The Annals of Statistics
1985, Vol. 13, No. 2, 827–831

AN ERROR BOUND FOR AN ASYMPTOTIC EXPANSION OF THE DISTRIBUTION FUNCTION OF AN ESTIMATE IN A MULTIVARIATE LINEAR MODEL

BY YASUNORI FUJIKOSHI

Hiroshima University

In this paper we consider asymptotic approximations to the distribution function $F(x)$ of a linear combination of an estimate in a multivariate linear model. A method is given for obtaining an asymptotic expansion $F_{s-1}(x)$ of $F(x)$ up to $O(n^{-s+1})$ and a bound c_s such that $|F(x) - F_{s-1}(x)| \le c_s$ uniformly in x and $c_s = O(n^{-s})$.

1. Introduction. Asymptotic expansions such as Edgeworth type expansions play an important part in the study of approximations to distributions. The general theory of the expansions has been studied by many authors: see for example Wallace, 1958; Chambers, 1967; Chibisov, 1972, 1973; Barndorff-Nielsen and Cox, 1979. However, it may be noted that little work on the explicit error bounds has been done for the approximations based on asymptotic expansions. In this paper we shall obtain an error bound for an asymptotic expansion of the distribution function of a linear combination of an estimate in a multivariate linear model.

Let $\mathbf{y}_1, \cdots, \mathbf{y}_N$ be a sample of size $N = n + 1$ from a p-variate normal distribution $N_p(B'\boldsymbol{\beta}, \Sigma)$, where B is a known $q \times p$ matrix of rank $q \le p$, $\boldsymbol{\beta}$ is a $q \times 1$ vector of unknown parameters and Σ is an unknown positive definite matrix. The model on the observations is called "growth curves" model and is a special case of the general MANOVA model due to Potthoff and Roy (1964). The Maximum Likelihood Estimate of $\boldsymbol{\beta}$ is given by $\hat{\boldsymbol{\beta}} = (BS^{-1}B')^{-1}BS^{-1}\bar{\mathbf{y}}$, where $\bar{\mathbf{y}} = (1/N)\sum_{j=1}^{N}\mathbf{y}_j$ and $S = \sum_{j=1}^{N}(\mathbf{y}_j - \bar{\mathbf{y}})(\mathbf{y}_j - \bar{\mathbf{y}})'$. We consider the distribution of

$$\xi = \sqrt{N}\mathbf{a}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})/\lambda \tag{1.1}$$

where $\mathbf{a}$ is a $q \times 1$ fixed vector ($\neq \mathbf{0}$) and $\lambda = \{\mathbf{a}'(B\Sigma^{-1}B')^{-1}\mathbf{a}\}^{1/2}$. Here ξ is standardized so that the distribution function $F(x)$ of ξ converges in law to the standard normal distribution function $\Phi(x)$. We treat the case $r = p - q > 0$, since the distribution of ξ in the case $r = 0$ is $N(0, 1)$. The exact distribution of $\hat{\boldsymbol{\beta}}$ has been studied by Gleser and Olkin (1972), who obtained three expressions for the density of $\hat{\boldsymbol{\beta}}$. However, it seems that the exact distribution of ξ would be very complicated if it could be obtained. The purpose of this paper is to find an asymptotic expansion $F_{s-1}(x)$ of $F(x)$ up to $O(n^{-s+1})$ and an error bound c_s such

Received February 1984; revised October 1984.
AMS 1980 *subject classifications*. Primary 62E20, secondary 62H10.
Key words and phrases. Error bound, asymptotic expansion, distribution function, linear combination of an estimate, multivariate linear model.

that

$$|F(x) - F_{s-1}(x)| \leq c_s \tag{1.2}$$

uniformly in x and $c_s = O(n^{-s})$. The result holds for any integer $s \geq 1$ such that $n - r - 2s + 1 > 0$.

2. A method for error estimation. Consider a random variable ξ with a special structure and let $F(x)$ be the distribution function of ξ. Here ξ may not be the same one as in (1.1). We shall make some assumptions on ξ, but in Section 3 it is shown that the ξ as in (1.1) satisfies all the assumptions required. Suppose that ξ is decomposed as

$$\xi = z - u. \tag{2.1}$$

We may regard u as a remainder when we approximate ξ by z. Let $G(x)$ be the distribution function of z. We make the following assumptions for some integer $s \geq 1$:

ASSUMPTION 1. z and u are independent.

ASSUMPTION 2. All the derivatives of $G(x)$ of order s and less are continuous and $\text{Sup}|G^{(s)}(x)| < \infty$, where $G^{(s)}(x)$ is the sth derivative of $G(x)$.

ASSUMPTION 3. $E[|u|^s] = \bar{m}_s < \infty$.

THEOREM 2.1. *Suppose that a random variable ξ is decomposed as in* (2.1) *and there exists an integer $s \geq 1$ such that Assumptions* 1–3 *hold. Then*

$$|F(x) - \tilde{F}_{s-1}(x)| \leq (1/s!)\bar{m}_s \text{Sup}_t |G^{(s)}(t)| \tag{2.2}$$

where

$$\tilde{F}_{s-1}(x) = G(x) + \textstyle\sum_{j=1}^{s-1} (1/j!) m_j G^{(j)}(x) \tag{2.3}$$

and $m_j = E[u^j]$, $j = 1, \cdots, s-1$.

PROOF. By Assumption 1, $F(x) = E_u\{G(x+u)\}$. The result is immediate on expanding $G(x+u)$ about $G(x)$.

In our application we can make further assumptions on z and u:

ASSUMPTION 4. z has the standard normal distribution, i.e., $G(x) = \Phi(x)$.

ASSUMPTION 5. The conditional distribution of u given a random matrix V is normal with mean zero and variance $h(V)$. Further, for some integer s,

$$h_s = E[\{h(V)\}^s] < \infty. \tag{2.4}$$

When $G(x) = \Phi(x)$, Assumption 2 is satisfied for any integer since $\lim_{|t|\to\infty} |\Phi^{(j)}(t)| = 0$. Let

$$\ell_s = \text{Sup}_t |\Phi^{(s)}(t)| < \infty. \tag{2.5}$$

Then it is easy to see that

$$\ell_1 = |\Phi^{(1)}(0)| = \frac{1}{\sqrt{2\pi}}, \quad \ell_2 = |\Phi^{(2)}(1)| = \frac{1}{\sqrt{2\pi e}}$$

$$\ell_3 = |\Phi^{(3)}(0)| = \frac{1}{\sqrt{2\pi}}, \quad \ell_4 = |\Phi^{(4)}(\sqrt{13-\sqrt{6}})| = \frac{1}{\sqrt{2\pi}} \times 1.38 \cdots$$

$$\ell_5 = |\Phi^{(5)}(0)| = \frac{3}{\sqrt{2\pi}}, \quad \ell_6 = |\Phi^{(6)}(\sqrt{\alpha})| = \frac{1}{\sqrt{2\pi}} \times 5.78 \cdots$$

where α is the minimum root of $\alpha^3 - 15\alpha^2 + 45\alpha - 15 = 0$. From Assumption 5 we obtain that for $j \le s$

$$m_j = E\{u^j\} = E_V E_{u|V}\{u^j\} = \begin{cases} h_{j/2} j!\{2^{j/2}(j/2)!\}^{-1}, & j \text{ even} \\ 0, & j \text{ odd.} \end{cases} \tag{2.6}$$

Substituting this result into the formula (2.2) obtained by replacing s by $2s$, we have the following theorem.

THEOREM 2.2. *Suppose that ξ is decomposed as in* (2.1) *and there exists an integer $s \ge 1$ such that Assumptions* 1, 4 *and* 5 *hold. Then*

$$|F(x) - F_{s-1}(x)| \le c_s = (1/2^s s!)\ell_{2s} h_s \tag{2.7}$$

where

$$F_{s-1}(x) = \Phi(x) + \textstyle\sum_{j=1}^{s-1} (1/2^j j!) h_j \Phi^{(2j)}(x). \tag{2.8}$$

We compare this result with the one obtained by the general method based on the characteristic function of ξ. Under the assumptions as in Theorem 2.2 the characteristic function of ξ can be expressed as

$$\phi(t) = \exp(-\tfrac{1}{2}t^2)E_V[\exp\{-\tfrac{1}{2}h(V)\}] = \phi_{s-1}(t) + R_s(t) \tag{2.9}$$

where

$$\phi_{s-1}(t) = \exp(-\tfrac{1}{2}t^2) \textstyle\sum_{j=0}^{s-1} (-\tfrac{1}{2}t^2)^j h_j/j!,$$

$$R_s(t) = \{(-\tfrac{1}{2}t^2)^s/s!\} \times \exp(-\tfrac{1}{2}t^2)E_V[\{h(V)\}^s \exp\{-\tfrac{1}{2}\theta t^2 h(V)\}]$$

and θ $(0 < \theta < 1)$ is the constant that appeared in the remainder term of Taylor's expansion of e^x. It is easy to see that the approximation $F_{s-1}(x)$ is obtained by inverting $\phi_{s-1}(t)$. If $h_j = O(n^{-j})(j = 1, \cdots, s)$, then it follows that $\mathrm{Sup}|F(x) - F_{s-1}(x)| = O(n^{-s})$, but the general theory on asymptotic expansions does not give explicit expressions for constants in the error bound. Theorem 2.2 gives an explicit error bound with the order of $O(n^{-s})$.

3. The distribution of ξ. We shall see that the random variable ξ defined by (1.1) satisfies Assumptions 1–5. Let $\delta = \sqrt{N}(\bar{y} - B'\beta)$. Then

$$\sqrt{N}(\hat{\beta} - \beta) = (BS^{-1}B')^{-1}BS^{-1}\delta \tag{3.1}$$

and δ is distributed as $N(\mathbf{0}, \Sigma)$. We define the variables z and u in (2.1) by

$$z = (1/\lambda)\mathbf{a}'(B\Sigma^{-1}B')^{-1}B\Sigma^{-1}\delta, \tag{3.2}$$

$$u = -(1/\lambda)\mathbf{a}'(BS^{-1}B')^{-1}BS^{-1}\{I - B'(B\Sigma^{-1}B')^{-1}B\Sigma^{-1}\}\delta. \tag{3.3}$$

Then z is distributed as $N(0, 1)$ and z and u are independent. The independence of z and u follows from the fact that δ and S are independent and $\{I - B'(B\Sigma^{-1}B')^{-1}B\Sigma^{-1}\}\Sigma^{1/2}(B\Sigma^{-1/2})' = 0$. The following Lemma has essentially been proved in Gleser and Olkin (1972).

LEMMA 3.1. *Let* $\mathbf{v}$: $r \times 1$ *and* W: $r \times r$ *be the independent random vector and matrix distributed as* $N_r(\mathbf{0}, I_r)$ *and* $W_r(I_r, n)$, *respectively, where* $W_r(I_r, n)$ *denotes the Wishart distribution of dimensionality* r *with* n *degrees of freedom and covariance matrix* I_r. *Then* u *is distributed like the random variable whose conditional distribution given* $V = [\mathbf{v}, W]$ *is normal with mean zero and variance*

$$h(V) = \mathbf{v}'W^{-1}\mathbf{v}. \tag{3.4}$$

Noting that $\mathbf{v}'W^{-1}\mathbf{v}$ is distributed as the ratio χ_r^2/χ_{n-r+1}^2 of two independent χ^2 variates (see, for example Anderson, 1958, page 106) we have

$$h_j = \frac{r(r+2)\cdots(r+2(j-1))}{(n-r-1)(n-r-3)\cdots(n-r-2j+1)} \tag{3.5}$$

if $n - r - 2j + 1 > 0$. Therefore we have the following theorem.

THEOREM 3.1. *Let* $F(x)$ *be the distribution function of the standardized statistic* ξ *defined by* (1.1). *Then it holds that if* $n - r - 2s + 1 > 0$,

$$|F(x) - F_{s-1}(x)| \le c_s \tag{3.6}$$

where $F_{s-1}(x)$ *and* c_s *are given by* (2.8) *and* (2.7) *with the coefficients* h_j *in* (3.5), *respectively.*

From the practical point of view the formula (3.6) for small s, especially $s = 1, 2$ is important. For the case $s = 1, 2$,

(1) If $n - r - 1 > 0$,

$$|F(x) - \Phi(x)| \le c_1 = \frac{1}{\sqrt{8\pi e}}\frac{r}{n-r-1}.$$

(2) If $n - r - 3 > 0$,

$$\left| F(x) - \left\{\Phi(x) + \frac{1}{2}\cdot\frac{r}{n-r-1}\Phi^{(2)}(x)\right\}\right| \le c_2 = \frac{1.38\cdots}{8\sqrt{2\pi}}\cdot\frac{r(r+2)}{(n-r-1)(n-r-3)}.$$

It may be noted that "$c_2 \le c_1$" does not always hold. If $(n - r - 3) \ge 0.57(r + 2)$, the inequality holds and hence the second approximation is recommended.

REFERENCES

ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

BARNDORFF-NIELSEN, O. and COX, D. R. (1979). Edgeworth and saddle-point approximations with statistical applications (with discussion). *J. Royal Statist. Soc.* **41** 279–312.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451. Corrigendum ibid **8** 1399.

CHAMBERS, J. M. (1967). On methods of asymptotic approximation for multivariate distributions. *Biometrika* **54** 367–383.

CHIBISOV, D. M. (1972). An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion. *Theor. Probab. Appl.* **17** 620–630.

CHIBISOV, D. M. (1973). An asymptotic expansion for a class of estimators containing maximum likelihood estimators. *Theor. Probab. Appl.* **18** 295–303.

GLESER, L. J. and OLKIN, I. (1972). Estimation for regression model with an unknown covariance matrix. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1** 541–568.

POTTHOFF, R. F. and ROY, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* **51** 313–326.

WALLACE, D. L. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29** 635–654.

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
HIROSHIMA UNIVERSITY
HIROSHIMA, JAPAN