The Annals of Statistics
1975, Vol. 3, No. 3, 741-746

ON EDGEWORTH EXPANSIONS WITH UNKNOWN CUMULANTS[1]

BY H. L. GRAY, W. A. COBERLY AND T. O. LEWIS

Southern Methodist University, Texas Tech University and NASA, Texas Tech University

In this paper a new method of approximating one distribution by another is introduced. The method is essentially a modification of the Edgeworth technique which eliminates the necessity of knowing the cumulants of the distributions involved.

1. Introduction. The representation of one distribution function in terms of another has been the subject of considerable research for some time. One of the most prominent of such representations is the Edgeworth expansion which is still a topic of interest (see [2] or [3]).

The most obvious application of such expansions is the calculation of a distribution function, F_1, by means of known values or more easily obtainable values of another distribution, F_2. However, their value for this purpose has been limited, primarily due to the difficulty in obtaining such expressions. Tables (see [2]) have been produced in a number of places to alleviate this problem but they are not readily available. Moreover, even if such tables were commonplace they would not eliminate the difficulty since the cumulants of the distributions F_1 and F_2 of rather high order are usually required.

In this paper it is shown how the general Edgeworth expansion can be utilized in such a way as to eliminate the requirement for knowing the cumulants without affecting the order of the error of approximation. In the particular cases considered this error is in fact often reduced. Thus they suggest that the expansion introduced in this paper may be preferable to the Edgeworth expansion even when the cumulants are known.

2. Preliminaries. Let $F(\ ;\lambda)$ and Φ be probability distribution functions with cumulants k_i and α_i respectively and let $\beta_i = k_i - \alpha_i$, where we shall assume for convenience that $\beta_1 = \beta_2 = 0$. In addition we assume

(1) $$\lim_{\lambda\to\infty} F(x;\lambda) = \Phi(x)$$

for all x in the support of $F(\ ;\lambda)$ and

(2) $$\beta_i = O(\lambda^{1-(i/2)}), \qquad i = 3, 4, \cdots.$$

Then the Edgeworth expansion in terms of $\Phi(x)$ corresponding to $F(x;\lambda)$ is defined by the following:

(3) $$F(x;\lambda) \sim \Phi(x) + \sum_{i=1}^{\infty} Q_i(x, \lambda),$$

Received October 1973; revised July 1974.

[1] This work was partially supported by ONR Contract No. N00014-72-A-0296.

AMS 1970 *subject classifications.* Primary E99; Secondary E20.

Key words and phrases. Edgeworth expansions, G-Transforms.

where the Q's are functions of λ and x determined by the β_i and the derivatives, $\{\Phi^{(m)}\}$, of Φ and satisfying the relation $Q_i(x, \lambda) = O(\lambda^{-i/2})$.

In many instances (see [1] or [3]) the correspondence in (3) is an equality or at least an asymptotic relation. In that event it is well known that

$$F(x; \lambda) - F_n(x; \lambda) = O(\lambda^{-(n+1)/2}) , \tag{4}$$

where

$$F_n(x; \lambda) = \Phi(x) + \textstyle\sum_{i=1}^{s} h_i(\lambda)\Phi^{(m_i)}(x) . \tag{5}$$

We shall tacitly assume that (4) holds through the remaining part of the paper.

If the cumulants associated with $F(\ ;\lambda)$ or Φ are difficult to calculate $F_n(x; \lambda)$ is difficult to obtain. We shall now introduce a transformation on (5) which will eliminate this difficulty when $F(\ ;\lambda)$ and Φ are known and hence produce a more easily obtainable approximation of $F(x; \lambda)$ in many instances.

3. A new approximation. In order to obtain the approximation we seek let us rewrite (4) in the form

$$F(x; \lambda) = \Phi(x) + \textstyle\sum_{i=1}^{k} g_i(\lambda)\Phi^{(m_i)}(x) + O(\lambda^{-(n+1)/2}) , \tag{6}$$

where (6) is just a rearrangement of (4) such that the m_i are distinct, $g_i(\lambda) \neq 0$ are the resulting coefficients, and k is the resulting number of distinct m_i.

Equation (6) suggests the following theorem which we state without proof. The proof can be found in [1].

THEOREM 1. *If $F(\ ;\lambda)$ and Φ are analytic functions in a domain R of the complex plane and if the functions $\{\Phi^{(m)}\}$ are bounded in R, then for each integer m*

$$F^{(m)}(x, \lambda) = \Phi^{(m)}(x) + \textstyle\sum_{i=1}^{k} g_i(\lambda)\Phi^{(m_i+m)}(x) + O(\lambda^{-(n+1)/2}) \tag{7}$$

uniformly as $\lambda \to \infty$.

Now let us rewrite (6) and (7) as follows:

$$\begin{aligned} \Phi(x) &= F(x; \lambda) - \textstyle\sum_{i=1}^{k} g_i(\lambda)\Phi^{(m_i)}(x) - O(\lambda^{-(n+1)/2}) \\ \Phi^{(m)}(x) - F^{(m)}(x; \lambda) &= -\textstyle\sum_{i=1}^{k} g_i(\lambda)\Phi^{(m_i+m)}(x) - O(\lambda^{-(n+1)/2}) \\ & \qquad m = 1, 2, \cdots . \end{aligned} \tag{8}$$

Treating $F(x; \lambda)$ and the $\{g_i(\lambda)\}$ as unknowns in equations (8) now leads us by Cramér's rule to define the approximation, $\hat{F}_n(x, \lambda)$, of $F(x; \lambda)$:

$$\hat{F}_n(x; \lambda) = \frac{H_k[\Phi(x), \psi_i(x); \Phi^{(m_i)}(x)]}{H_k[1, 0; \Phi^{(m_i)}(x)]} , \tag{9}$$

where

$$H_k[A, B_i; \Phi^{(m_i)}(x)] = \begin{vmatrix} A & B_1 & \cdots & B_k \\ \Phi^{(m_1)}(x) & \Phi^{(m_1+1)}(x) & \cdots & \Phi^{(m_1+k)}(x) \\ \vdots & & & \vdots \\ \Phi^{(m_k)}(x) & \Phi^{(m_k+1)}(x) & \cdots & \Phi^{(m_k+k)}(x) \end{vmatrix}$$

for all x such that the denominator is nonzero, where

$$\Psi_i(x, \lambda) = \Phi^{(i)}(x) - F^{(i)}(x, \lambda), \qquad i = 1, \cdots, k,$$

and k and n are defined by equation (6).

One should note that the approximation (9) does not require the cumulants associated with $F(\ ;\lambda)$ or Φ but makes use of the derivatives of the distribution functions instead. From the equations in (8) one would hope that $F(x;\lambda) - \hat{F}_n(x;\lambda) = O(\lambda^{-(n+1)/2})$. This turns out to be the case as the next theorem shows.

THEOREM 2. *If Theorem* 1 *holds and* $\hat{F}_n(x;\lambda)$ *is defined, then*

$$\hat{F}_n(x;\lambda) - F(x;\lambda) = O(\lambda^{-(n+1)/2}),$$

as $\lambda \to \infty$.

PROOF. From (6), (7) and (9)

$$\hat{F}_n(x;\lambda) = F(x;\lambda) + \frac{H_k[\varepsilon_0(x, n, \lambda), \varepsilon_i(x, n, \lambda); \Phi^{(m_i)}(x)]}{H_k[1, 0; \Phi^{(m_i)}(x)]}, \tag{10}$$

where

$$\varepsilon_i(x, n, \lambda) = F_n^{(i)}(x, \lambda) - F^{(i)}(x, \lambda) = O(\lambda^{-(n+1)/2}), \qquad i = 0, 1, \cdots, k,$$

and F_n is defined in (5).

Now since $\Phi(x)$ is not a function of λ we can write (10) in the form

$$\hat{F}_n(x;\lambda) = F(x;\lambda) + \textstyle\sum_{i=0}^k c_i(x)\varepsilon_i(x, n, \lambda) \tag{11}$$

where the $c_i(x)$ are not functions of λ. The theorem therefore follows by elementary properties of the order function.

Theorem 2 establishes the fact that $\hat{F}_n(x;\lambda)$ and $F_n(x;\lambda)$ are asymptotically equivalent. Whether or not this equivalence is representative of the relationship between $\hat{F}_n(x;\lambda)$ and $F_n(x;\lambda)$ for small n undoubtedly depends on the particular $\Phi(x)$ and $F(x;\lambda)$. In general, for small n, from the viewpoint of accuracy, neither approximation is preferred over the other and hence the choice of $\hat{F}_n(x;\lambda)$ or $F_n(x;\lambda)$ would normally be based on which is the easier to obtain. The two approximations are demonstrated in the following examples

EXAMPLE 1. Let $\Phi(x)$ be the $N(0, 1)$ cdf and let

$$F(x;\lambda) = \textstyle\int_{-\infty}^x \lambda^{\frac{1}{2}} g(t\lambda^{\frac{1}{2}} + \lambda)\, dt, \tag{12}$$

where

$$\begin{aligned} g(u) &= \frac{1}{\Gamma(\lambda)} u^{\lambda-1}e^{-u} \qquad && u > 0 \\ &= 0, && \text{elsewhere.} \end{aligned}$$

Then

$$\begin{aligned} k_i &= 0 \qquad && i = 1 \\ &= 1 && i = 2 \\ &= \lambda^{1-i/2}(i-1)! && i = 3, 4, \cdots. \end{aligned}$$

Moreover it is well known that $F(x; \lambda) \to \Phi(x)$ as $\lambda \to \infty$ for all real x. Further, since $\alpha_2 = 1$ and the remaining cumulants associated with Φ are zero, we have

$$\text{(13)} \qquad \begin{aligned} \beta_i &= 0\,, & i &= 1, 2 \\ &= \lambda^{1-(i/2)}(i-1)!\,, & i &= 3, 4, \cdots. \end{aligned}$$

Clearly the distributions here satisfy Theorem 2 and hence $\hat{F}_n(x; \lambda)$ and $F_n(x; \lambda)$ are asymptotically equivalent. Using (13), the Edgeworth expansion to $O(\lambda^{-(n+1)/2})$ can readily be calculated. On the other hand,

$$\begin{aligned} F^{(m)}(x; \lambda) &= \lambda^{m/2} D_u^{\,m-1} g(u)|_{u=x\lambda^{\frac{1}{2}}+\lambda} \\ &= \frac{\lambda^{m/2}}{\Gamma(\lambda)} \textstyle\sum_{i=0}^{m-1} (-1)^{m-1-i} \binom{m-1}{i} e^{-u} D_u^{\,i} u^{\lambda-1}|_{u=x\lambda^{\frac{1}{2}}+\lambda} \end{aligned}$$

where D_u denotes differentiation with respect to u. Thus both approximations can be easily calculated. A comparison of these approximations is given in Table 1 below. Both approximations seem adequate. However, $F_n(x, \lambda)$ appears better for small x, whereas $\hat{F}_n(x; \lambda)$ is usually better in the tails of the distribution. This is particularly significant if one realizes that $\hat{F}_n$ uses no more terms in the Edgeworth expansion than F_n and hence, even when the cumulants are easily obtained, $\hat{F}_n$ may be more useful.

As has already been noted $F_n(x; \lambda)$ can be difficult to obtain without a table of moments and a table of Edgeworth coefficients; however $\hat{F}_n(x; \lambda)$ can be obtained without such tables provided one can ascertain the order of coefficients of $\Phi^{(m)}(x)$ in an expansion of the general form

$$\text{(14)} \qquad F(x; \lambda) = \Phi(x) + \textstyle\sum_{m=1}^{\infty} c_m(\lambda) \Phi^{(m)}(x)\,.$$

These observations therefore seem to imply that $\hat{F}$ is a reasonable alternative to the Edgeworth approximation when the densities involved are known and in many cases it may be more attractive.

EXAMPLE 2. As a final example let us consider the Student t distribution. That is, suppose $\Phi(x)$ is the $N(0, 1)$ cdf and $F(x; \lambda)$ is the Student t cdf. Then the central moments associated with F are given by

$$\text{(13)} \qquad \begin{aligned} \mu_{2k} &= \lambda^k \Pi^{-\frac{1}{2}} \frac{\Gamma(k+\frac{1}{2})\Gamma((\lambda/2)-k)}{\Gamma(\lambda/2)} & 2k &< \lambda \\ \mu_{2k+1} &= 0 & 2k+1 &< \lambda\,, \end{aligned}$$

and μ_r does not exist for $r \geqq \lambda$.

Thus the Edgeworth expansion cannot be calculated for every λ. However, since $\hat{F}_n(x, \lambda)$ does not involve the moments, we can calculate it for each λ by simply utilizing the nonzero terms in the corresponding formal Edgeworth expansion to define the m_i in (6). Utilizing the relationship

$$\text{(14)} \qquad \begin{aligned} F^{(k+1)}(x, \lambda) = &-\frac{(\lambda+2k-1)x}{\lambda+x^2} F^{(k)}(x, \lambda) \\ &- \frac{(n-1)(\lambda+k-1)}{\lambda+x^2} F^{(k-1)}(x, \lambda)\,, \end{aligned}$$

$\hat{F}_n(x, \lambda)$ can then be easily calculated.

The accuracy of this approximation is considered in Table 2, which follows. It appears to be quite satisfactory for the values considered, which included some small λ, especially in the tails of the distribution.

A final remark should be made concerning (9). That is, one should note that we have assumed $H_k[1, 0; \Phi^{(m_i)}(x)] \neq 0$. However when that quantity is near zero some numerical problems in calculating (9) may arise. In fact the entry under $\lambda = 4$, $t = 1$, and $\hat{F}_2$ in Table 2 below is a good example of this remark. Thus in these regions some caution should be taken by the user.

TABLE 1

λ	x	F	$\hat{F}_1$	F_1	$\hat{F}_2$	F_2	$\hat{F}_4$	F_4
5	1	0.84748	0.84135	0.84135	0.84599	0.84941	0.84885	0.84770
	2	0.95902	0.97165	0.95310	0.96442	0.95850	0.95959	0.95804
	3	0.99069	0.99422	0.99335	0.99202	0.98849	0.99090	0.99178
	4	0.99812	0.99911	0.99967	0.99852	0.99866	0.99818	0.99783
15	1	0.84362	0.84135	0.84135	0.84325	0.84403	0.84381	0.84363
	2	0.96528	0.97075	0.96331	0.96692	0.96511	0.96538	0.96521
	3	0.99442	0.99575	0.99559	0.99476	0.99397	0.99444	0.99451
	4	0.99930	0.99957	0.99980	0.99037	0.99946	0.99931	0.99927
25	1	0.84276	0.84135	0.84315	0.84257	0.84296	0.84283	0.84276
	2	0.96763	0.97120	0.96645	0.96851	0.96753	0.96766	0.96760
	3	0.99552	0.99634	0.99629	0.99569	0.99531	0.99553	0.99554
	4	0.99955	0.99969	0.99983	0.99958	0.99963	0.99955	0.99954
100	1	0.84172	0.84135	0.84135	0.84169	0.84175	0.84172	0.84172
	2	0.97214	0.97318	0.97185	0.97228	0.97212	0.97214	0.97214
	3	0.99725	0.99746	0.99747	0.99727	0.99722	0.99725	0.99725
	4	0.99984	0.99986	0.99990	0.99984	0.99985	0.99984	0.99984

(x = standardized gamma values)

TABLE 2

λ	t	F	$\hat{F}_2$	$\hat{F}_4$
4	1	.81305	.70992	.81079
	2	.94194	.93634	.94704
	3	.98003	.97101	.97727
	4	.99193	.99609	.99476
6	1	.82204	.83597	.82120
	2	.95379	.94124	.95346
	3	.98800	.99502	.98621
	4	.99644	.99801	.99722
15	1	.83341	.83388	.83334
	2	.96803	.96766	.96798
	3	.99551	.99602	.99567
	4	.99942	.99959	.99947
20	1	.83537	.83558	.83534
	2	.97037	.97016	.97034
	3	.99646	.99674	.99653
	4	.99965	.99973	.99967

(t = non-standardized t values)

REFERENCES

[1] COBERLY, W. A. (1972). On Edgeworth expansions with unknown cumulants. Ph. D. dissertation, Texas Tech Univ.

[2] DRAPER, N. R. and TIERNEY, D. E. (1973). Exact formulas for additional terms in some important series expansions. *Comm. Statist.* **1** 495-524.

[3] HILL, G. W. and DAVIS, A. W. (1968). Generalized asymptotic expansions of Cornish-Fisher type. *Ann. Math. Statist.* **39** 1264-1273.

[4] KENDALL, M. G. and STUART, A. (1969). *The Advanced Theory of Statistics*, **1** (3rd ed.). Hafner, New York.

H. L. GRAY
DEPARTMENT OF STATISTICS
SOUTHERN METHODIST UNIVERSITY
DALLAS, TEXAS 75275

W. A. COBERLY
DEPARTMENT OF MATHEMATICS
TEXAS TECH
LUBBOCK, TEXAS 79409

T. O. LEWIS
DEPARTMENT OF MATHEMATICS
TEXAS TECH
LUBBOCK, TEXAS 79409