TABLE A1
Approximations to bootstrap percentage points of $\bar{X} - \mu$ based on a sample of $n = 10$ numbers with 1 d.p.

Probability	"Exact"	Via saddlepoint approximation		Normal approx.
		$\tilde{G}_s$	$\tilde{G}_1$	
0.0001	−6.34	−6.32	−6.3130	−8.46
0.0005	−5.79	−5.79	−5.7842	−7.48
0.001	−5.55	−5.53	−5.4223	−7.03
0.005	−4.81	−4.81	−4.8051	−5.86
0.01	−4.42	−4.44	−4.4331	−5.29
0.05	−3.34	−3.33	−3.3296	−3.74
0.10	−2.69	−2.69	−2.6863	−2.91
0.20	−1.86	−1.86	−1.8556	−1.91
0.80	1.80	1.79	1.7956	1.91
0.90	2.87	2.85	2.8516	2.91
0.95	3.73	3.74	3.7480	3.74
0.99	5.47	5.47	5.4765	5.29
0.995	6.12	6.12	6.1212	5.86
0.999	7.52	7.46	7.4634	7.03
0.9995	8.19	7.98	7.9889	7.48
0.9999	9.33	9.11	9.1145	8.46

compared in Table A1, which includes also the "exact" results obtained by Monte Carlo with 50,000 simulated samples, as well as normal approximation results obtained with the correct mean and variance for D. Calculations using $\tilde{G}_2$ never differ from those using $\tilde{G}_1$ by more than 3×10^{-4}.

The success of the saddlepoint approximation in this example extends to many bootstrap and permutation distributions, so long as we restrict ourselves to problems involving sums as in Daniels's papers. We are unaware of comparable saddlepoint approximations for general nonlinear statistics. For practical purposes the key result would be the analog of Reid's (28) for statistics T_n of the form

$$T_n = \theta + n^{-1} \sum a_j(X_j) + n^{-2} \sum\sum b_{jk}(X_j, X_k),$$

because many statistics are very well approximated by such an expression. The relevant approximation would apply, for example, to the bootstrap distributions of studentized linear estimates.

ADDITIONAL REFERENCES

DAVISON, A. C. and HINKLEY, D. V. (1988). Saddlepoint approximations in resampling methods. *Biometrika*. To appear.

EFRON, B. and TIBSHIRANI, R. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy (with discussion). *Statist. Sci.* **1** 54–77.

Comment

Luke Tierney

Professor Reid's paper is an excellent review of the use of saddlepoint methods in statistics. In this comment I would merely like to expand briefly on Professor Reid's discussion of the relation between saddlepoint approximations for sampling distributions and approximations to posterior moments and marginal densities based on Laplace's method.

As described, for example, in De Bruijn (1970) the basic Laplace method and saddlepoint method both involve approximating a number a_n defined as $a_n = \int f_n(y)\, dy$ for some function f_n when n is large. For Laplace's method the function and its arguments are real, for the saddlepoint method the function is complex and the integral is over a path in the complex plane. In both cases it is assumed that the behavior of the integral is determined by the behavior of f_n in the neighborhood of a particular point y_n. For Laplace's method y_n is a local maximum, for the saddlepoint approximation it is a saddlepoint.

The statistical applications of the saddlepoint approximation discussed by Professor Reid add some new features. Rather than approximate a single number a_n these methods approximate a density function $g_n(x)$ given as $g_n(x) = \int f_n(x, y)\, dy$. As n increases the density $g_n(x)$ becomes concentrated about some point x_n at rate $n^{-1/2}$. The point x_n represents the mean or some other measure of the center of the distribution with density $g_n(x)$. For each value of the argument of the density $g_n(x)$ the saddlepoint approximation involves the determination of the corresponding saddlepoint $y_n = y_n(x)$ of the function $f_n(x, \cdot)$.

This use of the saddlepoint approximation closely resembles the use of Laplace's method for computing approximate marginal posterior densities as described in Leonard (1982) and Tierney and Kadane (1986). In

Luke Tierney is Associate Professor, School of Statistics, University of Minnesota, 206 Church Street S.E., Minneapolis, Minnesota 55455.

that case the function f_n represents the joint density, g_n the marginal density. The point x_n represents the center of the marginal distribution g_n and $y_n = y_n(x)$ is the conditional mode of the distribution of y given x. Of course this approximate marginalization does not require f_n to be a posterior distribution. Phillips (1983) uses this approach to approximate the marginal sampling distribution of various econometric estimators.

The asymptotic properties of the saddlepoint approximation to sampling distributions and the Laplace approximation to marginal densities are very similar. Both yield approximations that have errors uniformly of order $O(n^{-1})$ on fixed neighborhoods of x_n. Both benefit from numerical renormalization in the sense that the absolute errors of the approximations are of order $O(n^{-3/2})$ in $n^{-1/2}$-neighborhoods of x_n. Another interpretation of this result is that the shapes of the densities $g_n(\cdot)$ are approximated to order $O(n^{-3/2})$ by both methods.

The approximation of posterior expectations by Laplace's method is somewhat different. A single number is to be approximated rather than a function. Direct application of Laplace's method yields the maximum likelihood estimate or the posterior mode as an approximation to the posterior mean. The error of this approximation is of order $O(n^{-1})$. More accurate approximations with an error of order $O(n^{-2})$ can be obtained by using higher order terms, as described by Lindley (1980), or by using different centers for the expansions of numerator and denominator integrals, as described in Tierney and Kadane (1986) and Tierney, Kass and Kadane (1987).

The approximate predictive densities discussed in Leonard (1982), Tierney and Kadane (1986) and Davison (1986) fall somewhere between marginal density and moment approximations. Because a predictive density is a density, its approximation would appear to be more closely related to the approximation of marginal and sampling densities. On the other hand, the predictive density at a particular point can be expressed as a posterior expectation. The result of applying second order expectation approximation, as in (4.1) of Tierney and Kadane (1986), is an approximation to the predictive density with an error of order $O(n^{-2})$. The order of this error term will generally not be improved by numerical renormalization. As a result I feel that these approximate predictive densities are more closely related to approximate expectations than to approximate marginal densities.

I hope that these comments have added to the discussion in Section 6 of Professor Reid's excellent paper.

ADDITIONAL REFERENCES

LEONARD, T. (1982). Comment on "A simple predictive density function" by M. Lejeune and G. D. Faulkenberry. *J. Amer. Statist. Assoc.* **77** 657–658.

LINDLEY, D. V. (1980). Approximate Bayesian methods. In *Bayesian Statistics* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 223–237. University Press, Valencia.

PHILLIPS, P. C. B. (1983). Marginal densities of instrumental variables estimators in the general single equation case. Cowles Foundation Paper No. 582.

Comment

Robert E. Kass

The world of asymptotics is beautiful and mysterious. Witness Stirling's approximation, and recall the first time you needed to use it. What explains the odd yet simple formula, you may have asked, and more, How is it that with one correction term it already achieves 99.95% accuracy in approximating factorials as small as 2? Marvel at Figure 1. But recognize, each time we consider a sample of size n to be part of an infinite sequence of observations, we are faced with an irony: limits do not depend on the first n values, yet they are able to inform us about the behavior of the sample. Our finite world seems tied to asymptopia, but how?

Second-order asymptotic results continue to produce this feeling of awe and amazement in those who aren't yet familiar with them. Nancy Reid's review not only tells the saddlepoint story, it also nicely demonstrates the similarity of method in applications to maximum likelihood and conditional inference, robust estimation and Bayesian analysis. My comment consists of (i) a brief description of the relationship between Laplace's method and the saddlepoint

Robert E. Kass is Associate Professor, Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213.