Biometrika (1992), 79, 3, *pp.* 651-3
Printed in Great Britain

Miscellanea

Interpolation rules suggested by asymptotic expansions

BY NIALL H. ANDERSON
Department of Statistics, University of Glasgow, Glasgow G12 8QW, U.K.

PETER HALL

Centre for Mathematics and its Applications, Australian National University, GPO Box 4, Canberra, ACT 2601, Australia

AND D. M. TITTERINGTON
Department of Statistics, University of Glasgow, Glasgow G12 8QW, U.K.

SUMMARY

We suggest a simple interpolation rule, based on Edgeworth and Cornish-Fisher expansions, which permits highly accurate numerical approximation in problems involving calculation of distributions or quantiles without requiring explicit calculation of terms in the expansions.

Some key words: Chi-squared distribution; Cornish-Fisher expansion; Cramér-von Mises statistic; Edgeworth expansion; Kolmogorov-Smirnov statistic; Linear interpolation; Quadratic interpolation; Student's t distribution.

1. INTRODUCTION

Edgeworth expansions are often used to provide explicit corrections for asymptotic approximations, as in recent work in testing and confidence procedures involving nonparametric methods (Pfanzagl, 1979; Hall, 1983; Abramovitch & Singh, 1985) and in time series, econometrics and related fields (Sargan, 1976; Phillips, 1978; Durbin, 1980; Taniguchi, 1985; Taniguchi & Maekawa, 1990). In the present note we point out a useful application of Edgeworth expansion methods in producing simple but accurate interpolation rules for deriving information from or for constructing statistical tables.

Suppose the function f is tabulated for discrete values n_i of its argument, and that we seek the value of f for an untabulated point, say n. We suggest that, should f admit an Edgeworth expansion in increasing powers of $n^{-\frac{1}{2}}$ or n^{-1}, then interpolation should be among values of $n^{\frac{1}{2}}f(n_i)$ or $n_if(n_i)$, respectively, rather than values of $f(n_i)$. This recommendation applies to approximations of both distribution functions and percentage points. In the latter case our reference to Edgeworth expansions should really be to Cornish-Fisher expansions.

We should stress that our method does not require calculation of any terms of an Edgeworth or Cornish-Fisher expansion. It uses only existing information available in statistical tables, and knowledge that the Edgeworth expansion exists. This existence has been established for large classes of statistical problems; see, for example, Sargan (1975), Phillips (1977), Bhattacharya & Ghosh (1978), Durbin (1980) and Sargan & Satchell (1986). Aside from advantages of greater accuracy compared with traditional interpolation rules, our method allows linear and quadratic interpolation in intervals such as (n_1, ∞), as we now show.

2. THE INTERPOLATION FORMULAE

We consider in detail only the cases of approximating to $nf(n)$ by a linear and to $n^2f(n)$ by a quadratic function. The case where the expansions are series in $n^{-\frac{1}{2}}$ is dealt with by replacing n by $n^{\frac{1}{2}}$ throughout.

Linear. For $n_1 < n < n_2$, use

$$f_L(n) = (n_2 f_2 - n_1 f_1)(n_2 - n_1)^{-1} + (f_1 - f_2)(n_2 - n_1)^{-1} n_1 n_2 n^{-1},$$

where $f_i = f(n_i)$. If $n_2 = \infty$, $f_L(n) = f_2 + (f_1 - f_2) n_1 n^{-1}$.

Quadratic. For $n_1 < n < n_3$, with $n_1 < n_2 < n_3$, use

$$f_Q(n) = a_0 + a_1 n^{-1} + a_2 n^{-2},$$

where

$$a_0 = (n_1 n_2^2 f_2 + n_2 n_3^2 f_3 + n_3 n_1^2 f_1 - n_1 n_3^2 f_3 - n_2 n_1^2 f_1 - n_3 n_2^2 f_2)\{(n_1 - n_2)(n_2 - n_3)(n_3 - n_1)\}^{-1},$$

$$a_1 = (n_1^2 f_1 - n_2^2 f_2)(n_1 - n_2)^{-1} - a_0(n_1 + n_2),$$

$$a_2 = n_1^2 f_1 - a_0 n_1^2 - a_1 n_1.$$

If $n_3 = \infty$, the only change is that $a_0 = f_3$.

3. APPLICATIONS

Skew distributions and their percentage points typically admit Edgeworth and Cornish–Fisher expansions which behave as series in $n^{-\frac{1}{2}}$. Examples include chi-squared or gamma distributions, where n equals the number of degrees of freedom or shape parameter, respectively. Symmetric distributions, and distributions of statistics for two-sided hypothesis tests, often involve expansions in n^{-1}. Student's t-distribution, and the distribution of the Cramér–von Mises and Kolmogorov statistics, are cases in point. Here n equals the number of degrees of freedom in the former case and sample size in the latter.

The effectiveness of the interpolation procedures can be illustrated by their ability to approximate the upper 5%, 2·5% and 1% points of various distributions. In the case of Student's t, linear interpolation based on the values for $n = 7, 20$ and ∞ degrees of freedom gives 3-figure accuracy for all intermediate n for the 5% point; for the $2\frac{1}{2}$% points, $n = 8, 20, \infty$ can be used, and for the 1% points, $n = 10, 20, \infty$. For the chi-squared distribution, the values for $n = 5, 20, \infty$ are enough for all three percentages. In the case of the one-sample Kolmogorov–Smirnov statistic for testing a simple null hypothesis, linear interpolation using the points for $n = 10, 20, \infty$ led to some errors of 1 unit in the third significant figure for large n, but this was corrected provided one used quadratic interpolation based on the points for $n = 10, 20, 50, 200, \infty$.

For the Cramér–von Mises statistic, we failed to unearth existing tables more extensive than those of Knott (1974). For the 5%, 2·5% and 1% points, we constructed a table for $n =$ 10(1)50(10)100, 150, 200, 1000 using quadratic interpolation based on the values for $n = 10, 20, 50, 200, \infty$ from Knott (1974). The values agree with the corresponding table based on linear interpolation, apart from a few discrepancies of one unit in the fourth significant figure in some of the 1% values. The table also agrees with Knott's (1974) values for $n = 1000$, to four significant figures.

ACKNOWLEDGEMENT

N.H. Anderson was supported by a Research Studentship and P. Hall by a Visiting Fellowship Research Grant, both provided by the U.K. Science and Engineering Research Council.

References

ABRAMOVITCH, L. & SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13**, 116–32.

BHATTACHARYA, R. N. & GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6**, 434–51.

DURBIN, J. (1980). Approximations for densities of sufficient estimators. *Biometrika* **67**, 311–33.

HALL, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11**, 569–76.

KNOTT, M. (1974). The distribution of the Cramér–von Mises statistic for small sample sizes. *J. R. Statist. Soc.* B **36**, 430–8.

PFANZAGL, J. (1979). Nonparametric minimum contrast estimators. *Selecta Statist. Canadiana* **5**, 105–40.

PHILLIPS, P. C. B. (1977). A general theorem in the theory of asymptotic expansions as approximations to the finite sample distributions of econometric estimators. *Econometrica* **45**, 1517–34.

PHILLIPS, P. C. B. (1978). Edgeworth and saddlepoint approximations in the first order non-circular autoregression. *Biometrika* **65**, 71–8.

SARGAN, J. D. (1975). Gram–Charlier approximations applied to t-ratios of K-class estimators. *Econometrica* **43**, 327–46.

SARGAN, J. D. (1976). Econometric estimators and their Edgeworth approximations. *Econometrica* **44**, 421–48.

SARGAN, J. D. & SATCHELL, S. E. (1986). A theorem of validity for Edgeworth expansions. *Econometrica* **54**, 189–213.

TANIGUCHI, M. (1985). An asymptotic expansion for the distribution of the likelihood ratio criterion for a Gaussian autoregressive moving average process under a local alternative. *Econometric Theory* **1**, 73–84.

TANIGUCHI, M. & MAEKAWA, K. (1990). Asymptotic expansions of the distributions of statistics related to the spectral density matrix in multivariate time series and their applications. *Econometric Theory* **6**, 75–96.

[*Received December* 1991. *Revised February* 1992]