp[\tfrac{1}{2} i\delta t\theta - \theta^2/(4N)]dt .$$

The density function of t can be obtained from that of Q and t_2 and is

$$(4.3.9) \qquad \frac{e^{-N\delta^2/4}}{\sqrt{\pi}} \frac{(1-t^2)^{\frac{1}{2}(p-3)}}{\Gamma(\frac{1}{2}[p-1])} \sum_{r=0}^{\infty} \frac{\Gamma(\frac{1}{2}[p+r])}{r!} (\sqrt{N}\,\delta t)^r.$$

For the sake of convenience we now change over from u_1 to the variable v_1 defined by the equation

$$(4.3.10) \qquad v_1 = (2N)^{\frac{1}{2}} u_1.$$

Let $\phi_{v_1}(\theta)$ denote the characteristic function of v_1.

Then,

$$\phi_{v_1}(\theta) = \phi_{u_1}([2N]^{\frac{1}{2}}\theta),$$

$$= \int_{-1}^{1} \exp[i(N/2)^{\frac{1}{2}}\delta t\theta - \theta^2/2]f(t)dt,$$

$$= \pi^{-\frac{1}{2}}\left\{\Gamma(\tfrac{1}{2}[p-1])\right\}^{-1} \exp(-\tfrac{1}{4}N\delta^2 - \tfrac{1}{2}\theta^2)$$

(4.3.11)

$$\int_{-1}^{1}\left[\sum_{r=0}^{\infty} \frac{\Gamma(\frac{1}{2}[p+r])}{r!} N^{\frac{1}{2}r}(\delta t)^r(1-t^2)^{\frac{1}{2}(p-3)}\right.$$

$$\exp[i(\tfrac{1}{2}N)^{\frac{1}{2}}\delta t\theta]dt.$$

Expanding the exponential factor of the integrand and integrating term by term, we obtain the following equation for $\phi_{v_1}(\theta)$:

(4.3.12) $$\phi_{v_1}(\theta) = \pi^{-\frac{1}{2}} \exp(-\tfrac{1}{4}N\delta^2 - \tfrac{1}{2}\theta^2)$$

$$\sum{}' \frac{\Gamma(\frac{1}{2}[p+r])\,\Gamma(\frac{1}{2}[r+m+1])}{\Gamma(\frac{1}{2}[p+r+m])} \frac{(N\delta^2)^{\frac{1}{2}(r+m)}(-\frac{1}{2}\theta^2)^{\frac{1}{2}m}}{r!\,m!},$$

where $\sum'$ denotes summation over all non-negative integral values of r and m such that $r+m$ is an even integer.

The inversion formula for characteristic functions now readily yields an expression for the density function of v_1. This expression is

$$\frac{1}{\pi} \exp\left(-\frac{1}{4} N\delta^2 - \frac{1}{2} v_1^2\right)$$

$$(4.3.13) \quad \Sigma' \frac{\Gamma(\frac{1}{2}[p+r])\,\Gamma(\frac{1}{2}[r+m+1])}{\Gamma(\frac{1}{2}[p+r+m])} \; \frac{(N\delta^2)^{\frac{1}{2}(r+m)}\, 2^{-\frac{1}{2}(m+1)}}{r!\;\; m!} \cdot H_m(v_1).$$

Here $H_r(x)$ denotes the Hermite polynomial of degree r ~~defined~~ by the equation

$$(4.3.14) \qquad \left(-\frac{d}{dx}\right)^r e^{-\frac{1}{2}x^2} = H_r(x)\, e^{-\frac{1}{2}x^2}.$$

<u>The asymptotic distribution of</u> $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$. As N_1 and N_2 tend to infinity, the distribution of

$$(4.3.15) \qquad (2/\sqrt{a_3})\,[e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}) - G(-\tfrac{1}{2}\delta)] / g(\tfrac{1}{2}\delta)$$

tends weakly to the normal distribution with zero mean and unit variance.

Again it may be better to use the asymptotic distribution of u_1 together with equation (4.3.3). The limiting distribution of $(2/\sqrt{a_3})(u_1 - \frac{1}{2}\delta)$ is the normal distribution with zero mean and unit variance. Hence we have

(4.3.16) $$\Pr\left[e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}) < z\right] \approx G\left(2\left[G^{-1}(z) + \tfrac{1}{2}\delta\right] / a_3^{\frac{1}{2}}\right).$$

4. CASE THREE : $\mu^{(1)}, \mu^{(2)}, \Sigma$ ALL UNKNOWN

In this case the classification procedure is to assign individuals to $P^{(1)}$ or $P^{(2)}$ according as

(4.4.1) $$(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}x' \lessgtr \tfrac{1}{2}(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}(\bar{x}^{(1)} + \bar{x}^{(2)})'.$$

Given $\bar{x}^{(1)}, \bar{x}^{(2)}$ and S, the probability of misclassifying an individual from $P^{(1)}$ is $1 - G(w_1)$, where

(4.4.2) $$w_1 = \frac{(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}(\tfrac{1}{2}[\bar{x}^{(1)} + \bar{x}^{(2)}] - \mu^{(1)})'}{\sqrt{[(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}\Sigma S^{-1}(\bar{x}^{(2)} - \bar{x}^{(1)})']}}.$$

We shall denote this probability by $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$. The exact distribution of $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ being complicated, we shall be content with giving its asymptotic distribution. As N_1, N_2 and n tend to infinity, the distribution of

$$(4.4.3)\qquad (2/\sqrt{a_3})[\, e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}) - G(-\tfrac{1}{2}\delta)] \,/g(\tfrac{1}{2}\delta)$$

tends (weakly) to the normal distribution with mean zero and variance unity. We have also, corresponding to equations (4.2.17) and (4.3.16), the equation

$$(4.4.4)\qquad \Pr\,[e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}; S) < z] \approx G(2[G^{-1}(z) + \tfrac{1}{2}\delta]/\sqrt{a_3}).$$

Chapter Five

EXTENSION OF THE RESULTS OF CHAPTER FOUR

1. INTRODUCTION

There are situations where the two kinds of errors are of unequal importance. In some cases it may even be possible to determine the different losses consequent on each type of mistake. Suppose c_{12} is the loss incurred in assigning an individual from $P^{(1)}$ to $P^{(2)}$ and c_{21} is the loss incurred in assigning an individual from $P^{(2)}$ to $P^{(1)}$. Suppose, further, that the a priori probability that the individual to be classified belongs to $P^{(k)}$ is $\pi^{(k)}$ $(k = 1, 2)$. Then, as we stated in chapter three, the classification procedure for which the expected loss is a minimum is that of assigning individuals to $P^{(1)}$ or $P^{(2)}$ according as

$$(\mu^{(2)} - \mu^{(1)}) \Sigma^{-1} x' \lesseqgtr \frac{1}{2} (\mu^{(2)} - \mu^{(1)}) \Sigma^{-1} (\mu^{(1)} + \mu^{(2)})' + c,$$

where

$$c = \log_e \frac{\pi^{(1)} c_{12}}{\pi^{(2)} c_{21}} .$$

The procedure we were considering till now was that corresponding to the value zero of c. A sufficient condition for c to be zero is that

$c_{12} = c_{21}$ and $\pi^{(1)} = \pi^{(2)}$, which is the case in many important problems.

The procedure mentioned above can be carried out only if all the parameters $\mu^{(1)}$, $\mu^{(2)}$ and Σ are known. If such is not the case, we may assign the individuals to $P^{(1)}$ or $P^{(2)}$ according as

$$(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}x' \lesseqgtr \frac{1}{2}(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}(\bar{x}^{(1)} + \bar{x}^{(2)})' + c.$$

The sampling fluctuations of $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ and $e_{21}(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ are again of interest. We shall in this chapter give indications of the changes to be made in some of the earlier formulae. They can be derived in the same way as the results of the earlier chapters.

2. THE EXTENDED RESULTS

The distribution of $e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)})$ is given by the equation

$$(5.2.1) \quad \Pr[e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = \Pr[u_1 > -G^{-1}(z)],$$

where

$$(5.2.2) \quad u_1 = \tfrac{1}{2}[(\bar{x}^{(2)} - \bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(2)} - \bar{x}^{(1)})']^{\frac{1}{2}}$$
$$+ [(\bar{x}^{(2)} - \bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(2)} - \bar{x}^{(1)})']^{-\frac{1}{2}}[(\bar{x}^{(2)} - \bar{x}^{(1)})\Sigma^{-1}(\bar{x}^{(1)} - \mu^{(1)})' + c$$

This equation shows that it is enough to obtain the distribution of u_1. Equation (4.3.5) of chapter four still gives the density function $h(u_1)$ of u_1, provided we take

$$F(u_1') = \int \frac{(2a_3)^{-\frac{1}{2}p}\, e^{-\frac{1}{2}\delta^2/a_3}}{a_2\delta\pi^{\frac{1}{2}}\,\Gamma(\frac{1}{2}p - \frac{1}{2})} [1 - \{u_1' - cQ^{-\frac{1}{2}} -$$

(5.2.3)

$$- \tfrac{1}{2}(a_1 - a_2)Q^{\frac{1}{2}}\}^2/(a_2\delta)^2]^{\frac{1}{2}(p-3)}$$

$$Q^{\frac{1}{2}p-1} \exp[-\frac{1}{2a_2a_3}(a_1Q - 2u_1'Q^{\frac{1}{2}} + 2c)]dQ.$$

If $N_1 \neq N_2$ and

$$q_1(u_1') = \frac{u_1'}{a_1-a_2} + \frac{a_2\,\delta}{|a_1-a_2|} - [(\frac{u_1'}{a_1-a_2} + \frac{a_2\,\delta}{|a_1-a_2|})^2 - \frac{2c}{a_1-a_2}]^{\frac{1}{2}},$$

$$q_2(u_1') = \frac{u_1'}{a_1-a_2} - \frac{a_2\,\delta}{|a_1-a_2|} - [(\frac{u_1'}{a_1-a_2} - \frac{a_2\,\delta}{|a_1-a_2|})^2 - \frac{2c}{a_1-a_2}]^{\frac{1}{2}},$$

$$q_3(u_1') = \frac{u_1'}{a_1-a_2} - \frac{a_2\,\delta}{|a_1-a_2|} + [(\frac{u_1'}{a_1-a_2} - \frac{a_2\,\delta}{|a_1-a_2|})^2 - \frac{2c}{a_1-a_2}]^{\frac{1}{2}},$$

$$q_4(u_1') = \frac{u_1'}{a_1-a_2} + \frac{a_2\,\delta}{|a_1-a_2|} + [(\frac{u_1'}{a_1-a_2} + \frac{a_2\,\delta}{|a_1-a_2|})^2 - \frac{2c}{a_1-a_2}]^{\frac{1}{2}},$$

the integration in (5.2.3) is over the interval $[q_3^2(u_1'), q_4^2(u_1')]$ if $c/(a_1 - a_2) \leq 0$, over the interval $[q_1^2(u_1'), q_4^2(u_1')]$ if $c/(a_1 - a_2) > 0$ and

$$\left(\frac{2c}{a_1 - a_2}\right)^{\frac{1}{2}} - \frac{a_2\,\delta}{|a_1 - a_2|} \leq \frac{u_1'}{a_1 - a_2} \leq \left(\frac{2c}{a_1 - a_2}\right)^{\frac{1}{2}} + \frac{a_2\,\delta}{|a_1 - a_2|},$$

and over the union of the intervals

$$[q_1^2(u_1'),\ q_2^2(u_1')] \quad \text{and} \quad [q_3^2(u_1'),\ q_4^2(u_1')], \quad \text{if} \quad c/(a_1 - a_2) > 0$$

and

$$\frac{u_1'}{a_1 - a_2} > \left(\frac{2c}{a_1 - a_2}\right)^{\frac{1}{2}} + \frac{a_2\,\delta}{|a_1 - a_2|};$$

if $c/(a_1 - a_2) > 0$, u_1' does not take values for which

$$\frac{u_1'}{a_1 - a_2} < \left(\frac{2c}{a_1 - a_2}\right)^{\frac{1}{2}} - \frac{a_2\,\delta}{|a_1 - a_2|}.$$

If $N_1 = N_2$ but $c \neq 0$, the upper and lower limits of integration in (5.2.3) are

$$\left[\max\left\{0,\ \frac{u_1'}{c} + \frac{a_2\,\delta}{|c|}\right\}\right]^{-2} \quad \text{and} \quad \left[\max\left\{0,\ \frac{u_1'}{c} - \frac{a_2\,\delta}{|c|}\right\}\right]^{-2}$$

respectively. [If both limits of integration are infinity, take the integral to be zero.] If $N_1 = N_2$ and $c = 0$, the limits of integration are zero and infinity.

Corresponding to equation (4.4.4) of Chapter four we have the equation

$$\Pr\,[e_{12}(\bar{x}^{(1)}, \bar{x}^{(2)}; S) < z]$$

$$(5.2.4) \qquad \approx G(2\delta^2[(\delta^2+2c)^2 N_1^{-1} + (\delta^2-2c)^2 N_2^{-1} + 2(2c\delta)^2 n^{-1}]^{-\frac{1}{2}} [G^{-1}(z) + c/\delta + \tfrac{1}{2}\delta]).$$

If we have more information about $P^{(1)}$ and $P^{(2)}$, we can obtain simpler results. Thus, if besides Σ, $\mu^{(1)}$ also is known, the distribution function of $e_{12}(\bar{x}^{(2)})$ is given by the equation

$$\Pr[e_{12}(\bar{x}^{(2)}) < z]$$

$$(5.2.5) \qquad \begin{aligned} &= \Pr[v < A_1(z)] + \Pr[v > A_2(z)], \text{ if } z < 1-G([2c]^{\frac{1}{2}}), \\ &= 1 \quad , \text{ if } z \geq 1 - G([2c]^{\frac{1}{2}}) \quad (\text{where } c \geq 0); \end{aligned}$$

where v is a noncentral chi-square with p degrees of freedom and noncentrality $\frac{1}{2} N_2\delta^2$;

$$(5.2.6) \qquad \begin{aligned} A_1(z) &= N_2[G^{-1}(z) + \left\{ [G^{-1}(z)]^2 - 2c \right\}^{\frac{1}{2}}]^2, \\ A_2(z) &= N_2[G^{-1}(z) - \left\{ [G^{-1}(z)]^2 - 2c \right\}^{\frac{1}{2}}]^2. \end{aligned}$$

Similarly, we have for the distribution function of $e_{21}(\bar{x}^{(2)})$ the equation

$$(5.2.7) \qquad \Pr\,[e_{21}(\bar{x}^{(2)}) < z] = \Pr\,[w < G^{-1}(z)],$$

where w is a random variable having the density function $h(w)$ defined below:

$$h(w) = \frac{e^{-N_2(\frac{1}{2}\delta^2-c)}}{K}\left[\int_{m_1(w)}^{m_2(w)} + \int_{m_3(w)}^{m_4(w)}\right][1-(w-\tfrac{1}{2}u-c/u)^2/\delta^2]^{\frac{1}{2}(p-3)}$$

$$u^{p-1}\, e^{-N_2uw}\, du, \quad \text{if} \quad w \geq \delta+(2c)^{\frac{1}{2}},$$

(5.2.8)

$$= \frac{1}{K}\, e^{-N_2(\frac{1}{2}\delta^2-c)} \int_{m_1(w)}^{m_4(w)} [1-(w-\tfrac{1}{2}u-c/u)^2/\delta^2]^{\frac{1}{2}(p-3)}\, u^{p-1}\, e^{-N_2uw}\, du,$$

$$\text{if} \quad (2c)^{\frac{1}{2}} - \delta \leq w < (2c)^{\frac{1}{2}} + \delta,$$

$$= 0 \quad , \quad \text{if} \quad w < (2c)^{\frac{1}{2}} - \delta \quad (\text{when } c \geq 0);$$

(5.2.9) $$m_1(w) = w + \delta - [(w + \delta)^2 - 2c]^{\frac{1}{2}};$$

$$m_2(w) = w - \delta - [(w - \delta)^2 - 2c]^{\frac{1}{2}};$$

$$m_3(w) = w - \delta + [(w - \delta)^2 - 2c]^{\frac{1}{2}};$$

$$m_4(w) = w + \delta + [(w + \delta)^2 - 2c]^{\frac{1}{2}};$$

K is the K of equation (4.2.13) of chapter four.

Observe that with probability one

(5.2.10) $$e_{12}(\bar{x}^{(2)}) \leq 1 - G([2c]^{\frac{1}{2}}) \qquad (c \geq 0), \text{ and}$$

(5.2.11) $$e_{21}(\bar{x}^{(2)}) \geq G([2c]^{\frac{1}{2}} - \delta) \qquad (c \geq 0).$$

If $c < 0$, we have

(5.2.12) $$\Pr\left[q_{12}(\bar{x}^{(2)}) < z\right] = \Pr\left[v > A_2(z)\right],$$

instead of equation (5.2.5), and for the density function $h(w)$ of w, the equation

$$h(w) = \frac{1}{K}\exp\left(cN_2 - \tfrac{1}{2}N_2\,\delta^2\right)$$

(5.2.13) $$\int_{m_3(w)}^{m_4(w)} \left[1-(w-\tfrac{1}{2}u-c/u)^2/\delta^2\right]^{\frac{1}{2}(p-3)} u^{p-1}\, e^{-N_2uw}\, du$$

instead of equation (5.2.8).

Chapter Six

EXTENSION OF THE RESULTS OF CHAPTER FIVE

1. INTRODUCTION

In chapters three to five we studied the sampling behaviour of the probability of misclassifying an individual from $P^{(1)}$ or $P^{(2)}$ when individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according as

$$(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}x' \lessgtr \frac{1}{2}(\bar{x}^{(2)} - \bar{x}^{(1)})S^{-1}(\bar{x}^{(1)} + \bar{x}^{(2)})' + c.$$

Here we inquire what is the probability of assigning to $P^{(k)}$ $(k=1,2)$ an individual from P. The methods employed are essentially the same as those of chapters three to five, but the results are less simple.

We shall denote the probability of assigning an individual to $P^{(k)}$ by $e_k(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$. If Σ is known, we use Σ, rather than S, and write $e_k(\bar{x}^{(1)}, \bar{x}^{(2)})$ for the probability of assigning the individual to $P^{(k)}$.

2. THE DISTRIBUTION OF $e_2(\mu^{(1)}, \bar{x}^{(2)})$

By a simple calculation we find that

(6.2.1) $$e_2(\mu^{(1)}, \bar{x}^{(2)}) = 1 - G(w_1),$$

where

(6.2.2) $$w_1 = \frac{1}{2}[(\bar{x}^{(2)} - \mu^{(1)})\Sigma^{-1}(\bar{x}^{(2)} - \mu^{(1)})']^{\frac{1}{2}}$$

$$+ [(\bar{x}^{(2)} - \mu^{(1)})\Sigma^{-1}(\bar{x}^{(2)} - \mu^{(1)})']^{-\frac{1}{2}}[(\mu^{(1)} - \mu)\Sigma^{-1}(\bar{x}^{(2)} - \mu^{(1)})' + c].$$

Therefore,

(6.2.3) $$\Pr[e_2(\mu^{(1)}, \bar{x}^{(2)}) < z] = \Pr[w_1 > -G^{-1}(z)].$$

We require the distribution of w_1.

If we set

(6.2.4) $$\beta_1 = (\mu^{(1)} - \mu)\Sigma^{-1}(\mu^{(1)} - \mu)',$$

(6.2.5) $$Q_1 = (\bar{x}^{(2)} - \mu^{(1)})\Sigma^{-1}(\bar{x}^{(2)} - \mu^{(1)})' \text{ and}$$

$$T_1 = [(\mu^{(1)} - \mu)\Sigma^{-1}(\bar{x}^{(2)} - \mu^{(1)})'] / \beta_1^{\frac{1}{2}},$$

(6.2.6) $$w_1 = \frac{1}{2} Q_1^{\frac{1}{2}} + (c + \beta_1^{\frac{1}{2}} T_1)/Q_1^{\frac{1}{2}}.$$

We shall suppose that $\mu \neq \mu^{(1)}$. [The case $\mu = \mu^{(1)}$ has been considered in chapter three]. Then $\beta_1 \neq 0$. Let

$$(6.2.7) \qquad \beta_2 = (\mu^{(1)}-\mu)\Sigma^{-1}(\mu^{(2)}-\mu^{(1)})', \; \beta_3 = \delta^2 - \beta_2^2 / \beta_1 \; (\neq 0),$$

$$T_1' = [(\mu^{(2)}-\mu^{(1)})\Sigma^{-1}(\bar{x}^{(2)}-\mu^{(1)})' - \beta_2 T_1 / \beta_1^{\frac{1}{2}}] / \beta_3^{\frac{1}{2}} .$$

The joint density of Q_1, T_1 and T_1' can be found by the method by which we found the joint density of Q_1 and t_1 in section two of chapter four and is

$$(6.2.8) \qquad \frac{(Q_1-T_1^2-T_1'^2)^{\frac{1}{2}p-2}}{\pi(2/N_2)^{\frac{1}{2}p}\,\Gamma(\frac{1}{2}p-1)} \exp\left[-\frac{N_2}{2}(Q_1-2\beta_4 T_1-2\beta_5 T_1' + \delta^2)\right],$$

where

$$(6.2.9) \qquad \beta_4 = \beta_2 / \sqrt{\beta_1} \; , \quad \beta_5 = \sqrt{\beta_3} \; .$$

From the joint density of Q_1, T_1 and T_1' we find the joint density of Q_1, T_1 and T_1'' $(= [Q_1 - T_1^2]^{-\frac{1}{2}} T_1')$. Integrating the joint density of Q_1, T_1 and T_1'' with respect to T_1'' from -1 to $+1$, we get the joint density of Q_1 and T_1 :

$$(6.2.10)\quad \frac{(Q_1 - T_1^2)^{\frac{1}{2}(p-3)}}{(2/N_2)^{\frac{1}{2}p}\sqrt{\pi}} \exp\left[-\frac{N_2}{2}(Q_1 - 2\beta_4 T_1 + \delta^2)\right]$$

$$\sum_{k=0}^{\infty} \frac{(\frac{1}{2} N_2\beta_5)^{2k} (Q_1 - T_1^2)^k}{\Gamma(k+1)\ \Gamma(\frac{1}{2}[p+2k-1])} .$$

From the joint density of Q_1 and T_1 we find the joint density of Q_1 and w_1. Integrating out Q_1 from the joint density of Q_1 and w_1, we get the density function of w_1. The density function of w_1 is

$$(6.2.11)\quad \int \frac{Q_1^{\frac{1}{2}}}{(2/N_2)^{\frac{1}{2}p}(\beta_1\pi)^{\frac{1}{2}}} \sum_{k=0}^{\infty} \frac{(\frac{1}{2} N_2\beta_5)^{2k}}{k!} \frac{[Q_1-(w_1Q_1^{\frac{1}{2}}-\frac{1}{2}Q_1-c)^2/\beta_1]^{\frac{1}{2}(p+2k-3)}}{\Gamma(\frac{1}{2}p+k-\frac{1}{2})}$$

$$\exp\left[-\frac{1}{2} N_2\left\{ Q_1-2\beta_4(w_1Q_1^{\frac{1}{2}}-\frac{1}{2}Q_1-c)/\beta_1^{\frac{1}{2}}+ \delta^2 \right\}\right]dQ_1,$$

where the region of integration is the interval

$$([\,w_1-\beta_1^{\frac{1}{2}} + \{(w_1-\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2, [w_1+ \beta_1^{\frac{1}{2}}+ \{(w_1+\sqrt{\beta_1})^2- 2c\}^{\frac{1}{2}}]^2)$$

if $c < 0$, the interval

$$([w_1+\beta_1^{\frac{1}{2}}-\{(w_1+\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2,\ [w_1+\beta_1^{\frac{1}{2}}+\{(w_1+\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2)$$

if $c \geq 0$ and $(2c)^{\frac{1}{2}} - \beta_1^{\frac{1}{2}} \leq w_1 \leq (2c)^{\frac{1}{2}} + \beta_1^{\frac{1}{2}}$, and the union of the intervals

$$([w_1-\beta_1^{\frac{1}{2}}-\{(w_1-\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2, [w_1+\beta_1^{\frac{1}{2}}-\{(w_1+\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2)$$

and

$$([w_1-\beta_1^{\frac{1}{2}}+\{(w_1-\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2,\ [w_1+\beta_1^{\frac{1}{2}}+\{(w_1+\sqrt{\beta_1})^2-2c\}^{\frac{1}{2}}]^2)$$

if $c \geq 0$ and $w_1 > (2c)^{\frac{1}{2}} + \beta_1^{\frac{1}{2}}$; if $c \geq 0$, $w_1 \geq (2c)^{\frac{1}{2}} - \beta_1^{\frac{1}{2}}$ with probability one.

3. THE DISTRIBUTION OF $e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$

(6.3.1) $$e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) = 1 - G(u), \text{ where}$$

(6.3.2) $$u = \frac{1}{2}\,[(\bar{x}^{(2)}-\bar{x}^{(1)})\,\Sigma^{-1}(\bar{x}^{(2)}-\bar{x}^{(1)})']^{\frac{1}{2}}$$

$$+ [(\bar{x}^{(2)}-\bar{x}^{(1)})\,\Sigma^{-1}(\bar{x}^{(2)}-\bar{x}^{(1)})']^{-\frac{1}{2}}$$

$$[(\bar{x}^{(2)}-\bar{x}^{(1)})\,\Sigma^{-1}(\bar{x}^{(1)}-\mu^{(1)})' + c].$$

Therefore,

$$\text{(6.3.3)} \qquad \Pr\,[e_2(\bar{\bar{x}}^{(1)}, \bar{\bar{x}}^{(2)}) < z] = \Pr\,[\,u > -G^{-1}(z)].$$

We require the distribution of u.

The distribution of u, given $(\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})$, is normal with variance $(N_1 + N_2)^{-1}$ and mean

$$\text{(6.3.4)} \quad w = \frac{1}{2}(a_1 - a_2)[(\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)}) \Sigma^{-1} (\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})']^{\frac{1}{2}}$$

$$+ [(\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)}) \Sigma^{-1} (\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})']^{-\frac{1}{2}}$$

$$[(a_1\mu^{(1)} + a_2\mu^{(2)} - \mu) \Sigma^{-1} (\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})' + c].$$

Let

$$\text{(6.3.5)} \qquad Q = (\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)}) \Sigma^{-1} (\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})',$$

$$\text{(6.3.6)} \qquad C_1 = (a_1\mu^{(1)} + a_2\mu^{(2)} - \mu) \Sigma^{-1} (a_1\mu^{(1)} + a_2\mu^{(2)} - \mu)',$$

$$\text{(6.3.7)} \qquad T = (a_1\mu^{(1)} + a_2\mu^{(2)} - \mu) \Sigma^{-1} (\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})' / C_1^{\frac{1}{2}}.$$

Then

$$\text{(6.3.8)} \qquad w = \frac{1}{2}(a_1 - a_2)Q^{\frac{1}{2}} + Q^{-\frac{1}{2}}(C_1^{\frac{1}{2}}\,T + c).$$

Let $h(u)$ denote the density function of u, and $F(w)$ of w. Then

(6.3.9) $$h(u) = \left(\frac{N_1+N_2}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} F(w) \exp\left[-\frac{1}{2}(N_1+N_2)(u-w)^2\right]dw.$$

Let

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

$$C_3 = \delta^2 - C_2^2 / C_1,$$

(6.3.11) $$T' = [(\mu^{(2)} - \mu^{(1)})\,\Sigma^{-1}(\bar{x}^{(2)} - \bar{x}^{(1)})' - C_1^{-\frac{1}{2}} C_2 T] / C_3^{\frac{1}{2}}.$$

To find the density function of w, first we find the joint density of Q, T and T'. If C_1 and C_2 are both different from zero, the joint density of Q, T and T' can be found by the method by which the joint density of Q_1 and t_1 was found in section two of chapter four and is

(6.3.12) $$\frac{(Q-T^2-T'^2)^{\frac{1}{2}p-2}}{\pi(2a_3)^{\frac{1}{2}p}\,\Gamma(\frac{1}{2}p-1)} \exp\left[-\frac{1}{2a_3}(Q-2C_4T-2C_5T' + \delta^2)\right],$$

where

(6.3.13) $$C_4 = C_2 / \sqrt{C_1}, \quad \text{and} \quad C_5 = \sqrt{C_3}.$$

From the joint density of Q, T and T' we find the joint density of Q, T and T'' $(= (Q - T^2)^{-\frac{1}{2}} T')$. Integrating the joint density of Q, T and T'' with respect to T'', we get the joint density of Q and T:

$$(6.3.14) \qquad \frac{(Q-T^2)^{\frac{1}{2}(p-3)}}{(2a_3)^{\frac{1}{2}p}\sqrt{\pi}} \exp\left[-\frac{1}{2a_3}(Q-2C_4T+\delta^2)\right] \sum_{k=0}^{\infty} \frac{(\frac{1}{2}C_5/a_3)^{2k}(Q-T^2)^k}{\Gamma(k+1)\,\Gamma(\frac{1}{2}[p+2k-1])}.$$

From the joint density of Q and T we find the joint density of Q and w. Integrating the joint density of Q and w with respect to Q ~~from zero to infinity~~, we get the density function of w. The density function of w is

$$(6.3.15) \qquad \frac{(2a_3)^{-\frac{1}{2}p}}{(C_1\pi)^{\frac{1}{2}}} \sum_{k=0}^{\infty} \frac{\left[Q-\left\{wQ^{\frac{1}{2}}-\frac{1}{2}(a_1-a_2)Q-c\right\}^2/C_1\right]^{\frac{1}{2}(p+2k-3)}}{k!\,\Gamma(\frac{1}{2}p+k-\frac{1}{2})}\left(\tfrac{1}{2}C_5/a_3\right)^{2k}$$

$$Q^{\frac{1}{2}} \exp\left[-\frac{1}{2a_3}\left\{Q-2C_4\left[wQ^{\frac{1}{2}}-\frac{1}{2}(a_1-a_2)Q-c\right]/C_1^{\frac{1}{2}}+\delta^2\right\}\right].$$

If $N_1 \neq N_2$ and

$$q_1(w) = \frac{w}{a_1-a_2} + \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} - \left[\left\{ \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} + \frac{w}{a_1-a_2} \right\}^2 - \frac{2c}{a_1-a_2} \right]^{\frac{1}{2}},$$

$$q_2(w) = \frac{w}{a_1-a_2} - \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} - \left[\left\{ \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} - \frac{w}{a_1-a_2} \right\}^2 - \frac{2c}{a_1-a_2} \right]^{\frac{1}{2}},$$

$$q_3(w) = \frac{w}{a_1-a_2} - \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} + \left[\left\{ \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} - \frac{w}{a_1-a_2} \right\}^2 - \frac{2c}{a_1-a_2} \right]^{\frac{1}{2}},$$

and

$$q_4(w) = \frac{w}{a_1-a_2} + \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} + \left[\left\{ \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} + \frac{w}{a_1-a_2} \right\}^2 - \frac{2c}{a_1-a_2} \right]^{\frac{1}{2}},$$

The integration in (6.3.15) is over the interval $[q_3^2(w), q_4^2(w)]$ if $c/(a_1 - a_2) \leq 0$, over the interval $[q_1^2(w), q_4^2(w)]$ if $c/(a_1 - a_2) > 0$ and

$$\left(\frac{2c}{a_1-a_2}\right)^{\frac{1}{2}} - \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|} < \frac{w}{a_1-a_2} \leq \left(\frac{2c}{a_1-a_2}\right)^{\frac{1}{2}} + \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|},$$

and over the union of the two intervals $[q_1^2(w), q_2^2(w)]$, $[q_3^2(w), q_4^2(w)]$ if $c/(a_1 - a_2) > 0$ and

$$\frac{w}{a_1-a_2} > \left(\frac{2c}{a_1-a_2}\right)^{\frac{1}{2}} + \frac{c_1^{\frac{1}{2}}}{2|a_1-a_2|}; \text{ if } c/(a_1 - a_2) > 0, w$$

does not take values for which

$$\frac{w}{a_1 - a_2} < \left(\frac{2c}{a_1 - a_2} \right)^{\frac{1}{2}} - \frac{C_1^{\frac{1}{2}}}{|a_1 - a_2|}.$$

If $N_1 = N_2$ but $c \neq 0$, the upper and lower limits of integration in (6.3.15) are respectively

$$\left[\max\left(0, \frac{w}{c} + \frac{C_1^{\frac{1}{2}}}{|c|}\right)\right]^{-2} \text{ and } \left[\max\left(0, \frac{w}{c} - \frac{C_1^{\frac{1}{2}}}{|c|}\right)\right]^{-2} .$$

[The integral is to be taken to be zero if both the limits of integration are infinity.] If $N_1 = N_2 = N$ and $c = 0$, the limits of integration in (6.3.15) are zero and infinity, and the imtegral simplifies to

(6.3.16)
$$\frac{e^{-N\delta^2/4}}{(C_1 \pi)^{\frac{1}{2}}} \sum_{k=o}^{\infty} \sum_{r=o}^{\infty} \Gamma\left(\tfrac{1}{2}p + k + \tfrac{1}{2}r\right)$$

$$N^k \frac{(NC_4^2)^{\text{[illegible]}} C_5^{2k} (1 - w^2/C_1)^{\frac{1}{2}p+k-3/2} (w/\sqrt{C})^r}{2^{2k} r! \, k! \, \Gamma\left(\frac{1}{2}p + k - \frac{1}{2}\right)} .$$

In this case we have also the equation

$$(6.3.17) \qquad h(u) = \frac{N^{\frac{1}{2}}}{\pi} \exp\left[-N(u^2 + \delta^2/4)\right]$$

$$\sum_{k=o}^{\infty} \sum{}' \frac{\Gamma(\frac{1}{2}p + k + \frac{1}{2}r)\, \Gamma(\frac{1}{2}[r+m+1])}{\Gamma(\frac{1}{2}p+\frac{1}{2}r\frac{1}{2}m+k)}$$

$$\frac{(2C_1)^{\frac{1}{2}m}\, C_4^r\, C_5^{2k}\, N^{k+\frac{1}{2}(r+m)}}{4^k r!\, m!\, k!} \; H_m(2^{\frac{1}{2}} N^{\frac{1}{2}} u)$$

corresponding to equation (4.3.13) of chapter four. [$\sum'$ denotes summation with respect to r and m over all pairs of non-negative integral values of r and m such that r+m is even.]

In deriving the joint density of Q, T and T', we had expressly assumed that neither C_1 nor C_3 was equal to zero. Let us consider the case of zero C_1. In this case T and T ' are not defined. This causes no difficulty. When $C_1 = o$, the distribution of u, given $\bar{x}^{(2)} - \bar{x}^{(1)}$, is normal with mean $\frac{1}{2}(a_1 - a_2)Q^{\frac{1}{2}} + cQ^{-\frac{1}{2}}$ and variance $(N_1 + N_2)^{-1}$. The random variable Q/a_3 is distributed as a non-central chi-square with p degrees of freedom and non-centrality $\delta^2/(2a_3)$. Therefore,

$$(6.3.18) \qquad h(u) = \left(\frac{N_1+N_2}{2\pi a_3^2}\right)^{\frac{1}{2}} \int_0^{\infty} \Delta_p\, (Q/a_3; \tfrac{1}{2}\delta^2/a_3)$$

$$\exp\left[-\frac{N_1+N_2}{2}\left\{u - \tfrac{1}{2}(a_1 - a_2)\sqrt{Q} - cQ^{-\frac{1}{2}}\right\}^2\right] dQ.$$

If $c = 0$, this simplifies to the equation

$$(6.3.19)\qquad h(u) = \left(\frac{N_1+N_2}{2\pi}\right)^{\frac{1}{2}} \frac{e^{-\frac{1}{2}\delta^2/a_3}}{a_3^{\frac{1}{2}p}} e^{-\frac{1}{2}(N_1+N_2)u^2} \sum_{r=0}^{\infty} k_r \frac{u^r}{r!},$$

where

$$(6.3.20)\qquad a = 1 + (N_1 - N_2)^2/(4N_1N_2) \text{ and}$$

$$(6.3.21)\qquad k_r = \left(\frac{a_3}{2a}\right)^{\frac{1}{2}r}(N_1-N_2)^r \sum_{m=0}^{\infty} \frac{\Gamma(\frac{1}{2}p+m+\frac{1}{2}r)}{m!\,\Gamma(\frac{1}{2}p+m)}\left(\frac{\frac{1}{2}\delta^2}{a a_3}\right)^m \quad (r=0,1,2,\ldots).$$

We also note that, if, further, $N_1 = N_2 (= N$, say), the marginal distribution of u is normal with zero mean and variance $(2N)^{-1}$.

A third case is that where $C_1 \neq 0$ but $C_3 = 0$. This is the case, if and only if there exists a scalar γ (say) such that $a_1\mu^{(1)} + a_2\mu^{(2)} - \mu = \gamma(\mu^{(2)} - \mu^{(1)})$. In this case

$$(6.3.22)\qquad w = \tfrac{1}{2}\sqrt{Q} + \gamma\, C_1^{\frac{1}{2}}\, t_2, \text{ where}$$

$$(6.3.23)\qquad t_2 = (\mu^{(2)} - \mu^{(1)})\, \Sigma^{-1}\, (\bar{x}^{(2)} - \bar{x}^{(1)})'/\delta.$$

Using the distribution of Q and t_2 obtained in section three of chapter four, we can determine the distribution of w. The result coincides with that obtained by putting $C_5 = 0$ in (6.3.15).

4. THE DISTRIBUTION OF $e_2(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$

If N_1, N_2 and n are large, we have

$$\Pr[e_2(\bar{x}^{(1)},\bar{x}^{(2)};S)< z] \approx G([d_1N_1^{-1}+d_2N_2^{-1}+d_3n^{-1}]^{-\frac{1}{2}}[G^{-1}(z)+\theta]), \tag{6.4.1}$$

where

$$\theta = \tfrac{1}{2}\delta + [(\mu^{(2)}-\mu^{(1)})\Sigma^{-1}(\mu^{(1)}-\mu)'+c]/\delta,$$

$$d_1=[\gamma_1(\mu^{(2)}-\mu^{(1)})+\mu^{(1)}-\mu]\Sigma^{-1}[\gamma_1(\mu^{(2)}-\mu^{(1)})+\mu^{(1)}-\mu]'/\delta^2,$$

$$d_2=[\gamma_2(\mu^{(2)}-\mu^{(1)})+\mu^{(1)}-\mu]\Sigma^{-1}[\gamma_2(\mu^{(2)}-\mu^{(1)})+\mu^{(1)}-\mu]'/\delta^2, \tag{6.4.2}$$

$$d_3=\frac{1}{\delta^2}\Big\{[(\mu^{(2)}-\mu^{(1)})\Sigma^{-1}[\tfrac{1}{2}(\mu^{(1)}+\mu^{(2)})-\mu-\gamma_3(\mu^{(2)}-\mu^{(1)})]'\Big\}^2$$
$$+[\tfrac{1}{2}(\mu^{(1)}+\mu^{(2)})-\mu-\gamma_3(\mu^{(2)}-\mu^{(1)})]\Sigma^{-1}$$
$$[\tfrac{1}{2}(\mu^{(1)}+\mu^{(2)})-\mu-\gamma_3(\mu^{(2)}-\mu^{(1)})]';$$

$$(6.4.3)\qquad \gamma_1 = -\tfrac{1}{2} - [(\mu^{(2)} - \mu^{(1)})\,\Sigma^{-1}(\mu^{(1)} - \mu)' + c]\,/\delta^2;$$

$$\gamma_2 = \tfrac{1}{2} - [(\mu^{(2)} - \mu^{(1)})\,\Sigma^{-1}(\mu^{(1)} - \mu)' + c]\,/\delta^2;$$

$$\gamma_3 = \left\{(\mu^{(2)} - \mu^{(1)})\,\Sigma^{-1}[\tfrac{1}{2}(\mu^{(1)} + \mu^{(2)}) - \mu]' + c\right\}/\delta^2.$$

Chapter Seven

THE STATISTICS U, R AND Z

1. INTRODUCTION

In chapters three to six we studied the sampling behaviour of the probability of assigning the individual to be classified to $P^{(1)}$ or $P^{(2)}$ when individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according as $W \lessgtr c$. Similar problems are present when the statistic used is U, R or Z. In the case of U, the methods of chapters three to six apply with minor modifications. With R and Z the problems are more difficult. This chapter will indicate the kind of results which can be obtained.

2. DISTRIBUTION OF THE PROBABILITY WHEN THE STATISTIC IS U.

Suppose individuals are assigned to $P^{(2)}$ when $U > c$. We shall let $e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ denote the probability of assigning an individual from P to $P^{(2)}$. If Σ is known, we use U_o, rather than U. We shall let $e_2'(\bar{x}^{(1)}, \bar{x}^{(2)})$ denote the probability of assigning an individual from P to $P^{(2)}$ when the individuals assigned to $P^{(2)}$ are those with $U_o > c$.

A simple calculation shows that

$$(7.2.1) \qquad e_2'\,(\bar{x}^{(1)}, \bar{x}^{(2)}) = 1 - G(w'), \text{ where}$$

$$(7.2.2) \qquad w' = \frac{c - (\bar{x}^{(2)} - \bar{x}^{(1)})\, \Sigma^{-1} \mu'}{\sqrt{\,[(\bar{x}^{(2)} - \bar{x}^{(1)})\, \Sigma^{-1} (\bar{x}^{(2)} - \bar{x}^{(1)})']}} .$$

Therefore,

$$(7.2.3) \qquad \Pr\,[e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = \Pr\,[w' > -G^{-1}(z)].$$

We require the distribution of w'.

Let

$$(7.2.4) \qquad \mu \neq 0,\ C_1 = \mu \Sigma^{-1} \mu',\ Q = (\bar{x}^{(2)} - \bar{x}^{(1)})\, \Sigma^{-1} (\bar{x}^{(2)} - \bar{x}^{(1)})' \text{ and}$$

$$T = [(\bar{x}^{(2)} - \bar{x}^{(1)})\, \Sigma^{-1} \mu'] \,/ \sqrt{C_1}.$$

Then

$$(7.2.5) \qquad w' = (c - C_1^{\frac{1}{2}}\, T)/\sqrt{Q} .$$

Expression (6.3.14) of chapter six is the joint density of Q and T, provided we take

$$(7.2.6)\qquad C_2 = (\mu^{(2)} - \mu^{(1)}) \Sigma^{-1} \mu', \quad C_3 = \delta^2 - C_2^2/C_1,$$

$$C_4 = C_2/\sqrt{C_1}, \quad C_5 = \sqrt{C_3} \text{ and}$$

C_1 as in equation (7.2.4). The density function of w' can be found from the joint density of Q and T and is

$$(7.2.7)\qquad \int \frac{(2a_3)^{-\frac{1}{2}p}}{(C_1\pi)^{\frac{1}{2}}} Q^{\frac{1}{2}} \exp[-\frac{1}{2a_3}\{Q-2C_4(c-Q^{\frac{1}{2}}w')/C_1^{\frac{1}{2}} + \delta^2\}]$$

$$\sum_{k=0}^{\infty} \frac{(\frac{1}{2}C_5/a_3)^k [Q-(c-Q^{\frac{1}{2}}w')^2/C_1]^{\frac{1}{2}(p+2k-3)}}{k!\,\Gamma(\frac{1}{2}p+k-\frac{1}{2})} dQ,$$

where the lower and upper limits of integration are respectively the minimum and maximum of

$$[\max\{0,(w'+C_1^{\frac{1}{2}})/c\}]^{-2} \text{ and } [\max\{0,(w'-C_1^{\frac{1}{2}})/c\}]^{-2}$$

if $c \neq 0$, and zero and infinity if $c = 0$. [If both the limits are infinity, take the integral to be zero.]

If $\mu = 0$, w' has the distribution of $c/(a_3\chi^{2'})^{\frac{1}{2}}$, where $\chi^{2'}$ is a non-central chisquare with degree of freedom p and non-centrality $\frac{1}{2}\delta^2/a_3$.

When $p = 1$, we have the following simpler results: if $c \geq 0$,

(7.2.8) $\Pr[e_2'(\bar{x}^{(1)},\bar{x}^{(2)}) < z] = G(\alpha_1) - G(\alpha_2)$, when $z < G(-|\mu|/\sigma)$,

$= 1$, when $z \geq G(|\mu|/\sigma)$,

$= G(\alpha_1)$ or $1 - G(\alpha_2)$ elsewhere

according as $\mu \gtrless 0$; if $c < 0$,

(7.2.9) $\Pr[e_2'(\bar{x}^{(1)},\bar{x}^{(2)}) < z] = 0$, when $z \leq G(-|\mu|/\sigma)$,

$= 1 + G(\alpha_2) - G(\alpha_1)$, when $z > G(|\mu|/\sigma)$,

$= G(\alpha_2)$ or $1 - G(\alpha_1)$ elsewhere,

according as $\mu \gtrless 0$; here

(7.2.10) $$\alpha_1 = \left[\frac{c}{\mu - \sigma G^{-1}(z)} - \mu^{(2)} + \mu^{(1)}\right] / (a_3^{\frac{1}{2}}\sigma),$$

and

(7.2.11) $$\alpha_2 = \left[\frac{c}{\mu + \sigma G^{-1}(z)} - \mu^{(2)} + \mu^{(1)}\right] / (a_3^{\frac{1}{2}}\sigma).$$

Equations (7.2.8) and (7.2.9) correspond to equation (3.2.2) for $e_2(\bar{\bar{x}}^{(1)}, \bar{\bar{x}}^{(2)})$ and can be established in the same way.

3. DISTRIBUTION OF THE PROBABILITY WHEN THE STATISTIC IS R

If $N_1 = N_2$, R is equivalent to W and the results of chapters three to six apply. In what follows we shall, therefore, assume that $N_1 \neq N_2$.

Assume that Σ is known. Whether the individuals assigned to $P^{(2)}$ should be those with large values of R_o or those with small values of R_o depends on whether $N_1 \gtrless N_2$. For the sake of definiteness we assume that $N_1 > N_2$. Suppose the individuals assigned to $P^{(2)}$ are those with $R_o > c$. We shall denote the probability, given $\bar{\bar{x}}^{(1)}$ and $\bar{\bar{x}}^{(2)}$, of assigning an individual from P to $P^{(2)}$ by $e_2'(\bar{\bar{x}}^{(1)}, \bar{\bar{x}}^{(2)})$.

The inequality $R_o > c$ is equivalent to the inequality

(7.3.1) $$V > c + \frac{1}{4} a_4 d^2 \, Q/a_3, \text{ where}$$

(7.3.2) $$V = [x - a_1\bar{\bar{x}}^{(1)} - a_2\bar{\bar{x}}^{(2)} - \tfrac{1}{2}d(a_4/a_3)^{\frac{1}{2}}(\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})]\, \Sigma^{-1}$$
$$[x - a_1\bar{\bar{x}}^{(1)} - a_2\bar{\bar{x}}^{(2)} - \tfrac{1}{2}d(a_4/a_3)^{\frac{1}{2}}(\bar{\bar{x}}^{(2)} - \bar{\bar{x}}^{(1)})]' \, .$$

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

$$(7.3.3) \qquad \tfrac{1}{2}[\mu - a_1\bar{x}^{(1)} - a_2\bar{x}^{(2)} - \tfrac{1}{2}d(a_4/a_3)^{\frac{1}{2}}(\bar{x}^{(2)} - \bar{x}^{(1)})]\,\Sigma^{-1}$$

$$[\mu - a_1\bar{x}^{(1)} - a_2\bar{x}^{(2)} - \tfrac{1}{2}d(a_4/a_3)^{\frac{1}{2}}(\bar{x}^{(2)} - \bar{x}^{(1)})]' = V' \text{ (say)}.$$

Therefore,

$$(7.3.4) \qquad e_2^*(\bar{x}^{(1)}, \bar{x}^{(2)}) = \int_{c+\frac{1}{4}a_4d^2Q/a_3}^{\infty} \Delta_p(\chi^2; V')\,d\chi^2 .$$

Hence,

$$(7.3.5) \qquad \Pr[e_2^*(\bar{x}^{(1)}, \bar{x}^{(2)}) < z] = \Pr[V' < L_p(c + \tfrac{1}{4}a_4d^2Q/a_3, z)].$$

To find $\Pr[V' < L_p(c + \frac{1}{4}a_4d^2Q/a_3\ \ z)]$, we first find

$\Pr[V' < L_p(c + \frac{1}{4}a_4d^2Q/a_3, z) \mid \bar{x}^{(2)} - \bar{x}^{(1)}]$.

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

$$(7.3.6) \qquad \tfrac{1}{2}(N_1+N_2)[\mu-a_1\mu^{(1)}-a_2\mu^{(2)}-\tfrac{1}{2}d(a_4/a_3)^{\frac{1}{2}}(\bar{x}^{(2)}-\bar{x}^{(1)})]\,\Sigma^{-1}$$

$$[\mu-a_1\mu^{(1)}-a_2\mu^{(2)}-\tfrac{1}{2}d(a_4/a_3)^{\frac{1}{2}}(\bar{x}^{(2)}-\bar{x}^{(1)})]' \quad (= V'', \text{say}).$$

Therefore,

$$(7.3.7) \qquad \Pr[V' < L_p(c + \tfrac{1}{4}\, a_4 d^2 Q/a_3\,, z) \mid \bar{x}^{(2)} - \bar{x}^{(1)}]$$

$$= \int_{2(N_1+N_2)L_p(c+\frac{1}{4}a_4d^2Q/a_3, z) \,\geq\, \chi^2 \,\geq\, 0} \Delta_p(\chi^2; V'')\,d\chi^2.$$

Now,

$$(7.3.8) \qquad V'' = \tfrac{1}{2}\,(N_1+N_2)[\,\tfrac{1}{4}(a_4/a_3)d^2Q + d(C_1a_4/a_3)^{\frac{1}{2}}T + C_1].$$

Therefore,

$$\int_{2(N_1+N_2)L_p(c+\frac{1}{4}\, a_4d^2Q/a_3, z) \,\geq\, \chi^2 \,\geq\, 0} \Delta_p(\chi^2; V'')\,d\chi^2$$

is a function of z, Q and T. Denote it by $\Pi(z; Q, T)$. Let $F(Q, T)$ denote the joint density of Q and T. Then

$$(7.3.9) \qquad \Pr \left[V' < L_p(c + \tfrac{1}{4} a_4 d^2 Q/a_3, \; z)\right]$$

$$= \int \int \Omega \,(z; \; Q, \; T) F(Q,T)\, dQ dT,$$

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

$$(7.3.10) \qquad \Pr \left[e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}) < z\right] = \int\int \Omega \,(z; Q,T) F(Q,T)\, dQ dT.$$

An expression for $F(Q,T)$ is (6.3.14) of chapter six, provided $C_1 \neq 0$; if $C_1 = 0$, $\Omega\,(z; \; Q, \; T)$ does not involve T and therefore in equation (7.3.10) $F(Q,T)$ may be replaced by the density function of Q; the density function of Q is $a_3^{-1} \Delta_p \left(Q/a_3; \; \frac{1}{2} \delta^2/a_3\right)$.

4. DISTRIBUTION OF THE PROBABILITY WHEN THE STATISTIC IS Z

Suppose individuals are assigned to $P^{(2)}$ when and only when

$$(7.4.1) \qquad \frac{N_1}{N_1+1} (x-\bar{x}^{(1)}) \Sigma^{-1} (x-\bar{x}^{(1)})' - \eta \frac{N_2}{N_2+1} (x-\bar{x}^{(2)}) \Sigma^{-1} (x-\bar{x}^{(2)})' > c.$$

Inequality (7.4.1) is equivalent to the inequality

$$(7.4.2)\qquad \left\{x-\bar{x}^{(1)}+\alpha(\bar{x}^{(2)}-\bar{x}^{(1)})\right\}\Sigma^{-1}\left\{x-\bar{x}^{(1)}+\alpha(\bar{x}^{(2)}-\bar{x}^{(1)})\right\}'$$

$$\gtrless \alpha(\alpha+1)Q+c'$$

accofding as

$$\frac{N_1}{N_1+1} \gtrless \frac{N_2\eta}{N_2+1},$$

$$(7.4.3)\qquad \alpha = \frac{N_2(N_1+1)\eta}{N_1(N_2+1)-N_2(N_1+1)\eta},$$

$$(7.4.4.)\qquad c' = \frac{c(N_1+1)(N_2+1)}{N_1(N_2+1)-N_2(N_1+1)\eta}.$$

For the sake of definiteness, we shall assume that

$$(7.4.5)\qquad \frac{N_1}{N_1+1} > \frac{N_2\eta}{N_2+1}.$$

[If $N_1/(N_1+1) = N_2\eta/(N_2+1)$, the procedure is equivalent to using W_c]. Given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, the probability of assigning an individual from P to $P^{(2)}$ is then the probability, given $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, of the inequality (7.4.2) with the upper inequality sign being satisfied. It is easily seen to be equal to

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

where

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

We shall denote this probability by $e_2'''(\bar{x}^{(1)}, \bar{x}^{(2)})$. By methods similar to those of the previous section we can show that

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

where

(7.4.9) $$\Gamma'(z;\,Q,\,T) = \int_{2(N_1+N_2)L_p(\alpha[\alpha+1]Q+c',\gamma)\geq\chi^2\geq 0} \Delta_p(\chi^2;\,v')\,d\chi^2;$$

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

(7.4.11) $$\beta = \alpha + a_2;$$

F(Q, T) is the function (6.3.14) of chapter six, provided $C_1 \neq 0$; C_1, C_2, C_3, C_4 and C_5 are as in equation (6.3.6), (6.3.10) and (6.3.13). If $C_1 = 0$, Γ' does not contain T and, therefore, in equation (7.4.8), $F(Q, T)$ may be replaced by the density function of Q; i.e. by

$$a_3^{-1} \Delta_p (Q / a_3 ; \tfrac{1}{2} \delta^2 / a_3).$$

Chapter Eight

EXPECTED VALUES

1. INTRODUCTION

In chapters three to seven we were concerned with the problem of deriving the distribution of the probability of assigning the individual to $P^{(1)}$ or $P^{(2)}$. This chapter will give the expected values of these probabilities. In sections two and three we shall consider the statistic W; the statistics U, R and Z will be considered in section four.

2. EXACT EXPRESSIONS

$$E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}; S) = \Pr\, (W \geq c). \tag{8.2.1}$$

In chapter two we have seen that W has the same distribution as

$$(n/\lambda^2)\left[\tfrac{1}{2}(N_1^{-1}-N_2^{-1})B_{11}+(a_3a_4)^{\frac{1}{2}}\left\{B_{12}-(n-p+2)^{-\frac{1}{2}}|B|^{\frac{1}{2}}t\right\}\right], \tag{8.2.2}$$

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

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

denotes its non-centrality matrix, with λ_{11}, λ_{12} and λ_{22} as given by equation (2.4.3) of chapter two; t follows Student's law with $n - p + 2$ degrees of freedom; χ^2 is a chi-square variable with $n - p + 1$ degrees of freedom. If $c \neq 0$, we can, therefore, write

$$(8.2.3) \qquad E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$$

$$= E \Pr\left(\chi^2 \lesseqgtr (n/c)\left[\tfrac{1}{2}(N_1^{-1} - N_2^{-1})B_{11} + (a_3 a_4)^{\frac{1}{2}}\{B_{12} - (n-p+2)^{-\frac{1}{2}}\ t\ |B|^{\frac{1}{2}}\}\right]\right),$$

where, on the right side, the upper or lower inequality sign is to be taken according as $c \gtrless 0$; the expectation on the right side is to be taken with respect to t and B. If $c = 0$, we can write

(8.2.4) $$E\, e_2 (\bar{x}^{(1)}, \bar{x}^{(2)} ; S)$$

$$= E\, \Pr(t < [(n-p+2)/(a_3 a_4 |B|)]^{\frac{1}{2}} [\tfrac{1}{2}(N_1^{-1} - N_2^{-1}) B_{11} + (a_3 a_4)^{\frac{1}{2}} B_{12}]),$$

where the expectation is to be taken with respect to B.

If we can assume that Σ is known, we can give more explicit results. In this case individuals are assigned to $P^{(1)}$ or $P^{(2)}$ according $W_0 \lessgtr c$. The expected probability of assigning an individual to $P^{(2)}$ is, therefore, the integral of the density function of W_0 from c to ∞. We have obtained the density function of W_0 in chapter two.

If, further, $c = 0$, we can obtain a simpler expression for $E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$ as follows.

(8.2.5) $$E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) = \Pr(W_0 \geq 0 \mid x \varepsilon P).$$

From chapter two we know that the distribution of W_0 is the same as that of $w_1/b_1 - w_2/b_2$ where w_1 and w_2 are independent non-central chi-squares, each having p degrees of freedom and non-centralities equal respectively to λ_1 and λ_2 of equation (2.4.6) of chapter two. Therefore,

$$(8.2.6) \qquad E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) = Pr[w_1/b_1 - w_2/b_2 > 0],$$

$$= Pr\,(w_1/w_2 > b_1/b_2),$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\Gamma(p+r+s)}{\Gamma(\frac{1}{2}p+r)\,\Gamma(\frac{1}{2}p+s)} \frac{\lambda_1^r\ \lambda_2^s}{r!\ s!}$$

$$\int_{b_1/b_2}^{\infty} \frac{u^{\frac{1}{2}p+r-1}}{(1+u)^{p+r+s}}\, du,$$

the integral of the density function* of (w_1/w_2) from (b_1/b_2) to ∞. The alternative form

* The density function of (w_1/w_2) is found on page 140 of [19].

(8.2.7) $E\, e_2\, (\bar{x}^{(1)}, \bar{x}^{(2)})$

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

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

of equation (8.2.6) may be found more convenient in view of the fact that tables of $I_\alpha\, (r, s)$ exist.

3. AN APPROXIMATION

In the previous section we obtained an exact expression for $E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$. We shall now give an approximation that enables us to evaluate $E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)})$ using only tables of $G(\alpha)$.

We start from the result (8.2.6). Now,

(8.3.1) $$\Pr\, (w_1/w_2 > b_1/b_2) = \Pr[(w_1/w_2)^{1/3} > (b_1/b_2)^{1/3}],$$

$$= \Pr[b_2^{1/3} w_1^{1/3} - b_1^{1/3}\, w_2^{1/3} > 0].$$

From [1] we know that if w is a non-central chi-square with

ν degrees of freedom and non-centrality λ, the variable $(w/r)^{1/3}$, where

$$r = \nu + 2\lambda \, ,$$

has approximately the normal distribution with expectation $1 - 2(1 + \theta)/(9r)$ and variance $2(1 + \theta)/(9r)$. $[\theta = 2\lambda/(\nu + 2\lambda).]$ Using this result, we obtain from equation (8.3.1) the equation

$$\text{(8.3.2)} \qquad E\, e_2(\bar{x}^{(1)}, \bar{x}^{(2)}) \approx G(a),$$

where

$$\text{(8.3.3)} \qquad a = \frac{2[b_1^{\frac{1}{3}}\, r_2^{-2/3}(1+\theta_2) - b_2^{\frac{1}{3}}\, r_1^{-\frac{2}{3}}(1+\theta_1)] + 9[(r_1 b_2)^{\frac{1}{3}} + (r_2 b_1)^{\frac{1}{3}}]}{(18)^{\frac{1}{2}}[b_1^{2/3} r_2^{-1/3}(1+\theta_2) + b_2^{2/3}\, r_2^{-1/3}(1+\theta_1)\,]^{\frac{1}{2}}} \, ;$$

$$\text{(8.3.4)} \qquad r_i = p + 2\lambda_i \qquad (i = 1, 2);$$

$$\text{(8.3.5)} \qquad \theta_i = 2\lambda_i \,/(p + 2\lambda_i) \quad (i = 1, 2).$$

The question, how good is the approximation ?, now arises. We should expect that the approximation involved is of the same order as that involved in approximating the distribution of the cube root of a non-central chi-square by the normal distribution. Numerical results given in [1] show that the normal approximation to the distribution of the cube-root of a non-central chi-square is fairly good.

4. THE STATISTICS U, R and Z

Results similar to those for W can be given for U, R and Z also. The equation corresponding to (8.2.3) is

(8.4.1) $$E\, e_2'\, (\bar{x}^{(1)}, \bar{x}^{(2)}; S)$$

$$= E \Pr[\, \chi^2 \lessgtr (n a_3^{\frac{1}{2}}/c)\{ B_{12} - (n-p+2)^{-\frac{1}{2}} |B|^{\frac{1}{2}}\, t \}\,],$$

if the individual is assigned to $P^{(2)}$ when $U > c$, is

(8.4.2) $$E\, e_2''\, (\bar{x}^{(1)}, \bar{x}^{(2)}; S)$$

$$= E \Pr[\, \chi^2 \lessgtr (n a_4/c)\{ B_{11} - dB_{12} + d(n-p+2)^{-\frac{1}{2}} |B|^{\frac{1}{2}}\, t \}\,],$$

if the individual is assigned to $P^{(2)}$ when $R > c$, and is

(8.4.3) $$E\, e_2'''\, (\bar{x}^{(1)}, \bar{x}^{(2)}; S)$$

$$= E \Pr[\chi^2 \lessgtr (n/c)(1-\eta)B_{11} + 2(n/c)(\eta - a_5 a_6 \eta)\{ B_{11} - (n - p + 2)^{-\frac{1}{2}} |B|^{\frac{1}{2}}\, t \}\,],$$

if the individual is assigned to $P^{(2)}$ when $Z > c$. [In equations (8.4.1) to (8.4.3), the upper inequality or the lower inequality is to be taken according as $c \gtrless 0$.] If $c = 0$, the equations are, respectively,

(8.4.4) $E\, e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}; S) = E\, Pr[t < (n-p+2)^{\frac{1}{2}} |B|^{-\frac{1}{2}} B_{12}],$

(8.4.5) $E\, e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}; S) = E\, Pr[t > (n-p+2)^{\frac{1}{2}} |B|^{-\frac{1}{2}} \{ B_{12} - B_{11}/d \}],$

(8.4.6) $E\, e_2'''(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$

$$= E\, Pr[t < (n - p + 2)^{\frac{1}{2}} |B|^{-\frac{1}{2}} \{ B_{12} + \frac{1}{2} (\eta - a_5 a_6 \eta)^{-\frac{1}{2}} (1 - \eta) B_{11} \}].$$

The distribution of the random variables involved in the above equations are the same as in the case of W except for the difference in the values of the parameters λ_{11}, λ_{12} and λ_{22}; in the case of $e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ the appropriate values of λ_{11}, λ_{12} and λ_{22} are given by equations (2.3.3), in the case of $e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$ by equations (2.5.3), and by equations (2.6.7) in the case of $e_2'''(\bar{x}^{(1)}, \bar{x}^{(2)}; S)$.

If $p = 1$ and individuals are assigned to $P^{(2)}$ when $U > 0$, we have the following simpler result:

$$(8.4.7) \qquad E\, e_2'(\bar{x}^{(1)}, \bar{x}^{(2)}, S) = G([\mu^{(1)} - \mu^{(2)}]/[a_3^{\frac{1}{2}}\sigma])G(-\mu/\sigma) + G([\mu^{(2)} - \mu^{(1)}]/[a_3^{\frac{1}{2}}\sigma])G(\mu/\sigma).$$

If $p=1$ and individuals are assigned to $P^{(2)}$ when $R > 0$,

$$(8.4.8) \qquad E\, e_2''(\bar{x}^{(1)}, \bar{x}^{(2)}; S) = G(a_{11}, a_{12}; \varrho) + G(a_{21}, a_{22}; \varrho),$$

where

$$(8.4.9) \qquad a_{11} = -a_{21} = (\mu - a_1\mu^{(1)} - a_2\mu^{(2)})/(a_3^{\frac{1}{2}}\sigma),$$

$$a_{12} = -a_{22} = \frac{\mu - a_1\mu^{(1)} - a_2\mu^{(2)} - d(a_4/a_3)^{\frac{1}{2}}(\mu^{(2)} - \mu^{(1)})}{a_4^{\frac{1}{2}}\sigma(1 + d^2)^{\frac{1}{2}}},$$

$$\varrho = (1 + d^2)^{-\frac{1}{2}}.$$

These results are similar to those for W in chapter three and can be derived in the same way.

The right members of equations (8.2.7) and (8.3.2) give — the one exactly and the other approximately— Pr ($U_o > 0$), Pr($R_o > 0$) or Pr ($Z_o > 0$) according as b_1, b_2 and λ_1, λ_2 are as in equations (2.3.5) and (2.3.6) or as in equations (2.5.5) and (2.5.6) or as in equations (2.6.13) to (2.6.16).

REFERENCES

[1] S.H.Abdel - Aty, 'Approximate formula for the percentage points and the probability integral of the non-central χ^2 distribution', Biometrika, Vol. 41 (1954), p. 538-540.

[2] T.W.Anderson, 'Classification by multivariate analysis', Psychometrika, Vol. 16 (1951), p. 31-50.

[3] --- An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York, 1958.

[4] M.M.Barnard, 'The secular variations of skull characters in four series of Egyptian skulls', Ann. Eugen., Vol. 6 (1935), p. 352-371.

[5] 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, Stanford Univ. Press, Stanford, California, 1960, p. 142-150.

[6] R.A.Fisher, 'The use of multiple measurements in taxonomic problems', Ann. Eugen., Vol. 7 (1936), p. 179-188.

[7] Herman Leon Harter, 'On the distribution of Wald's classification statistic', Ann. Math. Stat., Vol.25 (1951), p. 58-67.

[8] S. John, 'On the distribution of Wald's classification statistic when the dispersion matrix is known', Sankhyā, Vol. 21 (1959), p. 371-376.

[9] ---- 'On some classification problems - I', Sankhyā, Vol. 22 (1960), p. 301-308.

[10] ---- 'On some classification statistics', Sankhyā, Vol. 22 (1960), p. 309-316.

[11] S. John, 'Errors in discrimination', Ann. Math. Stat., Vol.32 (1961), p. (yet to appear)

[12] A. Kudo, 'The classificatory problem viewed as a two-decision problem', Memoirs of the Faculty of Science, Kyushu University, Series A, Vol. 13 (1959), p. 96-125.

[13] Karl Pearson, 'On the application of the Double Bessel Function $K_{\tau_1,\tau_2}(x)$ to statistical problems', Biometrika, Vol.25 (1933), p. 158-178.

[14] C.Radhakrishna Rao, 'The utilization of multiple measurements in taxonomic problems', J. Roy. Stat. Soc.,B, Vol.10,(1948), p. 159-193.

[15] ---- 'A general theory of discrimination when the information on alternative hypotheses is based on samples', Ann. Math. Stat., Vol. 25 (1954), p. 651-670.

[16] Rosedith Sitgreaves, 'On the distribution of two random matrices used in classification procedures', Ann. Math. Stat. Vol. 23, (1952), p. 263-270.

[17] L.J.Slater, Confluent Hyper-geometric Functions, First Edition, Cambridge, Univ. Press, 1960.

[18] Erling Sverdrup, 'Derivation of the Wishart distribution of the second order sample moments by straight-forward integration of a multiple integral', Skandinavisk Aktuarietedskrift, Vol. 30 (1947), p. 151-166.

[19] P.C.Tang, 'The power function of the analysis of variance tests with tables and illustrations of their use', Statistical Research Memoirs, Vol. 2 (1938), p. 126-157.

[20] A. Wald, 'On a statistical problem arising in the classification of an individual into one of two groups', Ann. Math. Stat. Vol.15 (1944), p.145-163.

[21] B.L.Welch, 'Note on discriminant functions', Biometrika, 31, (1939), p. 218-220.

[To be a complete bibliography on statistical taxonomy, this list would have to be extended to many times its length. We hope no published paper dealing with the four classification statistics considered in this thesis is missing from our list.]