ASYMPTOTIC APPROXIMATIONS TO DISTRIBUTIONS[1]

By David L. Wallace

University of Chicago

1. Introduction. The study of approximations to distributions formed a major part of statistical developments during the early part of this century and included important work by Charlier, Edgeworth, Pearson and numerous others. The principal problem was the approximation to empirical distributions by theoretical functions and the methods proposed consisted chiefly either of choosing an approximating function from some class of functions, such as the Pearson type distributions or the Gram-Charlier functions, or of choosing a transformation of the variable which would reduce the distribution to approximate normality.

With the increasing importance of statistical inference, interest in the original problem of approximating to empirical distributions virtually disappeared. But interest in approximations has continued because of the increasing number and complexity of theoretical distributions and the need for usable approximations to them. In addition to the direct use for approximate evaluation of the distribution functions or the quantiles of complicated distributions, approximations have been valuable in such problems as the Behrens-Fisher problem and in the investigation of robustness of standard tests of hypotheses.

There are several general approaches to distribution approximations. The one to which I restrict attention is that of finding asymptotic expansions—in which the errors of approximation approach zero as some parameter, typically a sample size, approaches infinity. Essentially, the method consists of finding improvements to the large sample approximations used throughout statistics. A variety of expansions have been developed for many problems and the approximations are amenable to theoretical as well as empirical study.

In a simple and common form, each function $F_n(x)$ in a sequence of functions is approximated by any partial sum of a series

$$\sum_{i=0}^{\infty} \frac{A_i(x)}{(\sqrt{n})^i}$$

and the errors satisfy the condition

$$\left| F_n(x) - \sum_{i=0}^{r} \frac{A_i(x)}{(\sqrt{n})^i} \right| \leqq \frac{C_r(x)}{(\sqrt{n})^{r+1}},$$

that is, the errors, using any partial sum, are of the same order of magnitude as the first neglected term. I call an asymptotic expansion valid to r terms if the

Received January 17, 1958; revised April 28, 1958.

[1] An Address presented on September 13, 1957 at the Atlantic City meeting of the Institute of Mathematical Statistics by invitation of the IMS Committee on Special Invited Papers. This paper was prepared with the partial support of the Office of Naval Research.

first $r + 1$ partial sums have this property, and valid uniformly in x if the bounds $C_r(x)$ do not depend on x. (The theory of asymptotic expansions is given, for example, by Erdelyi [28].)

There are a few points on the use of asymptotic expansions which have caused some confusion. Frequently, an expansion can be extended validly to infinitely many terms. For any fixed n, the infinite series may be convergent, but in statistical applications usually is not. The asymptotic property is a property of finite partial sums, and though the addition of the next term will for sufficiently large n improve the approximation, for any prescribed n it may not do so. Typically the bounds $C_r(x)$ increase rapidly with r, and for small n only the first few terms are improvements.

Ideally, sharp values of $C_r(x)$ should be known. (This is rare in statistical applications but common in applications to special functions like the gamma or Bessel functions.) Then successive terms could be added until the error bound reaches its minimum, giving the best guaranteed approximation, or an earlier sum used if the error is small enough. But asymptotic expansions, except where convergent, have the inherent limitation that there is a minimum error which limits the accuracy achievable.

For the aymptotic expansions used in statistics, the state of knowledge is much less satisfactory. Usually, only the order of magnitude of the errors is known, and only rarely are explicit bounds known—and these are far from sharp. Indeed, many expansions in common use have been obtained by formal operations with terms collected according to their order of magnitude, but without proof that the errors are of correct order. I call these *formal* asymptotic expansions and will try to indicate where they can be proved valid by careful but simple analysis.

The approximations discussed in this paper divide into two groups, the first consisting of approximations based ultimately on the central limit theorem and which use only the moments of the distribution to be approximated, and the second including various approximations using detailed information about the distribution.

2. The central limit theorem. The center of a large part of the asymptotic theory is the central limit theorem for sums of independent random variables. Let $\{X_n\}$ be a sequence of independent random variables. Denote by F_n the distribution function of the standardized sum

$$Y_n = \frac{\sum_{i=1}^{n} (X_i - E(X_i))}{\sqrt{\sum_{i=1}^{n} \operatorname{Var}(X_i)}} \tag{2.1}$$

and by Φ the unit normal distribution function. The central limit theorem then states that $\lim_{n\to\infty} F_n(x) = \Phi(x)$ for every fixed x, provided only that the means and variances are finite. If the $\{X_i\}$ are not identically distributed, an additional condition guaranteeing that the distributions are not too disbalanced is necessary (Lindeberg [51]).

The best possible general results on the order of magnitude of the errors in the central limit theorem were obtained during the 1940's by Berry [7], Esseen [29], [30], and Bergström [4], [5], [6]. Their results are of considerable interest and their methods are extremely important in much of asymptotic theory. The result for the sum of identically distributed random variables is that

$$\sup_x |F_n(x) - \Phi(x)| \leqq \frac{C\beta_3}{\sqrt{n}\,\sigma^3},$$

in which β_3 is the third absolute moment and σ^2 the variance of the component random variables. Several values for the constant C have been published, but only Berry's calculations have been published. Hsu [45] pointed out an error in Berry's calculation. This error can be corrected without affecting the result, but there is another more serious error. I have followed through the calculation and have found that 2.05 is a satisfactory replacement for the value 1.88 given by Berry. A more careful calculation would reduce this slightly. None of the other bounds suggested is as low as 2.05. Recent work of Esseen [31] has shown that as n approaches infinity, the minimum correct value of C approaches

$$\frac{\sqrt{10} + 3}{6} \cdot \frac{1}{\sqrt{2\pi}} \approx .41.$$

This value is achieved as n approaches infinity for a certain binomial distribution.

The bound holds also for sums of nonidentically distributed random variables, though the second and third moments enter in more complicated ways. Although the corrected Berry constant is the lowest known, the results of Esseen and Bergström are generally stronger because of the way that the second and third moments enter the bound.

All of the methods proceed by choosing as a kernel a distribution whose density function has a sharp maximum at the origin. A bound on the maximum difference of any two functions $F(x) - G(x)$ can be obtained from any bound on the convolution of this difference with the kernel distribution. The most common method of bounding the convolution has been to pass by Parseval's theorem to the characteristic functions and bound the resultant integral.

Much earlier, Lyapounov ([52], [53]) obtained a bound of order $\log n/\sqrt{n}$ for the central limit theorem error by using a normal distribution with variance of order $1/n$ for a kernel. Berry and Esseen were able to get the best result by choosing kernel distributions whose characteristic functions vanished outside a finite interval. The bounding then reduces to showing

$$\int_{-T}^{T} \frac{|f(t) - g(t)|}{|t|}\, dt = O\left(\frac{1}{T}\right)$$

where $1/T$ is the order of magnitude desired for the final result and where f and g are the characteristic functions of F and G respectively. For the central limit theorem, F is F_n, G is the normal distribution Φ and T is of order $\sqrt{n}$.

Bergström used the same choice as Lyapounov of a normal density for kernel, but he worked directly with the convolution integral. His method has proved valuable in extensions to the multivariate central limit theorems and he has proved ([5], [6]) that the error there is again of order $1/n^{\frac{1}{2}}$. The characteristic function techniques have not here been used successfully.

While the central limit theorem is very useful theoretically and often in practice, it is not always satisfactory. For small or moderate n, the errors of the normal approximation may be too large. Indeed, Berry's bound on the error is usually intolerable except for very large samples. Error bounds for special classes of distributions—chiefly the binomial and Poisson distributions—have been found by Uspensky [70] and others ([14], [33], [54], [55]).

3. Edgeworth series for sums. To obtain improvements and to prepare for later expansions, it will be convenient to develop a class of formal expansions sometimes known as the Charlier differential series [11]. In this formal development, the parameter n plays no role. The expansion is based on a distribution Ψ which need not be a normal distribution. Let ψ be its characteristic function and $\{\gamma_r\}$ its cumulants. Let F be the distribution to be approximated, f its characteristic function and $\{\kappa_r\}$ its cumulants. By the definition of the cumulants, the characteristic functions satisfy the formal identity

$$f(t) = \exp\left(\sum_{r=1}^{\infty}(\kappa_r - \gamma_r)\frac{(it)^r}{r!}\right)\psi(t). \tag{3.1}$$

If now, Ψ and all its derivatives vanish at the extremes of the range of x and exist for all x in that range, then by integration by parts, $(it)^r\psi(t)$ is the characteristic function of $(-1)^r\Psi^{(r)}(x)$. Introducing the differential operator D to represent differentiation with respect to x, the formal identity corresponds termwise in any formal expansion to the formal identity

$$F(x) = \exp\left(\sum_1^{\infty}(\kappa_r - \gamma_r)\frac{(-D)^r}{r!}\right)\Psi(x). \tag{3.2}$$

One can formally and apparently construct a distribution with prescribed cumulants by choosing Ψ and formally expanding.

The most important developing function $\Psi(x)$ is a normal distribution and with that choice, the formal expansion had been given earlier by Chebyshev [13], Edgeworth [27] and Charlier [10].

Chebyshev and Charlier proceeded by expanding and collecting terms according to the order of the derivatives. The resulting expansion is most commonly known as the Gram-Charlier A series and is identical with the formal expansion of $F - \Psi$ in Hermite orthogonal functions. It is a least squares expansion in derivatives of the normal integral Ψ with respect to a weight function which is the reciprocal of the normal density Ψ'. In this form, the expansion was developed and studied earlier by Chebyshev [12], Gram [41] and others.

The A-series converges for functions F whose tails approach zero faster than $\Psi'^{\frac{1}{2}}$ (see Szëgo [63] or Cramér [19]). Convergence obtains for all distributions on finite intervals but few others of any interest. The developing normal distribution is usually chosen to have the same mean and variance as the given distribution F.

This choice has no effect on convergence, though it clearly has a tremendous effect on the quality of approximation by the first few terms. Altogether, the convergence properties are of little value and the importance of the Gram-Charlier series arises from its properties as an inferior form of an asymptotic expansion.

The preferable development was done by Edgeworth as an improvement to the central limit theorem. Let the distribution to be approximated again be the distribution F_n of the standardized sum Y_n (eqn. 2.1) of independent random variables. Take the component random variables identically distributed with mean μ, variance σ^2, and higher cumulants $\{\sigma^r\lambda_r ; r \geqq 3\}$. Take the developing function Ψ to be the unit normal distribution function Φ. Then the cumulant differences in the formal identity (3.1) are

$$\kappa_1 - \gamma_1 = 0 = \kappa_2 - \gamma_2$$

$$\kappa_3 - \gamma_3 = \frac{\lambda_r}{n^{r/2-1}} \qquad r \geqq 3.$$

The Edgeworth series is obtained by collecting terms in the formal expansion according to powers of n, thus yielding a formal asymptotic expansion of the characteristic function of the form

$$f_n(t) = \left(1 + \sum_1^\infty \frac{P_r(it)}{n^{r/2}}\right)e^{-t^2/2}$$

with P_r a polynomial of degree $3r$ with coefficients depending on the cumulants of orders 3 through $r + 2$. If powers of Φ are interpreted as derivatives, the corresponding distribution function expansion is

$$F_n(x) = \Phi(x) + \sum_1^\infty \frac{P_r(-\Phi(x))}{n^{r/2}}.$$

It is important to note that every term beyond the normal approximation can be expressed as the product of the normal density and a polynomial in x. The first few terms of the expansion are:

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

In 1928, Cramér [20] proved the series valid uniformly in x, but gave no explicit bounds on errors. Apart from requiring that one more cumulant exist than used in any partial sum, the proof assumes that the characteristic function h of the component random variables satisfies the condition

$$\limsup_{|t|\to\infty} |h(t)| < 1. \tag{3.3}$$

This is satisfied if the component distribution has an absolutely continuous part. It is not satisfied for discrete distributions and the result then is generally not true.

The elementary proofs given later by Esseen [30] and Hsu [45] use the method developed for the central limit theorem bound and amount to showing that

$$\int_{-T}^{T} \frac{| f_n(t) - g_{n,k}(t) |}{|t|} dt = O\left(\frac{1}{T}\right)$$

with $T = c(n^{\frac{1}{2}})^k$ and with $g_{n,k}(t)$ the expansion of the characteristic function through terms of order $(1/n^{\frac{1}{2}})^{k-1}$ and using cumulants through order $k + 1$. Using a Maclaurin's expansion of the characteristic function, the integral up to $n^{\frac{1}{2}}$ is easily bounded by $(c_2\beta_{k+2})/T$ with the unknown distribution entering only through the absolute moment or order $k + 2$. An efficient determination of c_2 would be extremely difficult.

Using the Cramér condition (3.3) on the characteristic function, the integral from $n^{\frac{1}{2}}$ to T is easily bounded by c_3/T. But by this evaluation, the resulting bound c_3 depends on the unknown distribution through its characteristic function and this even more seriously prevents the determination of any numerically useful bounds.

Cramér [21] also proved the validity of the asymptotic expansion for sums of non-identically distributed random variables. The conditions are somewhat more restrictive. Cramér [20] showed that the termwise differentiated Edgeworth series is a valid expansion for the density function, provided the component random variables have a density function of bounded variation. Gnedenko and Kolmogorov [40] weaken this condition. They also present most of the work of Cramér and Esseen discussed here.

Esseen [30] studied the expansion problem when the Cramér condition (3.3) on the characteristic function is not satisfied. The error in using the first approximation

$$\Phi(x) - \frac{\lambda_3 \Phi^{(3)}(x)}{6\sqrt{n}}$$

is of smaller order than $1/n^{\frac{1}{2}}$ provided only that the third moment is finite and that the distribution is not a lattice distribution. If the distribution of the component random variables is lattice, i.e., takes all probability on a set of equally spaced points, a different expansion is available. The Edgeworth density function expansion is, except for a constant multiple, a valid expansion for the jumps at each possible point. The usual Edgeworth expansion for the distribution function can be modified by the addition of terms (discontinuous) so that the resultant expansion is a valid expansion, uniformly for all x. The corrections, when evaluated at the points half-way between possible values of the standardized sum, have no effect of order $1/n^{\frac{1}{2}}$, but do for all higher orders. Thus, for example, the usual Edgeworth series when applied to a binomial or Poisson distribution and evaluated only at half-integers is correct through order $1/n^{\frac{1}{2}}$ but needs a correction of order $1/n$.

Since the Gram-Charlier A series is only a rearrangement of the Edgeworth series, its asymptotic properties follow directly. Of course, many higher terms must be used before all terms of the desired order are included. What makes the

Gram-Charlier arrangement so bad in practice is that these extra terms involve cumulants of much higher order.

Multivariate Edgeworth series can be developed in complete analogy with the univariate expansions. Other than bounds for the normal approximation error, little theoretical work has been done with the multivariate expansions. Specifically, the multivariate Edgeworth series for sums of independent random variables has *not* been shown to be a valid asymptotic expansion. This would seem the most serious gap in theoretical knowledge of asymptotic approximations

4. General Edgeworth and Cornish-Fisher series. Many sample functions have distributions asymptotically normal for increasing sample size, but not all admit asymptotic expansions beyond the normal distribution term. Expansions can be constructed for functions, such as most functions of sample moments, behaving asymptotically as sums of independent random variables. To illustrate with the simplest example, let $H(\bar{X})$ be an arbitrary function, not depending on n, of the sample mean in a sample of size n from a population with cumulants $\{\mu, \sigma^2, \sigma^r\lambda_r\}$. The distribution of

$$W_n = \frac{\sqrt{n}(H(\bar{X}) - H(\mu))}{\sigma H'(\mu)} \tag{4.1}$$

is asymptotically unit normal, provided that $H'(\mu) \neq 0$ and H is smooth enough at μ. (The assumption $H'(\mu) \neq 0$ and its equivalent for functions of several moments rule out many interesting functions for which no general theory of asymptotic expansions is known.) Assume $H'(\mu) > 0$. Then the distribution function K_n of W_n is given by

$$\begin{aligned} K_n(x) &= P(W_n \leqq x) \\ &= P\left\{\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leqq \frac{\sqrt{n}\left(I\left[H(\mu) + \frac{x}{\sqrt{n}}\right] - \mu\right)}{\sigma}\right\} + O(n^{-p}) \end{aligned}$$

in which I denotes the uniquely defined function inverse to H near μ and all other solutions of the inequality are easily shown to be of higher order than any power of $1/n$. If the population satisfies the Cramér condition (3.3), the standardized mean has a valid Edgeworth expansion so that

$$K_n(x) = \Phi(u) + \sum_1^{k-1} \frac{P_r(-\Phi(u))}{n^{r/2}} + O(n^{-k/2})$$

in which

$$u = \frac{\sqrt{n}\left(I\left[H(\mu) + \frac{x}{\sqrt{n}}\right] - \mu\right)}{\sigma}.$$

Further, each derivative $\Phi^{(i)}(u)$ can be expanded in a Taylor series in $1/n^{\frac{1}{2}}$ about $n = \infty$, evaluating the derivatives of I at $H(\mu)$ from the derivatives of H at μ. If H is smooth enough at μ, and with some natural rearrangement of terms, a valid asymptotic expansion of the same general form as the Edgeworth series for sums is obtained. I call these series also Edgeworth series.

The construction would extend directly to the multivariate expansion of r functions of r sample moments if only the multivariate Edgeworth expansion for sums were valid. The expansion for the distribution of a single function of r moments could then be easily obtained as a marginal expansion.

I know of no literature on any of these expansions for general functions. Hsu [45] and Chung [17] proved respectively that the sample variance and the one sample t-statistic have valid expansions. (There are several errors in Chung's explicit expansion—equation (35).) Hsu proved several results needed for proofs for functions of any number of moments. But a very large amount of work was involved in completing the proof for each separate function. Hsu stated that students were working on other sample function, but I know of no others published except for a statement by Sun [62] that he had proved the result for the third moment about the mean and a proof by Hsu [46] for the expansion of the distribution of the ratios of two independent means.

A general result would be highly desirable or else an example of a statistic, smooth enough at the population value, but for which the series is not a valid asymptotic expansion to show that the construction described is not valid as generally as appears plausible.

The expansions can be obtained formally by a different approach using the Charlier differential series identity (3.2) and the classical so-called δ-method for calculating moments. Formally compute the moments and from them the cumulants of the statistic W_n of equation (4.1) by expanding $H(\bar{X})$ in a Taylor series in $\bar{X} - \mu$ and integrating term by term. The formal cumulant expansions for W_n are of the form:

$$\begin{aligned} \kappa_1(W_n) &= 0 + O\left(\frac{1}{\sqrt{n}}\right) \\ \kappa_2(W_n) &= 1 + O\left(\frac{1}{n}\right) \\ \kappa_r(W_n) &= O(n^{-r/2+1}) \qquad r > 2 \end{aligned}$$

so that the leading terms behave exactly as for standardized sums of random variables. If these formal cumulant expansions are substituted in the symbolic identity (3.2), using the unit normal as the developing function, and if the exponential operator is expanded formally and terms collected according to powers of $n^{\frac{1}{2}}$, the same expansion as previously constructed is obtained.

This latter method is almost always easier to use in practice, especially for functions of several moments. Most applications of Edgeworth series use this method or some slight variation of it, such as using exact or valid expansions for the moments, which are frequently obtainable.

The δ-method is often used to obtain formal asymptotic expressions for moments and cumulants of statistics. A few examples of such use—for various purposes—will be found in references [25], [26], [42], and [76]. The δ-method moment expansions are known to be valid in some special cases. If a function of sample moments is uniformly bounded by a power of the sample size, is smooth enough at the population moments, and if enough population moments (far more than apparently needed) exist, then the expansion can be proved valid by extending Cramér's proof [23, p. 354] for the leading terms of the mean and variance. Under severe distributional assumptions (for example, for functions of (normal theory) mean square variates—see section seven), the method can be shown valid. But there are also examples where the method is not valid, and a wide range of applications in between. However, as long as the moments are used only to get distribution approximations, it is generally plausible and sometimes known to be true that the distribution approximations are valid whether the moment expansions are or are not.

In many statistical applications, quantiles of a distribution are needed. From an Edgeworth expansion of a distribution function F_n, as asymptotic expansion for a quantile x of F_n in terms of the corresponding normal quantile z can be obtained by formal substitutions, Taylor expansions, and identification of coefficients of powers of n. The expansion is of the form

$$x = z + \frac{S_1(z)}{\sqrt{n}} + \frac{S_2(z)}{\sqrt{n}} + \cdots$$

in which the $\{S_i\}$ are polynomials. The reverse expansion

$$z = x + \frac{R_1(x)}{\sqrt{n}} + \frac{R_2(x)}{\sqrt{n}} + \cdots \tag{4.2}$$

is obtained as an intermediate step and is often useful in itself, giving an asymptotic transformation of a variate x with distribution F_n into a unit normal deviate. An expansion of the type (4.2) is often called a normalization formula. Numerically it serves the same purpose as the Edgeworth expansion but is often more convenient and possibly more accurate.

Cornish and Fisher [18] carried out these inversions, treating each cumulant of F_n according to the order of magnitude of its leading term as determined by the δ-method. For the expansion of x in terms of z, they table, for seven common probability levels, all the polynomials needed to obtain all terms through order $1/n^2$, that is, using up through the sixth cumulants.

For an absolutely continuous distribution, both of the inverted series, which I will call Cornish-Fisher series, can be proved to be valid asymptotic expansions for every probability level, whenever the initial Edgeworth series is valid. I know of no published proof of this, though Wasow's [73] proof of the invertability of a special class of distribution expansions can be modified and extended to work here.

The Edgeworth and Cornish-Fisher approximations have some faults which show up in the tails of the distribution. The distribution function approximations are not probability distributions and both monotonicity and the 0–1 range property are violated in parts of one or both tails. Similarly the quantile approximations are not always monotone in the probability levels. These troubles don't contradict the uniform validity of the Edgeworth expansion because it only refers to the absolute difference of two functions each approaching zero (or one) and not to the relative error. The validity of the Cornish-Fisher series is uniform for the probability level in each interior interval but the error increases as the level approaches 0 or 1.

Cramér [22], and others [15], [32], [61] in important work have investigated the relative accuracy of the central limit theorem approximation, but for the Edgeworth and Cornish-Fisher approximations, the importance of these tail difficulties at present must be determined from empirical evidence. Some different expansions constructed to eliminate the tail difficulty will be discussed in section six for several specific distributions.

There have been only a few numerical evaluations of the accuracy of these approximations, largely because of the difficulty of obtaining exact values for comparison.

In a major piece of unpublished work, Teichroew [65] has used the terms of the Cornish-Fisher series through n^{-5} to evaluate the quantiles of the normal theory chi-square distribution for a variety of degrees of freedom and probability levels. He has found that the accuracy of this approximation for four degrees of freedom, provided that the probability level is not in the extreme half of one percent, is at least three decimals with the accuracy improving rapidly as the degrees of freedom increase. Even for two degrees of freedom, the series is accurate to two decimals except in the extreme one percent. The series for the χ^2 is the most accurate application known.

For the standardized sums of samples of size ten from four symmetrical non-normal populations, Chand [9] compared the exact quantiles with the Cornish-Fisher approximations through orders $1/n$ and $1/n^2$. The latter gave better than three decimal accuracy and the former better than two decimal accuracy for probability levels ranging from $\frac{1}{2}\%$ to 25%.

Many more empirical studies of accuracy would be desirable including studies of the comparative accuracies of the Edgeworth expansion and the normalization expansion (4.2).

5. Investigations of robustness. Asymptotic expansions play an important part in investigations of the effect of deviations from normality (or other population) on the size and power of various tests. I use the null distribution of the one-sample t-statistic as an example. Denote by F_n and G_n respectively the general and normal population distributions of the t-statistic in samples of size n. Formal Edgeworth expansions of F_n and G_n can be obtained and for the t-statistic (but not otherwise) these have been proved valid Since the difference $F_n(x) - G_n(x)$ here is of interest, the difference of the two expansions provides a valid asymptotic expansion for the deviation in terms of powers of $1/n^{\frac{1}{2}}$ and of normal derivatives.

Effectively, the original Edgeworth series for F_n has been replaced by one in which the leading term is G_n (assumed known). The approximations to F_n are then exact for a normal population and greatly improved for "near normal" populations. Similar modifications of successively higher order terms might be expected to give improved accuracy, especially for small n. The possibility (quite generally) of using expansions

$$F_n(x) \sim \sum_{i=0} B_i(x) H_i(x, n),$$

asymptotically equivalent (at every partial sum) to

$$F_n(x) \sim \sum_{i=0} A_i(x) n^{-i/2}$$

is a powerful tool to permit improved accuracy of expansions. There is no theory on how to choose good functions H_i, but useful choices can often be made on heuristic grounds or on the basis of a few computations.

In the t-statistic example, the expansion of F_n in terms of successive derivatives of the normal theory t-distribution might appear natural. Geary [39] obtained such an expansion by formally applying the Charlier differential series (equation 3.2) with G_n as the generating distribution, collecting terms according to their orders or magnitude. The result can be proved asymptotically equivalent to the Edgeworth expansion and hence valid. Geary applies the same formal method to an F-statistic (though even the formal derivation of the Charlier identity is not valid) and Bartsch [3] applies the method to various t-type statistics.

In the most substantial investigations of this kind, Gayen ([35], [36], [37], [38]) has obtained a different asymptotically equivalent expansion for the distribution of t (as well as for two-sample t, the variance ratio, and the correlation coefficient). He has given extensive tables and graphs so his expansions are far more easily used than any alternative expansions. The expansions possess also a different asymptotic property.

There seem to be no comparisons of the quality of the several approximations. Seemingly, the only feasible method for proving the validity of any of these expansions is to show equivalence to the Edgeworth series and to prove it valid (if possible). This method would never lead to useful information on accuracy since the Edgeworth series is surely much less accurate than these modified expansions.

Although Gayen's expansions are asymptotically equivalent (in n) to the Edgeworth and other series, they have an additional property, not shared by the other series mentioned, of being a formal asymptotic expansion for any fixed finite n as the population "nonnormality" approaches zero. This is made definite by assuming that the population distribution itself can be expressed by an Edgeworth expansion in some unknown parameter m (i.e., that the population values themselves are the means of m independent "elementary errors"). The Gayen expansion is a formal asymptotic expansion in powers of $1/m^{\frac{1}{2}}$ (m does not need to be known to write down the series). This approach seems conceptually more relevant to robustness problems than asymptotic expansions in the sample size.

Theoretical study of the properties of these series would be desirable, as would some comparative computations on various approximations.

6. Quantile expansions for specific distributions. The expansions that have been considered have made use of only the moments or cumulants of a distribution. Many useful asymptotic approximations have been developed from analytic expressions for the density function of the distribution to be approximated. As practically the only distributions known analytically, normal theory distributions are the object of most of these expansions. However, the normal distribution does not here play the central role that it does in the Edgeworth theory.

Consider first the expansion of a quantile of one distribution of a convergent sequence in terms of the corresponding quantile of the limiting distribution or the reverse expansion. When the normal distribution is the limiting distribution, the results are necessarily exactly those given by the Cornish-Fisher expansions but use of the explicit analytic form greatly simplifies the derivation of higher order terms and proofs of validity.

Let $\{f_n\}$ be a sequence of density functions which converges to a density function ψ. The desired expansions are found as the solutions either for t or for z of the equation

$$\int_{-\infty}^{t} f_n(x)\,dx = \int_{-\infty}^{z} \psi(x)\,dx$$

or equivalently of the differential equation

$$f_n(t)\,\frac{dt}{dz} = \psi(z). \tag{6.1}$$

In 1923, Campbell [8] obtained a formal series solution of the differential equation for the quantiles of the χ^2 distribution in terms of those of the normal distribution. He carried the series to ten terms beyond the normal approximation. Teichroew [64] has tabled these polynomial terms and used them for the computation described in section four.

Hotelling and Frankel [44] followed the same procedure to get four correction terms for the transformation of a Student's t variate into a unit normal deviate and also for the transformation of a Hotelling's T^2 variate into a chi-square variate. They proved the validity of the expansions.

Wasow [73] has given conditions on a sequence of distributions with a normal limiting distribution such that these expansions can be validly obtained by the natural formal methods, and further that each term will be a polynomial in the variate.

The accuracy of these expansions decreases as the probability level becomes more extreme. Consider the transformation of Student's t to a normal deviate. It has the form

$$z = t\left[1 + \frac{P_2(t)}{n} + \frac{P_4(t)}{n^2} + \cdots\right]$$

in which the $\{P_i\}$ are even polynomials of the indicated order. Hotelling and Frankel observed empirically that the series is of no value for t^2 greater than n. Clearly, the expansion cannot be valid for t of the order of $n^{\frac{1}{2}}$ since, from the order of the polynomials, no term would approach zero with increasing n. The usefulness of the series for small n is severely limited.

To obtain expansions useful in the tails of the distribution, Teichroew [66] has considered a limiting process in which t and z both approach infinity with n. His results are rather spectacular.

Set $t = bn^{\frac{1}{2}} + u$ with b a constant for later choice and the variable u to be kept finite. Similarly, set $z = cn^{\frac{1}{2}} + v$. The choice $c = [\log (1 + b^2)^{\frac{1}{2}}]$ is forced by examining leading terms in the differential equation (6.1) relating z and t. The equation becomes an equation relating u and v and a formal expansion of v in terms of u is easily, though tediously, obtained:

$$v = p_1(u) + \frac{p_2(u)}{\sqrt{n}} + \frac{p_3(u)}{n} + \cdots .$$

The $\{p_i(u)\}$ are polynomials of the indicated order, respectively odd and even. The dependence on b is very complicated. The whole procedure can be reversed, treating c as fixed and getting a series for u in terms of v. In actual use, with a given value of t and n, b would be chosen so that u is made small or zero thus keeping the polynomial terms small. If u is made to be zero, all odd order polynomials vanish. For 1 degree of freedom and a selection of t values corresponding to tail probability levels ranging from $\frac{1}{4}$ to 10^{-6}, choosing b so that u is zero and using the first five non-zero terms, the approximation gives the equivalent normal deviate to better than two decimal places. The ordinary series is totally worthless.

The first term is of interest. Taking $u = 0$, $b = t/n^{\frac{1}{2}}$, it is

$$z_0 = \sqrt{n \log (1 + t^2/n)}.$$

This reduces to the usual normal approximation as n approaches ∞ with t fixed. By direct analysis, Wallace [72] has shown that for all $t > 0$ and $n \geqq 1$, it satisfies the bounds

$$-\frac{.37}{\sqrt{n}} \leqq z - z_0 \leqq 0.$$

Knowing that the first term is correct to the indicated order, the entire expansion can then be shown to be a valid asymptotic expansion, uniformly for u in any finite interval. No bounds are known beyond the first term.

Teichroew has treated the χ^2 distribution in the same way with the same spectacular results. Wallace has obtained a bound as with t for the first term approximation in the upper tail.

The method is applicable to many other distributions but I know of no further applications.

7. Laplace's method and studentization. Many calculations in statistics can be reduced to the evaluation of the expected value of some function of a mean-square variate:

$$E[f(v)] = c_n \int_0^\infty f(v)v^{(n/2)-1}e^{-nv/2}\, dv.$$

The integral here is a special case of the integral

$$\int g(u)e^{-nh(u)}\, du.$$

Its asymptotic evaluation by Laplace's method is very important in the theory of asymptotic expansions. If g and h are well-behaved functions, then for large n, the integral except in the neighborhood of the minimum of h is relatively negligible to an exponential order in n. Valid asymptotic expansions can be obtained.

This integral evaluation is an important part of the method of steepest descent ([28], p. 38) in which the path of integration, considered in the complex plane is chosen to pass through a minimum of h and in such a way that the absolute value of the exponential $e^{-nh(v)}$ falls off most rapidly from its maximum. The integral is then expanded by Laplace's method.

The method of steepest descent (and not just the Laplace integral evaluation) has been used by Daniels [24] to obtain some interesting expansions that generalize the Edgeworth expansions for sums. They have some superior properties but make use of explicit knowledge of the moment generating function.

In the expansion of $E[f(v)]$, a simple application of the δ-method is much more convenient than a straight application of Laplace's method (because of the constant c_n in the expansion for $E[f(v)]$). Expand $f(v)$ in a Taylor series about the population value of v (here equal to one) and integrate term by term. If the expectation exists for sufficiently large n and if f has bounded derivatives near one, then the expansion obtained is valid. Since the moments of v about its mean involve several powers of $1/n$, some rearrangement is needed to get an expansion of the form

$$E[f(v)] \sim \sum \frac{A_i}{n^i}.$$

But for this last step, the development would have gone as well using a root mean square variate $s = v^{\frac{1}{2}}$ as argument in the Taylor expansion and integration.

This expansion method and its natural extension to functions of several independent mean square variates are widely applicable in statistical work. They are unusually tractable for obtaining bounds on errors of approximation, but I am not aware of any such bounds.

One important application is to finding the distribution function H of a studentized statistic $Y/v^{\frac{1}{2}}$ in which Y/σ has the known distribution function G and v is an independent mean square estimate of the squared scale factor σ^2. Then

$$H(x) = E\left[G\left(\frac{x\sqrt{v}}{\sigma}\right)\right]$$

and its expansion is obtained as described. The terms in the expansion are all linear functions of the unstudentized distribution function G and its derivatives. The expansion was first obtained, in a different way, by Hartley [43]. Moriguti [56] developed the result as given here, except that he used the root mean square as argument with a consequent unnecessary complication. (His error bound (3.2) is incorrect).

Examples of the use of the expansion to get distributions of various studentized statistics are found in references [34], [58], [59], [60], and [69]. Ito [47] develops an example of a generalization to multivariate studentization.

8. The Behrens-Fisher problem. Another application is part of the development of what is to me the most interesting use of asymptotic expansions: the Welch solution for the Behrens-Fisher problem and the various extensions and analogous treatments of problems like finding confidence limits for variance components or for weighted averages when the weights must be estimated.

There have been a large number of papers attacking these problems, frequently repeating the same work ([1], [16], [57], [48], [49], [50], [67], [68], [71], [75], and others). Most of the work has consisted of formal expansions with no proofs that errors are really of their apparent order of magnitude and there has been some confusion as to what the expansions do provide. There have been a very few computations, and these very difficult, that indicate the accuracy of the approximations.

I consider in some detail a reduced form of the Behrens-Fisher problem. Let Y be normally distributed with mean μ and variance $\sum \lambda_i \sigma_i^2$ with $\{\lambda_i\}$ known positive constants and with the unknown variances σ_i^2 estimated by independent mean square variates s_i^2 respectively with n_i degrees of freedom. The problem is to find a test of the hypothesis $\mu = 0$, which has significance level α identically in the parameters σ_i^2.

The problem is already reduced to the sufficient statistics. Restrict it further by considering only one-sided tests of the form: reject if $Y > h(s_i^2, \alpha, \lambda_i, n_i)$ with h chosen so that $P(Y > h(s^2)) \equiv \alpha$. The Welch solution consists in an expansion

$$h(s^2) = h_0(s^2) + h_1(s^2) + h_2(s^2) + \cdots \tag{8.1}$$

in which $h_i(s^2)$ is of order n^{-i} in the degrees of freedom.

To my knowledge, it is still not known whether a non-randomized similar level α test exists. If there is no function h, the asymptotic expansion (8.1) cannot be valid. But the expansion is still of value because it provides tests that are asymptotically similar, that is, such that

$$P\left(Y > \sum_0^{r-1} h_i(s^2)\right) = \alpha + O(n^{-r}).$$

This interpretation of the asymptotic property was not clear in the original papers and was the source of confusion. From the large sample result that $Y/(\sum \lambda_i s_i^2)^{\frac{1}{2}}$ is asymptotically normally distributed the first term must be $h_0(s^2)$

$= z(\sum \lambda_i s_i^2)^{\frac{1}{2}}$ with z the level α normal quantile. Further terms are determined successively so that

$$P\left(Y \leqq \sum_0^{r-1} h_i(s^2)\right) = 1 - \alpha + O(n^{-r}),$$

For getting several terms the formal operator formula given by Welch is probably the most efficient procedure. The work is straightforward but Aspin [1] reported that 100 pages of detailed algebra were required to determine the term h_4.

I take a method that illustrates how a proof of validity could be given, but determine only one term. Suppose first that $h_1(s^2)$ is any function of the variances such that it and all its partial derivatives through order two are of order $1/n$. Let $Q(\sigma^2) = (\sum \lambda_i \sigma_i^2)^{\frac{1}{2}}$ and let

$$u_1(s^2) = \frac{h_0(s^2) + h_1(s^2)}{Q(\sigma^2)},$$

$$P(Y \leqq h_0(s^2) + h_1(s^2)) = E\{\Phi(u_1(s^2))\}.$$

To evaluate $E\{\Phi(u_1(s^2))\}$ expand $\Phi(u_1(s^2))$ in a Taylor series in $u_1(s^2) - z$ and integrate with respect to the distributions of the s_i^2.

$$E\{\phi(u_1(s^2))\} = 1 - \alpha + \phi(z)E[u_1(s^2) - z] + \frac{\phi'(z)}{2} E[u_1(s^2) - z]^2 + \frac{\phi''(z)}{6} E[u_1(s^2) - z]^3 + cE[u_1(s^2) - z]^4.$$

Each of these integrals here is of the form discussed in section seven and is validly expanded through formal Taylor expansions and termwise integration. Carrying out the process far enough to get all terms of order $1/n$, and remembering that h_1 and its derivatives are of order $1/n$, leads to the expression

$$P(Y \leqq h_0(s^2) + h_1(s^2)) = 1 - \alpha + \left[\phi(z)\frac{h_1(\sigma^2)}{Q(\sigma^2)} + \sum \frac{2\sigma_i^4}{n_i}\left\{\frac{z\phi(z)}{2Q(\sigma^2)}\left[\frac{\partial^2 Q(s^2)}{\partial(s_i^2)^2}\right]_{s^2=\sigma^2} + \frac{z^2\phi'(z)}{2Q^2(\sigma^2)}\left[\frac{\partial Q(s^2)}{\partial(s_i^2)}\right]^2_{s^2=\sigma^2}\right\}\right] + O\left(\frac{1}{n^2}\right).$$

If h_1 is chosen to make the $1/n$ term vanish, it clearly has all the assumed properties.

The determination and proof of validity for each additional term is essentially the same.

The first approximation is

$$z\sqrt{\Sigma \lambda_i s_i^2}\left[1 + \frac{(1 + z^2)}{4}\frac{\sum \frac{\lambda_i^2 s_i^4}{n_i}}{(\Sigma \lambda_i s_i^2)^2}\right].$$

It is particularly interesting because it is equivalent, through terms of order $1/n$, to a Student's t approximation using a degrees of freedom determined by the $\{s_i^2\}$ and the $\{\lambda_i\}$ that was proposed much earlier by Welch [74].

Welch (appendix to [2]) has computed the true significance levels obtained using the expansion through h_3 for two variances, each with 6 degrees of freedom, and using a nominal level of .05. He found that the variation from .05 does not exceed .0002. This result seems quite satisfactory, but several more computations would be helpful in view of the importance of the procedure.

The theory of the expansions used by Welch and others was given in a 1949 paper by Chernoff [16]. None of the papers written on these subjects take any notice of the Chernoff work. He gives conditions for validity of expansions in an asymptotic studentization procedure due to Wald. Although his detailed results are for one nuisance parameter, he illustrates the extension to several nuisance parameters by essentially the same construction as here indicated for the Welch solution of the Behrens-Fisher problem.

A straightforward application of the Chernoff results yields an asymptotic series solution for a confidence interval for a variance component γ, where estimates v_1 of $\sigma^2 + \gamma$ and v_2 of σ^2 are available. The expansion for this problem was first developed and proved valid by Moriguti [57].

Most notable of the other work along this line is the work of James ([48], [49], [50]), who has extended the Welch formal expansion to univariate and multivariate tests of general linear hypotheses with unknown and unequal variances and covariances as nuisance parameters.

9. Conclusion. I have by no means covered all the interesting and important work on asymptotic approximations and have not even considered any non-asymptotic approaches to approximations. I have discussed what are to me some of the interesting problems, attacks, and results. Much more work is needed, particularly theoretical and empirical studies of the qualities of the approximations.

REFERENCES

[1] A. A. Aspin, "An examination and further development of a formula arising in the problem of comparing two mean values," *Biometrika*, Vol. 35 (1948), pp. 88–96.

[2] A. A. Aspin, "Tables for use in comparisons whose accuracy involves two variances, separately estimated," with appendix by B. L. Welch, *Biometrika*, Vol. 36 (1949), pp. 290–296.

[3] G. Bartsch, "Confidence intervals for means," unpublished doctoral dissertation Johns Hopkins University (1957).

[4] H. Bergström, "On the central limit theorem," *Skand. Aktuarietidskr.*, Vol. 27 (1944), pp. 139–153.

[5] H. Bergström, "On the central limit theorem in the space R_k, $k > 1$," *Skand. Aktuarietidskr.*, Vol. 28 (1945), pp. 106–127.

[6] H. Bergström, "On the central limit theorem in case of not equally distributed random variables," *Skand. Aktuarietidskr.*, Vol. 32 (1949), pp. 37–62.

[7] A. C. Berry, "The accuracy of the Gaussian approximation to the sum of independent variates," *Trans. Amer. Math. Soc.*, Vol. 49 (1941), pp. 122–136.

[8] G. A. Campbell, "Probability curves showing Poisson's exponential summation," *Bell System Tech. J.*, Vol. 2 (1923), pp. 95–113.
[9] U. Chand, "Formulas for the percentage points of distributions of the arithmetic means in random samples from certain symmetrical universes," *J. Res. Nat. Bur. of Standards*, Vol. 43 (1949), pp. 79–80.
[10] C. V. L. Charlier, "Über das Fehlergesetz," *Ark. Mat. Astr. och Fys.*, Vol. 2 (1905–6), No. 8, 9 pp. (1905).
[11] C. V. L. Charlier, "Über die Darstellung willkürlicher Funktionen," *Ark. Mat. Astr. och Fys.*, Vol. 2 (1905–6), No. 20, 35 pp. (1906).
[12] P. L. Chebyshev, "Sur le developpement des fonctions à une seule variable," *Bull. Acad. Imp. Sci. St. Petersbourg*, Series 3, Vol. 1 (1860), pp. 193–202.
[13] P. L. Chebyshev, "Sur deux théorèmes relatifs aux probabilités," *Acta Math.*, Vol. 14 (1890), pp. 305–315.
[14] T. T. Cheng, "Normal approximation to the Poisson distribution and a proof of a conjecture of Ramanujan," *Bull. Amer. Math. Soc.*, Vol. 55 (1949), pp. 396–401.
[15] T. T. Cheng, "On asymptotic expansions connected with the sums of independent random variables," (In Chinese. English summary), *Acta Math. Sinica.*, Vol. 5 (1955), pp. 91–108.
[16] H. Chernoff, "Asymptotic studentization in testing of hypotheses," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 268–278.
[17] K. L. Chung, "The approximate distribution of Student's statistic," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 447–465.
[18] E. A. Cornish and R. A. Fisher, "Moments and cumulants in the specification of distributions," *Extrait de la Revue de l'Institute International de Statistique*, Vol. 4 (1937), pp. 1–14. Reprinted as paper No. 30 in R. A. Fisher, *Contributions to Mathematical Statistics*, New York, John Wiley, 1950.
[19] H. Cramér, "On some classes of series used in mathematical statistics," *Proceedings of the Sixth Scandinavian Congress of Mathematicians, Copenhagen*, (1925), pp. 399–425.
[20] H. Cramér, "On the composition of elementary errors," *Skand. Akturaietidskr.*, Vol. 11 (1928), pp. 13–74 and 141–180.
[21] H. Cramér, "Random Variables and Probability Distributions," *Cambridge Tracts in Mathematics*, No. 36 (1937).
[22] H. Cramér, "Sur un nouveau théorème-limite de la théorie des probabilités," *Actualités Sci. Ind.*, No. 736, Paris (1938).
[23] H. Cramér, *Mathematical Methods of Statistics*, Princeton: Princeton University Press (1946).
[24] H. E. Daniels, "Saddlepoint approximations in statistics," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 631–650.
[25] F. N. David, "The moments of the z and F distributions," *Biometrika*, Vol. 36 (1949), pp. 394–403.
[26] F. N. David and N. L. Johnson, "Statistical treatment of censored data. Part I," *Biometrika*, Vol. 41 (1954), pp. 228–240.
[27] F. Y. Edgeworth, "The law of error," *Cambridge Philos. Trans.*, Vol. 20 (1905), pp. 36–66 and 113–141.
[28] A. Erdelyi, *Asymptotic Expansions*, New York: Dover (1956).
[29] C. G. Esseen, "On the Liapounoff limit of error in theory of probability," *Ark. Mat. Astr. och Fys.*, Vol. 28A (1942), No. 9, 19 pp.
[30] C. G. Esseen, "Fourier analysis of distribution functions," *Acta Math.*, Vol. 77 (1945), pp. 1–125.
[31] C. G. Esseen, "A moment inequality with an application to the central limit theorem," *Skand. Aktuarietidskr.*, Vol. 39 (1956), pp. 160–170.

[32] W. FELLER, "Generalization of a probability limit theorem of Cramér," *Trans. Amer. Math. Soc.*, Vol. 54 (1943), pp. 361–372.

[33] W. FELLER, "On the normal approximation to the binomial distribution," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 319–329 and Vol. 21 (1950), 302.

[34] R. A. FISHER, "Expansion of 'Student's' integral in powers of n^{-1}," *Metron*, Vol. 5 (1925), No. 3, pp. 109–112.

[35] A. K. GAYEN, "The distribution of Student's t in random samples of any size from non-normal universes," *Biometrika*, Vol. 36 (1949), pp. 353–369.

[36] A. K. GAYEN, "The distribution of the variance ratio in random samples of any size from non-normal universes," *Biometrika*, Vol. 37 (1950), pp. 236–255.

[37] A. K. GAYEN, "Significance of difference between the means of two non-normal samples," *Biometrika*, Vol. 37 (1950), pp. 399–408.

[38] A. K. GAYEN, "The frequency distribution of the product moment correlation coefficient in random samples of any size from non-normal universes," *Biometrika*, Vol. 38 (1951), pp. 219–247.

[39] R. C. GEARY, "Testing for normality," *Biometrika*, Vol. 34 (1947), pp. 209–242.

[40] B. V. GNEDENKO AND A. N. KOLMOGOROV, *Limit Distributions for Sums of Independent Random Variables*, translated by K. L. Chung, Cambridge: Addison Wesley Publishing Co. (1954).

[41] J. P. GRAM, "Über die Entwickelung reeler Funktionen in Reihen mittelst der Methode der kleinsten Quadrate," *J. Reine Angew. Math.*, Vol. 94 (1883), pp. 41–73.

[42] J. B. S. HALDANE, "Normalization of frequency distributions," *Biometrika*, Vol. 29 (1937), pp. 392–404.

[43] H. O. HARTLEY, "Studentization," *Biometrika*, Vol. 33 (1944), pp. 173–180.

[44] H. HOTELLING AND L. R. FRANKEL, "The transformation of statistics to simplify their distribution," *Ann. Math. Stat.*, Vol. 9 (1938), pp. 87–96.

[45] P. L. HSU, "The approximate distribution of the mean and variance of a sample of independent variables," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 1–29.

[46] P. L. HSU, "On the approximate distribution of ratios," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 204–210.

[47] K. ITO, "Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1091–1105.

[48] G. S. JAMES, "The comparison of several groups of observations when the ratios of the population variances are unknown," *Biometrika*, Vol. 38 (1951), pp. 324–329.

[49] G. S. JAMES, "Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown," *Biometrika*, Vol. 41 (1954), pp. 19–43.

[50] G. S. JAMES, "On the accuracy of weighted means and ratios," *Biometrika*, Vol. 43 (1956), pp. 304–321.

[51] Y. W. LINDEBERG, "Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung," *Math. Z.*, Vol. 15 (1922), pp. 211–225.

[52] A. LYAPOUNOV, "Sur une proposition de la théorie des probabilités," *Bull. Acad. Imp. Sci. St. Petersbourg (Akad. Nauk SSSR)*, Ser. 5, Vol. 13 (1900), pp. 359–386.

[53] A. LYAPOUNOV, "Nouvelle forme du théorème sur la limite de probabilites," *Mem. Acad. Imp. Sci. St. Petersbourg (Akad. Nauk SSSR), Classe Phys. Math.*, Ser. 8, Vol. 12 (1901), No. 5, 12 pp.

[54] H. MAKABE, "A normal approximation to the binomial distribution," *Rep. Stat. Appl. Res. Un. Jap. Sci. Engrs.*, Vol. 4 (1955), pp. 47–53.

[55] H. MAKABE AND H. MORIMURA, "A normal approximation to the Poisson distribution," *Rep. Stat. Appl. Res. Un. Jap. Sci. Engrs.*, Vol. 4 (1955), pp. 37–46.

[56] S. MORIGUTI, "A note on Hartley's formula of studentization," *Rep. Stat. Appl. Res. Un. Jap. Sci. Engrs.*, Vol. 2 (1953), pp. 99–103.

[57] S. MORIGUTI, "Confidence limits for a variance component," *Rep. Stat. Appl. Res. Un. Jap. Sci. Engrs.*, Vol. 3 (1954), pp. 29–41.

[58] K. R. NAIR, "The studentized form of the extreme mean square test in the analysis of variance," *Biometrika*, Vol. 35 (1948), pp. 16–31.

[59] K. R. NAIR, "The distribution of the extreme deviate from the sample mean and its studentized form," *Biometrika*, Vol. 35 (1948), pp. 118–144.

[60] E. S. PEARSON AND H. O. HARTLEY, "Tables of the probability integral of the studentized range," *Biometrika*, Vol. 33 (1942), pp. 89–99.

[61] V. V. PETROV, "Extension of Cramér's limit theorem to non-identically distributed, independent quantities" (In Russian), Vestnik Leningrad Univ. (1953) No. 8, pp. 13–25.

[62] S. P. SUN, "On the successive approximation to the distribution of the third moment about the mean of independent variates," *Acad. Sinica Sci. Record*, Vol. 1 (1945), pp. 351–354.

[63] G. SZËGO, *Orthogonal Polynomials, American Mathematical Society Colloquium Publications*, Vol. 23 (1939).

[64] D. TEICHROEW, "A table of Campbell's polynomials for computation of probability points of the incomplete gamma distribution," Personal communication.

[65] D. TEICHROEW, "A table of millile probability points of the incomplete gamma distribution," Personal communication.

[66] D. TEICHROEW, "Generalized normalization polynomials for the t and incomplete gamma distributions," Personal communication.

[67] W. H. TRICKETT AND B. L. WELCH, "On the comparison of two means: further discussion of iterative methods for calculating tables," *Biometrika*, Vol. 41 (1954), pp. 361–374.

[68] W. H. TRICKETT, B. L. WELCH, AND G. S. JAMES, "Further critical values for the two-means problem," *Biometrika*, Vol. 43 (1956), pp. 203–205.

[69] S. URA, "On the power function of Welch's test procedure in the two sample problem," *Rep. Stat. Appl. Res. Un. Jap. Sci. Engrs.*, Vol. 4 (1955), pp. 1–13.

[70] J. V. USPENSKY, *Introduction to Mathematical Probability*, New York: McGraw-Hill (1937).

[71] A. WALD, "Testing the difference between the means of two normal populations with unknown standard deviations," *Selected Papers in Statistics and Probability by Abraham Wald*, pp. 669–695, New York: McGraw-Hill (1955).

[72] D. L. WALLACE, "Bounds for normal approximations to Student's and the chi-square distributions," Report SRC—80710Wc89, Statistical Research Center, University of Chicago (1958).

[73] W. WASOW, "On the asymptotic transformation of certain distributions into the normal distribution," *Proceedings of Symposia in Applied Mathematics VI, Numerical Analysis*, pp. 251–259, New York: McGraw-Hill (1956).

[74] B. L. WELCH, "The significance of the difference between two means when the population variances are unequal," *Biometrika*, Vol. 29 (1937), pp. 350–362.

[75] B. L. WELCH, "The generalization of 'Student's' problem when several different populations are involved," *Biometrika*, Vol. 34 (1947), pp. 28–35.

[76] E. B. WILSON AND M. M. HILFERTY, "The distribution of chi-square," *Proc. Nat. Acad. of Sci. U.S.A.*, Vol. 17 (1931), pp. 684–688.