Biometrika (1978), **65**, 1, *pp.* 91–8
Printed in Great Britain

Edgeworth and saddlepoint approximations in the first-order noncircular autoregression

BY P. C. B. PHILLIPS
Department of Economics, University of Birmingham

SUMMARY

The first-order noncircular autoregressive model is considered and Edgeworth and saddlepoint approximations are given for the distribution of the least squares estimator of the autoregressive coefficient. Numerical calculations are used to compare the two approximations with the exact distribution, found by numerical integration. Both approximations are unsatisfactory when the autoregressive coefficient is moderately large and the sample size small, the saddlepoint because it is undefined in the tail area, and the Edgeworth because it badly distorts tail area probabilities. When the sample size is larger, the saddlepoint approximation performs well and is capable of three decimal place accuracy, although it is still not available in tail areas when the autoregressive coefficient is large.

Some key words: Approximate sampling distribution; Asymptotic series; Autoregressive process; Edgeworth; Steepest descent; Tail area probability; Time series.

1. INTRODUCTION

Recently some useful general results and algorithms have appeared about the Edgeworth approximation to the distribution of statistics more general than standardized means. In his survey paper, Wallace (1958) suggested that Edgeworth type expansions could be constructed for quite general, smooth functions of sample moments of the underlying data. In fact, the idea of such general expansions was even considered in the seminal paper of Edgeworth (1905). This very long paper was in two parts and in the second Edgeworth examined a number of more general cases than the simple summation of independent random causes where his 'law of error' was applicable. Amongst these cases were multivariate statistics (pp. 116–20) and quite general functions, including multivariate functions, of random elements (pp. 120–6). However, at the time of Wallace's survey paper no rigorous theory was available to establish the validity of the Edgeworth series as a proper asymptotic series in such general situations.

Chambers (1967) went a long way towards filling this gap. He developed Edgeworth expansions for the distribution of multivariate statistics more general than standardized means, gave conditions for validity and algorithms for computation; he also derived expansions for the distribution of quite general vector functions of other multivariate statistics, such as sample moments of the underlying data, and gave computational algorithms in this case as well.

Recently the use and validity of Edgeworth type approximations in a very general setting have been considered in the econometric literature. A key paper is by Sargan (1975). He proved a theorem on the validity of Edgeworth expansions for sample distributions of quite general estimators and test statistics with limiting normal distributions, including all the usual simultaneous equations estimators and t ratio test statistics. While this theorem has great generality, it does not extend to models where there are lagged dependent variables amongst the regressors or where there are autoregressive errors. Extensions to such situations have recently been given by Sargan (1976) and Phillips (1977b). Sargan (1976) specializes the earlier results and algorithms of Chambers (1967) to derive general formulae which apply in such models; and an independent proof of the validity of the Edgeworth expansion in such cases is given by Phillips (1977b).

In the probability literature also, a number of important general theorems on the validity of Edgeworth series expansions have recently been established. Chibisov (1972), in particular, has proved the validity of an asymptotic expansion of the distribution of a multivariate statistic that can itself be represented in the form of an asymptotic series whose terms are polynomial functions of standardized means of independent, identically distributed random vectors. The recent survey article by Bhattacharya (1977) contains a similar theorem under somewhat weaker conditions than those used by Chibisov.

One difficulty with the use of the Edgeworth approximation in practical applications is that the approximation is often unsatisfactory in the tails. The errors in the approximation can be as large as the tail probabilities themselves and there is, usually, nothing to prevent the approximation taking on negative values or values greater than unity. These points were made by Daniels (1954) when he explored the use of the saddlepoint approximation in mathematical statistics. For this latter approximation does not suffer from the drawbacks of the Edgeworth approximation in the tail area. Later, Daniels (1956) applied the saddlepoint approximation to the distribution of serial correlation coefficients and considered, *inter alia*, the noncircular first-order autoregression.

Only very recently has the Edgeworth approximation been developed in the same context (Phillips, 1977a). The aim of the present paper is to consider the comparative performance of the two different approximations in this special case. The model we will use is the noncircular autoregression

$$y_t = \alpha y_{t-1} + u_t \quad (t = \ldots -1, 0, 1, \ldots), \tag{1}$$

where $|\alpha| < 1$ and the u_t are independently and identically distributed in $N(0, \sigma^2)$. We look, in particular, at the distribution of the least squares estimator of α in (1) given by

$$\hat{\alpha} = \sum_{t=1}^{T} y_t y_{t-1} \Big/ \sum_{t=1}^{T} y_{t-1}^2 \tag{2}$$

and we examine the performance of the two approximations in a wide region of the tail area. This should highlight the strengths and the weaknesses of the two approximations, at least in the present case.

2. The Edgeworth approximation

It is most convenient to work with the standardized statistic $T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}$ which has a limiting $N(0,1)$ distribution as $T \to \infty$. Earlier (Phillips, 1977a) the Edgeworth expansion of the distribution of $T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}$ was derived explicitly up to $O(T^{-\frac{1}{2}})$. We obtained there (p. 470)

$$\mathrm{pr}\{T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}} \leqslant x\} = I(x) + \frac{i(x)}{T^{\frac{1}{2}}} \frac{\alpha}{(1-\alpha^2)^{\frac{1}{2}}}(x^2+1) + O(T^{-1}), \tag{3}$$

where $i(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2}$ and

$$I(x) = \int_{-\infty}^{x} i(t)\, dt.$$

Unfortunately, if we wish to compute symmetric tail area probabilities from (3) the correction term of $O(T^{-\frac{1}{2}})$ cancels and we are left with the asymptotic normal approximation.

Thus

$$\text{pr}\{|T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}|>x\} = 1-I(x)+I(-x)+O(T^{-1}),$$

and to correct the asymptotic approximation here we must find the term of $O(T^{-1})$ in the expansion (3). This can be done by extending the derivations shown in the earlier paper by the present author (Phillips, 1977a). Details may be obtained from the author on request and the resulting expression is

$$\text{pr}\{T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}} \leqslant x\} = I(x)+i(x)\Big[T^{-\frac{1}{2}}\alpha(1-\alpha^2)^{-\frac{1}{2}}(x^2+1) + \frac{1}{4T}(1-\alpha^2)^{-1}\{(1-\alpha^2)x+(1+\alpha^2)x^3-2\alpha^2x^5\}\Big]+O(T^{-\frac{3}{2}}). \quad (4)$$

Neglecting the remainder of $O(T^{-\frac{3}{2}})$, we call the right-hand side of (4) the Edgeworth A approximation. In its present form, this approximation is not necessarily contained in the [0, 1] interval. But, the alternative representation

$$I\left\{x+\frac{1}{T^{\frac{1}{2}}}\alpha(1-\alpha^2)^{-\frac{1}{2}}+\frac{1}{4T}(1+\alpha^2)(1-\alpha^2)^{-1}x+\frac{1}{T^{\frac{1}{2}}}\alpha(1-\alpha^2)^{-\frac{1}{2}}x^2+\frac{1}{4T}(1+5\alpha^2)(1-\alpha^2)^{-1}\right\}+O(T^{-\frac{3}{2}}) \quad (5)$$

is the same up to $O(T^{-1})$ as can be verified by a Taylor series expansion (Phillips, 1977a; Sargan, 1976); and the alternative approximation based on (5) does lie in the [0, 1] interval, although it is not monotonic for all values of x. We call the approximation obtained by neglecting the remainder of $O(T^{-\frac{3}{2}})$ in (5) the Edgeworth B approximation. We can readily compute the approximations based on (4) and (5) for specified values of α and T and in § 4 below we report the results of some numerical computations.

3. Daniels's saddlepoint approximation

We write the least squares estimator of α in (2) as $\alpha = (y'C_1y)/(y'C_2y)$, where $y' = (y_0, \ldots, y_T)$,

$$C_1 = \begin{bmatrix} 0 & \frac{1}{2} & \ldots & 0 & 0 \\ \frac{1}{2} & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \frac{1}{2} \\ 0 & 0 & \ldots & \frac{1}{2} & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}. \quad (6)$$

Since $\hat{\alpha}$ is a ratio of quadratic forms in the underlying observations, an alternative point of departure in approximating the distribution of $\hat{\alpha}$ relies on Geary's extension of Cramér's result on the density function of a ratio of random variables. This was the approach taken by Daniels (1956). He dealt not with $\hat{\alpha}$ in the noncircular case but with the modified correlation coefficient

$$\alpha^* = \sum_{t=1}^{T} y_t y_{t-1} \Big/ \left(\tfrac{1}{2}y_0^2 + \sum_{t=1}^{T-1} y_t^2 + \tfrac{1}{2}y_T^2\right), \quad (7)$$

and derived a saddlepoint approximation to the probability density of α^*.

The estimator α^* has the advantage that $|\alpha^*| \leqslant 1$ and, because of the symmetry in the denominator, the use of (7) simplifies the analysis leading to the saddlepoint approximation. But, in more general situations it will not always be possible to find modified statistics such as (7); and it is, in any case, of interest to examine the small sample distribution of the unmodified least squares estimator in this model.

Daniels's method can be extended readily to derive a saddlepoint approximation to the density of $\hat{\alpha}$. The algebra closely parallels that of Daniels and I give only the final result here.

Details of the derivations can be obtained from the author on request. If we let $p_T(x)$ denote the density of $\hat{\alpha}$ then the approximation follows from

$$p_T(x) = \frac{(T-3)\,\Gamma(\frac{1}{2}T-\frac{1}{2})}{2\pi^{\frac{1}{2}}\,\Gamma(\frac{1}{2}T)} \frac{(1-\alpha^2)^{\frac{1}{2}}(1-x^2)^{\frac{1}{2}(T+1)}}{(1-ax)^{\frac{3}{2}}(1+\alpha x-2x^2)^{\frac{1}{2}}(1+\alpha^2-2x\alpha)^{\frac{1}{2}(T-1)}}\{1+O(T^{-1})\}. \tag{8}$$

A simple transformation then yields the density approximation for $T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}$ and tail area probabilities can be computed by numerical integration.

One difficulty that arises with the use of (8) is that the approximate density is undefined for values of x near the extremes of the region $(-1, 1)$. This is caused by the factor $(1+\alpha x-2x^2)^{\frac{1}{2}}$ in the denominator of (8). The density approximation becomes infinite outside the interval

$$\tfrac{1}{4}\alpha-(\tfrac{1}{16}\alpha^2+\tfrac{1}{2})^{\frac{1}{2}} < x < \tfrac{1}{4}\alpha+(\tfrac{1}{16}\alpha^2+\tfrac{1}{2})^{\frac{1}{2}} \tag{9}$$

and, as we shall see in the next section, this means that the approximation is not always available in a substantial part of the tails. Technically, this problem arises from the existence of a branch point of the integrand, in the integral defining $p_T(x)$, within the natural contour of integration. More precisely, the integrand has a branch point at $(2x-\alpha)^{-1}$ and a saddlepoint at x on the real axis. When $x > (2x-\alpha)^{-1} > 0$ or $1+\alpha x-2x^2 < 0$ the contour of integration cannot be simply transformed to pass through the saddlepoint. It is necessary to detour in a loop around the branch point and the path of integration is no longer a line of steepest descent through the saddlepoint. Moreover, the integrals in the neighbourhood of the branch point do not cancel, so that the new path of integration leads to an approximation of basically different form. Details of this work will be reported later.

4. Some numerical comparisons

The simplicity of the first-order autoregression and the form of the estimator $\hat{\alpha}$ have the advantage over more complicated time series models and estimators that the exact distribution of $\hat{\alpha}$ is relatively easy to compute. Numerical computations for the exact distribution were carried out using the technique of Imhof (1961). The Edgeworth A and B approximations were calculated directly from the formulae (4) and (5). Probabilities in the case of the saddlepoint approximation were computed by numerical integration over the appropriate interval using the form of (8); but care needed to be taken to ensure that the limits, (9), for which the approximation is defined were properly observed.

Various values of T and α were selected for our computations and Table 1 gives the exact and approximate tail probabilities

$$\text{pr}\{|T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}| > X\} \tag{10}$$

for X in the region $1{\cdot}0 \leqslant X \leqslant 3{\cdot}0$. The Edgeworth approximation is given in both forms A and B described earlier in § 2 and we also record the asymptotic normal approximation. The computations in Table 1 refer to α values of 0·2, 0·4, 0·6 and 0·8, and sample sizes of $T = 10$ and $T = 30$.

In a number of cases and, more particularly for the larger α values, the saddlepoint approximation is 'not available'. This means that the deviation X in (10) leads to a region in the tail of the distribution of $\hat{\alpha}$ which is outside the limits given in (9). Specifically when X is such that

$$\alpha + X\left(\frac{1-\alpha^2}{T}\right)^{\frac{1}{2}} \geqslant \tfrac{1}{4}\alpha+(\tfrac{1}{16}\alpha^2+\tfrac{1}{2})^{\frac{1}{2}}$$

then the saddlepoint approximation is undefined. This happens more frequently as T becomes small and α approaches unity; see, for example, Table 1(d), $T = 10$, $\alpha = 0{\cdot}8$.

Table 1. *Exact values and approximations for* $\text{pr}\{|T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}|>X\}$

(a) $T = 10$, $\alpha = 0{\cdot}2$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·2987	0·2931	0·2939	0·3173	0·2926
1·10	0·2513	0·2451	0·2462	0·2713	0·2440
1·20	0·2091	0·2023	0·2038	0·2301	0·2005
1·30	0·1720	0·1647	0·1666	0·1936	0·1620
1·40	0·1400	0·1321	0·1345	0·1651	0·1283
1·50	0·1126	0·1043	0·1071	0·1336	0·0990
1·60	0·0896	0·0809	0·0842	0·1095	0·0737
1·70	0·0706	0·0616	0·0653	0·0891	0·0512
1·80	0·0551	0·0460	0·0500	0·0718	0·0274
1·90	0·0427	0·0335	0·0377	0·0574	n.a.
2·00	0·0328	0·0239	0·0286	0·0455	n.a.
2·10	0·0251	0·0165	0·0205	0·0357	n.a.
2·20	0·0191	0·0110	0·0148	0·0278	n.a.
2·30	0·0144	0·0071	0·0105	0·0214	n.a.
2·40	0·0108	0·0043	0·0073	0·0163	n.a.
2·50	0·0081	0·0025	0·0050	0·0124	n.a.
2·60	0·0060	0·0013	0·0033	0·0093	n.a.
2·70	0·0044	0·0006	0·0022	0·0069	n.a.
2·80	0·0033	0·0002	0·0014	0·0051	n.a.
2·90	0·0024	0·0000	0·0008	0·0037	n.a.
3·00	0·0018	0·0000	0·0005	0·0026	n.a.

(b) $T = 10$, $\alpha = 0{\cdot}4$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3009	0·2931	0·2946	0·3173	0·2792
1·10	0·2547	0·2460	0·2479	0·2713	0·2298
1·20	0·2141	0·2045	0·2067	0·2301	0·1854
1·30	0·1790	0·1685	0·1709	0·1936	0·1451
1·40	0·1492	0·1380	0·1401	0·1615	0·1065
1·50	0·1241	0·1124	0·1141	0·1336	n.a.
1·60	0·1032	0·0915	0·0922	0·1095	n.a.
1·70	0·0860	0·0746	0·0740	0·0891	n.a.
1·80	0·0717	0·0613	0·0588	0·0718	n.a.
1·90	0·0598	0·0510	0·0463	0·0574	n.a.
2·00	0·0498	0·0431	0·0360	0·0455	n.a.
2·10	0·0412	0·0372	0·0276	0·0357	n.a.
2·20	0·0340	0·0326	0·0207	0·0278	n.a.
2·30	0·0279	0·0291	0·0153	0·0214	n.a.
2·40	0·0228	0·0262	0·0110	0·0163	n.a.
2·50	0·0185	0·0239	0·0077	0·0124	n.a.
2·60	0·0149	0·0218	0·0052	0·0093	n.a.
2·70	0·0120	0·0198	0·0034	0·0069	n.a.
2·80	0·0095	0·0179	0·0021	0·0051	n.a.
2·90	0·0075	0·0161	0·0013	0·0037	n.a.
3·00	0·0059	0·0142	0·0007	0·0026	n.a.

(c) $T = 10$, $\alpha = 0{\cdot}6$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3082	0·2931	0·2965	0·3173	0·02475
1·10	0·2655	0·2482	0·2514	0·2713	n.a.
1·20	0·2293	0·2100	0·2121	0·2301	n.a.
1·30	0·1990	0·1782	0·1779	0·1936	n.a.
1·40	0·1736	0·1526	0·1483	0·1615	n.a.
1·50	0·1519	0·1327	0·1226	0·1336	n.a.
1·60	0·1332	0·1178	0·1002	0·1095	n.a.
1·70	0·1169	0·1071	0·0805	0·0891	n.a.
1·80	0·1025	0·0997	0·0634	0·0718	n.a.
1·90	0·0899	0·0947	0·0487	0·0574	n.a.
2·00	0·0787	0·0913	0·0363	0·0455	n.a.
2·10	0·0688	0·0888	0·0261	0·0357	n.a.
2·20	0·0600	0·0866	0·0180	0·0278	n.a.
2·30	0·0522	0·0841	0·0118	0·0214	n.a.
2·40	0·0453	0·0811	0·0074	0·0163	n.a.
2·50	0·0392	0·0774	0·0043	0·0124	n.a.
2·60	0·0338	0·0729	0·0024	0·0093	n.a.
2·70	0·0290	0·0678	0·0012	0·0069	n.a.
2·80	0·0249	0·0621	0·0005	0·0051	n.a.
2·90	0·0213	0·0561	0·0002	0·0037	n.a.
3·00	0·0181	0·0498	0·0000	0·0026	n.a.

(d) $T = 10$, $\alpha = 0{\cdot}8$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3444	0·2931	0·2992	0·3173	n.a.
1·10	0·3090	0·2556	0·2531	0·2713	n.a.
1·20	0·2783	0·2279	0·2104	0·2301	n.a.
1·30	0·2515	0·2098	0·1705	0·1936	n.a.
1·40	0·2280	0·2006	0·1332	0·1615	n.a.
1·50	0·2071	0·1991	0·0994	0·1336	n.a.
1·60	0·1884	0·2040	0·0699	0·1095	n.a.
1·70	0·1716	0·2132	0·0458	0·0891	n.a.
1·80	0·1565	0·2250	0·0275	0·0718	n.a.
1·90	0·1427	0·2375	0·0148	0·0574	n.a.
2·00	0·1302	0·2488	0·0070	0·0455	n.a.
2·10	0·1188	0·2576	0·0028	0·0357	n.a.
2·20	0·1083	0·2628	0·0009	0·0278	n.a.
2·30	0·0987	0·2638	0·0002	0·0214	n.a.
2·40	0·0899	0·2602	0·0000	0·0163	n.a.
2·50	0·0819	0·2521	0·0000	0·0124	n.a.
2·60	0·0745	0·2400	0·0000	0·0093	n.a.
2·70	0·0678	0·2246	0·0000	0·0069	n.a.
2·80	0·0616	0·2066	0·0000	0·0051	n.a.
2·90	0·0559	0·1868	0·0000	0·0037	n.a.
3·00	0·0507	0·1662	0·0000	0·0026	n.a.

EDGE A, Edgeworth A. EDGE B, Edgeworth B. SADDLE, saddlepoint approximation.
n.a. Not available; the approximate density is not defined.

Table 1. *Exact values and approximations for* $\mathrm{pr}\{|T^{\frac{1}{2}}(\hat{\alpha}-\alpha)/(1-\alpha^2)^{\frac{1}{2}}|>X\}$
(continued)

(e) $T = 30$, $\alpha = 0·2$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3100	0·3092	0·3093	0·3173	0·3106†
1·10	0·2634	0·2625	0·2627	0·2713	0·2640†
1·20	0·2218	0·2208	0·2210	0·2301	0·2223†
1·30	0·1850	0·1839	0·1841	0·1936	0·1854†
1·40	0·1528	0·1517	0·1519	0·1615	0·1531†
1·50	0·1249	0·1238	0·1241	0·1336	0·1252†
1·60	0·1012	0·1000	0·1004	0·1095	0·1014†
1·70	0·0811	0·0799	0·0804	0·0891	0·0813†
1·80	0·0644	0·0632	0·0637	0·0718	0·0645†
1·90	0·0507	0·0494	0·0499	0·0574	0·0507†
2·00	0·0395	0·0383	0·0388	0·0455	0·0395†
2·10	0·0305	0·0293	0·0298	0·0357	0·0304†
2·20	0·0234	0·0222	0·0227	0·0278	0·0233†
2·30	0·0178	0·0166	0·0171	0·0214	0·0176†
2·40	0·0134	0·0123	0·0127	0·0163	0·0132†
2·50	0·0101	0·0091	0·0094	0·0124	0·0098†
2·60	0·0076	0·0066	0·0068	0·0093	0·0073†
2·70	0·0056	0·0048	0·0049	0·0069	0·0053†
2·80	0·0042	0·0034	0·0035	0·0051	0·0038†
2·90	0·0031	0·0025	0·0025	0·0037	0·0027†
3·00	0·0023	0·0018	0·0017	0·0026	0·0019†

(f) $T = 30$, $\alpha = 0·4$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3103	0·3092	0·3094	0·3173	0·3112†
1·10	0·2640	0·2628	0·2631	0·2713	0·2649†
1·20	0·2228	0·2216	0·2218	0·2301	0·2236†
1·30	0·1866	0·1852	0·1855	0·1936	0·1873†
1·40	0·1551	0·1536	0·1539	0·1615	0·1557†
1·50	0·1280	0·1265	0·1267	0·1336	0·1285†
1·60	0·1051	0·1035	0·1037	0·1095	0·1055†
1·70	0·0858	0·0843	0·0842	0·0891	0·0861†
1·80	0·0698	0·0683	0·0680	0·0718	0·0700†
1·90	0·0567	0·0553	0·0547	0·0574	0·0568†
2·00	0·0460	0·0447	0·0437	0·0455	0·0459†
2·10	0·0373	0·0362	0·0348	0·0357	0·0371†
2·20	0·0302	0·0294	0·0275	0·0278	0·0299†
2·30	0·0245	0·0240	0·0217	0·0214	0·0241†
2·40	0·0198	0·0196	0·0170	0·0163	0·0193
2·50	0·0160	0·0162	0·0132	0·0124	n.a.
2·60	0·0130	0·0134	0·0102	0·0093	n.a.
2·70	0·0105	0·0112	0·0078	0·0069	n.a.
2·80	0·0084	0·0093	0·0059	0·0051	n.a.
2·90	0·0068	0·0078	0·0044	0·0037	n.a.
3·00	0·0054	0·0065	0·0032	0·0026	n.a.

(g) $T = 30$, $\alpha = 0·6$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3111	0·3092	0·3096	0·3173	0·3127†
1·10	0·2657	0·2636	0·2641	0·2713	0·2672†
1·20	0·2257	0·2234	0·2238	0·2301	0·2270†
1·30	0·1910	0·1884	0·1886	0·1936	0·1920†
1·40	0·1612	0·1585	0·1582	0·1615	0·1620†
1·50	0·1361	0·1333	0·1322	0·1336	0·1365†
1·60	0·1151	0·1123	0·1102	0·1096	0·1151†
1·70	0·0976	0·0951	0·0916	0·0891	0·0971†
1·80	0·0831	0·0811	0·0759	0·0718	0·0815†
1·90	0·0710	0·0698	0·0627	0·0574	n.a.
2·00	0·0609	0·0607	0·0515	0·0455	n.a.
2·10	0·0523	0·0534	0·0421	0·0357	n.a.
2·20	0·0449	0·0474	0·0341	0·0278	n.a.
2·30	0·0385	0·0423	0·0273	0·0214	n.a.
2·40	0·0330	0·0379	0·0217	0·0163	n.a.
2·50	0·0282	0·0340	0·0169	0·0124	n.a.
2·60	0·0240	0·0305	0·0130	0·0093	n.a.
2·70	0·0204	0·0272	0·0098	0·0069	n.a.
2·80	0·0173	0·0241	0·0073	0·0051	n.a.
2·90	0·0146	0·0211	0·0053	0·0037	n.a.
3·00	0·0123	0·0184	0·0038	0·0026	n.a.

(h) $T = 30$, $\alpha = 0·8$

X	Exact	EDGE A	EDGE B	Normal	SADDLE
1·00	0·3159	0·3092	0·3107	0·3173	0·2984
1·10	0·2739	0·2661	0·2669	0·2713	0·2535
1·20	0·2384	0·2294	0·2285	0·2301	0·2096
1·30	0·2088	0·1990	0·1951	0·1936	n.a.
1·40	0·1844	0·1745	0·1661	0·1615	n.a.
1·50	0·1641	0·1554	0·1407	0·1336	n.a.
1·60	0·1468	0·1410	0·1182	0·1095	n.a.
1·70	0·1317	0·1305	0·0983	0·0891	n.a.
1·80	0·1183	0·1229	0·0805	0·0718	n.a.
1·90	0·1063	0·1174	0·0647	0·0574	n.a.
2·00	0·0955	0·1132	0·0508	0·0455	n.a.
2·10	0·0857	0·1097	0·0389	0·0357	n.a.
2·20	0·0759	0·1061	0·0289	0·0278	n.a.
2·30	0·0689	0·1022	0·0207	0·0214	n.a.
2·40	0·0617	0·0976	0·0142	0·0163	n.a.
2·50	0·0552	0·0923	0·0094	0·0124	n.a.
2·60	0·0493	0·0862	0·0059	0·0093	n.a.
2·70	0·0441	0·0795	0·0035	0·0069	n.a.
2·80	0·0393	0·0722	0·0019	0·0051	n.a.
2·90	0·0350	0·0647	0·0010	0·0037	n.a.
3·00	0·0312	0·0572	0·0000	0·0026	n.a.

EDGE A, Edgeworth A. EDGE B, Edgeworth B. SADDLE, saddlepoint approximation.
n.a. Not available; the approximate density is not defined.
† 'Saddlepoint approximation' as close or closer to 'exact' than Edgeworth A.

When the saddlepoint approximation is closer to the exact probability than the Edgeworth A approximation we have inserted a † in the final column of the Table. From Tables 1(a)–(d) we see that, when $T = 10$, the saddlepoint approximation is not so good as the Edgeworth A. Indeed, for this value of T the saddlepoint approximation is not available over a wide range of X values. But the Edgeworth A is itself far from being a good approximation in this case: when $T = 10$ and $\alpha = 0{\cdot}8$ the Edgeworth approximation produces values which are frequently over twice the size of the true probabilities; note also that the asymptotic normal approximation is here just as unsatisfactory, by heavily underestimating the true probabilities. Only for the smaller values of α is the Edgeworth approximation really satisfactory; and for $\alpha = 0{\cdot}2$ and $0{\cdot}4$ we note that it is generally superior to the asymptotic normal approximation.

The situation changes as T increases. For $T = 30$ and the smaller α values (Tables 1(e) and (f)) the saddlepoint approximation is very accurate where it is defined. In Table 1(e) we see that it dominates the Edgeworth A over the whole range of X; and for most values of X it is accurate up to three decimal places. When $T = 30$ and $\alpha = 0{\cdot}4$ we see that the saddlepoint approximation is not quite as good: it is superior to the Edgeworth A over much of the region but it is not available in the tails around 1%. For $T = 30$ and the larger α values, Tables 1(g) and 1(h), we find that the saddlepoint approximation continues to do well where it is defined; it is generally more accurate than the Edgeworth A for $\alpha = 0{\cdot}6$ but is undefined for 20% of the tail area when $\alpha = 0{\cdot}8$.

Further calculations were undertaken for larger T values but are not reported here in detail. We found that the saddlepoint approximation continued to be superior for the smaller α values but was still undefined for a region of the tail when α was large: specifically for 2% of the tail when $\alpha = 0{\cdot}4$ and $T = 50$; and 13% of the tail when $\alpha = 0{\cdot}8$ and $T = 50$.

5. Conclusion

Approximations to the distribution of the least squares estimator in the model (1) which perform well over the whole range of the distribution including the tail seem difficult to obtain. Our computations show that the saddlepoint approximation can produce very accurate results, certainly for sample sizes as large as 30. But this approximation has the disadvantage that it is not defined over a sizeable region of the tail for values of the autoregressive parameter greater than 0·4. The Edgeworth approximation also suffers drawbacks for these values of the autoregressive parameter. We found in this case that the errors on the approximation can be as large or larger than the true probabilities themselves for a wide region of the tail when the sample size is small. It would, therefore, appear that neither approximation is satisfactory when the autoregressive coefficient is moderately large and the sample size is small.

I am grateful to Michael Prior for assistance in programming some of the computations reported in the paper, to the S.S.R.C. for support and to the referee for helpful comments. Some of the numerical results in this paper were first recorded in a discussion paper by the author on large deviation limit theory and associated asymptotic expansions for large deviations.

References

Bhattacharya, R. N. (1977). Refinements of the multidimensional Central Limit Theorem and applications. *Ann. Prob.* **5**, 1–27.

Chambers, J. M. (1967). On methods of asymptotic approximation for multivariate distributions. *Biometrika* **54**, 367–83.

CHIBISOV, D. M. (1972). An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion. *Theory Prob. Applic.* **17**, 620–30.

DANIELS, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **25**, 631–50.

DANIELS, H. E. (1956). The approximate distribution of serial correlation coefficients. *Biometrika* **43**, 169–85.

EDGEWORTH, F. Y. (1905). The law of error. *Trans. Camb. Phil. Soc.* **20**, 36–65; 113–41.

IMHOF, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 419–26.

PHILLIPS, P. C. B. (1977a). Approximations to some finite sample distributions associated with a first order stochastic difference equation. *Econometrica* **45**, 463–86.

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

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 the Edgeworth approximation. *Econometrica* **44**, 421–48.

WALLACE, D. L. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29**, 635–54.

[*Received May* 1977. *Revised October* 1977]