ASYMPTOTIC EXPANSIONS

Asymptotic expansions of functions are useful in statistics in three main ways. Firstly, conventional asymptotic expansions of special functions are useful for approximate computation of integrals arising in statistical calculations. An example given below is the use of Stirling's approximation to the gamma function. Second, asymptotic expansions of density or distribution functions of estimators or test statistics can be used to give approximate confidence limits for a parameter of interest or p-valuesfor an hypothesis test. Use of the leading term of the expansion as an approximation leads to confidence limits and p-values based on the limiting form of the distribution of the statistic, whereas use of further terms often results in more accurate inference. Third, asymptotic expansions for distributions of estimators or test statistics may be used to investigate properties such as the efficiency of an estimator or the power of a test. The first two of these are discussed in this entry, which is a continuation of Serfling [35]. The third is discussed in the entry Asymptotics, Higher Order*.

An asymptotic expansion of a function is a re-expression of the function as a sum of terms adjusting a base function, expressed as follows

$$f_n(x) = f_{0n}(x)\{1 + b_{1n}g_1(x) + b_{2n}g_2(x) + \ldots + b_{kn}g_k(x) + O(b_{k+1,n})\} \quad (n \to \infty). \tag{1}$$

The sequence $\{b_{kn}\} = \{1, b_{1n}, b_{2n}, \ldots\}$ determines the asymptotic behavior of the expansion: in particular how the re-expression approximates the original function. Usual choices of $\{b_{kn}\}$ are $\{1, n^{-1/2}, n^{-1}, \ldots\}$ or $\{1, n^{-1}, n^{-2}, \ldots\}$; in any case it is required that $b_{kn} = o(b_{k-1,n})$ as $n \to \infty$. For sequences of constants $\{a_n\}$, $\{b_n\}$, we write $a_n = o(b_n)$ if $a_n/b_n \to 0$ as $n \to \infty$ and $a_n = O(b_n)$ if a_n/b_n remains bounded as $n \to \infty$. The notation $o_p(\cdot)$, $O_p(\cdot)$ is useful for sequences of random variables $\{Y_n\}$: Y_n is $o_p(a_n)$ if Y_n/a_n converges in probability to 0 as $n \to \infty$, and is $O_p(a_n)$ if $|Y_n/a_n|$ is bounded in probability as $n \to \infty$.

Asymptotic expansions are used in many areas of mathematical analysis. Three helpful textbooks on asymptotic expansions are Bleistein and Handelsman [9], Jeffreys [25] and DeBruijn [18]. An important feature of asymptotic expansions is that they are not in general convergent series and taking successively more terms from the right hand side of (1) is not guaranteed to improve the approximation to the left hand side. In the study of asymptotic expansions in analysis, emphasis is typically on $f_n(x)$ as a function of n, with x treated as an additional parameter, and n considered as the argument of the function. In these treatments it is usual to let n be real or complex, and the notation $f(z;x)$ or $f(z)$ is more standard. The functions $g_j(\cdot)$ that we have used in (1) are then just constants (in z), and the sequence $\{b_{kn}\}$ is typically $\{z^{-k}\}$ if $z \to \infty$ or $\{z^k\}$ if $z \to 0$.

An asymptotic expansion is used to provide an approximation to the function $f_n(x)$ by taking the first few terms of the right hand side of (1): for example we might write

$$f_n(x) \doteq f_{0n}(x)\{1 + b_{1n}g_1(x)\}.$$

Although the approximation is guaranteed to be accurate only as $n \to \infty$, it is often quite accurate for rather small values of n. It will usually be of interest to investigate the accuracy of the approximation for various values of x as well, an important concern being the range of values for x over which the error in the approximation is uniform.

In statistics the function $f_n(x)$ is typically a density function, a distribution function, or a moment or cumulant generating function, for a random variable computed from a sequence of random variables of length n. For example $f_n(x)$ could be the density of the standardized mean $\bar{X}_n$, say, of n independent, identically distributed random variables X_i: $\bar{X}_n = \sum X_i/n$. In this case the function $f_{0n}(x)$ is the limiting density for the standardized version of $\bar{X}_n$, usually the normal density function $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$: see Asymptotic Normality*.

Asymptotic expansions of special functions

A familiar example of an asymptotic expansion is the expansion of the gamma function $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ by Stirling's formula given, for example, in Abramowitz and Stegun [1, 6.1.37]:

$$\Gamma(z) = e^{-z}z^{z-\frac{1}{2}}(2\pi)^{\frac{1}{2}}\left(1 + \frac{1}{12z} + \frac{1}{288z^2} + O(z^{-3})\right) \quad (z \to \infty \quad \text{in} \quad |\arg z| < \pi). \qquad (2)$$

The leading term of the right hand side is Stirling's approximation to the gamma function. There are similar expansions given in Abramowitz and Stegun [1, Ch. 6] for $\log\Gamma(z)$ and its first two derivatives, the digamma and trigamma functions. The approximations given by the first several terms in these expansions are used, for example, in computing the maximum likelihood estimator and its asymptotic variance for a sample from a gamma density with unknown shape parameter.

Another example is the asymptotic expansion for the tail of the standard normal cumulative distribution function:

$$1-\Phi(z)=\int_z^\infty (2\pi)^{-1/2}\exp(-x^2/2)dx=\phi(z)z^{-1}\left(1-\frac{1}{z^2}+\frac{3}{z^4}+O(z^{-6})\right).$$

The approximation $\{1-\Phi(z)\}/\phi(z)\doteq z^{-1}$ is often called Mill's ratio.

The asymptotic expansions given above are examples of expansions obtained using Laplace's method*. Laplace's method is also very useful for deriving approximations to integrals arising in Bayesian inference.

Edgeworth and saddlepoint expansions

For statistics which are asymptotically normally distributed, the Edgeworth expansion* for density or distribution functions gives a useful and readily computed approximation. Such statistics are typically either sample means, or smooth functions of sample means, and the Edgeworth expansion for the distribution function of a standardized sample mean is given in Serfling [35]. We assume for simplicity that $X_1,\dots,X_n$ are independent and identically distributed. Define $S_n=(\bar{X}_n-n\mu)/(n^{1/2}\sigma)$, where μ and σ^2 are the mean and variance of X_i. The Edgeworth expansion for the density of S_n is

$$f_n(s)=\phi(s)\left[1+\frac{1}{n^{1/2}}(\lambda_3/6)h_3(s)+\frac{1}{n}\{(\lambda_4/24)h_4(s)+(\lambda_3^2/72)h_6(s)\}+O(n^{-3/2})\right] \quad (4)$$

where λ_3 and λ_4 are the third and fourth cumulants of $(X_i-\mu)/\sigma$, and $h_j(s)=(-1)^j\phi^{(j)}(s)/\phi(s)$ is the jth Hermite polynomial.

Note that the expression of (4) suggests the use of the first three terms as an approximation to the exact density, with the remaining terms absorbed into the expression $O(n^{-3/2})$. The full expansion for the distribution function is given in Serfling [35]. The Edgeworth expansion is derived in many textbooks, cf. the references in Serfling [35] as well as Feller [22, Ch. 16], McCullagh, [30, Ch. 6] and Barndorff-Nielsen and Cox [5, Ch. 4].

The Edgeworth approximation is quite accurate near the center of the density. In particular at $s = \mu$ the relative error in using the normal approximation is $O(n^{-1})$, and that in using the approximation suggested by (4) is $O(n^{-2})$, because the odd-order Hermite polynomials are 0 at $s = 0$. For large values of $|s|$, though, the approximation is often inaccurate for fixed values of n, as the polynomials oscillate substantially as $|s| \to \infty$. A particular difficulty is that the approximation to $f_n(s)$ may in some cases take negative values.

A different type of asymptotic expansion for the density of a sample mean is given by the saddlepoint expansion. Let the cumulant generating function of X_i be denoted by $K(t)$. The saddlepoint expansion for the density of $\bar{X}$ is defined by

$$f_n(\bar{x}) = \frac{1}{\sqrt{2\pi}} \left\{ \frac{n}{|K''(\hat{z})|} \right\}^{1/2} \exp\left[n\{K(\hat{z}) - \hat{z}\bar{x}\}\right] \left\{ 1 + \frac{3\lambda_4(\hat{z}) - 5\lambda_3^2(\hat{z})}{24n} + O(n^{-2}) \right\} \qquad (5)$$

where $\hat{z}$ is called the saddlepoint and is defined by $K'(\hat{z}) = \bar{x}$, and the rth cumulant function $\lambda_r(z) = K^{(r)}(z)/\{K''(z)\}^{r/2}$. Note that this is an asymptotic expansion in powers of n^{-1}. The next term in the expansion is a complicated expression involving cumulant functions up to order 6. The leading term of (5) is the saddlepoint approximation* to the density of $\bar{X}_n$. This expression is always positive, but will not usually integrate to exactly one, so in practice is renormalized. The renormalization improves the order of the approximation:

$$f_n(\bar{x}) = c \left\{ \frac{n}{|K''(\hat{z})|} \right\}^{1/2} \exp\left[n\{K(\hat{z}) - \hat{z}\bar{x}\}\right] \left\{ 1 + O(n^{-3/2}) \right\}. \qquad (6)$$

Evaluating the saddlepoint approximation requires knowledge of the cumulant generating function $K(z)$. Approximations based on estimating $K(z)$ by estimating the first four cumulants are discussed in Easton and Ronchetti [21], Wang [41] and Cheah, Fraser and Reid [10].

The saddlepoint approximation to the distribution function of $\bar{X}$ can be obtained by integrating (6) or by applying the saddlepoint technique and the result, due to Lugannani and Rice [29], is

$$F_n(\bar{x}) \doteq \Phi(r) + \phi(r) \left(\frac{1}{r} - \frac{1}{q} \right) \qquad (7)$$

where

$$
\begin{aligned}
r &= \operatorname{sign}(q)[2n\{K(\hat{z}) - \hat{z}\bar{x}\}]^{1/2} \\
q &= \hat{z}\{K''(\hat{z})\}^{1/2}.
\end{aligned}
$$

Approximation (7) is often surprisingly accurate throughout the range of $\bar{x}$, except near $\bar{x} = \mu$ or $r = 0$, where it should be replaced by its limit as $r \to 0$: $F_n(\mu) \doteq (1/2) - (1/6)\lambda_3(0)/\sqrt{2\pi n}$. Approximation (7) has relative error $O(n^{-1})$ for all $\bar{x}$ and $O(n^{-3/2})$ for the so-called moderate deviation region $\bar{x} - \mu = O(n^{-1/2})$. The approximation can be expressed in an asymptotically equivalent form by defining $r^* = r + r^{-1}\log(q/r)$: the approximation

$$
F_n(\bar{x}) \doteq \Phi(r^*), \tag{8}
$$

originally due to Barndorff-Nielsen [4], is asymptotically equivalent to (7).

Approximations (5) and (7) were derived in Daniels [13] and Lugannani and Rice [29] respectively, using the saddlepoint technique of asymptotic analysis. Daniels [15] exemplifies the derivation of (7). Both Kolassa [27] and Field and Ronchetti [23] provide rigorous derivations of (5) using the saddlepoint method. General discussions of the saddlepoint method can be found in Bleistein and Handelsman [19] or Courant and Hilbert [11]. The approximations can also be derived from the Edgeworth expansion, cf. Barndorff-Nielsen and Cox [5, Ch. 4] where (5) is called the tilted Edgeworth expansion.

The Edgeworth and saddlepoint approximations for distribution functions discussed here apply to continuous random variables, and adjustments to the approximations are needed in the case that the variables X_i take values on a lattice. The details are provided in Kolassa [27, Chs. 3, 5].

For vector X_i of length m, say, multivariate versions of the Edgeworth and saddlepoint density approximations are readily available. The multivariate Edgeworth approximation requires for its expression a definition of generalized Hermite polynomials, which are conceptually straightforward but notationally complex. A brief account is given in Reid [33], adapted from McCullagh [30, Ch. 5]. The multivariate version of the saddlepoint approximation is the same as (6), with $K(z) = \log E \exp(z^T x)$, $\hat{z}\bar{x}$ interpreted as a scalar product of the two vectors $\hat{z}$ and $\bar{x}$, and $|K''(\hat{z})|$ interpreted as the determinant of the $p \times p$ matrix of second derivatives of the cumulant generating function. The distribution function approximation (7) is only available for the univariate case, but an approximation to the conditional distribution function $\Pr(\bar{X}_{(1)} \leq \bar{x}_{(1)} | \bar{x}_{(2)})$ is derived in Skovgaard [36], and extended in Wang [39] and Kolassa [27, Ch. 7]. The form of this approximation has proved very useful for inference about scalar parameters in the presence of nuisance parameters.

The Edgeworth and saddlepoint approximations are both based on a limiting normal distribution for the statistic in question. Some statistics may have a limiting distribution which is not normal; in particular sample maxima or minima usually have limiting distributions of the extreme value form. For such statistics a series expansion of the density in which the basic function corresponds to the limiting density may be of more practical interest. In principle this is straightforward, and for example McCullagh [30, Ch. 5] derives Edgeworth-type expansions using arbitrary basic functions and the associated orthogonal polynomials. Examples of saddlepoint approximations based on non-normal limits are discussed in Jensen [26] and Wood, Booth and Butler [42].

Stochastic asymptotic expansions

It is often very convenient in deriving asymptotic results in statistics to use stochastic asymptotic expansions, which are analogues of (1) for random variables. For a sequence of random variables $\{Y_n\}$ a stochastic asymptotic expansion is expressed as

$$Y_n = X_0 + X_1 b_{1n} + \ldots + X_k b_{kn} + O_p(b_{k+1,n}) \tag{9}$$

where $\{X_0, X_1, \ldots\}$ have a distribution not depending on n. Stochastic asymptotic expansions are discussed in Cox and Reid [12] and in Barndorff-Nielsen and Cox [5, Ch. 3]. As an example, Cox and Reid [12] show that if Y_n follows a chi-squared distribution with n degrees of freedom, then

$$(Y_n - n)/\sqrt{2n} = X_0 + \frac{1}{n^{1/2}} \frac{\sqrt{2}}{3}(X_0^2 - 1) + O_p(n^{-1}),$$

where X_0 is a standard normal random variable. The relationship between stochastic asymptotic expansions and expansions for the corresponding distribution functions is discussed in Cox and Reid [12]. Expansions similar to (9) where the distributions of $X_0, X_1, \ldots$ are only asymptotically free of n are very useful in computing asymptotic properties of likelihood based statistics.

A expansion closely related to (9) but usually derived in the context of the Edgeworth expansion for the distribution function of the sample mean is the Cornish-Fisher expansion* for the quantile of the distribution function. As described in Serfling [35], an expansion for the value z_α satisfying $F_n(s_\alpha) = 1 - \alpha$ can be obtained by a reversion of the Edgeworth expansion. The result is

$$s_\alpha = z_\alpha + \frac{1}{6\sqrt{n}}(z_\alpha^2 - 1)\lambda_3 + \frac{1}{24n}(z_\alpha^3 - 3z_\alpha)\lambda_4 - \frac{1}{36n}(2z_\alpha^3 - 5z_\alpha)\lambda_3^2 + O(n^{-3/2}),$$

where z_α satisfies $\Phi(z_\alpha) = 1 - \alpha$. Other asymptotic expansions for $F_n(s)$ lead to alternate expansions for s_α, and in particular the r^* approximation given at (8) can be derived from the saddlepoint expansion for the distribution function of $\bar{X}_n$.

Applications to parametric inference

In recent years there has been considerable development of statistical theory based on the use of higher order approximations derived from asymptotic expansions. Likelihood-based inference, or inference from parametric models, has developed particularly rapidly, although the approximations are very useful in other contexts as well. Some examples of this will now be sketched.

Suppose that $X = (X_1, \ldots, X_n)$ is a sample from a parametric model that is an exponential family *, i.e. is of the form

$$f(x; \theta) = \exp\{\theta^T x - b(\theta) - d(x)\} \tag{10}$$

where X and θ take values in $\mathbb{R}^m$, say. The minimal sufficient statistic is $S = s(X) = \sum X_i$, and the maximum likelihood estimator of θ is a one-to-one function of S: $c'(\hat{\theta}) = S/n$. Thus the saddlepoint approximation for S given in (7) can be used to give an approximation to the density for $\hat{\theta}$, which takes the form

$$f_n(\hat{\theta};\theta) \doteq c|b''(\hat{\theta})|^{1/2}\exp\{(\theta-\hat{\theta})'s - nb(\theta) + nb(\hat{\theta})\}.$$

If we denote the log-likelihood function for θ based on x by $\ell(\theta;x)$, and the observed Fisher information function $-\partial^2\ell(\theta)/\partial\theta\theta^T$ by $j(\theta)$ we can re-express this approximation as

$$f_n(\hat{\theta};\theta) \doteq c|j(\hat{\theta})|^{1/2}\exp\{\ell(\theta;x) - \ell(\hat{\theta};x)\} \tag{11}$$

This approximation to the density for the maximum likelihood estimator is usually known as Barndorff-Nielsen's approximation, or, following Barndorff-Nielsen, the p^* approximation. Although it has been used here to illustrate the saddlepoint approximation in an exponential family, it turns out that approximation (11) is valid quite generally. This was exemplified and illustrated in Barndorff-Nielsen [2,3] and several subsequent papers. A review of the saddlepoint approximation and the literature on the p^* formula through 1987 is given in Reid [33]. A general proof and discussion of the interpretation of the p^* formula is given in Skovgaard [38]. Chs. 6 and 7 of Barndorff-Nielsen and Cox [6] provide an extensive discussion of the p^* formula and its applications in parametric inference. The validity of (11) in more general models requires the existence of a one-to-one transformation from the minimal sufficient statistic to $(\hat{\theta}, a)$, where a is a complementary statistic with a distribution either exactly or approximately (in a specific sense) free of θ; such statistics are called exact or approximate ancillary* statistics. The right hand side of (11) then approximates the conditional distribution of $\hat{\theta}$, given a.

An illustration of the cumulative distribution function approximation (7) in the exponential family is also instructive. Suppose in (10) that $m = 1$. Then (7) provides an approximation to the distribution function for the maximum likelihood estimate which is simply

$$F_n(\hat{\theta};\theta) \doteq \Phi(r) + \phi(r)\left(\frac{1}{r} - \frac{1}{q}\right) \doteq \Phi\{r - r^{-1}\log(r/q)\} \doteq \Phi(r^*) \tag{12}$$

where

$$
\begin{aligned}
r &= \operatorname{sign}(\hat{\theta}-\theta)[2\{\ell(\hat{\theta})-\ell(\theta)\}] \\
q &= (\hat{\theta}-\theta)|j(\hat{\theta})|^{1/2}
\end{aligned}
$$

are the signed square root of the log-likelihood ratio statistic, and the standardized maximum likelihood estimator, respectively, and $r^* = r + r^{-1}\log(q/r)$.

As with the p^* approximation, approximation (12) holds much more generally, with q replaced by a sometimes complicated statistic that depends on the underlying model and in particular on the exact or approximate ancillary statistic required for the validity of (11) in general models. Furthermore, (12) can be applied to marginal and conditional distributions for the maximum likelihood estimate of a parameter of interest, after nuisance parameters have been eliminated via a marginal or conditional likelihood. A recent accessible reference is Pierce and Peters [32]. The approximation due to Skovgaard [36] is an important ingredient in this development. The r^* approximation, which for general families is due to Barndorff-Nielsen [4], is discussed in Barndorff-Nielsen and Cox [6, Ch. 6].

As an illustration of stochastic asymptotic expansions in likelihood based inference, consider Taylor series expansion of the score equation $\ell'(\hat{\theta}) = 0$ (assuming for simplicity that this uniquely defines the maximum likelihood estimator):

$$0 = \ell'(\theta) + (\hat{\theta}-\theta)\ell''(\theta) + (1/2)(\hat{\theta}-\theta)^2\ell'''(\theta) + \dots .$$

Reversion of this expansion gives an expansion for maximum likelihood estimator that can be expressed as

$$\sqrt{n}(\hat{\theta}-\theta) = \frac{Z_1(\theta)}{i(\theta)} + \frac{1}{\sqrt{n}}\left[\frac{Z_2(\theta)Z_1(\theta)}{\{i(\theta)\}^2} + \frac{Z_1^3(\theta)\rho_3(\theta)}{2\{i(\theta)\}^3}\right] + O_p(n^{-1}), \tag{14}$$

where $Z_1 = n^{-1/2}\ell'(\theta)$ and $Z_2 = n^{-1/2}\{\ell''(\theta) - ni(\theta)\}$, $i(\theta) = n^{-1}E\{\ell'(\theta)\}^2$, $\rho_3(\theta) = n^{-1}E\{\ell'''(\theta)\}$. The random variables Z_1 and Z_2 are $O_p(1)$ and have mean zero. In expansions of this type it is much easier to keep track of the orders of various terms using these standardized variables Z; this notation is originally due to D.R. Cox, and is extensively used as well in McCullagh [30, Ch. 7].

Expansion (14) is a type of stochastic asymptotic expansion, although strictly speaking the distributions for Z_1, Z_2 are only asymptotically free of n. The leading term of (14) gives the usual asymptotic normal approximation for the maximum likelihood estimator, and the next order term is useful for deriving refinements of this. For example, it is readily verified that the expected value of $\hat{\theta}$ has the expansion

$$E(\hat{\theta}) = \theta + n^{-1}\{i'(\theta) + \rho_3(\theta)/2\}/\{i(\theta)\}^2 + O(n^{-2}),$$

and that $\mathrm{var}\,(\hat{\theta}) = \{ni(\theta)\}^{-1} + O(n^{-2})$.

The multivariate version of (14) is given in McCullagh [30, Ch. 7], as are extensions to the nuisance parameter case, and several illustrations of the use of these expansions. One particularly relevant application is the substitution of (14) into a Taylor series expansion of the log-likelihood ratio statistic $w(\theta) = 2\{\ell(\hat{\theta}) - \ell(\theta)\}$ to obtain an expansion of both the density and the expected value of $w(\theta)$. These expansions lead to the results

$$Ew(\theta) = m\{1 + b(\theta)/n + O(n^{-2})\}$$

and

$$w(\theta)/\{1 + b(\theta)/n\} = X_m^2\{1 + O(n^{-3/2})\} \tag{15}$$

where m is the dimension of θ and X_m^2 is a random variable following a χ^2 distribution on m degrees of freedom. The improvement of the approximation to the distribution of the log-likelihood ratio statistic given by (15) is called Bartlett correction, after Bartlett [7], where the correction was derived for a particular case, that of testing the equality of several normal variances. It is a multivariate analogue of the improvement of the normal approximation to the signed square root given in (7). Expansion (15) is originally due to Lawley [28]: for details of the derivation see McCullagh [30, Ch. 7], Barndorff-Nielsen and Cox [6, Ch. 6], Bickel and Ghosh [8], and DiCiccio and Martin [19].

Approximations using the Edgeworth and saddlepoint expansions are also useful for statistics that are not derived from a likelihood-based approach to inference. Edgeworth expansions for more general statistics are discussed in Serfling [35] and in considerable generality in Pfanzagl [31]. Skovgaard [38] considers formulations for the density of minimum contrast estimators that lead to the p^* approximation. Saddlepoint approximation to the density of M-estimators is discussed in Daniels [14] and Field and Ronchetti [23]. Application of the saddlepoint approximation to the bootstrap is introduced in Davison and Hinkley [17], and explored further in Daniels and Young [16], DiCiccio, Martin and Young [20], Wang [40] and Ronchetti and Welsh [34].

A somewhat different application of asymptotic expansions in parametric inference is the use of the technqiues outlined here to obtain asymptotic expansions for the efficiency of estimators, and the power function of test statistics. One purpose of this is to provide a means for choosing among various estimators or test statistics that have the same efficiency or power to a first order of asymptotic theory. Helpful surveys of these types of results are given in Skovgaard [37], Ghosh [24]: see also the entry Asymptotics, Higher Order*.

References

[1] Abramowitz, M. and Stegun, I.A. (1965). *Handbook of mathematical functions*. Dover, New York.

[2] Barndorff-Nielsen, O.E. (1980). Conditionality resolutions. *Biometrika* **67**, 293–310. (Introduction to the p^* formula and proof of exactness for transformation models.)

[3] Barndorff-Nielsen, O.E. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika* **70**, 343–365. (Easier to read than [2], includes derivation of p^* from the saddlepoint approximation.)

[4] Barndorff-Nielsen, O.E. (1986) Inference on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika* **73**, 307–322. (First paper on the use of r^* approximation.)

[5] Barndorff-Nielsen, O.E. and Cox, D.R. (1990). *Asymptotic techniques for use in statistics*. Chapman and Hall, London.(Presents detailed development of many asymptotic techniques that have proved useful for obtaining higher order approximations in statistics, and includes a great number of useful examples. A very valuable reference text for this area. This article has been very strongly influenced by this text.)

[6] Barndorff-Nielsen, O.E. and Cox, D.R. (1994). *Inference and Asymptotics*. Chapman

and Hall, London.(Follows on from [5], but emphasizes the role of higher order asymptotics in parametric inference, as outlined in the last section of this article.)

[7] Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proc. Roy. Soc.* A **160**, 268–282.

[8] Bickel, P. J. & Ghosh, J. K. (1990). A decomposition for the likelihood ratio statistic and Bartlett correction – a Bayesian argument *Ann. Statist.* **18**, 1070–1090.

[9] Bleistein, N. and Handelsman, R.A. (1975). *Asymptotic Expansions of Integrals.* Holt, Rinehart and Winston, New York.

[10] Cheah, P.K., Fraser, D.A.S., and Reid, N. (1993). Some alternatives to Edgeworth. *Canad. J. Statist.* **21**, 131–138.

[11] Courant, R. and Hilbert D. (1950). *Methods of Mathematical Physics.* J. Wiley & Sons, New York.

[12] Cox, D.R. and Reid, N. (1987). Approximations to non-central distributions. *Canad. J. Statist.* **15**, 105–114.

[13] Daniels, H.E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **25**, 631–650. (The first derivation of the saddlepoint approximation for the density of the sample mean in the statistical literature. Very clearly written and important for any study of saddlepoint expansions. A helpful, lengthier, treatment is given in [27].)

[14] Daniels, H.E. (1983). Saddlepoint approximations for estimating equations. *Biometrika* **70**, 89–96.

[15] Daniels, H.E. (1987). Tail probability approximations. *Int. Statist. Rev.* **55**, 37–48. (Derivation of distribution function approximations using the saddlepoint technique. Illustration of the Lugannani and Rice approximation and two closely related approximations. Very clearly written. A helpful, lengthier, treatment is given in [27].)

[16] Daniels, H.E. and Young, G.A. (1991). Saddlepoint approximation for the studentized mean, with an application to the bootstrap. *Biometrika* **78**, 169–79.

[17] Davison, A.C. and Hinkley, D.V. (1988). Saddlepoint approximations in resampling methods. *Biometrika* **75**, 417–31.

[18] DeBruijn, N.G. (1970). *Asymptotic Methods in Analysis.* 2nd ed. North-Holland,

Amsterdam. Reprinted by Dover, 1981.

[19] DiCiccio, T.J. and Martin, M.A. (1993) Simple modifications for signed roots of likelihood ratio statistics. *J. R. Statist. Soc.* B **55**, 305–316.

[20] DiCiccio, T.J., Martin, M.A., and Young, G.A. (1992). Fast and accurate approximation to double bootstrap confidence intervals. *Biometrika* **79**, 285–96.

[21] Easton, G.S. and Ronchetti, E. (1986). General saddlepoint approximations with applications to L-statistics. *J. Am. Statist. Assoc.* **81**, 420–430.

[22] Feller, W. (1970). *Introduction to Probability Theory and Applications*, Vol. II. J. Wiley & Sons, New York.

[23] Field, C.A. and Ronchetti, E. (1990). *Small Sample Asymptotics.* Institute of Mathematical Statistics, Hayward. (Emphasizes the use of exponential tilting in constructing approximations for statistics derived from sample means, with special emphasis on robust estimators.)

[24] Ghosh, J.K. (1994). *Higher Order Asymptotics.* Institute of Mathematical Statistics, Hayward.

[25] Jeffreys, H. (1962). *Asymptotic Approximations.* Oxford University Press.

[26] Jensen, J.L. (1986). Inference for the mean of a gamma distribution with unknown shape parameter. *Scand. J. Statist.* **13**, 135–151.

[27] Kolassa, J.E. (1994). *Series Approximation Methods in Statistics.* Springer-Verlag, New York.. (A very detailed derivation of saddlepoint and Edgeworth expansions that includes a lot of background material. Very helpful for understanding how higher order approximations are derived.)

[28] Lawley, D. (1956). A general method for approximating to the distribution of the likelihood ratio criterion. *Biometrika* **43**, 295–303.

[29] Lugannani, R. and Rice, S.O. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables. *Adv. Appl. Probab.* **12**, 475–490. (Derivation of the approximation given in equation (7).)

[30] McCullagh, P. (1987). *Tensor Methods in Statistics.* Chapman and Hall, London.(Introduces tensor calculus for use in multivariate statistical calculations. Derives multivariate Edgeworth and saddlepoint expansions, and considers application of asymptotic ex-

pansions and stochastic asymptotic expansions in parametric statistical inference.)

[31] Pfanzagl, J. (1985). *Asymptotic Expansions for General Statistical Models.* Springer-Verlag, New York.(Rigorous development of expansions for efficiency of statistical estimators and power of statistical tests. A comparison of this approach and the one emphasized in this article is given in [37].)

[32] Pierce, D.A. and Peters, D. (1992) Practical use of higher order asymptotics for multi-parameter exponential families (with discussion). *J. R. Statist. Soc.* B **54**, 701–38.

[33] Reid, N. (1988). Saddlepoint methods and statistical inference (with discussion). *Statist. Science* **3**, 213–238. (A review of the saddlepoint approximation and its relationship to the p^* formula, and the uses of the p^* formula in inference.)

[34] Ronchetti, E. and Welsh, A. (1994). Empirical saddlepoint approximations for multivariate M-estimators. *J. R. Statist. Soc.* B **56**, 313–326.

[35] Serfling, R.J. (1982). Asymptotic expansion. in *Encyclopedia of Statistical Sciences* **1**, 137–138 . J. Wiley & Sons, New York.

[36] Skovgaard, I.M. (1987). Saddlepoint expansions for conditional distributions. *J. Appl. Probab.* **24**, 875–87.

[37] Skovgaard, I.M. (1989). A review of higher order likelihood inference. *Bull. Inter. Statist. Inst.: proceedings of the 47th Session* **53**, 331–350.

[38] Skovgaard, I.M. (1990) On the density of minimum contrast estimators. *Ann. Statist.* **18**, 779–89. (A very helpful derivation of the p^* formula given in Section 3.)

[39] Wang, S. (1990a). Saddlepoint approximation for bivariate distributions. *J. Appl. Probab.* **27**, 586–97.

[40] Wang, S. (1990b). Saddlepoint approximations in resampling. *Ann. Inst. Stat. Math.* **42**, 115–31.

[41] Wang, S. (1992). General saddlepoint approximations in the bootstrap. *Stat. Probab. Letters* **13**, 61–66.

[42] Wood, A.T.A., Booth, J.G. and Butler, R.W. (1993). Saddlepoint approximations to the cdf of some statistics with non-normal limit distributions. *J. Am. Statist. Assoc.* **88**, 680–686.

(ASYMPTOTIC NORMALITY
ASYMPTOTICS, HIGHER ORDER
CORNISH-FISHER EXPANSION
EDGEWORTH EXPANSION
LAPLACE'S METHOD
SADDLEPOINT APPROXIMATIONS)

CITATIONS OF "ASYMPTOTIC EXPANSIONS" IN OTHER PARTS OF ESS

Nancy Reid