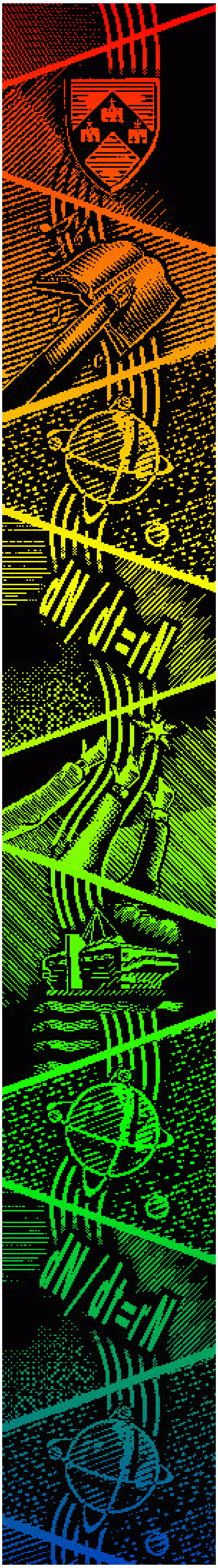

THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 2000/58

Edgeworth Expansions in Gaussian Autoregression

by

Patrick Marsh

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Edgeworth Expansions in Gaussian Autoregression

Patrick Marsh*

Department of Economics and

Department of Mathematics

University of York

U.K.

November 27, 2000

Abstract

We consider the construction of valid Edgeworth expansions for statistics arising in the context of Gaussian autoregression. By exploiting the properties of exponential families (to which these models belong), validity, *of any order*, is routinely established for a wide class of statistics.

*Address for correspondence: Department of Economics, University of York, Heslington, York, YO10 5DD, U.K. e-mail: pwnm1@york.ac.uk.

1 Introduction

The application of higher-order asymptotic techniques for autocorrelation estimators in Gaussian autoregression has received considerable attention. Examples include derivation of $o(n^{-1})$ Edgeworth series contained in Phillips (1977), Ochi (1983), Satchel (1984), Bose (1988), Taniguchi (1991) and more recently Kakizawa (1999), while Phillips (1978) and Durbin (1980) (for the circular case) derive variants of the saddlepoint approximation. For the purposes of this paper, we suppose that a sample $y = (y_1, ..., y_N)'$ was generated by the following Gaussian $AR(p)$ process,

$$\begin{aligned} y_i &= \alpha_1 y_{i-1} + \alpha_2 y_{i-2} + ... + \alpha_p y_{i-p} + \varepsilon_i \qquad (1) \\ \varepsilon_i &\sim IN(0, \sigma^2) \quad ; \quad y_{-p+1} = y_{-p+2} = ... = y_0 = 0, \end{aligned}$$

and we assume that the roots of the polynomial $(z^p + \alpha_1 z^{p-1} + ... + \alpha_p)$ lie inside the unit circle, so that (1) is stationary.

The major stumbling block for applications of higher-order asymptotic tools in models such as (1) is, as one might expect, demonstrating validity. In general, let X_N be a $k \times 1$ random vector, having cumulants κ^{I_v}, for $v = 1, 2, ...$, (for details of the 'index notation' see McCullagh (1987), Chapter 2) and suppose that as $N \to \infty$, $X_N \to_d N_k(0, \kappa^{I_2})$. Expansion and term-by-term inversion of the cumulant generating function of X_N yields an Edgeworth series for the distribution, $F_N(x) = \Pr(X_N < x)$, as

$$\hat{F}_N(x) = \Phi(x; \kappa^{I_2}) - \phi(x; \kappa^{I_2}) \left\{ \sum_{j=3}^{b} c_{j,N}(\kappa)\, q_j(x) \right\}, \qquad (2)$$

where the $c_{j,N}(\kappa)$ are coefficients involving the sample size and the cumulants of X_N, the $q_j(x)$ are the tensoral Hermite polynomials and Φ and ϕ represent the k-dimension normal CDF and PDF respectively. The series in (2) is valid if

$$\sup_x \left| F_N(x) - \hat{F}_N(x) \right| = o(N^{-(b-2)/2}).$$

Although necessary conditions for validity are well established, (see Bhattacharya and Ghosh (1978) or Durbin (1980)), to demonstrate validity in non-i.i.d. data settings the following is sufficient (for example, see Hall (1992), Chapter 2 and Taniguchi (1991), p.14).

Assumption 1 Let $\kappa_X^{I_j}$ be the j^{th} cumulant of X_N, then

$$\kappa_X^{I_j} = N^{-(j-2)/2} \sum_{l=1}^{\infty} \kappa_{X,l}^{I_j} N^{-(l-1)/2}, \tag{3}$$

where the cumulant coefficients $\kappa_{X,l}^{I_j}$ are free of N and we also assume $\kappa_{X,1}^{I_1} = 0$.

For the special case of the autoregressive model (1), with $p = 1$, validity up to order $o(N^{-1})$ (as, for example, in Phillips (1977) and Kakizawa (1999)) is demonstrated by finding low order cumulant coefficients, and showing that, for $j = 1, .., 4$,

$$\left| \kappa_X^{I_j} - N^{-(j-2)/2} \kappa_{X,1}^{I_j} - N^{-j/2} \kappa_{X,2}^{I_j} \right| = o(N^{-j/2}).$$

In this paper we demonstrate that in order to prove validity, such calculations are in fact not necessary. That is by fully exploiting the properties of the model, validity, up to order $o(N^{-(b-2)/2})$, not just $o(N^{-1})$ is assured, for a wide class of statistics arising from this model. Moreover the approach of this paper seems more straightforward and general than that of Taniguchi and Watanabe (1994), who derive $o(N^{-1})$ approximations for the MLE in curved exponential models. The key is the fact that joint distributions of samples generated by Gaussian autoregression are simply members of the exponential family and consequently the majority of inference will be conducted through simple functions of the sufficient statistic. Therefore, in order to establish validity we need only prove validity for (a) the sufficient statistic itself and (b) simple functions of it. Once validity has been established for Edgeworth series a fuller range of higher order techniques such as Bootstrap (Hall (1992)) and transformation methods (Niki and Konishi (1986)) may be used.

2 Main Results

Let $y = (y_1, .., y_N)$ have the usual curved exponential density (see for example, Barndorff-Nielsen and Cox (1989)), viz.

$$f(y;\theta) = \exp\{t_N'\eta - K_N(\eta) + h(t_N)\}, \tag{4}$$

with k-dimension sufficient statistic $t_N = t(y)$ and canonical parameter $\eta = \eta(\theta)$, a smooth function of the d-dimension parameter θ and cumulant function $K_N(\eta) = K_N(\eta(\theta))$. Before specialising to autoregression we examine the validity of asymptotic expansions of the type (2), for the distribution of t_N. Let $s_N = N^{1/2}t_N$, and denote the cumulants of s_N by $\kappa_s^{I_j}$, then s_N is also minimal sufficient and (4) may be reparameterised, with $\gamma = \gamma(\theta)$, as

$$f(y;\theta) = \exp\{s_N'\gamma - K_N(\gamma) + h^*(s_N)\}.$$

Moreover, the cumulant generating function of t_N is $K_t(\lambda) = K_N(\eta + i\lambda) - K_N(\eta)$, and so t_N has cumulants

$$\kappa_t^{I_j} = (i)^{-j} \left. \frac{\partial^j K_N(\eta + i\lambda)}{\partial\lambda_1^{j_1}..\partial\lambda_k^{j_k}} \right|_{\lambda=0} = (i)^{-j} N^{-j/2} \left. \frac{\partial^j K_N(\gamma + i\lambda)}{\partial\lambda_1^{j_1}..\partial\lambda_k^{j_k}} \right|_{\lambda=0} = N^{-j/2}\kappa_s^{I_j},$$

with $\sum_{l=1}^k j_l = j$, so that if $\kappa_s^{I_j} = O(N)$ for all j, then $\kappa_t^{I_j} = O(N^{-(j-2)/2})$, which satisfies Assumption 1, with $\kappa_{t,l}^{I_j} = 0$ for $l \geq 2$. That is validity for the distribution of t_N follows if the cumulants of $N^{1/2}t_N$ are all $O(N)$.

Now, let $g_N = g(x_N)$ be a $m \times 1$ function of the sufficient statistics $x_N = s_N/N = (x_1, ..., x_k)$, satisfying the following assumption.

Assumption 2 $g(x_N)$ is v times differentiable, where $v > b$, with derivatives

$$g_{I_j} = \frac{\partial^j g(x_N)}{\partial x_1^{i_1}..\partial x_k^{i_k}} \quad ; \quad \sum_{l=1}^k i_l = j$$

with the g_{I_v} continuous in a neighbourhood of $\tau = \kappa_s^{I_1} = E[x_N]$ and all minors of g_{I_1} bounded away from zero.

Under Assumption 2, g_N permits the following stochastic expansion

$$g^r = \bar{g}_{I_0}^r + \sum_{j=1}^v \bar{g}_{I_v}^r Z^{I_v} + O_p(N^{-(v+1)/2}) \quad ; \quad r = 1, .., m, \qquad (5)$$

where $\bar{g}_{I_v}^r = \left. \frac{\partial^v g(x_N)}{\partial x_1^{i_1}..\partial x_k^{i_k}} \right|_{x_N = \tau}$ and $Z = (x_N - \tau) = O_p(N^{-1/2})$. We consider approximating the distribution of the standardised statistic

$$h_N = V_N^{-1/2}\left(g(x_N) - g(\tau)\right),$$

where $V_N = var[g(x_N)]$.

Theorem 1 *Assume that the cumulants of* $s_N = N^{1/2}t_N$ *are* $O(N)$ *and that Assumption 2 holds for* $g_N = g(x_N)$ *and let the distribution of* h_N *be* $F_N(h)$*, with Edgeworth series approximation, analogous to (2),* $\hat{F}_N(h)$*, then*

$$\sup_h \left| F_N(h) - \hat{F}_N(h) \right| = o(N^{-(b-2)/2}).$$

Proof. The proof of Theorem 1 is given in the appendix.

A consequence of Theorem 1 is that the only condition required for validity is the existence of a sufficient statistic with cumulants of order $O(N)$. Hence, in Gaussian autoregressions, this is the only condition which needs to be checked. As a comparison, if there exists such a sufficient statistic, then Assumptions 1 and 3 in the approach of Taniguchi and Watanabe (1994) are immediately satisfied.

Specifically, the set of $k = (p+1)(p+2)/2$ statistics, $s_N = (s_1, .., s_k)'$, is the minimal sufficient statistic, see Anderson (1994, p.358), where

$$s_N = \begin{array}{llll} \left\{\sum_{i=p+1}^{T-p} y_i^2, & y_1^2 + y_N^2, & .., & y_p^2 + y_{N-p+1}^2, \right. \\ \sum_{i=p+1}^{T-p+1} y_i y_{i-1}, & y_1 y_2 + y_{N-1} y_N, & .., & y_{p-1} y_p + y_{N-p+1} y_{N-p+2}, \\ \vdots & & & \\ \left. \sum_{i=p+1}^{T} y_i y_{i-p} \right\} & & & \end{array} \qquad (6)$$

Since under normality $y \sim N(0, \Sigma(\theta))$, where $\theta = (\alpha_1, .., \alpha_p, \sigma^2)$ and

$$\Sigma(\theta)^{-1} = \sum_{j=1}^{k} \eta_j(\theta) A_j,$$

for a sequence of constant matrices A_j and smooth functions of θ, $\eta_j(\theta)$ (see van Garderen (1997) and Anderson (1994), Section 6), then the joint density of y may be written

$$f(y;\theta) = \exp\left\{-\frac{1}{2}N^{1/2}t_N'\eta(\theta) + \frac{1}{2}\ln\left|\sum_{j=1}^{k}\eta_j(\theta)A_j\right| - \frac{N}{2}\ln 2\pi\right\}, \tag{7}$$

where $\eta(\theta) = (\eta_1(\theta), .., \eta_k(\theta))'$. Consequently, for the Gaussian autoregression model defined by (1), we have the following theorem.

Theorem 2 *Let the minimal sufficient statistic s_N for model (1), be defined as in (6), then the cumulants of s_N satisfy $\kappa_s^{I_j} = O(N)$ for all j.*

Proof. Theorem 2 is proved in the appendix.

Applying, first Theorem 2, then Theorem 1, yields the following results. First, we may approximate, *to any order*, the density of the sufficient statistics via an appropriate Edgeworth series. Second, we may then transform to any function, say g_N, of those sufficient statistics, provided only that Assumption 2 holds, and approximate *to any order*, the density of g_N. As a consequence, the results of the papers mentioned in the first paragraph of the introduction may be obtained, in principle, as special cases of the results here.

3 Application

Although we have proved validity of higher-order Edgeworth approximations for Gaussian autoregression, validity in itself, is by no means a guarantee of reasonable accuracy. For comparison with previous studies, take the simplest AR(1) process; $y_i = \alpha y_{i-1} + \varepsilon_i$, $\varepsilon_i \sim N(0,1)$ and $i = 1, .., N$, with sufficient statistics

$$s = \begin{pmatrix} s_1 \\ \\ s_2 \end{pmatrix} = \begin{pmatrix} \sum_i y_{i-1}^2 \\ \\ \sum_i y_i y_{i-1} \end{pmatrix},$$

with mean vector $E(s) = \tau = (\tau_1, \alpha\tau_1)$, where $\tau_1 = (N - (1-\alpha^{2N})/(1-\alpha^2))/(1-\alpha^2)$. We will consider the distribution of the z-transformation, see Taniguchi (1991), of the (bias corrected) MLE for α, $\tilde{\alpha} = \frac{N+1}{N}\frac{s_2}{s_1}$

$$h_N(\tilde{\alpha}) = \sqrt{N}\left(\log\left(\frac{1+\tilde{\alpha}}{1-\tilde{\alpha}}\right) - \log\left(\frac{1+\alpha}{1-\alpha}\right)\right).$$

Notice that although $\log\left(\frac{1+\tilde{\alpha}}{1-\tilde{\alpha}}\right)$ is defined only for $\tilde{\alpha} \in (-1,1)$, the probability $1 - \Pr\left[\tilde{\alpha} \in (-1,1)\right]$ is of exponentially small order in N, and hence does not affect our calculations here. Denote the cumulants of h_N by κ_h^j, $j = 1, 2, ..$, then after some algebra, the cumulant coefficients in (3) are

$$\begin{aligned} \kappa_{h,2}^1 &= \frac{-\alpha}{\sqrt{(1-\alpha^2)}} \quad ; \quad \kappa_{h,1}^2 = 1 \quad ; \quad \kappa_{h,1}^3 = 0 \\ \kappa_{h,2}^2 &= \frac{3\alpha^2}{(1-\alpha^2)} \quad ; \quad \kappa_{h,1}^4 = \frac{(2-6\alpha^2)}{(1-\alpha^2)} \end{aligned}$$

Then, applying Theorems 1 and 2 an Edgeworth approximation to $\Pr[h_N \leq h] = F(h)$ is

$$\hat{F}(h) = \Phi(h) - \phi(h)\left(\sum\nolimits_{j=3}^{b} N^{-(j-2)/2} a_j(h)\right) + o(N^{-(b-2)}), \tag{8}$$

with Φ and ϕ the standard normal CDF and PDF, respectively and, for example,

$$\begin{aligned} a_3(h) &= \kappa^1_{h,2} \\ a_4(h) &= \frac{1}{2}((\kappa^1_{h,2})^2 + \kappa^2_{h,2})H_1(h) + \frac{1}{24}\kappa^4_{h,1}H_3(h) \end{aligned} \tag{9}$$

where $H_j(h)$ is the j^{th} Hermite polynomial. Since higher-order Edgeworth approximations are prone to non-monotone behaviour, caused by the highest-order Hermite polynomial in the expansion, see Niki and Konishi (1986), then removing the term involving $\kappa^3_{h,1}$, minimises the risk of non-monotonicity.

To illustrate the problem of non-monotonicity and to highlight the care needed when constructing higher-order asymptotic approximations we will examine the distribution of both $\tilde{\alpha}$ and $h_N(\tilde{\alpha})$. In particular, the empirical distributions of $\tilde{\alpha}$ and $h_N(\tilde{\alpha})$ were simulated for $N = 50$ and $\alpha = 0.9$ with $100,000$ replications. Then $o(N^{-1})$ approximations were constructed for $\tilde{\alpha}$ (using the expansion contained in Ochi (1983) and also Kakizawa (1999), p.346) and for $h_N(\tilde{\alpha})$ upon substitution of (9) into (8). Comparisons between the simulated and approximate distributions are contained in Figures 1 and 2 in the appendix.

Importantly, from Theorems 1 and 2, both approximations are valid, but that for the MLE $\tilde{\alpha}$, even in a moderate sample size, is clearly an unsuitable base for inference about α. Therefore, if we wish to use higher-order asymptotic approximations for, for example, one-sided confidence intervals for unknown parameters, then it is crucial that we apply such techniques to suitable statistics.

Acknowledgments: This paper is derived from my PhD thesis, 'Higher-Order Asymptotics for Econometric Estimators and Tests', for which thanks for patient and helpful supervision go to Grant Hillier. Comments by Karim Abadir, Francesco Bravo, Giovanni Forchini, Soren Johansen, Paul Marriott, Peter Phillips, Mark Salmon and Steve Satchell, participants at 'Differential Geometric Methods in Econometrics', held at EUI, Florence and by an anonymous referee proved most helpful. Financial support in the form of a Leverhulme Special Research Fellowship is gratefully acknowledged.

References

Anderson, T.W. (1994): *The Statistical Analysis of Time Series.* Wiley Classics Library, Wiley, New York.

Barndorff-Nielsen, O.E. and Cox , D. (1989): *Asymptotic Techniques for use in Statistics.* Chapman and Hall, London.

Bhattacharya, R.N. and Ghosh, J.K. (1978): On the validity of the formal Edgeworth expansion. *Annals of Statistics*, 6, 434-451.

Bose, A. (1988): Edgeworth correction by bootstrap in autoregressions. *Annals of Statistics,* 16, 1709-22.

Durbin, J. (1980): Approximations for the densities of sufficient estimates. *Biometrika*, 77, 311-333.

Hall, P. (1992): *The Bootstrap and Edgeworth Expansion*, Springer series in statistics, Springer-Verlag.

Kakizawa, Y. (1999): Valid Edgeworth expansions of some estimators and bootstrap confidence intervals in first-order autoregression. *Journal of Time Series Analysis*, 20, 343–359.

McCullagh, P. (1987): *Tensor Methods in Statistics.* Chapman and Hall, London.

Niki, N. and Konishi, S. (1986): Effects of transformations in higher order asymptotic expansions. *Annals of the Institute of Statistical Mathematics*, 38, 371-383.

Ochi, Y. (1983): Asymptotic expansions for the distribution of an estimator in the first-order autoregressive process. *Journal of Time Series Analysis*, 4, 57–67.

Phillips, P.C.B. (1977): Approximations to some finite Sample distributions associated with a first-order stochastic difference equation. *Econometrica*, 45, 463-485.

——— (1978): Edgeworth and saddlepoint approximations in a first order non-circular autoregression. *Biometrika,* 65, 91-98.

Satchell, S. (1984): Approximation to the finite sample distribution for non-stable first-order stochastic difference equation. *Econometrica,* 52, 1271-1289.

Taniguchi, M. (1991): *Higher Order Asymptotic Theory for Time Series Analysis.* Lecture notes in statistics, Springer-Berlin.

Taniguchi, M. and Watanabe, Y. (1994): Statistical Analysis of Curved Probability Densities. *Journal of Multivariate Analysis*, 48, 228-248.

van Garderen, K-J. (1997): Curved exponential models in statistics. *Econometric Theory*, 15, 771-790.

Appendix

Proof of Theorem 1

The proof requires only that we show that Assumption 1 holds for h_N. Without loss of generality, assume $\tau = g(\tau) = 0$ and let $g = g_N$ and $x = x_N$, the joint density of y may be written, for $\beta = \beta(\theta)$, as

$$f(y;\theta) = \exp\{\beta' x - K_N(\beta) + r(x)\}$$

and the characteristic function of g is

$$\begin{aligned} M_g(\zeta) &= \int_{\mathbb{R}^k} \exp\{i\zeta' g + \beta' x - K_N(\beta) + r(x)\} dx \\ &= e^{-K_N(\beta)} \int_{\mathbb{R}^k} \exp\{i\zeta' g + \beta' x + r(x)\} dx. \end{aligned} \tag{10}$$

The integral on the RHS of (10) is equal to the complex Laplace transform:

$$\mathcal{L}\left(\exp\{i\zeta' g\} \exp\{r(x)\}\right), \tag{11}$$

where $\mathcal{L}(u(x)) = \int_{\mathbb{R}^k} u(x) e^{-\eta' x} dx$, denotes the Laplace transform operator. Now we know

$$\mathcal{L}(\exp\{r(x)\}) = \exp\{K_N(\beta)\},$$

and expanding $\exp\{i\zeta' g\}$ on account of (5),

$$\exp\{i\zeta' g\} = 1 + \sum\nolimits_{j=1}^{v} b_{I_j} x^{I_j} + O_p(N^{-(v+1)/2}),$$

where $x^{I_j} = x^{i_1} .... x^{i_k}$, and the b_{I_j} are $O(1)$ tensor coefficients in the elements of $i\zeta$ and the derivatives of $g(.)$, then

$$\mathcal{L}\left(\left\{1 + \sum\nolimits_{j=1}^{v} b_{I_j} x^{I_j}\right\} \exp\{r(x)\}\right) = \sum\nolimits_{j=0}^{v} (-1)^j b_{I_j} \frac{\partial^j \exp\{K_N(\beta)\}}{\partial\beta_1^{i_1} .. \partial\beta_k^{i_k}}, \tag{12}$$

where in (12), $b_{I_0} = 1$, and we have used

$$\mathcal{L}\left(x^{I_v}\exp\{r(x)\}\right) = (-1)^v \frac{\partial^v \exp\{K_N(\beta)\}}{\partial\beta_1^{i_1}..\partial\beta_k^{i_k}}$$

and the linearity and continuity of the Laplace transform. Consequently, the Laplace transform (11) is (noting that Cramér's condition is satisfied automatically in the exponential case, and hence truncation of the integral series is permitted)

$$\mathcal{L}\left(\exp\{i\zeta' g\}\exp\{r(x)\}\right) = \sum\nolimits_{j=0}^{v}(-1)^j b_{I_j}\frac{\partial^j \exp\{K_N(\beta)\}}{\partial\beta_1^{i_1}..\partial\beta_k^{i_k}} + O(N^{-(v+1)/2}),$$

and so,

$$M_g(\zeta) = 1 + e^{-K_N(\beta)}\sum\nolimits_{j=1}^{v}(-1)^j b_{I_j}\kappa_x^{I_j} + O(N^{-(v+1)/2}), \tag{13}$$

that is $M_g(\lambda)$ permits a series expansion in terms of the cumulants of x. Assuming $v > b$, taking logs, and expanding term by term, the cumulant generating function of g is

$$K_g(\zeta) \sim \sum\nolimits_{k=1}^{\infty}(-1)^{k-1}\left(e^{-K_N(\beta)}\sum\nolimits_{j=1}^{v}(-1)^j b_{I_j}\kappa_x^{I_j}\right)^k,$$

where $\sim$ implies asymptotic equivalence of order $O(N^{-(v+1)/2})$. Since $\kappa_x^{I_2} = O(N^{-1})$, then $V_N \sim N^{-1/2}V$, where $V = O(1)$, and so the cumulant generating function of h_N is

$$K_{h_N}(\zeta) \sim \sum\nolimits_{k=1}^{\infty}(-1)^{k-1}\left(e^{-K_N(\gamma)}\sum\nolimits_{j=1}^{v}(-1)^j N^{j/2} d_{I_j}\kappa_x^{I_j}\right)^k,$$

where $d_{I_j} = V^{H_j, I_j} b_{H_j}$ so that, since by definition, $e^{-K_N(\gamma)} = O(1)$ and $\kappa_x^{I_v} = O(N^{-v+1})$, then

$$\left.\frac{\partial^b K_{h_N}(\zeta)}{\partial\zeta_1^{i_1}..\partial\zeta_k^{i_k}}\right|_{\zeta=0} \sim N^{-(b-2)/2}\sum_{l=1}^{v} d_{I_j}^{*} N^{-(l-1)/2},$$

where the $d^*_{I_j}$ are $O(1)$, as required. ■

Proof of Theorem 2

The cumulant generating function of the minimal sufficient statistic is

$$K_t(\lambda) = K_N(\eta + i\lambda) - K_N(\eta),$$

and from (7) the cumulant function is

$$K_N(\eta) = -\frac{1}{2}\ln\left|\sum\nolimits_{j=1}^{k}\eta_j A_j\right|, \tag{14}$$

and hence the cumulants are given by

$$\kappa_t^{I_j}(\theta) = -\frac{(i)^{-j}}{2}\left.\frac{\partial^j \ln\left|\sum_{j=1}^{k}(\eta_j + i\lambda_j)A_j\right|}{\partial\lambda_1^{j_1}..\partial\lambda_k^{j_k}}\right|_{\lambda=0}, \quad \sum\nolimits_{i=1}^{k} j_i = j,$$

where $\theta = \{\alpha_1, .., \alpha_p, \sigma^2\}$ forms the natural parameter in the $AR(p)$ model. Defining $\sum_{j=1}^{k}(\eta_j + i\lambda_j)A_j = A(\lambda)$ say, then note the following identities

$(a) : \frac{\partial \ln|A(\lambda)|}{d\lambda_j} = Tr\left[A^{-1}(\lambda)\frac{\partial A(\lambda)}{d\lambda_j}\right]$ $\quad (b) : \frac{\partial A^{-1}(\lambda)}{d\lambda_j} = -A^{-1}(\lambda)\frac{\partial A(\lambda)}{d\lambda_j}A^{-1}(\lambda)$

$(c) : \frac{\partial Tr[A(\lambda)]}{d\lambda_j} = Tr\left[\frac{\partial A(\lambda)}{d\lambda_j}\right]$ $\quad (d) : \lim_{\lambda\to 0} A^{-1}(\lambda) = \Sigma(\theta).$

Applying (a), (b) and (c) gives

$$\frac{\partial^j \ln\left|\sum_{j=1}^{k}(\eta_j + i\lambda_j)A_j\right|}{\partial\lambda_1^{j_1}..\partial\lambda_k^{j_k}} \propto Tr\left[\prod_{v=1}^{k}\left(A^{-1}(\lambda)A_v\right)^{j_v}\right]$$

and hence from (d)

$$\left.\frac{\partial^j \ln\left|\sum_{j=1}^{k}(\eta_j + i\lambda_j)A_j\right|}{\partial\lambda_1^{j_1}..\partial\lambda_k^{j_k}}\right|_{\lambda=0} \propto Tr\left[\prod_{v=1}^{k}\left(\Sigma(\theta)A_v\right)^{j_v}\right].$$

Since the A_v are constant matrices then

$$Tr\left[\prod_{v=1}^{k}(\Sigma(\theta)A_v)^{j_v}\right]=\sum\nolimits_{v=1}^{k}r_v(N)l_v(\theta),$$

where $r_v(N)=rank\,(\Sigma(\theta)A_v)^{j_v}$ and the $l_v(\theta)$ are polynomials in θ, and are thus $O(1)$ for all v and J. Consequently, defining

$$r_{\max}(\theta)=\max_{v,J}\ rank\,(\Sigma(\theta)A_v)^{j_v}$$

then

$$Tr\left[\prod\nolimits_{v=1}^{k}(\Sigma(\theta)A_v)^{j_v}\right]\leq\sum\nolimits_{v=1}^{k}r_{\max}(N)l_v(\theta),$$

and further, since $r_{\max}(N)=(N-2)$, then

$$Tr\left[\prod\nolimits_{v=1}^{k}(\Sigma(\theta)A_v)^{j_v}\right]\leq(N-2)\sum\nolimits_{v=1}^{k}l_v(\theta)=O(N),$$

and the theorem is proved. ■

Figures

Fig. 1: Simulated (solid) and $o(N^{-1})$ Edgeworth

approximation (dotted) for the distribution of $\tilde{\alpha}$

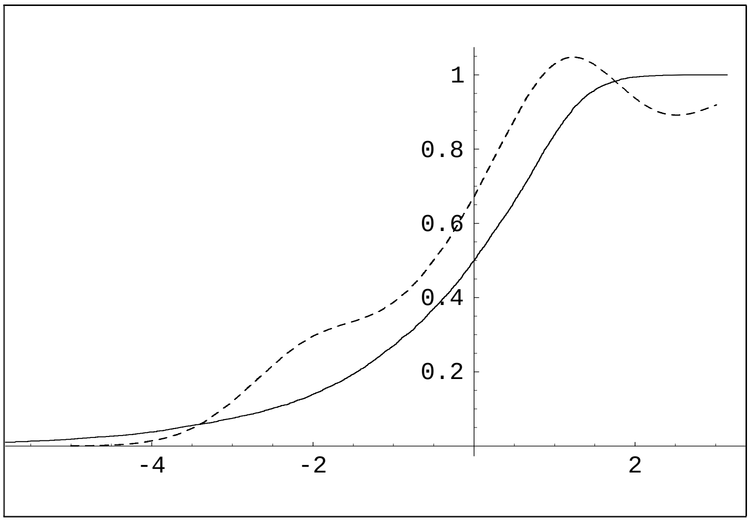

Fig. 2: Simulated (solid) and $o(N^{-1})$ Edgeworth

approximation (dotted) for the distribution of $h_N(\tilde{\alpha})$