Bartlett correction of the unit root test in autoregressive models[1]

BY B. NIELSEN

Institute of Mathematical Statistics,
Universitetsparken 5, 2100 Copenhagen Ø, Denmark
bentn@math.ku.dk

23 June 1995

SUMMARY

The usual conditions for a Bartlett correction are not fulfilled when testing for a unit root in a Gaussian autoregressive model. However, by expanding the moments of the Likelihood ratio it is shown that the Bartlett correction does improve the asymptotic distribution approximation.

Some key words: Bartlett correction, Gaussian autoregressive model, unit root, moment expansions.

1. INTRODUCTION

When testing a simple hypothesis in i.i.d. situations the likelihood ratio is asymptotically χ^2 distributed. Bartlett (1937) suggested a scalar correction in small samples and later Lawley (1956) proved that this correction improves the distributional approximation since the rth cumulant of the likelihood ratio then has the form

$$\kappa_r(LR_n) \simeq m_r \left(1 - \frac{d}{n}\right)^r, \qquad (1)$$

where m_r is the cumulant of the limit distribution. Since then this idea has been studied extensively and also in a more general framework.

The test for unit root in an autoregressive model does not fit into this framework. Under the null the process is a random walk, the information is stochastic even in the limit, the likelihood ratio distribution is not asymptotic χ^2 and so forth. The results in this paper show that even though the relation (1) does not hold either, but rather

$$\kappa_r(LR_n) \simeq m_r \left(1 - \frac{d_r}{n}\right), \qquad (2)$$

the Bartlett correction does improve the asymptotic distribution in the sense

$$\kappa_r\left(\frac{LR_n}{1 - \frac{d_1}{n}}\right) \simeq m_r \left(1 - \frac{d_r^B}{n}\right), \qquad (3)$$

[1] The paper does also appear as preprint from IMS

where the absolute value of $d_r^B = d_r - r \cdot d_1$ is smaller than d_r at least for the first four cumulants.

2. The model

The model considered is given by a first order autoregressive equation,

$$\Delta X_i = \alpha X_{i-1} + \varepsilon_i, \qquad i = 1, \ldots, n$$

conditional on the initial value, $X_0 = 0$, where the innovations are independently mean zero Gaussian distributed with variance ω^2 and the parameters (α, ω^2) vary freely. In this model the likelihood ratio test statistic for the unit root, $\alpha = 0$, is given by

$$Q = -n \log(1 - S) \tag{4}$$

where

$$S = \frac{\left(\sum_{i=1}^n X_{i-1} \Delta X_i\right)^2}{\sum_{i=1}^n X_{i-1}^2 \sum_{i=1}^n \Delta X_i^2}.$$

The statistic, S is scale invariant whereas $\sum \Delta X_i^2$ is complete and sufficient under the null and thus S and $\sum \Delta X_i^2$ are independent. Moreover the product of these, Q_{SUB} say, is the likelihood ratio for the unit root in the submodel where the innovation variance is known. Now, let $\omega^2 = 1$ in the remainder of the paper.

The model is a curved exponential family and the Laplace transform of the minimal sufficient statistic is the determinant of a large matrix and can be rewritten by solving a second order difference equation, cf. White (1958). However, in this context an approximation is more applicable than the exact Laplace transform,

$$E \exp\left(-a \frac{1}{n^2} \sum_{i=1}^n X_{i-1} - b \frac{1}{2n} X_n^2 - c \frac{1}{2n} \sum_{i=1}^n \Delta X_i^2\right) \simeq \exp\left(-\frac{c^*}{2}\right)\left(\cosh\sqrt{2a^*} + \frac{\sinh\sqrt{2a^*}}{\sqrt{2a^*}} b^*\right)^{-\frac{1}{2}}$$

where

$$\begin{aligned} a^* &= a\left(1 - n^{-1}c + n^{-2}(c^2 - a/6)\right), \\ b^* &= b - n^{-1}(a + bc) + n^{-2}(bc^2 + ac - ab/3), \\ c^* &= c - n^{-1}c^2/2 + n^{-2}c^3/3. \end{aligned}$$

White reported the marginal result where $b = -c$ in correspondence with the relation $2 \sum X_i \Delta X_{i-1} = X_n^2 - \sum \Delta X_i^2$.

By a convexity argument the Laplace transform converges in a neighbourhood of zero. This implies weak convergence of the corresponding statistic, cf. Curtiss (1942). The statistics of interest here, Q, S, Q_{SUB}, have the same limit distribution which is described only indirectly by the Laplace transform above. The maximum likehood estimator of the autoregressive parameter is, however, a simpler function of the sufficient statistic and its asymptotic distribution is known explicitly (Abadir, 1993).

3. Moment expansions

The statistics S and Q_{SUB} are ratios of powers of the sufficient statistic and hence their cumulant can be found from the joint Laplace transform using the following wellknown idea. That is, for random variables Y, Z with Laplace transforms defined in a neighbourhood of zero and where Z is positive

$$EY = -\lim_{t\to 0}\frac{\partial}{\partial t}Ee^{-tY} \qquad \text{and} \qquad E\frac{1}{Z} = \int_0^\infty Ee^{-tZ}dt$$

with the latter equality following by Tonelli's Theorem. The cumulants of the likelihood ratio, Q, are found by expanding the logarithm. It turns out that the cumulants can be expressed in terms of integrals of the type

$$\int_0^\infty \frac{x^{2k-1}}{\sqrt{\cosh x}}dx \qquad \text{and} \qquad \int_0^\infty \frac{\tanh^{2k}x}{x\sqrt{\cosh x}}dx \qquad \text{for } k\in\mathbf{N}$$

which are solved by numerical integration. As an example the limit mean equals

$$\lim_{n\to\infty} EQ = \frac{1}{4}\int_0^\infty \frac{1}{\sqrt{\cosh x}}\left(3\frac{\tanh^2 x}{x} + x\right)dx - 1.$$

The independence of S and $\sum \Delta X_i^2$ made a check of tedious calculations possible and simulations support the numerical results, cf. the Appendix.

Table 1 shows the leading terms of the first five cumulants. These are compared with those of the χ^2_1 distribution, of a scaled χ^2_1 distribution, and of a fitted Γ distribution. This Γ distribution with form 0.5869 and scale 1.945 determined from the first two moments fits remarkably well and its quantiles are reported in Table 2.

The expansions of the first moments were found to be

$$EQ \simeq 1.142\left(1 - n^{-1}0.01712 + n^{-2}0.8327\right), \tag{5}$$

$$EQ_{SUB} \simeq 1.142\left(1 - n^{-1}1.559 + n^{-2}3.685\right). \tag{6}$$

For the mean in the full model the numerical value of the third order term is more important than the second order term for $n \leq 48$. So in applications the full model Bartlett correction is recommended to include the third order term. This is, however, not the case for the submodel. Table 3 shows second order terms of the moments before and after Bartlett correction, based on the usual first two terms of the mean, cf. (3). The second order terms do not disappear by the correction but are certainly reduced. As a curiosity, the second order term of the mean of the test in the submodel, $1.142 \cdot 1.559/n \simeq 1.781/n$, equals the first order term of the bias of maximum likelihood estimator, cf. Abadir (1995).

	$\kappa_1 = \mu$	$\kappa_2 = \sigma^2$	κ_3	κ_4	κ_5
Asymptotic	1.142	2.221	8.640	50.94	402.7
fitted Γ	1.142	2.221	8.639	50.41	392.2
scaled χ^2	1.142	2.606	11.90	81.52	744.5
χ^2	1	2	8	48	384

Table 1. Leading terms of the cumulant expansion, compared with a fitted Γ distribution and the χ^2 distribution. The same asymptotic mean and variance was found by Larsson (1994).

50 %	80 %	90 %	95 %	97.5 %	99 %
0.5911	1.882	2.984	4.141	5.333	6.944

Table 2. Quantiles of the fitted Γ distribution.

	$\kappa_2 = \sigma^2$	κ_3	κ_4
Q	0.2503	1.184	0.2183
Q Bartlett corrected	0.2161	1.133	0.1498
Q_{SUB}	3.251	5.644	6.145
Q_{SUB} Bartlett corrected	0.1324	0.9664	0.09148

Table 3. Second order terms of the cumulants, cf. d_k and d_k^B in (2) and (3).

[Figure 1 about here]

4. Bartlett correction in practice

The finite sample distributions of the two likelihood ratios are quite different. In Figure 1 simulated quantiles are compared with an approximation of the asymptotic distribution in Q-Q plots. For practical purposes these distributions are indistinguishable for 8 observations of the full model and 100-200 observations of the submodel. Moreover, the full model test is oversized whereas the submodel test is undersized. Together with the signs of the expansion (5) this implies that the full model test is doing slightly worse together with the second order Bartlett correction than without. In Table 4 simulated p-values are given for an asymptotic 5 % test, and Bartlett corrected tests using respectively second and third order terms of the mean.

	Q			Q_{SUB}		
n	∞	n^{-1}	n^{-2}	∞	n^{-1}	n^{-2}
3	8.4 %	8.5 %	7.3 %	3.0 %	12.9 %	4.0 %
6	5.2 %	5.3 %	5.0 %	3.2 %	6.5 %	4.9 %
25	4.9 %	4.9 %	4.9 %	4.3 %	5.0 %	5.0 %

Table 4. p-values for tests based on the asymptotic quantile and Bartlett corrected quantiles including respectively the second and the third order terms. The p-values were simulated using 10^6 repetitions and the Γ approximated asymptotic 95 % quantile.

5. The second order autoregression

Now, consider the second autoregressive model as given by

$$\Delta X_i = \alpha X_{i-1} + \beta \Delta X_{i-1} + \varepsilon_i .$$

The likelihood ratio for the unit root hypothesis is basically of the form (4) but with the terms of S being corrected for lagged differences into S_2, say. Larsson (1994) expanded the first moment in this situation and found

$$EQ_{AR2} \simeq EQ + n^{-1} R(\beta)$$

where $R(0) = 1.241$ and $R(\beta) \simeq R(0)$ fairly accurately for $|\beta \leq 0.3|$ and $n > 20$.

Acknowledgements

This paper was written while I visited Nuffield College, Oxford and I gratefully acknowledge discussion with D. R. Cox in that period.

References

Abadir, K. M. (1993), *The limiting distribution of the autocorrelation coefficient under a unit root*, Annals of Statistics **21**, 1058–1070.

Abadir, K. M. (1995), *Unbiased estimation as a solution to testing for random walks*, Economic Letters **47**, 263–268.

Bartlett, M. S. (1937), *Properties of sufficiency and statistical tests*, Proceedings of the Royal Society of London, Series A **160**, 268–282.

Curtiss, J. H. (1942), *A note on the theory of moment generating functions*, Annals of Mathematical Statistics **13**, 430–433.

Larsson, R. (1994), *Bartlett corrections for unit root test statistics*, Preprint, I.M.S., Univ. of Copenhagen.

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

[Figure 1 is not available from WWW]

Figure 1. From left the solid curves (—) are Q-Q plots for the distribution of Q for $n = 3, 4, 6, 8$ plotted against the asymptotic distribution. From right the dashed curves (- -) are Q-Q plots for Q_{SUB} for $n = 3, 12, 25, 50$. The final sample distributions were simulated using 10^5 repetitions and the asymptotic distribution is approximated by a Γ distribution.

Appendix

The reported numerical results can easily be checked by a graphical presentation based on simulations. The moment expansions are of the form

$$EX = a + n^{-1}b + n^{-2}c$$

where the moments are estimated by $\widehat{EX}$, say, using 10^7 repetitions. The residual function

$$f(n^{-2}) = \widehat{EX} - a - n^{-1}b$$

is approximately linear as $1/n^2 \to 0$. Moreover, for the first two moments the third order approximations c/n^2 have been calculated and appear as tangents to the residual function in the figures.

The figures are based on the following momentexpansions:

Figure A.1:	$\widehat{EQ} = 1.142 - n^{-1}0.01954 + n^{-2}0.9506.$
Figure A.2:	$\widehat{EQ_{SUB}} = 1.142 - n^{-1}1.781 + n^{-2}4.206.$
Figure A.3:	$\widehat{EQ^2} = 3.524 - n^{-1}0.6004 + n^{-2}0.5681.$
Figure A.4:	$\widehat{EQ^2_{SUB}} = 3.524 - n^{-1}11.29 + n^{-2}48.72.$
Figure A.5:	$\widehat{EQ^3} = 17.73 - n^{-1}12.34.$
Figure A.6:	$\widehat{EQ^3_{SUB}} = 17.73 - n^{-1}92.32.$
Figure A.7:	$\widehat{EQ^4_{SUB}} = 124.2 - n^{-1}70.98.$
Figure A.8:	$\widehat{EQ^4_{SUB}} = 124.2 - n^{-1}814.7.$

In Figure A.3 the tangent is almost horisontal while the residual function is curved. Thus in Figure A.9 the third order residual

$$g(n^{-3}) = \widehat{EQ^2} - a - n^{-1}b - n^{-2}c$$

is plotted against $1/n^3$ and it seems to be linear as $1/n^3 \to 0$.

[The figures are not available from WWW]