EDGEWORTH EXPANSIONS IN STATISTICS: SOME RECENT DEVELOPMENTS

by

Shanti S. Gupta
Purdue University

and

S. Panchapakesan
Southern Illinois University

Technical Report #82-18

Department of Statistics
Purdue University
(Revised December 1982)

*This research was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

EDGEWORTH EXPANSIONS IN STATISTICS: SOME RECENT DEVELOPMENTS *

by

Shanti S. Gupta
Purdue University

and

S. Panchapakesan
Southern Illinois University

1. Introduction. The study of approximations to distributions has been an important part of statistical investigations since the early part of this century. Charlier, Chebyshev, and Edgeworth were notable contributors among others. Though the interest initially was the approximation to empirical distributions by theoretical functions, the focus quickly shifted to approximate evaluation of distribution functions or quantiles of complicated distributions. Such approximations are very useful in statistical inference, especially in the investigation of robustness of standard tests of hypotheses and of estimators.

Consider a sequence of statistics, $\{T_N\}$, $N \geq 1$, where N usually denotes the sample size. The distribution function (d.f.) F_N of T_N is said to possess an asymptotic expansion valid to (r+1) terms if functions $A_0,\ldots,A_r$ can be found such that

$$(1.1)\qquad \left|F_N(x) - A_0(x) - \sum_{j=1}^{r} \frac{A_j(x)}{N^{j/2}}\right| = o(N^{-r/2}).$$

The expansion is said to be uniformly valid to (r+1) terms if

$$(1.2)\qquad \sup_x \left|F_N(x) - A_0(x) - \sum_{j=1}^{r} \frac{A_j(x)}{N^{j/2}}\right| = o(N^{-r/2}).$$

It is to be noted that, in defining the above concepts, Wallace (1958) requires the remainder to be $O(N^{-(r+1)/2})$. Our definition is in accord with

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Erdelyi (1956) and is used by Bickel (1974) and many other authors. An expansion which is valid to just one term gives an ordinary limit theorem. The higher order terms are of practical interest. The basic approximation A_0 can often be greatly improved by taking one or two additional terms of the expansion. Further, these expansions are useful and necessary to discriminate procedures equivalent to first order. This point has been clearly brought out by Hodges and Lehmann (1970).

Expansions of the type (1.1) are, in general, known as Edgeworth expansions. The related expansions of F_N^{-1} are the Cornish-Fisher expansions. Of course, there are some other expansions of interest such as the expansions for density functions and frequency functions of lattice random variables. In some cases, better approximations can be obtained based on saddle point method; however, this involves a deep knowledge of the characteristic function of F_N which may not be usually available. Our interest here is mainly confined to Edgeworth and Cornish-Fisher expansions in which $A_0(x) = \Phi(x)$, the standard normal distribution function (as, in fact, the case was, when they were first introduced). It is perhaps appropriate at this point to make a comment about the Edgeworth series and the related Charlier's A-series. Charlier's paper was published in 1905 and Edgeworth's in 1907. However, Gnedenko and Kolmogorov (1968, Chapter 8) point out that both these types of expansions appear already in the work of Chebyshev. For further historical information about these expansions, we refer to Cramér (1972) and Särndal (1971).

Another problem that has always been of interest is to obtain suitable bounds for $\sup_x |F_N(x) - A_0(x)|$ or in other words, to determine the rate of convergence to A_0. These bounds are called Berry-Esseen bounds. For basic contributions to this problem, one should refer to Berry (1941), Esseen (1942, 1945), and Bergström (1944, 1945, 1949). In the course of this paper, we will refer to some recent results regarding these bounds.

Significant contributions to the theory of asymptotic expansions were made by Rao (1960) who obtained Edgeworth expansions and Berry-Esseen bounds for sums of independent random vectors. These and other related developments in the field including the results of B. von Bahr, R. N. Bhattacharya and A. Bikjalis are discussed in Bhattacharya and Rao (1976). For excellent accounts of the theory of Edgeworth expansions for sums of independent random variables, one should also refer to Cramér (1962), Gnedenko and Kolmogorov (1968), Feller (1971) and Petrov (1975).

There are also three fine expository papers by Wallace (1958), Bickel (1974), and Pfanzagl (1980). Wallace has given a nice summary of the state of the art regarding asymptotic approximations to distributions. He has also discussed the uses of these approximations in problems such as the Behrens-Fisher problem and in the investigation of robustness of standard tests of hypotheses. Bickel (1974) surveys more recent work on Edgeworth expansions for M-estimates, rank tests and some other statistics arising in nonparametric models. Pfanzagl (1980), while developing a general parametric statistical theory based on asymptotic methods, has discussed the available literature on stochastic and Edgeworth expansions. In particular, he has discussed the numerical accuracy of results based on Edgeworth expansions.

In reviewing the development of Edgeworth expansions in statistics, we do not attempt to be exhaustive - neither in coverage nor in details. We restrict our attention to some of the recent developments. These relate to transformations of Edgeworth series (Section 2), Cornish-Fisher expansions

(Section 3), Berry-Esseen bounds and Edgeworth expansions for test-statistics such as linear rank statistics, U-statistics, and linear combinations of order statistics (Section 4), expansions for minimum contrast estimators and Fisher-consistent estimators (Section 5), Edgeworth and Cornish-Fisher expansions in selection and ranking problems (Section 6), and asymptotic expansions for statistics with nonnormal limiting distributions (Section 7).

2. Edgeworth Series for Sums and Tranformations of Series. Let $\{X_n\}$, $n \geq 1$, be a sequence of independent and identically distributed (i.i.d.) random variables and let F_n denote the d.f. of the standardized sum $Y_n = \sum_{i=1}^{n} (X_i-\mu)/\sigma\sqrt{n}$, where $\mu = E(X_1)$ and $\sigma^2 = V(X_1)$. Now, letting the cumulants $\kappa_r = \sigma^r\lambda_r$, $r \geq 3$, the first few terms of the Edgeworth expansion are given by

$$F_n(x) = \Phi(x) - \lambda_3 \frac{\Phi^{(3)}(x)}{6\sqrt{n}} + \frac{1}{n}\left[\frac{\lambda_4\Phi^{(4)}(x)}{24} + \frac{\lambda_3^2\ \Phi^{(6)}(x)}{72}\right] + \ldots,$$

where $\Phi^{(r)}(x)$ denotes the rth derivative of $\Phi(x)$ with respect to (w.r.t.) x. Cramér (1928) proved that the series is valid uniformly in x provided that one more cumulant exists than used in any partial sum and the characteristic function $\psi(t)$ of X_1 satisfies the condition

$$\limsup_{|t| \to \infty} |\psi(t)| < 1. \tag{2.1}$$

This condition (commonly now called the Cramér's condition) is satisfied if the distribution of X_1 has an absolutely continuous part. For discrete distributions, the condition is not satisfied and in this case the result is not generally valid; however, a different expansion is available. Takeuchi and Akahira (1977) have given Edgeworth expansion of F_n when moments do not necessarily exist but when the density can be approximated by rational functions.

The theory of Edgeworth expansions for sums of independent random variables is fairly well-developed and for excellent accounts one can refer to any of the books mentioned in this regard in the previous section.

Suppose we have a sequence of distributions of random vectors approximated to a certain order of accuracy by an Edgeworth series. Bhattacharya and Ghosh (1978) proved that such an expansion may be transformed by a sequence of smooth functions of the corresponding random vectors to yield a valid Edgeworth expansion of the resulting sequence of distributions. This latter expansion may be obtained by the so-called delta method in which the moments of a function of a random vector are formally calculated from a Taylor series expansion of the function. The conditions needed to be satisfied are essentially concerned with the derivatives of the functions. Skovgaard (1981a) gives a generalization of the Bhattacharya-Ghosh theorem. These results provide the mechanism for Edgeworth expansions of distributions of statistics in non-standard cases. Skovgaard (1981b) has used this approach to obtain Edgeworth expansion of the distribution of maximum likelihood estimators in the general (non i.i.d.) case.

3. Cornish-Fisher Expansions. Typically, in many statistical applications, we require quantiles of a distribution. Starting with an Edgeworth expansion of a distribution function F_n, Cornish and Fisher (1937) obtained an asymptotic expansion of a quantile x of F_n in terms of the corresponding normal quantile z by means of formal substitutions, Taylor expansion and identification of powers of n. The expansion is of the form

$$x = z + \frac{P_1(z)}{\sqrt{n}} + \frac{P_2(z)}{n} + \frac{P_3(z)}{n^{3/2}} + \frac{P_4(z)}{n^2} + \dots \tag{3.1}$$

where the $P_i(z)$ are polynomials in z. We refer to $P_i(z)$ as the ith adjustment. Cornish and Fisher (1937) have tabled, for nine probability levels, all the polynomials needed to obtain all terms through the fourth adjustment based on the sixth cumulant. Fisher and Cornish (1960) extended the formulae and tables to the sixth adjustment based on the eighth cumulant for ten (the earlier nine and one more) probability levels. These are further extended to order n^{-4} by Draper and Tierney (1973) and to order n^{-6} by Hill (1964).

Now, we can also expand z in terms of x in a form similar to (3.1). This is known as the normalizing expansion. In fact, this is obtained as an intermediate step in getting (3.1). It is also useful in itself as it provides an asymptotic transformation of a random variable X with distribution F_n into a standard normal random variable. Both expansions are referred to as Cornish-Fisher expansions. For an absolutely continuous distribution both expansions are valid for every probability level whenever the initial Edgeworth series is valid. This can be proved by adapting the proof of Wasow (1956) for the invertability of a special class of distribution expansions. Incidentally, Hill and Davis (1968) obtained formal expansions which generalize Cornish-Fisher relations to an arbitrary analytic Φ.

The Edgeworth and Cornish-Fisher expansions require the knowledge of the cumulants of the distributions involved. Gray, Coberly and Lewis (1975) showed how the general Edgeworth expansion can be suitably utilized to eliminate the requirement of knowing the cumulants without affecting the order of the error of approximation. McCune and Gray (1975) used this result along with the expansions of Hill and Davis (1968) to obtain Cornish-Fisher type expansions with unknown cumulants. However, these new expansions lack simplicity for applications. Using a technique analogous to that used by Gray, Coberly and Lewis (1975), a simpler expression was obtained by McCune (1977).

Finally, it should be noted that the Edgeworth and Cornish-Fisher expansions are known to have some deficiencies which show up in tails of the distribution. The approximations for $F_n(x)$ are not probability distributions. Further, the monotonicity property as well as the zero-one range property are violated in parts of either or both tails. Correspondingly, the approximations for quantiles are not always monotonic in the probability level.

4. Berry-Esseen Bounds and Edgeworth Expansions for Certain Test Statistics. As we have already mentioned, a reasonably complete theory of Berry-Esseen bounds and Edgeworth expansions is available for sums of independent random variables and vectors. In recent years, Berry-Esseen bounds and asymptotic expansions have been obtained for several statistics occurring in statistical estimation and tests which are of a different structure. In this section, we discuss these results for linear combination of order statistics, spacings, simple linear rank statistics and U-statistics.

In obtaining these results for statistics T_N in these situations, two methods have been used. One method is to obtain a stochastic expansion for T_N; in other words, T_N is approximated sufficiently accurately by a statistic T_N' which has a simpler structure. The desired results are then proved for T_N' instead of T_N. In many cases that were first considered, T_N' is a smooth function of a sum of independent random vectors and the problem is solved by the use of the classical theory. In some other cases, T_N' is of a different type. An important example of such T_N' is a U-statistic. This situation arises, for example, when we are dealing with one-sample linear rank statistic and linear combinations of order statistics.

Another technique used by Albers, Bickel and van Zwet (1976) and Bickel and van Zwet (1978) to obtain Edgeworth expansions for the one- and two-sample problems is based on conditioning. With the right conditioning, it turns out easier to obtain an Edgeworth expansion for the conditional d.f. of the linear rank statistic T_N. One can then obtain an expansion for the unconditional d.f. of T_N by taking the expected value.

Albers, Bickel and van Zwet (1976) gave a rigorous proof of Edgeworth expansion for the d.f. of the general linear rank statistic for the one-sample problem under the null hypothesis as well as general alternatives. They have also shown that the expansion can be greatly simplified by considering contiguous location alternatives and smooth scores. Numerical aspects of these expansions are considered in Albers (1974). Albers (1979) specialized the general results to certain contiguous nonparametric alternatives and smooth scores. Bickel and van Zwet (1978) provided the Edgeworth expansion for the general two-sample linear rank statistic under the null hypothesis as well as contiguous location alternatives. The expansion under the null hypothesis has been obtained also by Robinson (1978). Rogers (1971) obtained an expansion under the null hypothesis in the special case of the two-sample Wilcoxon statistic but his proof appears to be in error. A Berry-Esseen bound in this special case was obtained by Stoker (1954).

Suppose that $X_1, X_2, \ldots, X_N$ are independent random variables with density functions $f_1, f_2, \ldots, f_N$, respectively. Let R_j be the rank of X_j when the X_i are arranged in increasing order. For sequences of real numbers $c_1, \ldots, c_N$ and $a_1, \ldots, a_N$, we define a simple linear rank statistic by

$T_N = \sum_{j=1}^{N} c_j a_{R_j}$. Here the a_i are the scores. By taking $c_1 = \ldots = c_m = 0$, $c_{m+1} = \ldots = c_N = 1$ for $1 < m < N$, we get the two-sample rank statistic. The behavior of the characteristic function of a suitably standardized T_N for the large values of the argument has been investigated by van Zwet (1980). Using this result, an Edgeworth expansion with remainder $o(N^{-1})$ was obtained for simple linear rank statistics under the null hypothesis by Does (1981) who considered scores a_j generated by a function $J(t)$, $0 < t < 1$, using either $a_j = J(j/(N+1))$ or $a_j = E(J(U_{j,N}))$, $j = 1, \ldots, N$, where $U_{j,N}$ denotes the jth order statistic in a random sample of size N from the uniform distribution on (0,1). His theorem holds for a wide class of functions J which includes the normal quantile function, thus allowing unbounded scores.

Berry-Esseen bounds of order $O(N^{-\frac{1}{2}})$ for simple linear rank statistics under the null-hypothesis have been obtained by von Bahr (1976) and Ho and Chen (1978) as a by-product of their results for statistics of the form $\sum_{j=1}^{N} X(j,R_j)$ where $X = \{X(i,j): 1 \le i,j \le N\}$ is a square matrix of random variables with independent row vectors. Hušková (1977) has obtained the same bound under the null-hypothesis as well as under contiguous alternatives. All these results were obtained for bounded scores. Recently, Does (1982b) has obtained the same bound under the null-hypothesis for a wide class of scores generating functions as in the case of his Edgeworth expansion mentioned above. For these results of Does as well as asymptotic expansions under contiguous alternatives, reference can also be made to Does (1982a).

Suppose that $X_{1,N} \le X_{2,N} \le \cdots \le X_{N,N}$ are order statistics corresponding to $X_1, X_2, \ldots, X_N$ which are i.i.d. with distribution F. Let $T_N = N^{-1} \sum_{j=1}^{N} c_{j,N} X_{j,N}$, where the $c_{j,N}$ are known real numbers (weights). Statistics of the form T_N are linear combinations of order statistics. In this case, Berry-Esseen bounds are given by Bjerve (1977) and Helmers (1977, 1981a,b). Edgeworth expansion for T_N with smooth weight functions is obtained by Helmers (1980) under the assumption that the underlying distribution function possesses a finite fourth moment and a certain local smoothness property. His proof uses Theorem 4.1 of van Zwet (1977) which provides a bound for the characteristic function of T_N. For the special case in which the underlying distribution is uniform, the above results of Bjerve, Helmers and van Zwet are substantially weaker than the result of van Zwet (1979).

The results of Helmers (1980) for linear combination of order statistics do not include trimmed means. However, Bjerve (1974) has shown that trimmed means admit asymptotic expansion. His proof employs a special property of trimmed means and thus does not apply to more general trimmed linear combinations. Expansion in the general case has been obtained by Helmers (1979).

These and related results are summarized in Helmers (1981b). A brief but clear account of the results of Bjerve (1977) and Helmers (1977) is given by van Zwet (1977).

Suppose that $U_1, U_2, \ldots$ is a sequence of i.i.d. random variables which are uniformly distributed on (0,1). For $N = 1,2,\ldots$, let $U_{1,N} \leq U_{2,N} \leq \cdots \leq U_{N,N}$ be the order statistics. Let $U_{0,N} = 0$ and $U_{N+1,N} = 1$. The spacings are defined by $D_{iN} = U_{i,N} - U_{i-1,N}$, $i = 1,\ldots,N+1$. Let g_N: $(0,\infty) \to \mathbb{R}$, $N \geq 1$, be a sequence of measurable functions and define $T_N = \sum_{i=1}^{N+1} g_N((N+1)D_{iN})$, $N \geq 1$. Does and Helmers (1980) have obtained an Edgeworth expansion for T_N. They have also established a Berry-Esseen bound of order $O(N^{-\frac{1}{2}})$ for normalized T_N but under conditions that are hard to check. The same bound is obtained by Does and Klaassen (1981) under conditions which are easier to check.

Another important class of statistics are known as U-statistics. In finding a T'_N which stochastically approximates T_N at hand, a possible situation is that T'_N is a U-statistic of order k ($1 \leq k \leq N$). In other words,

$$T'_N = \sum_{1 \leq i_1} \cdots \sum_{< \cdots < i_k \leq N} h(X_{i_1}, \ldots, X_{i_k}),$$

where $X_1, \ldots, X_N$ are i.i.d. random variables. One-sample linear rank statistics and linear combinations of order statistics are known examples of such T_N. Berry-Esseen type results establishing convergence to normality of the d.f. of a standardized T'_N at the rate of $N^{-\frac{1}{2}}$ have been given by Bickel (1974), Chan and Wierman (1977) and Callaert and Janssen (1978). These authors discuss U-statistics of order 2; however, as pointed out by these authors, all their results hold for U-statistics of any fixed order and also hold in the multi-sample case [see Janssen (1978)].

To describe the results for U-statistics of order 2, define

$$U_N = \sum_{1 \leq i < j \leq N} h(X_i, X_j),$$

where the kernel h is a symmetric function of two variables with $Eh(X_1,X_2) = 0$. Let σ_N^2 denote the variance of U_N. Asymptotic normality of $\sigma_N^{-1}U_N$ was proved by Hoeffding (1948) under the sole condition that $Eh^2(X_1,X_2)$ exists. Grams and Serfling (1973) showed that $\sup_x|P[\sigma_N^{-1}U_N \leq x] - \Phi(x)| = O(N^{-r/(2r+1)})$ when $Eh^{2r} < \infty$; this gives $O(N^{-\frac{1}{2}+\epsilon})$ when h has finite moments of all orders. An order bound of exactly $O(N^{-\frac{1}{2}})$ was obtained by Bickel (1974) assuming that the kernel h is bounded. The same order bound was obtained by Chan and Wierman (1977) assuming only the existence of the fourth moment of h and by Callaert and Janssen (1978) under the assumption that the third absolute moment of the kernel exists. Recently, Helmers and van Zwet (1982) relaxed this moment condition even further by requiring that $E|h(X_1,X_2)|^p < \infty$ for some $p \geq 5/3$, $Eg^2(X_1) > 0$, and $E|g(X_1)|^3 < \infty$, where $g(x) = E(h(X_1,X_2)|X_1 = x)$. A Berry-Esseen theorem for U-statistics when the sample size is random is given by Ahmad (1980) when the random size is independent of the observations. A special case of Ahmad's result is a Berry-Esseen theorem for random sums which is also discussed by Landers and Rogge (1976); however, the latter authors do not assume the independence between the sample observations and the sample size but assume more stringent conditions on $\{N_n\}$, the sequence of positive integer-valued random variables whose values are the sample sizes. The latest of these results is that of Callaert and Aerts (1982). Their theorem holds for U-statistics when the sample size is random and includes the earlier results of Landers and Rogge (1976), Ahmad (1980), and Helmers and van Zwet (1982). When σ_N^2 is unknown, a Berry-Esseen bound of order $O(N^{-\frac{1}{2}})$ is given by Callaert and Veraverbeke (1982) for the studentized U-statistic $N^{\frac{1}{2}}S_N^{-1}U_N$ where $S_N/N^{\frac{1}{2}}$ is the jackknife estimator [see, for example, Miller (1968)] of σ_N^2. An

Edgeworth expansion for the d.f. of $\sigma_N^{-1} U_N$ with remainder $o(N^{-1})$ has been obtained by Callaert, Janssen and Veraverbeke (1979) assuming the existence of the fifth absolute moment of h and some regularity conditions on the kernel. The same result is established by these authors in another paper (1980) under a set of less restrictive but less tractable set of conditions.

Recently, Efron (1979) gave a "bootstrap" method for setting confidence intervals and estimating significance levels. The bootstrap procedure is a resampling procedure to approximate the distribution of a function of the observations and the underlying distribution. The approximation is called the bootstrap distribution of the quantity. Bickel and Freedman (1981) have obtained some asymptotic results regarding bootstrap pivotal quantities. Singh (1981) has obtained Edgeworth expansions with remainder $o(n^{-1/2})$ for the bootstrap approximation of the standardized sample mean in the lattice as well as non-lattice case, under the assumption that the absolute third moment exists. He has also obtained the convergence rates of the bootstrap approximation of the distributions of sample quantiles.

Finally, for the class of Neyman's $C(\alpha)$ tests, Chibisov (1972a) obtained the term of order $N^{-\frac{1}{2}}$ in the normal approximation for the distributions of the test statistics. In another paper, he (1972b) derived asymptotic expansions for the distributions of some test criteria for testing compound hypotheses.

5. Asymptotic Expansions for Estimators. In this section, we discuss asymptotic expansions for a number of statistics occurring in parametric models such as maximum likelihood (m.ℓ.) estimators and, more generally, minimum contrast (m.c.) estimators. Of course, there could be tests based on these estimators and thus could have been included in the previous section. But they are considered here to emphasize their roles in estimation. Most of these results are reviewed by Pfanzagl (1980).

Linnik and Mitrofanova (1965) obtained asymptotic expansions for the distributions of m.ℓ. estimators. Mitrofanova (1967) extended the result for vector parameters. As pointed out by Pfanzagl (1973a, p. 998) and Bickel (1974, p. 11), her proof contains serious gaps. Chibisov (1973 a,b) extended the work of Linnik and Mitrofanova (1965) to a wider class of m.c. estimators for a single parameter under progressively much weaker conditions. Pfanzagl (1973a) also deals with Edgeworth expansions for m.c. estimators for a single parameter. Earlier, Michel and Pfanzagl (1971) and Pfanzagl (1971) established the order of the error of the normal approximation for m.c. estimators. In another paper, Pfanzagl (1973b) showed that, for m.c. estimators of vector parameters, the approximation by normal distribution holds with an error of order $O(N^{-\frac{1}{2}})$ uniformly over the class of all convex sets. For a discussion of the results of Linnik and Mitrofanova (1965), Pfanzagl (1971, 1973a), and Michel and Pfanzagl (1971), reference should be made to Bickel (1974).

Skovgaard (1981b) has discussed the method of computing the Edgeworth expansions of the distributions of m.ℓ. estimators in the non i.i.d. case. He has also given the first four terms of the corresponding stochastic expansion.

Edgeworth expansions of distributions of Fisher-consistent estimators for curved exponential family of parent distributions (assumed to be dominated by the Lebesgue measure) are obtained by Ghosh, Sinha and Subrahmanyam (1979).

This result leads to second-order efficiency of the m.ℓ. estimator w.r.t. any bounded, bowl shaped loss function. The formal Edgeworth expansions here are not valid without the assumption of the dominating Lebesgue measure.

Another type of results, though not concerned directly with the distributions of estimators, is the rate of convergence to normality of stopping times associated with sequential estimation. One such result is given by Ghosh (1980).

6. Edgeworth and Cornish-Fisher Expansions in Selection and Ranking Procedures. We will restrict our attention here to the so-called subset selection procedures. For an overall view of the theory, we refer to Gupta and Panchapakesan (1979) and Gupta and Huang (1981). A typical problem associated with many classical subset selection rules is the evaluation of the constant d, $d > 0$, such that $P[\max_{1\le i\le k}(X_i - X_0) \le d] = P^*$ where P^* is specified and $X_0, X_1, \ldots, X_k$ are i.i.d. random variables with the real line as the support, or the evaluation of the constant c, $c > 1$, such that $P[\max_{1\le i\le k}(X_i/X_0) \le c] = P^*$ where $X_0, X_1, \ldots, X_k$ are i.i.d. random variables with the positive real axis as the support. We will now discuss one such problem.

Let $X_0, X_1, \ldots, X_p$ be i.i.d. normal random variables with mean μ and variance σ^2. Let s_ν^2 be an estimator of σ^2 which is independent of the X_i such that $\nu s_\nu^2/\sigma^2$ has a chi-square distribution with ν degrees of freedom. Define $Y = (X_{[p]} - X_0)/s_\nu$, where $X_{[p]} = \max(X_1, \ldots, X_p)$. The statistic Y arises in the problem of selecting a subset of normal populations (with a common unknown variance σ^2) which contains the "best" population with probability at least P^*. This problem was studied by Gupta (1956). For implementation of his procedure, we need to evaluate appropriate percentage points of the distribution of Y. This statistic also appears in a few other selection problems [see Gupta and Sobel (1957)]. The probability $P[Y < y]$ can be evaluated by using an Edgeworth expansion of standardized Y. For evaluating the percentage points of Y, we can use Cornish-Fisher expansions. We will illustrate this method of evaluation and discuss some numerical results.

For $r < \nu$, we can write μ_r', the rth moment of Y in the form

$$\mu_r' = E[(\frac{s_\nu}{\sigma})^{-r}]\, E[(\frac{X_{[p]}-X}{\sigma})^r] = A_{-r} \sum_{j=0}^{[\frac{r}{2}]} \binom{r}{2j} \frac{(2j)!}{2^j j!}\, a_{p,r-2j}, \tag{6.1}$$

where $[r/2]$ is the integral part of $r/2$, $a_{p,i}$ is the ith moment of the largest of p independent standard normal random variables, and

$$A_\beta = E[(\frac{s_\nu}{\sigma})^\beta] = \frac{\Gamma(\frac{\nu+\beta}{2})}{(\frac{\nu}{2})^{\beta/2}\, \Gamma(\frac{\nu}{2})} \tag{6.2}$$

provided that $\beta > -\nu$.

Let σ^2 be the variance of Y. Let $y(P^*)$ and $y_s(P^*)$ denote the 100 P^* percentage points of the distributions of Y and $Y_s = (Y-\mu_1')/\sigma$, respectively. Let $\alpha_r = \kappa_r/\sigma^r$, where κ_r is rth cumulant of Y. The Cornish-Fisher expansion for $y_s(P^*)$ is given by

$$\begin{aligned} y_s(P^*) = z(P^*) &+ [\alpha_3 I_c] + [\alpha_4 I_d + \alpha_3^2 I_{c^2}] + [\alpha_5 I_e + \alpha_3\alpha_4 I_{cd} + \alpha_3^3 I_{c^3}] + \\ &[\alpha_6 I_f + \alpha_3\alpha_5 I_{ce} + \alpha_4^2 I_{d^2} + \alpha_3^2\alpha_4 I_{c^2d} + \alpha_3^4 I_{c^4}] + \\ &[\alpha_7 I_g + \alpha_3\alpha_6 I_{cf} + \alpha_4\alpha_5 I_{de} + \alpha_3^2\alpha_5 I_{c^2e} + \alpha_3\alpha_4^2 I_{cd^2} + \\ &\alpha_3^3\alpha_4 I_{c^3d} + \alpha_3^5 I_{c^5}] + [\alpha_8 I_h + \alpha_3\alpha_7 I_{cg} + \\ &\alpha_4\alpha_6 I_{df} + \alpha_5^2 I_{e^2} + \alpha_3^2\alpha_6 I_{c^2f} + \alpha_3\alpha_4\alpha_5 I_{cde} + \\ &\alpha_4^3 I_{d^3} + \alpha_3^3\alpha_5 I_{c^3e} + \alpha_3^2\alpha_4^2 I_{c^2d^2} + \alpha_3^4\alpha_4 I_{c^4d} + \alpha_3^5 I_{c^5}] + \ldots, \end{aligned} \tag{6.3}$$

where $z(P^*)$ is the standard normal deviate corresponding to P^* and I_c, I_d, I_{c^2},... are tabulated in Table II of Fisher and Cornish (1960) for $P^* = .5, .75, .90, .975, .99, .995, .9975, .999$ and $.9995$. Now, $y(P^*) = \mu_1' + \sigma y_s(P^*)$. For calculating μ_r' using (6.1), the values of $a_{p,i}$ are available in Ruben (1954) for $i = 1(1)10$ and $p = 1(1)50$.

Gupta and Sobel (1957) tabulated the $y(P^*)$-values correct two decimal places using four adjustment terms. Their table ranges over the following values: $P^* = .75, .90, .95, .975, .99$; $p = 1,4,9(1)15(2)19(5)39,49$; $\nu = 15(1)20,24,30,36,40,48,60,80,100,120,360,\infty$. The $y(P^*)$-values can also be obtained from the tables of Krishnaiah and Armitage (1966) for $P^* = .95, .99$; $p = 1(1)10$ and $\nu = 5(1)35$. These latter tables were computed using Gauss-Hermite and Gauss-Laguerre quadrature formulas. The two tables agree well in all the common cases.

Another method of evaluating the percentage points of Y is based on the result of Hartley (1943-46), who obtained the probability $P[Y \leq y]$ as a solution of certain difference-differential equation. This result can be stated in the form

$$(6.4) \qquad P[Y \leq y] = I_0(y) + \frac{1}{4\nu}[y^2\{I_2(y)-I_0(y)\} - yI_1(y)] + \frac{1}{16\nu^2}[\frac{1}{2}y^4\{3I_0(y)-6I_2(y) + I_4(y)\} - \frac{1}{3}y^3\{I_3(y)-3I_1(y)\} - \frac{1}{2}y^2\{I_2(y)-I_0(y)\} + \frac{1}{2}yI_1(y)],$$

where

$$(6.5) \qquad I_j(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(t-y)^2}{2}}(t-y)^j\,\Phi^p(t)dt.$$

It should be noted that $I_0(\sqrt{2}H)$ is the probability that the maximum of p equally correlated standard normal variates with the correlation equal to 0.5 does not exceed H. This probability for several values of H and the values of H corresponding to several probability levels are tabulated by Gupta, Nagel and Panchapakesan (1973).

In Table 1 below we give the values of $y(P^*)$ obtained in different ways for $P^* = .90, .95$; $p = 1, 4, 9$ and $\nu = 17, 18, 20, 60$. For each combination of the values of P^*, p and ν, the first three entries correspond to the values obtained by taking four, six and eight adjustment terms, respectively, in the Cornish-Fisher expansion. The last entry is obtained by evaluating $P[Y \leq y]$ in (6.4) using Gauss-Hermite quadrature formula and bisection method. All these computations except those based on Cornish-Fisher expansion with four adjustment terms are new with this paper.

TABLE 1

VALUES OF $y(P^*)$ OBTAINED BY CORNISH-FISHER EXPANSION WITH 4, 6, AND 8 ADJUSTMENT TERMS (LINES 1, 2 AND 3) AND HARTLEY'S EXPANSION WITH TWO ADJUSTMENT TERMS (LINE 4)

$P^* = .90$

ν \ p	1	4	9
17	1.89	2.75	3.18
	1.88824	2.75879	3.19306
	1.86438	2.53829	2.62070
	1.88570	2.74894	3.18010
18	1.88	2.74	3.17
	1.88323	2.74729	3.17778
	1.86846	2.59753	2.78203
	1.88147	2.74023	3.16854
20	1.87	2.72	3.15
	1.87525	2.72949	3.15419
	1.86881	2.65339	2.95164
	1.87433	2.72557	3.14908
60	1.83	2.64	3.04
	1.83257	2.64050	3.03672
	1.83256	2.64018	3.03602
	1.83257	2.64048	3.03670

TABLE 1 (cont.)

P* = .95

ν \ p	1	4	9
17	2.46 2.47450 2.41414 2.46047	3.29 3.32350 3.30553 3.29995	3.72 3.76197 3.90291 3.73373
18	2.45 2.46248 2.42389 2.45258	3.28 3.30146 3.30025 3.28516	3.70 3.73310 3.84884 3.71421
20	2.44 2.44463 2.42682 2.43929	3.25 3.26891 3.27674 3.26042	3.67 3.69093 3.76500 3.68176
60	2.36 2.36268 2.36264 2.36267	3.12 3.12065 3.12092 3.12065	3.50 3.50159 3.50225 3.50162

We can see from Table 1 some general indications about the effect of increasing the number of adjustment terms. It appears that

(1) the approximations based on four adjustment terms are generally quite good, certainly to two decimal places;

(2) adding additional terms may not produce desirable results for small values of ν, large values of p, and P* values closer to 1;

(3) as ν gets larger than 20, the terms in the expansion stabilize faster and more so for small p; and

(4) the values obtained by using Hartley's result with two adjustment terms seem to be adequate.

7. Concluding Remarks. In the previous sections, we discussed asymptotic expansions for the distribution functions of random variables which have a limiting normal distribution. However, there are test statistics of practical importance which have nonnormal limiting distributions, for example, the statistics that arise in likelihood ratio, goodness of fit, Kolmogorov-Smirnov, Cramér-von Mises and Pearson chi-square tests. We will briefly refer to expansions in some of these cases.

Let X have a continuous d.f. F(x) and let $F_n(x)$ denote the empirical distribution corresponding to n independent observations on X. Define $D_n^+ = \sup_{-\infty<x<\infty} \sqrt{n}\,(F_n(x)-F(x))$ and $D_n = \sup_{-\infty<x<\infty} \sqrt{n}\,|F_n(x)-F(x)|$. Smirnov (1944) studied the distribution of D_n^+ and showed that, for $0 < x < O(n^{1/6})$,

$$\Phi_n^+(x) = P[D_n^+<x] = 1 - e^{-2x^2}[1+ \frac{2x}{3\sqrt{n}}] + O(\frac{1}{n}). \tag{4.1}$$

This result was improved by Li-Tsian Chan (1955) who added two more terms for the same range of x; he showed that

$$\Phi_n^+(x) = 1 - e^{-2x^2}[1+ \frac{2x}{3\sqrt{n}} + \frac{2x^2}{3n}(1- \frac{2x^2}{3}) + \frac{4x}{9n^{3/2}}(\frac{1}{5} - \frac{19x^2}{15} + \frac{2x^4}{3})] + O(n^{-2}). \tag{4.2}$$

These results are discussed in Gnedenko, Koroluk and Skorokhod (1961). Later, Lauwerier (1963) gave the following expansion for $\Phi_n^+(x)$:

$$\Phi_n^+(x) = 1 - \frac{n!}{n^n e^{-n}\sqrt{2\pi n}}\, e^{-2x^2} \sum_{j=0}^{\infty} f_j(H)n^{-j/2}, \tag{4.3}$$

where the symbolic expression $f_j(H)$ stands for a polynomial $f_j(t)$ in which the powers t^m are replaced by the Hermite polynomials $H_m(2x)$. The polynomials f_j are determined by certain generating series. Li-Tsian Chan (1956) obtained also an expansion for the d.f. of D_n with a remainder $O(n^{-2})$. Another reference in this connection is Borovkov (1970), who treated these as special cases of more general problems of first passage times of random walks.

For the likelihood ratio criterion, Hayakawa (1977) gave a formal expansion. This was later justified by the results of Chandra and Ghosh (1980), who gave valid expansions for the likelihood ratio statistic as well as Wald's and Rao's statistics under contiguous alternatives. The general results of Chandra and Ghosh are obtained under a set of conditions in an earlier paper of theirs (1979) besides Cramér's condition and smoothness conditions on moments.

Robinson (1980) has obtained an asymptotic expansion for permutation tests with several samples. Let V_n denote the standardized sum of squares of the means of r+1 random samples of sizes $s_0, s_1, \ldots, s_r$ $(n = s_0 + \ldots + s_r)$ taken without replacement from n numbers. The asymptotic expansion for the d.f. of V_n has the d.f. of a chi-square random variable with r degrees of freedom as the first term and has an approximation error which is generally of smaller order than n^{-1}.

An excellent exposition of Edgeworth and saddle-point approximations for the densities of sums of independent random vectors has been given by Barndorff-Nielson and Cox (1979). Besides an earlier paper by Daniels (1954), some recent papers dealing with saddle-point approximations are Phillips (1978) and Daniels (1980).

Finally, it is important to remind ourselves that there is no useful numerical bound for the error term. Any general error bound which might become available is expected to be a gross overestimate. The only way to assess the degree of numerical accuracy is to study the actual errors in particular examples.

Acknowledgements

We are grateful to Ms. Regina Becker and Mr. Joong Kweon Sohn for all programming and computational help.

Thanks are also due to a referee for some useful suggestions.

REFERENCES

Ahmad, I. A. (1980). On the Berry-Esseen theorem for random U-statistics. Ann. Statist. 8, 1395-1398.

Albers, W. (1974). Asymptotic Expansions and the Deficiency Concept in Statistics. Tract No. 58, Mathematical Centre, Amsterdam.

Albers, W. (1979). Asymptotic expansions for the power of one-sample tests against contiguous nonparametric alternatives. Proc. Second Prague Symposium on Asymptotic Statistics, Amsterdam: North-Holland, pp. 105-117.

Albers, W., Bickel, P. J. and van Zwet, W. R. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. Ann. Statist. 4, 108-157. Correction. Same Annals 6 (1978), 1170-1171.

Bahr, B. von (1976). Remainder term estimate in a combinatorial limit theorem. Z. Wahrscheinlichkeitstheorie verw. Gebeite 35, 131-139.

Barndorff-Nielson, O. and Cox, D. R. (1979). Edgeworth and saddle-point approximations with statistical applications (with discussion). J. Roy. Statist. Soc. Ser. B 41, 279-312.

Bergström, H. (1944). On the central limit theorem. Skand. Aktuarietidskr. 27, 139-153.

Bergström, H. (1945). On the central limit theorem in the space R_k, $k > 1$. Skand. Aktuarietidskr. 28, 106-127.

Bergström, H. (1949). On the central limit theorem in case of not equally distributed random variables. Skand. Aktuarietidskr. 32, 37-62.

Berry, A. C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. Trans. Amer. Math. Soc. 49, 122-136.

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. Ann. Statist. 6, 434-451. Correction: same Annals 8 (1980), 1399.

Bhattacharya, R. N. and Rao, Ranga R. (1976). Normal Approximation and Asymptotic Expansions. New York: John Wiley.

Bickel, P. J. (1974). Edgeworth expansions in nonparametric statistics. Ann. Statist. 2, 1-20.

Bickel, P. J. and van Zwet, W. R. (1978). Asymptotic expansions for the power of distributionfree tests in the two-sample problem. Ann. Statist. 6, 937-1004.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. Ann. Statist. 9, 1196-1217.

Bjerve, S. (1974). Error Bounds and Asymptotic Expansions for Linear Combinations of Order Statistics. Ph.D. Thesis, Department of Statistics, University of California, Berkeley.

Bjerve, S. (1977). Error bounds for linear combinations or order statistics. Ann. Statist. 5, 357-369.

Borovkov, A. (1970). Limit theorems for random walks with boundaries. Proc. Sixth Berkeley Symp. Math. Statist. Vol. III, 19-30.

Callaert, H. and Aerts, M. (1982). Exact approximation order in the central limit theorem for random U-statistics. Tech. Report, Limburgs Universitair Centrum, Diepenbeek.

Callaert, H. and Janssen, P. (1978). The Berry-Esseen theorem for U-statistics. Ann. Statist. 6, 417-421.

Callaert, H., Janssen, P. and Veraverbeke, N. (1979). Some asymptotic results for the distribution function of U-statistics. Proc. Second Prague Symposium on Asymptotic Statistics, Amsterdam: North-Holland, pp. 105-117.

Callaert, H., Janssen, P. and Veraverbeke, N. (1980). An Edgeworth expansion for U-statistics. Ann. Statist. 8, 299-312.

Callaert, H. and Veraverbeke, N. (1981). The order of the normal approximation for a studentized U-statistic. Ann. Statist. 9, 194-199.

Chan, Li-Tsian (1955). The exact statistical distribution of N. V. Smirnov and its asymptotic expansion. Shuxue Jinshan 1, 775-790 (In Chinese). Russian translation in the Mathematika Collection, Vol. 4 (1960), 121-134.

Chan, Li-Tsian (1956). The exact statistical distribution of A. N. Kolmogorov and its asymptotic expansion. Acta Math. Sinica 6, 55-79 (In Chinese). Russian translation in the Mathematika Collection, Vol. 4 (1960), 135-159.

Chan, Y-K. and Wierman, J. (1977). On the Berry-Esseen theorem for U-statistics. Ann. Prob. 5, 136-139.

Chandra, T. K. and Ghosh, J. K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables. Sankhyā Ser. A 41, 22-47.

Chandra, T. K. and Ghosh, J. K. (1980). Valid asymptotic expansions for the likelihood ratio and other statistics under contiguous alternatives. Sankhyā Ser. A 42, 170-184.

Charlier, C.V.L. (1905). Über das Fehlergesetz. Ark. Mat. Astr. Fys. 2, No. 8.

Chibisov, D. M. (1972a). On the normal approximation for a certain class of statistics. Proc. Sixth Berkeley Symp. Math. Statist. Prob. 1, 153-174.

Chibisov, D. M. (1972b). Asymptotic expansions for the distributions of some statistical criteria for testing compound hypotheses. Theor. Prob. Appl. 17, 563-565.

Chibisov, D. M. (1973a). An asymptotic expansion for a class of estimators containing maximum likelihood estimators. Theor. Prob. Appl. 18, 295-303.

Chibisov, D. M. (1973b). An asymptotic expansion for the distribution of sums of a special form with an application to minimum contrast estimates. Theor. Prob. Appl. 18, 649-661.

Cornish, E. A. and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. Rev. de l'Inst. Int. de Stat. 5, 307-320.

Cramér, H. (1928). On the composition of elementary errors. Skand. Aktuarietidskr. 11, 13-74 and 141-180.

Cramér, H. (1962). Random Variables and Probability Distributions, 2nd Ed. Cambridge University Press.

Cramér, H. (1972). On the history of certain expansions used in mathematical statistics. Biometrika 59, 205-207.

Daniels, H. E. (1954). Saddlepoint approximation in statistics. Ann. Math Statist. 25, 631-650.

Daniels, H. E. (1980). Exact saddle-point approximations. Biometrika 67, 59-63.

Does, R.J.M.M. (1981). An Edgeworth expansion for simple linear rank statistics under the null-hypothesis. Report 78/81, Mathematical Centre, Amsterdam.

Does, R.J.M.M. (1982a). Higher Order Asymptotics for Simple Linear Rank Statistics. Tract No. 151, Mathematical Centre, Amsterdam.

Does, R.J.M.M. (1982b). Berry-Esseen theorems for simple linear rank statistics under the null-hypothesis. Ann. Prob. 10, 982-991.

Does, R.J.M.M. and Helmers, R. (1980). Edgeworth expansions for functions of uniform spacings. Report SW 55/80. Mathematical Centre, Amsterdam. To appear in Proc. of the Colloquium on Nonparametric Statistical Inference, (Ed. I. Vincze), Budapest.

Does, R.J.M.M. and Klaassen, C.A.J. (1981). A Berry-Esseen theorem for functions of uniform spacings. Report SW 76/81, Mathematical Centre, Amsterdam.

Draper, N. R. and Tierney, D. E. (1973). Exact formulas for additional terms in some important series expansions. Commun. Statist. 1, 495-524.

Edgeworth, F. Y. (1907). On the representation of statistical frequency by a series. J. Roy. Statist. Soc. 70, 102-106.

Erdelyi, A. (1956). Asymptotic Expansions, New York: Dover.

Esseen, C. G. (1942). On the Liapounoff limit of error in theory of probability. Ark. Mat. Astr. och Fys. 28A, No. 9, 19 pp.

Esseen, C. G. (1945). Fourier analysis of distribution functions. Acta Math. 77, 1-125.

Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol. 2, 2nd Ed. New York: John Wiley.

Fisher, R. A. and Cornish, E. A. (1960). The percentile points of distributions having known cumulants. Technometrics 2, 209-225. Errata: 2, 523.

Ghosh, J. K., Sinha, B. K. and Subrahmanyam, K. (1979). Edgeworth expansions for Fisher-consistent estimators and second order efficiency. Calcutta Statist. Assoc. Bull. 28, 1-18. Addendum. 28, 171.

Ghosh, M. (1980). Rate of convergence to normality for random means: applications to sequential estimation. Sankhyā Ser. A 42, 231-240.

Gnedenko, B. V. and Kolmogorov, A. N. (1968). Limit distributions for Sums of Independent Random Variables, 2nd Edition. Cambridge, Mass: Addison-Wesley.

Gnedenko, B. V., Koroluk, V. S. and Skorokhod, A. V. (1961). Asymptotic expansions in probability theory. Proc. Fourth Berkeley Symp. Math. Statist. Prob. Vol. II, 153-170.

Grams, W. F. and Serfling, R. J. (1973). Convergence rates for U-statistics and related statistics. Ann. Statist. 1, 153-160.

Gray, H. L., Coberly, W. A. and Lewis, T. O. (1975). On Edgeworth expansions with unknown cumulants. Ann. Statist. 3, 741-746.

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeo. Ser. No. 150, Institute of Statistics, Univ. of North Carolina, Chapel Hill, North Carolina.

Gupta, S. S. and Huang, D. Y. (1981). Multiple Decision Theory: Recent Developments. Lecture Notes in Statistics, Vol. 6, New York: Springer-Verlag.

Gupta, S. S., Nagel, K. and Panchapakesan, S. (1973). On the order statistics from equally correlated normal random variables. Biometrika 60, 403-413.

Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. New York: John Wiley.

Gupta, S. S. and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. Ann. Math. Statist. 28, 957-967.

Hartley, H. O. (1943-46). Studentization or the elimination of standard deviation of the parent population from the random sample distribution of statistics. Biometrika 33, 173-180.

Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. Ann. Inst. Statist. Math. 29, 359-378.

Helmers, R. (1977). The order of the normal approximation for linear combinations of order statistics with smooth weight functions. Ann. Prob. 5, 940-953.

Helmers, R. (1979). Edgeworth expansions for trimmed linear combinations of order statistics. Proc. Second Prague Symposium on Asymptotic Statistics, Amsterdam: North-Holland, pp. 221-232.

Helmers, R. (1980). Edgeworth expansions for linear combinations of order statistics with smooth weight functions. Ann. Statist. 8, 1361-1374.

Helmers, R. (1981a). A Berry-Esseen theorem for linear combinations of order statistics. Ann. Prob. 9, 342-347.

Helmers, R. (1981b). Edgeworth Expansions for Linear Combinations of Order Statistics. Tract No. 105, Mathematical Centre, Amsterdam.

Helmers, R. (1982). A Berry-Esseen bound for studentized U-statistics. Tech. Report, Mathematical Centre, Amsterdam.

Helmers, R. and van Zwet, W. R. (1982). The Berry-Esseen bound for U-statistics. In Statistical Decision Theory and Related Topics-III, Vol. 1 (Ed. S. S. Gupta and J. Berger), New York: Academic Press, pp. 497-512.

Hill, G. W. (1964). Cornish-Fisher polynomials for asymptotic expansions to order n^{-6} of distributions specified by cumulants. Working paper No. 25, Graduate School of Business, Stanford University, 72 pp.

Hill, G. W. and Davis, A. W. (1968). Generalized asymptotic expansions of Cornish-Fisher type. Ann. Math. Statist. 39, 1264-1273.

Ho, S. T. and Chen, L.Y.H. (1978). An L_p bound for the remainder in a combinatorial central limit theorem. Ann. Prob. 6, 231-249.

Hodges, J. L. and Lehmann, E. L. (1970). Deficiency. Ann. Math. Statist. 41, 783-801.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, 293-325.

Hušková, M. (1977). The rate of convergence of simple linear rank statistics under hypothesis and alternatives. Ann. Statist. 5, 658-670.

Janssen, P. (1978). De Berry-Esseen stelling en een asymptotische ontwikkeling voor U-statistieken. Ph.D. Thesis, University of Leuven.

Krishnaiah, P. R. and Armitage, J. V. (1966). Tables for multivariate t distribution. Sankhyā Ser. B 28, 21-56.

Landers, D. and Rogge, L. (1976). The exact approximation order in the central limit theorem for random summation. Z. Wahrscheinlichkeitstheorie und verw. Gebiete 36, 264-283.

Lauwerier, H. A. (1963). The asymptotic expansion of the statistical distribution of N. V. Smirnov. Z. Wahrscheinlichkeitstheorie und verw. Gebeite 2, 61-68.

Linnik, Yu. V. and Mitrofanova, N. M. (1965). Some asymptotic expansions for the distribution of the maximum likelihood estimate. Sankhyā Ser. A 27, 73-82.

McCune, E. D. (1977). On Cornish-Fisher expansions as a ratio of determinants without cumulants. Commun. Statist.-Theor. Meth. A6, 243-250.

McCune, E. D. and Gray, H. L. (1975). On Cornish-Fisher expansions with unknown cumulants. Commun. Statist. 4, 1043-1055.

Michel, R. and Pfanzagl, J. (1971). The accuracy of the normal approximation for minimum contrast estimates. Z. Wahrscheinlichkeitstheorie verw. Gebeite 18, 73-84.

Miller, R. G. (1968). Jackknifing variances. Ann. Math. Statist. 39, 567-582.

Mitrofanova, N. N. (1967). An asymptotic expansion for the maximum likelihood estimate of vector parameters. Theor. Prob. Appl. 12, 364-372.

Petrov, V. V. (1975). Sums of Independent Random Variables. Berlin: Springer.

Pfanzagl, J. (1971). Berry-Esseen bound for minimum contrast estimates. Metrika 17, 82-91.

Pfanzagl, J. (1973a). Asymptotic expansions related to minimum contrast estimators. Ann. Statist. 1, 993-1026. Corrections: 2 (1974), 1357-1358.

Pfanzagl, J. (1973b). The accuracy of the normal approximation for estimates of vector parameters. Z. Wahrscheinlichkeitstheorie verw. Gebeite 25, 171-198.

Pfanzagl, J. (1980). Asymptotic expansions in parametric statistical theory. Chapter 1 in Developments In Statistics, Vol. 3 (Ed. P.R. Krishnaiah), New York: Academic Press.

Phillips, P. C. B. (1978). Edgeworth and saddlepoint approximations in the first-order noncircular autoregression. Biometrika 65, 91-98.

Rao, Ranga R. (1960). Some Problems in Probability Theory. D. Phil. Thesis, Calcutta University.

Robinson, J. (1978). An asymptotic expansion for samples from a finite population. Ann. Statist. 6, 1005-1011.

Robinson, J. (1980). An asymptotic expansion for permutation tests with several samples. Ann. Statist. 8, 851-864.

Rogers, W. F. (1971). Exact null distributions and asymptotic expansions for rank test statistics. Technical Report, Stanford University.

Ruben, H. (1954). On the moments of order statistics in samples from normal populations. Biometrika 41, 200-227.

Särndal, C. E. (1971). The hypothesis of elementary errors and the Scandinavian School in statistical theory. Biometrika 58, 375-391.

Singh, K. (1981). On the asymptotic accuracy of Effron's bootstrap. Ann. Statist. 9, 1187-1195.

Skovgaard, I. M. (1981a). Transformation of an Edgeworth expansion by a sequence of smooth functions. Scand. J. Statist. 8, 207-217.

Skovgaard, I. M. (1981b). Edgeworth expansions of the distributions of maximum likelihood estimators in the general (non i.i.d.) case. Scand. J. Statist. 8, 227-236.

Smirnov, N. V. (1944). An approximation to the distribution laws of random quantities determined by empirical data. Uspechi Mat. Nauk. 10, 179-206.

Stoker, D. J. (1954). An upper bound for the deviation between the distribution of Wilcoxon's test statistic for the two sample problem and its limiting normal distribution for finite samples, I, II. Indag. Math. 16, 599-606, 607-614.

Takeuchi, K. and Akahira, M. (1977). Extension of Edgeworth type expansion of the distribution of the sums of i.i.d. random variables in non-regular cases. Ann. Inst. Statist. Math. 29, 397-406.

van Zwet, W. R. (1977). Asymptotic expansions for the distribution functions of linear combinations of order statistics. In Statistical Decision Theory and Related Topics-II (Ed. S. S. Gupta and D. S. Moore), New York: Academic Press, pp. 421-437.

van Zwet, W. R. (1979). The Edgeworth expansion for linear combinations of uniform order statistics. Proc. Second Prague Symposium on Asymptotic Statistics, Amsterdam: North-Holland, pp. 93-101.

van Zwet, W. R. (1980). On the Edgeworth expansion for the simple linear rank statistic. Report SW 70/80, Mathematical Centre, Amsterdam.

Wallace, D. L. (1958). Asymptotic approximations to distributions. Ann. Math. Statist. 29, 635-654.

Wasow, W. (1956). On the asymptotic transformation of certain distributions into the normal distribution. Proceedings of Symposia in Applied Mathematics Vol. VI, Numerical Analysis, New York: McGraw-Hill, pp. 251-259.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	Technical Report #82-18
2. GOVT ACCESSION NO.	
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	Edgeworth Expansions in Statistics: Some Recent Developments
5. TYPE OF REPORT & PERIOD COVERED	Technical
6. PERFORMING ORG. REPORT NUMBER	Technical Report #82-18
7. AUTHOR(s)	Shanti S. Gupta and S. Panchapakesan
8. CONTRACT OR GRANT NUMBER(s)	ONR N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS	Purdue University Department of Statistics West Lafayette, IN 47907
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS	Office of Naval Research Washington, DC
12. REPORT DATE	June 1982
13. NUMBER OF PAGES	28
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Asymptotic expansions, Edgeworth and Cornish-Fisher expansions, linear rank statistics, U-statistics, linear combinations of order statistics, minimum contrast estimators, Fisher-consistent estimators, selection and ranking statistics, Kolmogorov-Smirnov statistics, likelihood ratio criterion.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Asymptotic expansions for distribution functions are of great importance in statistical inference. The basic aspects and some recent developments regarding Edgeworth expansions and Cornish-Fisher expansions are discussed in Sections 2 and 3, respectively. In recent years, substantial progress has been made in obtaining Berry-Esseen bounds and Edgeworth expansions for test statistics such as linear rank statistics, U-statistics, linear (trimmed and untrimmed) combinations of order statistics. An account of these results is given in Section 4 followed by expansions for minimum contrast estimators and Fisher-consistent estimators

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(Section 5). Some specific applications of Cornish-Fisher expansions to problems in selection and ranking are considered in Section 6. The last section includes a discussion on expansions for distributions of random variables which have nonnormal limiting distributions, such as Kolmogorov-Smirnov statistics and likelihood ratio statistics.