MONASH

MONASH UNIVERSITY

AUSTRALIA

ON THE APPROXIMATION OF SADDLEPOINT EXPANSIONS IN STATISTICS

Offer Lieberman

Working Paper No. 17/91

December 1991

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 959 5

ON THE APPROXIMATION OF SADDLEPOINT EXPANSIONS IN STATISTICS

Offer Lieberman

Working Paper No. 17/91

December 1991

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

ABSTRACT

The exact saddlepoint approximation as developed by Daniels [4], is an extremely accurate method for approximating probability distributions. Recent applications of the technique to densities of statistics of interest have been hindered by the requirement of explicit knowledge of the cumulant generating function, and the need to obtain an analytic solution to the saddlepoint defining equation. In this paper we show the conditions under which any approximation to the saddlepoint is justified, and suggest a solution that does not affect the usual merits of the exact expansion. We illustrate with an approximate saddlepoint expansion of the Durbin-Watson test statistic.

* I am grateful to Grant Hillier for suggesting the problem and for numerous helpful discussions. I am also thankful to Randolph Tan and Merran Evans for their comments, and to John Small for his assistance with the data.

1. INTRODUCTION

In recent years there has been renewed interest in saddlepoint approximations in statistics. Given that the limiting distributions of many statistics are notoriously inaccurate in small samples, and that the corresponding exact densities are difficult to derive analytically, it seems that this approximation is endowed with a superior small sample accuracy over other comparable asymptotic expansions, see for example, Reid [17].

The Edgeworth expansion in particular, which only requires knowledge of the first few moments of the distribution and (hence) is relatively easy to implement, is renowned for poor performance in the tail area of the distribution, where the Hermite polynomials may be unbounded, and can in fact assign negative values to the approximate density. Its truncation at the r^{th} term gives an absolute error of $O(n^{-r/2})$.

On the other hand, the saddlepoint expansion (which was first introduced by Daniels [4], and then revived by Barndorff-Nielsen & Cox [2]), appears to perform well in the tail area and has a relative error of $O(n^{-1})$, so that convergence to the limiting distribution is quicker than that of the Edgeworth expansion, and small sample performance is superior to the latter approximation.

The relative merits of the saddlepoint approximation raised hopes for good small sample asymptotic approximations, and consequently a considerable body of literature has been dedicated to the issue by, among others, Phillips [14], Reid [17], Barndorff-Nielsen [1], Barndorff-Nielsen & Cox [3], McCullagh [13]. However, exact applications of the approximation appear to have been at a stand-still for a while, possibly due to its major drawback compared with the Edgeworth expansion, namely, the requirement of explicit knowledge of the cumulant generating function.

In practice, even full knowledge of the latter does not guarantee an analytical solution to the saddlepoint defining equation, and instead, numerical methods if applied, are often computationally expensive, and may inflict severe violations to the usual features of the expansion. The motivation for this paper is to provide a theoretically justified approach to an approximation of the saddlepoint expansion which leaves the qualitative merits of the exact expansion almost unaffected, and further, saves computing time considerably. In section 2 we introduce notation and the approximate saddlepoint expansion. A practical example via the Durbin-Watson test statistic follows in section 3, and section 4 concludes the paper.

2. THE APPROXIMATE SADDLEPOINT EXPANSION

Let $X_1,\ldots,X_n$ be independent, identically distributed random variables, each with a density function $f_{\mathbf{x}}(x)$, moment generating function $M_{\mathbf{x}}(\lambda)$, and cumulant generating function $K_{\mathbf{x}}(\lambda)=\ell n M_{\mathbf{x}}(\lambda)$. Assume that $M_{\mathbf{x}}(\lambda)$ exists in some non-vanishing interval that contains the origin. Also assume that the first two moments of X, μ and σ^2, are fixed, and that higher order cumulants are bounded. Let $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$. Following the approach taken by Barndorff-Nielsen and Cox [2], we embed the density of $\bar{X}$, $f_{\bar{\mathbf{x}}}(\bar{x})$, in the exponential family

$$f_{\bar{\mathbf{x}}}(\bar{x}:\lambda) = \exp\{n[\lambda\bar{x}-k_{\mathbf{x}}(\lambda)]\}\ f_{\bar{\mathbf{x}}}(\bar{x}), \qquad (1)$$

where $\bar{X}$ denotes the random variable defined above, and $\bar{x}$ is a given value of $\bar{X}$ at which the density is evaluated. We now proceed by fitting an Edgeworth expansion to the conjugate density, $f_{\bar{x}}(\bar{x}:\lambda)$, such that

$$f_{\bar{x}}(\bar{x}:\lambda) = [k''_{\mathbf{x}}(\lambda)/n]^{-1/2}\phi(z)\left\{1 + \frac{\rho_3(\lambda)H_3(z)}{6\sqrt{n}} + \frac{1}{n}\left[\frac{\rho_4(\lambda)H_4(z)}{24} + \frac{\rho_3^2(\lambda)H_6(z)}{72}\right] + O(n^{3/2})\right\}, \tag{2}$$

where

$$z = \frac{\bar{x}-E_\lambda(\bar{\mathbf{x}})}{[\mathrm{var}_\lambda(\bar{\mathbf{x}})]^{1/2}} = \sqrt{n}\,\frac{\bar{x}-k'_{\mathbf{x}}(\lambda)}{[k''_{\mathbf{x}}(\lambda)]^{1/2}}, \tag{3}$$

$\phi(z)$ is the standard normal PDF, $\rho_r(\lambda) = k_{\mathbf{x}}^{(r)}(\lambda)/(k''_{\mathbf{x}}(\lambda))^{r/2}$,

$r \geq 3$, and $H_j(z)$ is the j^{th} degree Hermite polynomial defined as

$$H_j(z) = (-1)^j \frac{\partial^j \phi(z)}{\partial z^j} \Big/ \phi(z).$$

In particular, the Hermite polynomials appearing in (2) are

$$\begin{aligned} H_3(z) &= z^3 - 3z \\ H_4(z) &= z^4 - 6z^2 + 3 \\ H_6(z) &= z^6 - 15z^4 + 45z^2 - 15. \end{aligned} \tag{4}$$

At this point, we choose $\lambda = \tilde{\lambda}$ such that

$$k'_{\mathbf{x}}(\tilde{\lambda})-\bar{x}=0, \tag{5}$$

which implies (from (3)) that $z = 0$, hence $H_3(z) = 0$. With this choice of $\tilde{\lambda}$, we substitute (2) into (1) and rearrange to get

$$f_{\bar{\mathbf{x}}}(\bar{x}) = \left[2\Pi k''_{\mathbf{x}}(\tilde{\lambda})/n\right]^{-1/2} \exp\left\{n[k_{\mathbf{x}}(\tilde{\lambda}) - \tilde{\lambda}\bar{x}]\right\}\left\{1 + \frac{1}{n}\left[\frac{\rho_4(\tilde{\lambda})}{8} - \frac{5\rho_3^2(\tilde{\lambda})}{24}\right] + O(n^{-2})\right\} \quad (6)$$

(6) is the saddlepoint approximation for the density of $\bar{X}$, which was first derived by Daniels [4], who applied the complex variable technique of steepest descents to the Fourier inversion formula. In the same paper, Daniels proved that the saddlepoint, $\tilde{\lambda}$, is real and unique.

In practice, even if known, the cgf is often algebraically complicated and the saddlepoint defining equation, (5), cannot be solved analytically for $\tilde{\lambda}$. Numerical methods such as Newton-Raphson, have been suggested. See for example Easton and Ronchetti [9], or DiCiccio *et al.* [6]. The problem is that there are almost always numerical errors in these procedures, and they can take considerable time to compute. Any approximation to $\tilde{\lambda}$, say by $\hat{\lambda}$, such that the two points are in the same neighbourhood, but still with $\tilde{\lambda} \neq \hat{\lambda}$, implies that $k'_{\mathbf{x}}(\hat{\lambda}) \neq \bar{x}$, hence z and $H_j(z)$, $j \geq 3$ and odd, will no longer vanish in the construction of (6). Unless $\hat{\lambda}$ is suitably chosen, the approximate saddlepoint expansion with $\hat{\lambda}$ replacing $\tilde{\lambda}$ will have an error of $O(n^{-1/2})$ rather than $O(n^{-1})$, and the leading term in (6) will become

$$\left[k''_{\mathbf{x}}(\hat{\lambda})/n\right]^{-1/2} \phi(z) \exp\left\{n\left[k_{\mathbf{x}}(\hat{\lambda}) - \hat{\lambda}\bar{x}\right]\right\} \quad (7)$$

These consequences can be circumvented by choosing λ such that $z = O(n^{-1/2})$, which implies that

$$H_j(z) = \begin{cases} O(n^{-1/2}) & \text{if } j \text{ is odd} \\ O(1) & \text{if } j \text{ is even.} \end{cases} \quad (8)$$

Although $H_3(z)$ is no longer zero, it is $O(n^{-1/2})$ and the correction factor involving it in (2) becomes $O(n^{-1})$. Further, in view of (8) and as $z \neq 0$, $H_4(z)$, $H_6(z)$ are no longer 3 and -15 respectively, but still $O(1)$. Lower order correction terms retain steps of $O(n^{-1})$. For a detailed discussion on the Hermite tensors, the reader is referred to McCullagh [13], ch.5.

From (3), the requirement of $z = O(n^{-1/2})$ is equivalent to choosing λ such that

$$\frac{\bar{x} - k_{\mathbf{x}}'(\lambda)}{[k_{\mathbf{x}}''(\lambda)]^{1/2}} = O(n^{-1}). \tag{9}$$

On expansion of the derivatives of the cgf about the origin, (9) becomes

$$\frac{\bar{x} - [\mu+\sigma^2\lambda+k_3\lambda^2/2+k_4\lambda^3/6+\ldots]}{[\sigma^2+k_3\lambda+k_4\lambda^2/2+\ldots]^{1/2}} . \tag{10}$$

Clearly, from the numerator of (10), the saddlepoint is a function of $\bar{x}-\mu$. Daniels [5] observed that for $\bar{X}$ in the range $\bar{x}-\mu=O(n^{-1/2})$, where most statistical applications are, $\tilde{\lambda}=O(n^{-1/2})$. For fixed $\bar{x}-\mu$, it is $O(1)$. Reid [17], section 2, further discusses the saddlepoint as a function of $\bar{x}-\mu$. The order of magnitude of the approximate saddlepoint is in effect the key to its justification.

Now let $\lambda = \hat{\lambda} = \dfrac{\bar{x}-\mu}{\sigma^2}$. This choice is not arbitrary as will be shown a little later on. Substitution of $\hat{\lambda}$ into (10) gives

$$\frac{\dfrac{k_3}{2}\left(\dfrac{\bar{x}-\mu}{\sigma^2}\right)^2 + \dfrac{k_4}{6}\left(\dfrac{\bar{x}-\mu}{\sigma^2}\right)^3+\ldots}{\left[\sigma^2 + k_3\left(\dfrac{\bar{x}-\mu}{\sigma^2}\right) + \dfrac{k_4}{2}\left(\dfrac{\bar{x}-\mu}{\sigma^2}\right)^2+\ldots\right]^{1/2}} . \tag{11}$$

When $\bar{X}$ is in the range $\bar{x}-\mu=O(n^{-1/2})$, the highest order term in the numerator is $\frac{k_3}{2}\left(\frac{\bar{x}-\mu}{\sigma^2}\right)^2 = O(n^{-1})$. The order of magnitude of successive terms decrease in powers of $n^{1/2}$. In the denominator, the highest order term is $\sigma^2 = O(1)$, and so the whole expression in (11) is $O(n^{-1})$. Hence (9) is $O(n^{-1})$ and $z = O(n^{-1/2})$, as required. In the far tails of the distribution, $\frac{\bar{x}-\mu}{\sigma^2} = O(1)$, hence (11) is $O(1)$. Unless the saddlepoint defining equation is solved exactly, the approximation will be justified only for the range $\bar{x}-\mu = O(n^{-1/2})$.

It has already been mentioned that the choice of $\hat{\lambda} = \frac{\bar{x}-\mu}{\sigma^2}$ is not coincidential. As $\tilde{\lambda}$ locates the minimum of $f(\lambda) = k_x(\lambda) - \lambda\bar{x}$, the Newton-Raphson algorithm

$$\lambda_{n+1} = \lambda_n - \frac{f'(\lambda_n)}{f''(\lambda_n)} ,$$

with $\lambda_0 = 0$, gives $\lambda_1 = \hat{\lambda} = \frac{\bar{x}-\mu}{\sigma^2}$. The applied researcher who has implemented this procedure, may feel somewhat relieved to know that the end of the first iteration is sufficient for z to be $O(n^{-1/2})$, but only for $\bar{X}$ in the range $\bar{x}-\mu = O(n^{-1/2})$. For the normal distribution, $k_x(\lambda) = \mu\lambda + \sigma^2\lambda^2/2$, and so $\hat{\lambda}$ is in fact the exact saddlepoint.

In general, this choice of $\hat{\lambda}$ is not expected to yield reliable results as soon as $\bar{x}$ moves away from the mean (at $\bar{x} = \mu$, $\tilde{\lambda} = \hat{\lambda} = 0$), as information on higher order cumulants is ignored. Instead, an easily implemented choice of an approximate saddlepoint, starts by a series reversion of the saddlepoint defining equation, namely

$$\tilde{\lambda} = \frac{\bar{x}-\mu}{\sigma^2} - \frac{k_3}{2\sigma^2}\left(\frac{\bar{x}-\mu}{\sigma^2}\right)^2 + \left[\frac{k_3^2}{2\sigma^4} - \frac{k_4}{6\sigma^2}\right]\left(\frac{\bar{x}-\mu}{\sigma^2}\right)^3 + \dots \qquad (12)$$

For then, approximation of $\tilde{\lambda}$ by the first three terms in this series, and substitution into (10) gives

$$\frac{\left\{\frac{5k_3}{4\sigma^4}\left[\frac{k_3^2}{2\sigma^2}-\frac{k_4}{3}\right]+\frac{k_5}{24\sigma^2}\right\}\left(\frac{\bar{x}-\mu}{\sigma^2}\right)^4+O\left(\frac{\bar{x}-\mu}{\sigma^2}\right)^5}{(\sigma^2+\ldots)^{1/2}} . \quad (13)$$

As the highest order term in the numerator of (13) involves $\left(\frac{\bar{x}-\mu}{\sigma^2}\right)^4$, this expression is $O(n^{-2})$ for $\bar{x}-\mu = O(n^{-1/2})$, implying that z, $H_3(z)$ are $O(n^{-3/2})$. The saddlepoint expansion (6) becomes

$$f_{\bar{\mathbf{x}}}(\bar{x}) = \left[k''_{\mathbf{x}}(\lambda^*)/n\right]^{-1/2} \phi(z) \exp\left\{n\left[k_{\mathbf{x}}(\lambda^*) - \lambda^*\bar{x}\right]\right\}$$

$$\left\{1 + \frac{1}{n}\left[\frac{\rho_4(\lambda^*)H_4(z)}{24} + \frac{\rho_3^2(\lambda^*)H_6(z)}{72}\right] + O(n^{-2})\right\} , \quad (14)$$

where λ^* is the RHS of (12), truncated after the third term, and the $O(n^{-2})$ term in (14) incorporates in it the factor $\rho_3(\lambda^*)H_3(z)/6\sqrt{n}$. This choice of $\lambda = \lambda^*$ serves as an excellent approximation, as will be demonstrated in the next section, and the computation time of the saddlepoint expansion with it is reduced dramatically, as the line search for $\tilde{\lambda}$ is avoided. To obtain the approximate tail area probability, $Q_n(\bar{x}) = \int_{\bar{x}}^{\max \bar{x}} f_{\bar{\mathbf{x}}}(\bar{x})d\bar{x}$, we can either integrate (14) numerically, or much more elegantly follow one of the approaches discussed in Daniels [5]. For $\lambda = \lambda^*$ and $\bar{X}$ in the range $\bar{x}-\mu = O(n^{-1/2})$, an expansion of the exponent in the inversion formula

$$Q_n(\bar{x}) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{n\{k_{\mathbf{x}}(\lambda) - \lambda\bar{x}\}} \frac{d\lambda}{\lambda} \qquad (c > 0), \quad (15)$$

about λ^* and allowing for $k'_{\mathbf{x}}(\lambda^*) \neq \bar{x}$, leads to the approximate tail area probability

$$Q_n(\bar{x}) = H(-V^*) + e^{n[k_x(\lambda^*)-\lambda^*\bar{x}]} + 1/2(v^*)^2 \Bigg\{ \Big[H(v^*) - \Phi(v^*) \Big]$$

$$\left[1 - \frac{\rho_3(\lambda^*)(v^*)^3}{6n^{1/2}} + \frac{1}{n}\left(\frac{\rho_4(\lambda^*)(v^*)^4}{24} + \frac{\rho_3^2(\lambda^*)(v^*)^6}{72} \right) \right]$$

$$+ \phi(v^*) \left[\frac{\rho_3(\lambda^*)}{6n^{1/2}} \left((v^*)^2 - 1 \right) - \frac{1}{n} \left[\frac{\rho_4(\lambda^*)}{24} \left((v^*)^3 - v^* \right) \right. \right.$$

$$\left.\left. + \frac{\rho_3^2(\lambda^*)}{72} \left((v^*)^5 - (v^*)^3 + 3v^* \right) \right] \right] \Bigg\} \left\{ 1 + 0(n^{-3/2}) \right\} , \quad (16)$$

where $v^* = \lambda^*(nK''_x(\lambda^*))^{1/2}$, $H(v^*) = 0, 1/2, 1$ when $v^* < 0, = 0, > 0$ respectively, and $\Phi(v^*)$ is the standard normal CDF. See appendix 2 for rationale.

3. PRACTICAL EXAMPLE: THE DURBIN-WATSON TEST STATISTIC

Consider the classical linear regression model, $Y = X\beta + u$, where Y is $(T \times 1)$, X is $(T \times K)$, non stochastic and of rank $K(<T)$, β is $(K \times 1)$ and fixed, and u is a $(T \times 1)$ disturbance vector. The Durbin-Watson test statistic for a first order autocorrelation in the disturbances is given by

$$d = \frac{u'MAMu}{u'Mu} , \quad (17)$$

where $M = I_T - X(X'X)^{-1}X'$, and A is the first differencing matrix

$$A = \begin{bmatrix} 1 & -1 & \cdots & & 0 \\ -1 & 2 & & & \\ & 2 & & & \\ & & \ddots & & \\ & -1 & & 2 & -1 \\ 0 & & & -1 & 1 \end{bmatrix}$$

There are a number of other statistics of the general form of d and so the discussion is not restricted to the Durbin-Watson test statistic alone. Given that the exact density of (17) is unknown, most applied econometricians either still use the traditional bounds test due to Durbin and Watson [7], or calculate exact CDF values by the Imhof [11] routine. Henshaw [10] derived the Beta approximation for the distribution of this statistic, and a recent algorithm which avoids the eigenvalue approach has been proposed by Shively *et al.* [15]. Except for the bounds test, all these methods require numerical integration.

Under the Null hypothesis of no autocorrelation, $u \sim N(0, \sigma^2 I_T)$. The cgf of d is then

$$K_d(\lambda) = \ell n \sum_{j=0}^{\infty} \frac{\lambda^j (1/2)_j}{j!(m/2)_j} C_{[j]}(L) , \qquad (18)$$

where $m = T-K$, $(a)_k \equiv a(a+1)\cdots(a+k-1)$, $L=MA$, and $C_{[j]}(L)$ is the zonal polynomial of L, see Smith [16] equation (3.7). A tabulation of zonal polynomials up to sixth degree is given by James [12].

From (5) and (18), the saddlepoint defining equation is

$$\frac{\sum_{j=1}^{\infty} \frac{\tilde{\lambda}^{j-1}(1/2)_j}{(j-1)!(m/2)_j} C_{[j]}(L)}{\sum_{j=0}^{\infty} \frac{\tilde{\lambda}^{j}(1/2)_j}{j!(m/2)_j} C_{[j]}(L)} = \bar{d} , \qquad (19)$$

where $\bar{d}$ is a given value of d. An analytical solution for $\tilde{\lambda}$ in (19) is evidently difficult. This is a typical situation of having to approximate $\tilde{\lambda}$ by one of the methods discussed above. In addition, we have the practical problem of (18) being an infinite power series, but this can be conveniently overcome by a quadruple approximation of $K_d(\lambda)$, so that

$$\hat{K}_d(\lambda) = \mu\lambda + \sigma^2\lambda^2/2 + K_3\lambda^3/6 + K_4\lambda^4/24. \qquad (20)$$

The error is uniformly $O(n^{-1})$, see Easton & Ronchetti [9]. The first four cumulants are readily derived from (18) and are reported in appendix 3. They agree with the first four moments appearing in Henshaw [10].

We compare three $Q_1(\bar{d})$ values in Table 1, appendix 1, where $Q_1(\bar{d}) = \int_{\bar{d}}^{\max a_i} f_d(\bar{d})\, d\bar{d}$ and $\max a_i$, being the largest eigenvalue of L, is the upper bound of $\bar{d}$. The first is obtained by the Imhof [11] procedure and is regarded as "exact". The other two correspond to the approximate saddlepoint tail area probability (16), with λ^* as the approximate saddlepoint, and once again (16) with $\tilde{\lambda}_0$ as the saddlepoint, where $\tilde{\lambda}_0$ is obtained by the Newton-Raphson algorithm, satisfying $\hat{K}'_d(\tilde{\lambda}_0) - \bar{d} < 10^{-6}$.

Strictly speaking, even if $\hat{K}'_d(\tilde{\lambda}_0) = \bar{d}$, $\tilde{\lambda}_0$ is only a true saddlepoint if K_r, $r \geq 5$, are all zero. The numerical results will suggest that the advantage in the line search over a direct application of (16) with λ^* in terms of additional accuracy is minimal.

A range of data sets, covering a variety of econometric scenarios, have been included in this study to account for the well known dependence of the cumulants of d on the X matrices. The regressors in each data set are:

Watson: $(\tilde{a}_2+\tilde{a}_{T-1})/\sqrt{2}$ and $(\tilde{a}_3+\tilde{a}_{T-2})/\sqrt{2}$, where $\tilde{a}_i$ are the non-zero eigenvectors of the A matrix, $i=1,\dots,T-1$.

Spirit: Log real income per head and log relative price of spirits in the U.K., commencing 1870. As in Durbin & Watson [8].

POP: Household population, Households, and Household headship ratios, all in 1966 and 1971, in Australia.

Trend: A linear time trend.

An intercept was included in all the above. The author has also tried other data sets which have been used in previous studies of the Durbin-Watson test statistic, but they all exhibit very little variability in the cumulants and in the $Q_1(\bar{d})$ values. The X matrices selected for demonstration here, are those with the greatest such variability.

The main results are summarised in Table 1. Table 2 provides the first four cumulants of d for the different X matrices. A few issues are to be attended. First, although only an approximate cgf has been used, both saddlepoint approximations are generally accurate to the third decimal place. The comparative advantage of the approximation with $\tilde{\lambda}_0$ over the one with λ^* is minimal, whereas the extra CPU time designated for the line search for the former is substantial. Second, accuracy increases with T, and when considering the fact that the Imhof routine requires the calculation of the eigenvalues of a (T×T) matrix, as well as a costly numerical integration, the comparative speed of the approximate saddlepoint expansion over Imhof increases with T very rapidly. Third, additional accuracy can be obtained by normalization of either approximations, but this requires knowledge of max a_i, as well as numerical integration. The costs have already been mentioned.

When T is small, both approximations can break down in the tails. It reflects through $\hat{K}''_d(\tilde{\lambda}_0)$ or $\hat{K}''_d(\lambda^*)$ becoming negative. The true cgf though, $K_d(\lambda)$, is convex. See for example McCullagh's [13] (section 6.2) treatment of the Legendre transformation of $K_{\mathbf{x}}(\lambda)$, or Daniels [4] section 6. It suggests that higher order cumulants than 4 are significantly different from zero, at least when T is small. The conventional 95% level for testing does not pose any difficulties, even for sample sizes smaller than 20. On the pure basis of simplicity and computational advantage then, the approximate saddlepoint tail area probability which does not require eigenvalues, numerical integration and line searches, has very much to commend it.

4. CONCLUSION

Costly numerical methods are often required to solve the saddlepoint defining equation. Unless the error in the approximation to $\tilde{\lambda}$ is such that $z = O(n^{-1/2})$, $r \geq 1$, none of the advantages of the saddlepoint expansion over the Edgeworth expansion is retained. The approximate saddlepoint, λ^*, only requires knowledge of the first four cumulants of the distribution, not the form of the cgf. For $\bar{X}$ in the range $\bar{x}-\mu = O(n^{-1/2})$, where most of the statistical interest lies, the approximate saddlepoint expansion with λ^* is theoretically justified, accurate, and very quick to compute.

REFERENCES

1. Barndorff-Nielsen, O.E., On a formula for the distribution of the Maximum Likelihood estimator. Biometrika 70 (1983): 343-365.

2. Barndorff-Nielsen, O.E. & D.R. Cox, Edgeworth and Saddlepoint approximations with statistical applications (with discussion). Journal of the Royal Statistical Society Series B 41 (1979): 279-312.

3. Barndorff-Nielsen, O.E. & D.R. Cox, Asymptotic techniques for use in Statistics. New York: Chapman and Hall LTD, 1989.

4. Daniels, H.E. Saddlepoint approximations in statistics. Annals of Mathematical Statistics, 25 (1954): 635-650.

5. Daniels, H.E., Tail probability approximations. International Statistical Review, 55 (1987): 37-48.

6. Di Ciccio, T.J., Martin, M.A. & G.A. Young. Fast and accurate double Bootstrap confidence intervals. Technical report No. 369, Statistics Department, Standord University.

7. Durbin, J. & G.S. Watson, Testing for serial correlation in least Squares Regression, I. Biometrika 37 (1950): 409-428.

8. Durbin, J. & G.S. Watson, Testing for serial correlation in least squares regression, II, Biometrika, 38 (1951): 159-178.

9. Easton, G.S. & E. Ronchetti, General Saddlepoint approximations with applications to L Statistics. Journal of the American Statistical Association 81 (1986): 420-430.

10. Henshaw, R.C., Testing single-equation least squares regression models for autocorrelated disturbances. Econometrica 34 (1966): 646-660.

11. Imhof, P.J., Computing the distribution of quadratic forms in normal variables. Biometrika 48 (1961): 419-426.

12. James, A.T., Distributions of matrix variates and Latent roots derived from normal samples. Annals of Mathematical Statistics 35 (1964): 475-501.

13. McCullagh, P., Tensor methods in statistics. London: Chapman and Hall LTD, 1987.

14. Phillips, P.C.B., Edgeworth and Saddlepoint approximations in first order non-circular autoregression. Biometrika 65 (1978): 91-98.

15. Shively, T.S., Ansley, C.F. & R. Kohn. Fast evaluation of the distribution of the Durbin-Watson and other invariant test statistics in Time Series Regression. Journal of the American Statistical Association 85 (1990), No. 411: 676-685.

16. Smith, M.D., On the expectation of a ratio of quadratic forms in normal variables. Journal of Multivariate Analysis 31 (1989): 244-257.

17. Reid, N., Saddlepoint methods and statistical inference. Statistical Science Vol. 3, No. 2 (1988): 213-238.

APPENDIX 1

Table 1: Tail area probabilities

		T=40			T=20		
Data	Imhof	$\bar{d}$	SP	APSP	$\bar{d}$	SP	APSP
Watson	.9990	1.3353	.9899	.9897	1.1722	.9899	.9896
	.9500	1.5240	.9498	.9496	1.4054	.9500	.9496
	.9000	1.6274	.8998	.8997	1.5346	.8998	.8996
Spirit	.9900	1.3911	.9898	.9896	1.2045	.9944	.9903
	.9500	1.5930	.9498	.9497	1.4822	.9499	.9500
	.9000	1.7038	.9000	.8999	1.6376	.9005	.9008
Pop	.9900	1.3854	.9898	.9897			
	.9500	1.5926	.9499	.9499	1.3978	.9515	.9480
	.9000	1.7059	.9001	.9001	1.5456	.8989	.8983
Trend	.9900	1.3440	.9898	.9896			
	.9500	1.5443	.9497	.9494	1.4104	.9490	.9485
	.9000	1.6545	.8998	.8997	1.5598	.8995	.8992

$\bar{d}$: DW values.
SP: "exact saddlepoint" (Newton-Raphson).
APSP: Approximate saddlepoint.

Table 2: DW cumulants

Data	T	k1	k2	k3	k4
Watson	40	2.0000	.0835	-4.0E-09	-.0007
	20	2.0000	.1298	-6.0E-10	-.0022
spirit	40	2.1014	.0942	-.0008	-.0012
	20	2.1938	.1798	-.0071	-.0075
pop	40	2.1101	.0971	-.0014	-.0013
	20	2.2511	.1860	-.0119	-.0079
trend	40	2.0524	.0946	-.0004	-.0012
	20	2.1095	.1772	-.0029	-.0078

APPENDIX 2

The tail area probability is given by

$$Q_n(\bar{x}) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{n[k_x(\lambda) - \lambda\bar{x}]}\,\frac{d\lambda}{\lambda} \quad (c > 0) \qquad (2.1)$$

Using Daniels' [5] approach, expansion of the exponent in (2.1) about λ^*, and allowing for $k_x'(\lambda^*) \neq \bar{x}$, we have

$$Q_n(\bar{x}) = \frac{1}{2\pi i}\, e^{n[k_x(\lambda^*) - \lambda^*\bar{x}]}\int_{c-i\infty}^{c+i\infty} e^{1/2nk_x''(\lambda^*)(\lambda-\lambda^*)^2}\exp\Big\{n\Big[(k_x'(\lambda^*)-\bar{x})$$

$$\times (\lambda-\lambda^*) + \frac{1}{6}k_x^{(3)}(\lambda^*)(\lambda-\lambda^*)^3 + \frac{1}{24}k_x^{(4)}(\lambda^*)(\lambda-\lambda^*)^4 + \ldots\Big]\,\frac{d\lambda}{\lambda}$$

$$= \frac{1}{2\pi i}\, e^{n[k_x(\lambda^*) - \lambda^*\bar{x}]}\int_{c-i\infty}^{c+i\infty} e^{1/2nk_x''(\lambda^*)(\lambda-\lambda^*)^2}\Big\{1 + n\Big[(k_x'(\lambda^*)-\bar{x})$$

$$\times (\lambda-\lambda^*) + \frac{1}{6}k_x^{(3)}(\lambda^*)(\lambda-\lambda^*)^3 + \frac{1}{24}k_x^{(4)}(\lambda^*)(\lambda-\lambda^*)^4 + \ldots\Big]$$

$$+ \frac{n^2}{2}\Big[\Big(k_x'(\lambda^*) - \bar{x}\Big)^2(\lambda-\lambda^*)^2 + \frac{1}{36}\Big(k_x^{(3)}(\lambda^*)(\lambda-\lambda^*)^3\Big)^2$$

$$+ \frac{1}{3}\Big(k_x'(\lambda^*) - \bar{x}\Big)k_x^{(3)}(\lambda^*)(\lambda-\lambda^*)^4 + \frac{1}{12}(k_x'(\lambda^*) - \bar{x})$$

$$\times k_x^{(4)}(\lambda^*)(\lambda-\lambda^*)^5 + \ldots\Big] + \frac{n^3}{6}\Big(k_x'(\lambda^*) - \bar{x}\Big)^3\Big(\lambda-\lambda^*\Big)^3 + \ldots\Big\}\,\frac{d\lambda}{\lambda}\,.$$

(2.2)

Now, write $V = \lambda\Big(nk_x''(\lambda^*)\Big)^{1/2}$, $V^* = \lambda^*\Big(nk_x''(\lambda^*)\Big)^{1/2}$, then (2.2) becomes

$$e^{n[k_x(\lambda^*)-\lambda^*\bar{x}] + 1/2\,(V^*)^2}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{1/2\,V^2-VV^*}\left\{1 + n\left[\frac{k'_x(\lambda^*)-\bar{x}}{(nk''_x(\lambda^*))^{1/2}}\right.\right.$$

$$\times\ (V - V^*) + \frac{\rho_3(\lambda^*)}{6n^{3/2}}(V - V^*)^3 + \frac{\rho_4(\lambda^*)}{24n^2}(V - V^*)^4 + \ldots\Bigg] + \frac{n^2}{2}$$

$$\times\left[\frac{(k'_x(\lambda^*)-\bar{x})^2}{nk''_x(\lambda^*)}(V - V^*)^2 + \frac{\rho_3^2(\lambda^*)}{36n^3}(V - V^*)^6 + \frac{\rho_3(\lambda^*)(k'_x(\lambda^*)-\bar{x})}{3n^2(k''_x(\lambda^*))^{1/2}}(V - V^*)^4\right.$$

$$\left.\left. + \frac{\rho_4(\lambda^*)(k'_x(\lambda^*)-\bar{x})}{12n^{5/2}(k''_x(\lambda^*))^{1/2}}(V - V^*)^5 + \ldots\right] + \frac{n^{3/2}(k'_x(\lambda^*)-\bar{x})^3}{(k''_x(\lambda^*))^{3/2}}(V - V^*)^3 + \ldots\right\}\frac{dV}{V}$$

$$(2.3)$$

From (13), when $\bar{X}$ is in the range $\bar{x}-\mu = 0(n^{-1/2})$,

$\dfrac{k'_x(\lambda^*)-\bar{x}}{(k''_x(\lambda^*))^{1/2}} = 0(n^{-2})$. Rearranging (2.3) in decreasing powers of n, we then have

$$e^{n[k_x(\lambda^*)-\lambda^*\bar{x}] + 1/2\,(V^*)^2}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{1/2\,V^2-VV^*}\left\{1 + \frac{\rho_3(\lambda^*)}{6n^{1/2}}(V - V^*)^3\right.$$

$$\left. + \frac{1}{n}\left[\frac{\rho_4(\lambda^*)}{24}(V - V^*)^4 + \frac{\rho_3^2(\lambda^*)}{72}(V - V^*)^6\right] + \ldots\right\}\frac{dV}{V}. \qquad (2.4)$$

Except for an obvious change in notation, and for λ^* replacing the true saddlepoint, $\tilde{\lambda}$, (2.4) is identical to equation (3.7) in Daniels [5]. Hence, from equations (3.11) and (5.3) of that paper, the desired result follows.

APPENDIX 3

Let $L = MA$, and denote the trace of a matrix by tr. Then under the Null hypothesis of no autocorrelation, the first four cumulants of the Durbin-Watson test statistic are:

$$k_1 = \mu = \frac{trL}{m}$$

$$k_2 = \sigma^2 = 2\left\{\frac{mtrL^2 - (trL)^2}{m^2(m+2)}\right\}$$

$$k_3 = 8\left\{\frac{m^2trL^3 - 3mtrLtrL^2 + 2(trL)^3}{m^3(m+2)(m+4)}\right\}$$

$$k_4 = \Big\{12\left\{m^3[4trL^4 + (trL^2)^2] - 2m^2[8trLtrL^3 + trL^2(trL)^2]\right.$$

$$\left. + m[24trL^2(trL)^2 + (trL)^4] - 12(trL)^4\right\} / \; m^4(m+2)(m+4)(m+6)\Big\} - 3k_2^2.$$