The Annals of Probability
1983, Vol. 11, No. 4, 1028–1036

CHI SQUARED APPROXIMATIONS TO THE DISTRIBUTION OF A SUM OF INDEPENDENT RANDOM VARIABLES

BY PETER HALL

Australian National University

We suggest several Chi squared approximations to the distribution of a sum of independent random variables, and derive asymptotic expansions which show that the error of approximation is of order n^{-1} as $n \to \infty$. The error may be reduced to $n^{-3/2}$ by making a simple secondary approximation.

1. Introduction. The central limit theorem for sums of independent random variables is of great significance in mathematics and statistics, since it provides a very simple and tractable approximation to a wide range of complicated distributions. This importance is reflected in the vast literature on rates of convergence, which supplies information on the order of the approximation and on the factors which influence its accuracy. For example, it follows from Chebyshev-Edgeworth-Cramér expansions that if the summand distribution is skew rather than symmetric then the central limit approximation in a sample of size n is of order $n^{-1/2}$. If the distribution were symmetric then this approximation could be as accurate as order n^{-1}.

It seems natural to approximate a sum of independent, skewed random variables by another skewed sum. Such an approximation should be valid even in the case of a discrete distribution, such as the binomial, provided an appropriate continuity correction is incorporated. Perhaps the best known example of a skewed sum is the Chi squared distribution. Thus, we are led to approximate the distribution of a sum of independent random variables by a Chi squared distribution. We shall give formal descriptions of several versions of this approximation, and show that the Chi squared approximation is of order n^{-1} rather than $n^{-1/2}$.

Of course, the Chi squared distribution is itself asymptotically normal, and so for very large samples the Chi squared approximation is close to the normal approximation. Our thesis is that in many circumstances, the Chi squared distribution provides a good penultimate approximation to the distribution of a sum of independent random variables. The concept of penultimate approximations in statistics is by no means new. It was employed more than half a century ago by R. A. Fisher and L. H. C. Tippett to improve on the approximation to normal extremes by an extreme value distribution. The point we wish to make is that in many circumstances, such as the construction of hypothesis tests, distributional approximations have definite advantages over approximations via asymptotic expansions. For example, suppose the distribution of a (standardized) statistic T_n admits the expansion

$$P(T_n \le x) = \Phi(x) + n^{-1/2}\psi(x)\phi(x) + o(n^{-1/2}), \tag{1.1}$$

where Φ is the standard normal distribution function and $\phi = \Phi'$. Often we wish to choose x_0 such that $P(T_n \le x_0) \simeq \alpha$, for a fixed, predetermined level α. Direct application of (1.1) to this problem involves considerable "trial-and-error" computation, to obtain a number x_0 satisfying $\Phi(x_0) + n^{-1/2}\psi(x_0)\phi(x_0) = \alpha$. An indirect but more efficient approach is to observe from (1.1) that

$$P\{T_n \le x - n^{-1/2}\psi(x)\} = \Phi(x) + o(n^{-1/2}). \tag{1.2}$$

Received April 1982; revised January 1983.

AMS 1980 *subject classifications*. Primary 60F05, 60G50; secondary 62E20.

Key words and phrases. Approximation, asymptotic expansion, central limit theorem, Chi squared, rate of convergence, sums of independent random variables.

We could take $x_0 = x_1 - n^{-1/2}\psi(x_1)$, where x_1 satisfies $\Phi(x_1) = \alpha$. However, the approximation (1.2) does not hold *uniformly* in x, and can be particularly poor out in the tails, which is precisely where it is usually required. The distributional approximation suggested in this paper takes the form $P\{T_n \le x_n(\alpha)\} = \alpha + o(n^{-1/2})$, *uniformly* in α, and does not suffer the drawbacks cited above.

Our study of Chi squared approximations falls naturally into two parts; first of all, the case where the third moment of the underlying distribution is known, and secondly, where it is unknown. These situations are handled in Sections 2 and 3, respectively, and proofs of results from Section 2 are deferred until Section 4. In both cases we assume that the underlying variance is known, either because of some parametric knowledge about the form of the distribution or because of practical experience with the distribution in the past. This assumption is often satisfied in practice. Indeed, normal approximations (rather than Chi squared approximations) under this condition are taught in most elementary statistics courses. See for example Section 9.9, page 230, and Exercises 14–16, page 235 of Freund (1979). For a different approach to normal approximations, see Hall (1983).

It is worth remarking that our Chi squared approximations are really gamma approximations, since half a Chi squared random variable on n degrees of freedom is gamma with parameter $\frac{1}{2}n$. However, gamma tables are less readily available that Chi squared tables, and so it is more practical to study Chi squared approximations.

2. The case where third moments are known. Let $Y_1, Y_2, \cdots$ be independent random variables with mean μ and variance σ^2 (known). Confidence intervals for μ, or hypothesis tests about μ, are usually based on the standardized statistic,

$$\textstyle\sum_1^n (Y_i - \mu)/\sigma.$$

Therefore we may simplify our problem by considering the random variables $X_i = (Y_i - \mu)/\sigma$, $i \ge 1$, instead of the Y_i's. The X_i's have zero mean and unit variance. Let us assume that they also have finite fourth moment $\mu_4 = E(X^4)$, and set $\mu_3 = E(X^3)$. (Here X is a random variable with the same distribution as X_1.) In this section we shall assume that μ_3 is known. Without loss of generality we may take $\mu_3 \ge 0$, since the contrary case may be handled by replacing X_i by $-X_i$ for $i \ge 1$. We define $S_n = \sum_{j=1}^n X_j$, and let ϕ denote the standard normal density function.

The usual normal approximation to $n^{-1/2n}$ consists of regarding $n^{-1/2}S_n$ as normal $N(0, 1)$. The Chi squared approximation is carried out as follows. Let $\nu = \nu(n) = [8n/\mu_3^2]$, which can be taken as either the integer part of $8n/\mu_3^2$ or the integer nearest to $8n/\mu_3^2$. Consider $n^{-1/2}S_n$ as having the same distribution as $T_\nu \equiv (2\nu)^{-1/2}(\chi_\nu^2 - \nu)$, where χ_ν^2 has the Chi squared distribution on ν degrees of freedom. In the case $\mu_3 = 0$, corresponding to $\nu = \infty$, T_ν is taken to have the standard normal distribution. Let $z_\nu^+(\alpha)$ be the upper $(1-\alpha)$-level critical point for T_ν; that is, $P\{T_\nu \le z_\nu^+(\alpha)\} = 1 - \alpha$, for $0 < \alpha < 1$. Write $z = z(\alpha)$ for $z_\infty^+(\alpha)$, the upper $(1-\alpha)$-level critical point of the standard normal distribution. The order of the Chi squared approximation is described by the following theorem, which covers the case of a smooth distribution.

THEOREM 1. *Suppose $E(X^4) < \infty$ and $\mu_3 \ge 0$, and that the distribution of X satisfies Cramér's continuity condition,*

$$\text{(C)} \qquad \limsup_{t\to\infty} |E(e^{itX})| < 1.$$

Then

$$P\{n^{-1/2}S_n \le z_\nu^+(\alpha)\} = (1-\alpha) + n^{-1}(1/48)z(z^2-3)(3\mu_3^2 - 2\mu_4 + 6)\phi(z) + o(n^{-1})$$

uniformly in $0 < \alpha < 1$, as $n \to \infty$.

Interestingly, in the special case $\alpha = 0.04163+$, corresponding to $z(\alpha) = \sqrt{3}$, it follows

TABLE 1
Approximations to the value of x satisfying $P\{\chi_n^2(\lambda) \le x\} = 1 - \alpha$. (The approximations x_A, x_P, x_N, x_{NC} are defined following Theorem 1; x_0 is the exact x.)

$\alpha = 0.10$; λ small:

n	10	15	20	25
λ	2.935	3.599	4.161	4.658
x_0	20.483	27.488	34.170	40.647
$x_A - x_0$	−0.113	−0.124	0.099	0.074
$x_P - x_0$	0.712	0.609	0.545	0.501
$x_N - x_0$	−0.325	−0.347	−0.360	−0.368
$x_{NC} - x_0$	0.183	0.152	0.132	0.119

$\alpha = 0.10$; λ large:

n	10	15	20	25
λ	9.432	11.189	12.677	13.992
x_0	29.588	37.697	45.315	52.620
$x_A - x_0$	−0.149	−0.057	0.050	−0.045
$x_P - x_0$	−0.160	−0.604	0.759	0.493
$x_N - x_0$	−0.415	−0.424	−0.428	−0.431
$x_{NC} - x_0$	0.154	0.134	0.121	0.111

from Theorem 1 that

$$P\{n^{-1/2}S_n \le z_\nu^+(\alpha)\} = (1 - \alpha) + o(n^{-1}).$$

Under the slightly more severe moment condition $E(|X|^5) < \infty$, the remainder $o(n^{-1})$ in Theorem 1 may be sharpened to $O(n^{-3/2})$.

As an application of Theorem 1 we shall derive a central Chi squared approximation to the noncentral Chi squared distribution. If $\chi_n^2(\lambda)$ denotes a variable with the noncentral Chi squared distribution on n degrees of freedom and with noncentrality parameter λ, we may write

$$\{\chi_n^2(\lambda) - (n + \lambda)\}/\{2(n + 2\lambda)\}^{1/2} = n^{-1/2} \textstyle\sum_1^n X_i,$$

where the X_i's have zero mean, unit variance and third moment $\mu_3 = 2^{3/2}(1 + 3\lambda/n)(1 + 2\lambda/n)^{-3/2}$. An application of the preceding theory suggests the approximation

$$P\{\chi_n^2(\lambda) \le x\} \simeq P[\chi_\nu^2(0) \le \nu + (n + 2\lambda)(n + 3\lambda)^{-1}\{x - (n + \lambda)\}],$$

where $\nu = [(n + 2\lambda)^3(n + 3\lambda)^{-2}]$. This differs from the commonly used Chi squared approximation, which is due to Patnaik (1949) and takes the form

$$P\{\chi_n^2(\lambda) \le x\} \simeq P\{\chi_{\nu'}^2(0) \le (n + \lambda)(n + 2\lambda)^{-1}x\},$$

where $\nu' = n + [\lambda^2(n + 2\lambda)^{-1}]$.

Suppose it is desired to find x such that $P\{\chi_n^2(\lambda) \le x\} = 1 - \alpha$, for predetermined α and λ. The approximation described by Theorem 1 suggests taking $x = x_A \equiv n + \lambda + (n + 3\lambda)(n + 2\lambda)^{-1}(\xi - \nu)$, where $P\{\chi_\nu^2(0) \le \xi\} = 1 - \alpha$; Patnaik's approximation suggests $x = x_P \equiv (n + 2\lambda)(n + \lambda)^{-1}\eta$, where $P\{\chi_{\nu'}^2(0) \le \eta\} = 1 - \alpha$; the normal approximation suggests $x = x_N \equiv n + \lambda + \{2(2\lambda + n)\}^{1/2}\zeta$, where $\Phi(\zeta) = 1 - \alpha$; and the normal approximation with correction for skewness (see (1.2)) suggests

$$x = x_{NC} \equiv n + \lambda + \{2(n + 2\lambda)\}^{1/2}\zeta + \{2(n + 3\lambda)/3(n + 2\lambda)\}(\zeta^2 - 1).$$

There is little to choose between these methods from the point of view of simplicity. Their performances are compared in Table 1, which suggests that the approximation x_A is superior to the other three. (In Table 1, exact values x_0 equal the upper 2.5% points of $\chi_n^2(0)$ in the case of small λ, and 0.1% points in the case of large λ. The λ values were taken from power tables for the Chi squared test prepared by Haynam, Govindarajulu and Leone (1970).)

High orders of approximation may be achieved for many other distributions, provided we make a simple secondary approximation. This is demonstrated by the following result.

THEOREM 2. *Assume the conditions of Theorem* 1, *and let* $\hat{\mu}_4 = \hat{\mu}_4(X_1, \cdots, X_n)$ *be an estimate of* μ_4 *which satisfies*

$$P(|\hat{\mu}_4 - \mu_4| > \delta) = o(n^{-1}) \tag{2.1}$$

as $n \to \infty$, *for all* $\delta > 0$. *Then for each* $\varepsilon > 0$,

$$P\{n^{-1/2}S_n \le z_\nu^+(\alpha) + n^{-1}(1/48)z(z^2 - 3)(2\hat{\mu}_4 - 3\mu_3^2 - 6)\} = 1 - \alpha + o(n^{-1}) \tag{2.2}$$

uniformly in $\varepsilon < \alpha < 1 - \varepsilon$, *as* $n \to \infty$.

Again, the term $o(n^{-1})$ may be sharpened to $O(n^{-3/2})$ under more stringent moment conditions. One candidate for the estimator $\hat{\mu}_4$ is $\hat{\mu}_4 = n^{-1}\sum_1^n X_j^4$, and then it follows from Theorem 27, page 283 of Petrov (1975) that condition (2.1) holds if $E(X^8) < \infty$. Incidentally, (2.2) remains true if $z = z(\alpha)$ is replaced by $z_\nu^+(\alpha)$.

Perhaps the best illustration of the use of the Chi squared approximation for μ_3 known is the case of the binomial distribution. In this situation the summands are distributed on a lattice, and condition (C) no longer holds. We must set up a little more theory. Since a continuity correction should be incorporated into the normal approximation in the case of a lattice distribution, it is necessary to state the lattice results slightly differently from Theorems 1 and 2.

Assume that X takes values only in the set $a + d\mathbb{Z}$, where a is a real number, $d > 0$ is the maximal span of the lattice and $\mathbb{Z} = \{\cdots, -1, 0, 1, \cdots\}$. Then $n^{-1/2}S_n$ takes only values of the form $x = (na + md)/n^{1/2}$, where $m \in \mathbb{Z}$. We continue to assume that X has zero mean and unit variance, and set $\nu = [8n/\mu_3^2]$.

THEOREM 3. *Suppose* $E(X^4) < \infty$ *and* $\mu_3 \ge 0$, *and that the distribution of* X *is lattice as defined above. Then*

$$\begin{aligned} P(n^{-1/2}S_n \le x) &= P\left\{(2\nu)^{-1/2}(\chi_\nu^2 - \nu) \le x + \frac{d}{2n^{1/2}}\right\} \\ &\quad + n^{-1}x\left\{\frac{1}{24}d^2 + \frac{1}{48}(x^2 - 3)(3\mu_3^2 - 2\mu_4 + 6)\right\}\phi(x) + o(n^{-1}) \end{aligned} \tag{2.3}$$

uniformly in x *of the form* $(na + md)/n^{1/2}$, *where* $m \in \mathbb{Z}$, *as* $n \to \infty$.

The term $d/2n^{1/2}$ appearing on the right hand side of (2.3) is the correction for continuity; see Yates (1934) and Pearson (1947, page 147). As an example, suppose $\{p(1 - p)\}^{1/2}S_n + np$ is binomial Bi(n, p), where $0 < p \le 1/2$. Then $\mu_3 = (1 - 2p)/\{p(1 - p)\}^{1/2} \ge 0$, $\mu_4 = (1 - 3p + 3p^2)/p(1 - p)$, $d = 1/\{p(1 - p)\}^{1/2}$, and the expansion (2.3) becomes

$$\begin{aligned} P(n^{-1/2}S_n \le x) &= P\left\{(2\nu)^{-1/2}(\chi_\nu^2 - \nu) \le x + \frac{d}{2n^{1/2}}\right\} \\ &\quad + n^{-1}\frac{1}{48p(1 - p)}x(x^2 - 1)\phi(x) + o(n^{-1}). \end{aligned}$$

To illustrate the application of this result, let us calculate $P(Y \le 2)$ where Y is binomial Bi(10, 0.1). The exact probability equals 0.9298, the normal approximation with continuity correction gives 0.9431, and the Chi squared approximation gives 0.9276. Other lattice distributions to which this approximation can be applied include the Poisson, the negative binomial and the Pascal.

3. The case where third moments are unknown. We adapt the notation introduced in Section 2, and so we regard $X, X_1, X_2, \cdots$ as independent, identically distributed random variables with finite fourth moment μ_4, zero mean, unit variance and third moment μ_3. In this section we assume that μ_3 is unknown, and we estimate it using

$$\hat{\mu}_3 = n^{-1} \sum_1^n (X_j - \bar{X})^3,$$

where $\bar{X} = n^{-1} \sum_1^n X_j$.

Let $-z_\nu^-(\alpha)$ be the lower $(1-\alpha)$-level critical point for $(2\nu)^{-1/2}(\chi_\nu^2 - \nu)$. That is,

$$P\{(2\nu)^{-1/2}(\chi_\nu^2 - \nu) \le -z_\nu^-(\alpha)\} = \alpha.$$

Define the integer-valued random variable N by $N = [8n/\hat{\mu}_3^2]$, and set $z_N(\alpha) = z_N^+(\alpha)$ if $\hat{\mu}_3 \ge 0$; $z_N^-(\alpha)$ otherwise. Recall that $z = z(\alpha)$ is the upper $(1-\alpha)$-level critical point of the standard normal distribution. Our next result is an analogue of Theorem 1.

THEOREM 4. *Suppose $E(X^6 |\log|X||^{3+\eta}) < \infty$ for some $\eta > 0$, and the joint distribution of (X, X^3) satisfies Cramér's continuity condition,*

$$\limsup_{|s|+|t|\to\infty} |E \exp(itX + isX^3)| < 1.$$

Then

$$(3.1) \quad \begin{aligned} &P\{n^{-1/2}S_n \le z_N(\alpha)\} \\ &\quad = (1-\alpha) + n^{-1}(1/48)z\{3(z^2-3)\mu_3^2 + 2(3z^2-1)(\mu_4-3)\}\phi(z) + o(n^{-1}) \end{aligned}$$

uniformly in $0 < \alpha < 1$, as $n \to \infty$.

Note that in the case of any distribution with zero skewness and kurtosis, the term of order n^{-1} in (3.1) vanishes. Therefore the Chi squared approximation will not be seriously in error when the underlying distribution is, in fact, normal.

We should comment on the moment condition imposed in Theorem 4. In a sense, the expansion (3.1) is a Chebyshev-Edgeworth-Cramér expansion for a function of a vector of sums of independent random variables. As such, it could have been derived in part by using results on asymptotic expansions; see for example Bhattacharya and Rao (1976) or Bhattacharya and Ghosh (1978). However, this would have entailed very restrictive moment conditions. To achieve a term of order n^{-1} it is necessary to assume that the vector has finite fourth moments, and since one element of the vector is $\sum_1^n X_j^3$, this would require the assumption that $E|X|^{12} < \infty$. To avoid this imposition we use a longer argument, involving non-standard truncations.

If the uniformity in (3.1) is required only on $(\varepsilon, 1-\varepsilon)$, the logarithmic factor in the moment condition may be dropped, as the following theorem shows. This result is an analogue of Theorem 2.

THEOREM 5. *Assume the conditions of Theorem 4, except that the constraint $E(X^6|\log|X||^{3+\eta}) < \infty$ may be replaced by $E(X^6) < \infty$. Then the expansion* (3.1) *holds uniformly in $\varepsilon < \alpha < 1-\varepsilon$ as $n \to \infty$, for each $\varepsilon > 0$. Furthermore, if $\hat{\mu}_4 = \hat{\mu}_4(X_1, \cdots, X_n)$ satisfies condition* (2.1), *then*

$$P[n^{-1/2}S_n \le z_N(\alpha) + n^{-1}(1/48)z\{3(3-z^2)\hat{\mu}_3^2 + 2(1-3z^2)(\hat{\mu}_4 - 3)\}] = 1 - \alpha + o(n^{-1})$$

uniformly in $\varepsilon < \alpha < 1-\varepsilon$, as $n \to \infty$.

If $E(X^8) < \infty$ then (2.1) is satisfied with $\hat{\mu}_4 = n^{-1} \sum_1^n (X_j - \bar{X})^4$. The remainders $o(n^{-1})$ in Theorems 4 and 5 may be reduced to $O(n^{-3/2})$ under more stringent moment conditions.

4. Proofs. The Symbol C throughout denotes a positive generic constant. The proofs of Theorems 4 and 5 are rather long, and at the request of the editors they have been deleted from the present paper, to be published elsewhere (Hall, 1982).

PROOF OF THEOREM 1. The usual Chebyshev-Edgeworth-Cramér expansion of the distribution of $n^{-1/2}S_n$ may be written as

$$P(n^{-1/2}S_n \le x) = \Phi(x) + n^{-1/2}(1/6)\mu_3(1 - x^2)\phi(x) + n^{-1}(1/72)x\{(10x^2 - x^4 - 15)\mu_3^2 + 3(3 - x^2)(\mu_4 - 3)\}\phi(x) + o(n^{-1}) \tag{4.1}$$

uniformly in x, where Φ and ϕ are the standard normal distribution and density functions, respectively. See for example Theorem 4, page 169 of Petrov (1975). The following lemma, whose proof is given after the proof of Theorem 1, provides an expansion of $z_n^+(\alpha)$.

LEMMA 1. *Let $z = z(\alpha)$ be the solution of the equation $1 - \Phi(z) = \alpha$, and set*

$$y_n = y_n(\alpha) = z + n^{-1/2}\frac{2^{1/2}}{3}(z^2 - 1) + n^{-1}\frac{1}{18}z(z^2 - 7) + n^{-3/2}p_1(z) + n^{-2}p_2(z),$$

where p_1 and p_2 are polynomials. We may choose p_1 and p_2 such that for all $\beta, \gamma > 0$, we have $z_n^+(\alpha) = y_n(\alpha) + O(n^{\beta-5/2})$ uniformly in $\gamma n^{-\beta} \le \alpha \le 1 - \gamma n^{-\beta}$, as $n \to \infty$.

Let $x_n = (2\beta \log n)^{1/2}$ where $\beta > 0$, and note that

$$1 - P\{(2n)^{-1/2}(\chi_n^2 - n) \le x_n\} \sim 1 - \Phi(x_n) \sim (4\pi\beta \log n)^{-1/2}n^{-\beta}$$

as $n \to \infty$. (The first asymptotic equivalence follows from a result on large deviation probabilities; see for example Theorem 1, page 218 of Petrov, 1975). Since $z_n^+(\alpha)$ is the solution of the equation

$$1 - P\{(2n)^{-1/2}(\chi_n^2 - n) \le z_n^+(\alpha)\} = \alpha,$$

then if $\alpha = \alpha_n = \gamma n^{-\beta}$ we must necessarily have $z_n^+(\alpha_n) < (2\beta \log n)^{1/2}$ for large n, and similarly, $z_n^+(1 - \alpha_n) > -(2\beta \log n)^{1/2}$ for large n. Therefore

$$\sup_{\alpha_n<\alpha<1-\alpha_n} | z_n^+(\alpha) | < (2\beta \log n)^{1/2} \tag{4.2}$$

for large n. We may now deduce from Lemma 1 that, if $\beta = 1 + \delta$ for a small positive δ,

$$z_n^+(\alpha) = z + n^{-1/2}\frac{2^{1/2}}{3}(z^2 - 1) + n^{-1}\frac{1}{18}z(z^2 - 7) + O(n^{-1-\eta})$$

uniformly in $\alpha_n \le \alpha \le 1 - \alpha$, for some $\eta > 0$. On taking $x = z_n^+(\alpha)$ in (4.1), and constructing Taylor expansions of the functions on the right hand side about the point z, we find that

$$\begin{aligned} &P\{n^{-1/2}S_n \le z_\nu^+(\alpha)\} \\ &= \left[\Phi(z) + \left\{\nu^{-1/2}\frac{2^{1/2}}{3}(z^2 - 1) + \nu^{-1}\frac{1}{18}z(z^2 - 7)\right\}\phi(z) - \frac{1}{2}\left\{\nu^{-1/2}\frac{2^{1/2}}{3}(z^2 - 1)\right\}^2 z\phi(z)\right] \\ &\quad + n^{-1/2}\frac{1}{6}\mu_3\left[(1 - z^2)\phi(z) + \left\{\nu^{-1/2}\frac{2^{1/2}}{3}(z^2 - 1)\right\}z(z^2 - 3)\phi(z)\right] \\ &\quad + n^{-1}\frac{z}{72}\left\{(10z^2 - z^4 - 15)\mu_3^2 + 3(3 - z^2)(\mu_4 - 3)\right\}\phi(z) + o(n^{-1}) \\ &= \Phi(z) + n^{-1}\frac{1}{48}z(z^2 - 3)(3\mu_3^2 - 2\mu_4 + 6)\phi(z) + o(n^{-1}) \end{aligned}$$

uniformly in $\alpha_n \le \alpha \le 1 - \alpha_n$. This proves Theorem 1 for $\alpha_n \le \alpha \le 1 - \alpha_n$.

To complete the proof we shall treat the case $\alpha < \alpha_n$. The case $\alpha > 1 - \alpha_n$ may be handled similarly. Recall that $\alpha_n = \gamma n^{-\beta}$, where $\beta = 1 + \delta > 1$. The argument leading to (4.2) may be repeated to show that $z(\alpha_n) \sim (2\beta \log n)^{1/2}$, and so for all sufficiently large n,

$$\inf_{\alpha<\alpha_n} z(\alpha) > \{(2+\delta)\log n\}^{1/2}.$$

Therefore

$$\sup_{\alpha<\alpha_n} | \alpha - n^{-1}(1/48) z(z^2-3)(3\mu_3^2 - 2\mu_4 + 6)\phi(z) | = o(n^{-1})$$

as $n \to \infty$. And it follows from the expansion (4.1) that for large n,

$$\sup_{\alpha<\alpha_n}[1 - P\{n^{-1/2}S_n \le z_\nu^+(\alpha)\}] \le 1 - P[n^{-1/2}S_n \le \{(2+\delta/2)\log n\}^{1/2}] = o(n^{-1}).$$

Therefore

$$\sup_{\alpha<\alpha_n} | P\{n^{-1/2}S_n \le z_\nu^+(\alpha)\} - (1-\alpha) - n^{-1}(1/48) z(z^2-3)(3\mu_3^2 - 2\mu_4 + 6)\phi(z) | = o(n^{-1})$$

as $n \to \infty$, completing the proof of Theorem 1.

Proof of Lemma 1. By the usual expansion of the distribution function of a sum of independent random variables,

$$\begin{aligned} &P\{(2n)^{-1/2}(\chi_n^2 - n) \le y\} \\ &\quad = \Phi(y) + n^{-1/2}\frac{2^{1/2}}{3}(1-y^2)\phi(y) + n^{-1}\frac{y}{18}(11y^2 - 2y^4 - 3)\phi(y) \qquad (4.3) \\ &\quad\quad + n^{-3/2}q_1(y)\phi(y) + n^{-2}q_2(y)\phi(y) + O(n^{-5/2}) \end{aligned}$$

uniformly in $-\infty < y < \infty$, as $n \to \infty$, where q_1 and q_2 are polynomials. Replace y by $y_n = z + y_n'$, say, and expand the functions on the right hand side in Taylor series about z. For example, expanding $\Phi(y_n)$ we obtain

$$\Phi(y_n) = \Phi(z) + y_n'\phi(z) - \tfrac{1}{2}y_n'^2 z\phi(z) + \tfrac{1}{6}y_n'^3\xi_1(z) + \tfrac{1}{24}y_n'^4\xi_2(z) + \tfrac{1}{120}y_n'^5\xi_3(z+\theta y_n'),$$

where $0 < \theta < 1$ and the functions ξ_i all have the form $r_i\phi$ for polynomials r_i. Now, $\phi(z+\theta y_n')/\phi(z) = \exp(-\theta y_n' z - \tfrac{1}{2}\theta^2 y_n'^2)$, and so for any choice of the polynomials p_1 and p_2, $\phi(z+\theta y_n')/\phi(z)$ is bounded uniformly in $|z| \le \log n$ and $|\theta| \le 1$, as $n \to \infty$. Therefore for polynomials q_3 and q_4 not depending on p_2,

$$\begin{aligned} \Phi(y_n) &= \Phi(z) + n^{-1/2}\frac{2^{1/2}}{3}(z^2-1)\phi(z) + n^{-1}\frac{z}{18}(5z^2 - 2z^4 - 9)\phi(z) \\ &\quad + n^{-3/2}\{p_1(z) + q_3(z)\}\phi(z) + n^{-2}\{p_2(z) + q_4(z)\}\phi(z) + O(n^{-5/2}) \end{aligned}$$

uniformly in $|z| \le \log n$. The polynomial q_4 depends on p_1, but q_3 does not. Carrying out an expansion of this type for each of the terms on the right in (4.3), and collecting the terms, we may deduce that for polynomials q_5 and q_6,

$$\begin{aligned} P\{(2n)^{-1/2}(\chi_n^2 - n) \le y_n\} &= \Phi(z) + n^{-3/2}\{p_1(z) + q_5(z)\}\phi(z) \\ &\quad + n^{-2}\{p_2(z) + q_6(z)\}\phi(z) + O(n^{-5/2}) \end{aligned}$$

uniformly in $|z| \le \log n$. Neither q_5 nor q_6 depends on p_2, and only q_6 depends on p_1. Therefore if we define $p_1 = -q_5$ and $p_2 = -q_6$, and recall that

$$\Phi(z) = 1 - \alpha = P\{(2n)^{-1/2}(\chi_n^2 - n) \le z_n^+\},$$

we see that

$$P\{(2n)^{-1/2}(\chi_n^2 - n) \le y_n\} = P\{(2n)^{-1/2}(\chi_n^2 - n) \le z_n^+\} + O(n^{-5/2}) \tag{4.4}$$

uniformly in values of α for which $|z| \le \log n$.

The argument which we used to derive the inequality (4.2) may be used to show that with $\alpha_n = \gamma n^{-\beta}$, we have $z(\alpha_n) < (2\beta \log n)^{1/2}$ for large n. A slightly longer proof will demonstrate that $y_n(\alpha_n) < (2\beta \log n)^{1/2}$ for large n. Therefore

$$\max\{y_n(\alpha_n), z_n^+(\alpha_n)\} < (2\beta \log n)^{1/2} \tag{4.5}$$

for large n. Set $\delta_n(\alpha) = z_n^+(\alpha) - y_n(\alpha)$, and let f_n be the density of the random variable $(2n)^{-1/2}(\chi_n^2 - n)$. Then for each α,

$$P\{(2n)^{-1/2}(\chi_n^2 - n) \le z_n^+(\alpha)\} = P\{(2n)^{-1/2}(\chi_n^2 - n) \le y_n(\alpha)\} + \delta_n(\alpha) f_n\{u_n(\alpha)\},$$

where $u_n(\alpha)$ lies between $z_n^+(\alpha)$ and $y_n(\alpha)$. It follows from (4.5) that $u_n(\alpha) < (2\beta \log n)^{1/2}$ for all $\alpha \ge \alpha_n$ and all large n, and similarly it may be proved that $u_n(\alpha) > -(2\beta \log n)^{1/2}$ for all $\alpha \le 1 - \alpha_n$ and all large n. Consequently

$$\begin{aligned} \sup_{\alpha_n \le \alpha \le 1-\alpha_n} |\delta_n(\alpha)| \le [\sup_{\alpha_n \le \alpha \le 1-\alpha_n} | P\{(2n)^{-1/2}(\chi_n^2 - n) \le z_n^+(\alpha)\} \\ - P\{(2n)^{-1/2}(\chi_n^2 - n) \le y_n(\alpha)\} |]\{\inf_{|u|<(2\beta \log n)^{1/2}} f_n(u)\}^{-1}. \end{aligned} \tag{4.6}$$

A local limit theorem for f_n, such as Theorem 14, page 206 of Petrov (1975), allows us to deduce that

$$\inf_{|u|<(2\beta \log n)^{1/2}} f_n(u) \sim \phi\{(2\beta \log n)^{1/2}\} = (2\pi)^{-1/2} n^{-\beta}.$$

When this result is substituted into (4.6), and the result (4.4) used to estimate the numerator on the right hand side of (4.6), we see that

$$\sup_{\alpha_n \le \alpha \le 1-\alpha_n} |\delta_n(\alpha)| = O(n^{-5/2}),$$

as required.

PROOF OF THEOREM 2. Let $p_n(\alpha)$ denote the probability on the left in (2.2), let E stand for the event $\{|\hat{\mu}_4 - \mu_4| > \delta\}$ where $0 < \delta < \mu_4$, and observe that if $z(z^2 - 3) \ge 0$,

$$\begin{aligned} P\{n^{-1/2}S_n \le z_\nu^+(\alpha) + n^{-1}(1/48)z(z^2 - 3)(2\mu_4 - 2\delta - 3\mu_3^2 - 6)\} - P(E) \le p_n(\alpha) \\ \le P\{n^{-1/2}S_n \le z_\nu^+(\alpha) + n^{-1}(1/48)z(z^2 - 3)(2\mu_4 + 2\delta - 3\mu_3^2 - 6)\} + P(E). \end{aligned}$$

Therefore it suffices to prove that

$$\begin{aligned} P\{n^{-1/2}S_n \le z_\nu^+(\alpha) + n^{-1}(1/48)z(z^2 - 3)(2\mu_4 \pm 2\delta - 3\mu_3^2 - 6)\} \\ = (1 - \alpha) + n^{-1}(1/48)z(z^2 - 3)(\pm 2\delta)\phi(z) + o(n^{-1}) \end{aligned}$$

uniformly in $\varepsilon < \alpha < 1 - \varepsilon$, as $n \to \infty$. This may be achieved using the argument in the proof of Theorem 1. The case $z(z^2 - 3) < 0$ is treated similarly.

PROOF OF THEOREM 3. It may be deduced from Theorem 4, page 61 of Esséen (1945) or Theorem 23.1, page 238 of Bhattacharya and Rao (1976) that

$$P(n^{-1/2}S_n \le x) = \Phi(x) + n^{-1/2}\psi_1(x) + n^{-1}\psi_2(x) + o(n^{-1}) \tag{4.7}$$

uniformly in x of the form $(na + md)/n^{1/2}$, $m \in \mathbb{Z}$, where

$$\psi_1(x) = \{\tfrac{1}{2}d + \tfrac{1}{6}\mu_3(1 - x^2)\}\phi(x)$$

and

$$\psi_2(x) = x[-(\tfrac{1}{12})\, d^2 + (\tfrac{1}{12})\, d\mu_3(x^2 - 3) + \tfrac{1}{72}\{(10x^2 - x^4 - 15)\mu_3^2 + 3(3 - x^2)(\mu_4 - 3)\}]\phi(x).$$

The expansion (4.3) shows that

$$P\{(2\nu)^{-1/2}(\chi_\nu^2 - \nu) \le x + d/2n^{1/2}\}$$
$$= \Phi(x) + n^{-1/2}\{\tfrac{1}{2}d + \tfrac{1}{6}\mu_3(1 - x^2)\}\phi(x)$$
$$+ n^{-1}x\{-\tfrac{1}{8}d^2 + \tfrac{1}{12}\, d\mu_3(x^2 - 3) + (\tfrac{1}{144})\mu_3^2(11x^2 - 2x^4 - 3)\}\phi(x) + o(n^{-1})$$

uniformly in x, and the desired result follows on subtracting this expansion from (4.7).

Acknowledgment. The paper has benefited from the editors' suggestions.

REFERENCES

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451. Corrigendum *ibid* **8** (1980) 1399.

BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

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

FISHER, R. A. and TIPPETT, L. H. C. (1928). Limiting forms of the frequency of the largest or smallest member of a sample. *Proc. Cambridge Phil. Soc.* **24** 180–190.

FREUND, J. E. (1979). *Modern Elementary Statistics*, 5th Edn. Prentice-Hall, Englewood Cliffs, New Jersey

HALL, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11** 569–576.

HALL, P (1982). A derivation of chi square approximations to the distribution of a sum of independent random variables. Unpublished manuscript.

HAYNAM, G. E., GOVINDARAJULU, Z. and LEONE, F. C. (1970). Tables of the cumulative non-central chi-square distribution. In *Selected Tables in Mathematical Statistics* vol. I, Eds. H. L. Harter and D. B. Owen. Markham, Chicago.

PATNAIK, P. R. (1949). The non-central χ^2- and F-distributions and their applications. *Biometrika* **36** 202–232.

PEARSON, E. S. (1947). The choice of statistical tests illustrated on the interpretation of data classed in a 2 × 2 table. *Biometrika* **34** 139–167.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, Berlin.

YATES, F. (1934). Contingency tables involving small numbers, and the χ^2 test. *J. Roy. Statist. Soc. Suppl.* **1** 217–235.

DEPARTMENT OF STATISTICS
THE AUSTRALIAN NATIONAL UNIVERSITY
P. O. BOX 4
G. CANBERRA ACT 2601
AUSTRALIA