International Journal of Statistics and Probability; Vol. 12, No. 5; September 2023
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Delta Method Confidence Intervals for Linear Regression Processes With Long-memory Disturbances

Mosisa Aga

Correspondence: Department of Mathematics, Auburn University at Montgomery, Montgomery, AL, USA

Received: May 16, 2023 Accepted: August 23, 2023 Online Published: September 22, 2023

doi:10.5539/ijsp.v12n5p12 URL: https://doi.org/10.5539/ijsp.v12n5p12

Abstract

This paper provides third and fourth-order coverage probability errors of delta method confidence intervals (CIs) for the covariance parameters of a time series generated by a linear regression model with strongly dependent errors. The CIs are based on the plug-in maximum likelihood (PML) estimators. Bounds have been established on the coverage probability errors of one-and two-sided delta method CIs based on the plug-in log-likelihood (PLL) function under some sets of conditions on the regression coefficients, the spectral density function, and the parameter values. It is shown that the the fourth order delta method CIs in the case of linear regression model with Gaussian, stationary and strongly dependent errors have coverage probability errors of $O(n^{-1})$ and that of the third-order has errors of $O(n^{-1/2})$ which is the same order of magnitude asymptotically as in the independent and identically distributed (iid) case.

Keywords: Confidence interval, delta method, Edgeworth expansion, Gaussian process, linear regression model, long memory process, maximum likelihood estimator, plug-in likelihood function

AMS (2000) Subject Classification Number: Primary: 62M10

1. Introduction

Let $\{\omega_i, i \geq 1\}$ be a discrete time stationary and Gaussian process with unknown mean μ_0 and spectral density $f_\theta(\lambda)$ for $\lambda \in (-\pi, \pi)$, where $\theta = (d, \theta_2, \ldots, \theta_r)' \in \mathbb{R}^r$ and

$$f_\theta(\lambda) = O(|\lambda|^{-2d-\delta}) \tag{1.1}$$

as $|\lambda| \downarrow 0, \forall\delta > 0$, $d \in (0, 1/2)$, and $\theta_1 = d$. The parameter d is generally termed as the fractional 'differencing' operator and when $0 < d < 1/2$, the process whose spectral density satisfies (1.1) is known as a long-memory process with long-memory parameter d. The most well known model for long-memory processes satisfying (1.1) above is the the autoregressive fractionally integrated moving average ARFIMA (p,d,q) process introduced by Hosking (1980) and Grangr et al. (1981) and defined by

$$\Phi(B)X_t = \psi(B)(1 - B)^{-d}\epsilon_t, \tag{1.2}$$

where $\Phi(B) = 1 + \Phi_1 B + ... + \Phi_p B^p$ and $\Psi(B) = 1 + \Psi_1 B + ... + \Psi_q B^q$ are autoregressive and moving-average operators, $\Phi(B)$ and $\Psi(B)$ have no common roots, $d \in (0, \frac{1}{2})$, $(1 - B)^{-d}$ is defined by the binomial formula $(1 - B)^{-d} = \sum_{j=0}^{\infty} \eta_j B^j$, where

$$\eta_j = \frac{\Gamma(j + d)}{\Gamma(j + 1)\Gamma(d)}, \tag{1.3}$$

and Γ is the gamma function, and ϵ_t is a white noise sequence with finite variance. The auto-covariance function of the process in (1.2) is slow decaying leading to the non-summability that corresponds to a pole at the origin in the spectral density function. In general, a process with spectral density satisfying

$$f_\theta(\lambda) \sim |\lambda|^{-\alpha(\theta)} A_\theta(\lambda) \tag{1.4}$$

as $\lambda \to \infty$, with $0 < \alpha(\theta) < 1$, where $\sim$ indicates that the ratio of the left and right sides tend to 1 and $A_\theta(\lambda)$ is slowly varying at 0 in the sense that $\lambda^\delta A_\theta(\lambda)$ is bounded for every δ, is referred to as strongly dependent or long memory by Beran (1992, 1994), Robinson (1994, 1995), Baillie (1996), Taqqu (1986) and others. The ARFIMA (p,d,q) given in (1.2) satisfies (1.4) with $\alpha(\theta) = 2d$ because its spectral density function can be written as:

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}|2\sin\frac{\lambda}{2}|^{-2d}\frac{|\Psi(e^{-i\lambda})|^2}{|\Phi(e^{-i\lambda})|^2} \tag{1.5}$$

and it can be shown that

$$f_\theta(\lambda) \sim \frac{\sigma^2}{2\pi}\frac{|\Psi(1)|^2}{|\Phi(1)|^2} \tag{1.6}$$

for $|\lambda| \to 0$ (see Palma (2007), pp 47.)
Andrews et al. (2006) have determined the coverage probability errors of the delta method confidence intervals for covariance parameters of the process $\{\omega_t\}$ described in (1.1) above. In this paper we establish the coverage probability errors of the delta method confidence intervals for the covariance parameters of a linear regression process whose error terms are the long-meory process $\{\omega_t\}$ by imposing an additional condition on the regression coefficients and a mild condition on the spectral density function.

Our model, which we call *the Gaussian long-memory linear regression model* is described as follows. Let $\beta_0 = \left(\beta_1, \beta_2, \ldots, \beta_p\right)'$ be a vector of deterministic, but unknown real numbers. Let $\{Y_i, i \geq 1\}$ be a process such that: $Y_i = Z_i\beta_0 + \omega_i$, where $Z_i = \left(z_{i1}, z_{i2}, \ldots, z_{ip}\right)$ are non-stochastic regressors for $i = 1, 2, ..., n$ and ω_i are the stationary Gaussian processes mentioned above. With out loss of generality we can assume that the mean of the ω_i is zreo.
Now, the process $\{Y_i, i \geq 1\}$ is non-stationary because the mean of Y_i is $\mu_i = EY_i = Z_i\beta_0$, which varies with i. However, if $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ is an observed sample of size n and $\mathcal{W} = (\omega_1, \omega_2, \ldots, \omega_n)'$, then clearly the covariance matrix of $\mathbf{Y}$ is the same as that of $\mathcal{W}$.
Let $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ be the true mean of $\mathbf{Y}$. Then, the least square estimate (LSE) $\hat{\beta} = \left(\hat{\beta_1}, \hat{\beta_2}, \ldots, \hat{\beta_p}\right)'$ of β_0 is given by $\hat{\beta} = \mathcal{V}^{-1}Z'Y$, where $Z = (z_{ij})$ for $i = 1, ..., n$ and $j = 1, ..., p$ denote the design matrix of our regression model and

$$\mathcal{V} = Z'Z \tag{1.7}$$

is a $p \times p$ matrix. Thus, the estimator of μ is $\hat{\mu} = (\hat{\mu_1}, \ldots, \hat{\mu_n})$, where $\hat{\mu_t} = Z_t\hat{\beta}$, $t = 1, 2, ..., n$. We shall assume the rank of Z is p. Then, the matrix $\mathcal{V}$ (and hence $\mathcal{V}^{-1}$) is symmetric and positive definite.

The $n \times n$ (Toeplitz) covariance matrix corresponding to $f_\theta(\lambda)$ is denoted by $\Gamma_n(f_\theta)$ and has (j, k) element defined by:

$$\Gamma_n(f_\theta)_{j,k} = \int_{-\pi}^{\pi} e^{i(j-k)\lambda} f_\theta(\lambda) d\lambda. \tag{1.8}$$

The log-likelihood function is

$$\Lambda_n(\theta, \mu) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln(\det(\Gamma_n(f_\theta))) - \frac{1}{2}(Y - \mu)'\Gamma_n^{-1}(f_\theta)(Y - \mu). \tag{1.9}$$

We refer to $\Lambda_n(\theta, \hat{\mu})$, where $\hat{\mu}$ is replaced for μ in (1.9) above, as the *plug-in log-likelihood* (PLL) function. Let $A_n = Z\mathcal{V}^{-1}Z'$ and let $K_n = I_n - A_n$, where I_n is the $n \times n$ identity matrix. We note that there exists an $n \times p$ matrix E such that

$$A_n = EE'. \tag{1.10}$$

Using the fact that $(Y - \hat{\mu})' = Y'K_n$, the PLL function can now be written as

$$\Lambda_n(\theta, \hat{\mu}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln(\det(\Gamma_n(f_\theta))) - \frac{1}{2}Y'K_n\Gamma_n^{-1}(f_\theta)K_nY. \tag{1.11}$$

Let $\hat{\Theta}_n$ denote the set of solutions to the first order conditions of the PLL function. That is, $\frac{\partial}{\partial\theta}\Lambda_n(\hat{\theta}_n, \hat{\mu}) = 0$ for all $\hat{\theta}_n \in \hat{\Theta}_n$. If no solution to this condition exists, then we define $\hat{\Theta}_n$ to contain values that maximize the PLL function. Let $\hat{\theta}_n$ denote an element of $\hat{\Theta}_n$. We call $\hat{\theta}_n$ a *plug-in maximum likelihood* (PML) estimator of the true parameter θ_0. By Theorem 2.1 of Dahlhaus (1989), the asymptotic covariance matrix of a consistent PML estimator $\hat{\theta}_n$ is $\Sigma(\theta_0)$, where $\Sigma(\theta) = \left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta}\ln(f_\theta(\lambda))\frac{\partial}{\partial\theta'}\ln(f_\theta(\lambda))d\lambda\right]^{-1}$. A consistent estimator of $\Sigma(\theta_0)$ is $\Sigma(\hat{\theta}_n)$, provided that $f_\theta(\lambda)$ is smooth with respect to θ. Our goal is to formulate the structure of delta method test statistic and confidence intervals of the true parameter θ_0 of the process using $\Sigma(\hat{\theta}_n)$ and then determine the magnitude of the coverage probability errors.
The remainder of the paper proceeds as follows. Section 2 presents some preliminaries and lists a set of assumptions. Section 3 describes the delta method and provides bounds on the third and fourth-order coverage probability errors of one and two-sided delta method confidence interval estimates.

2. Background Preliminaries and Assumptions

2.1 A Brief Description of Delta Method CIs and Tests

Let θ_h denote some element of Θ, the parameter space. Let $\theta_{0,r}$, $\theta_{h,r}$, and $\hat{\theta}_{n,r}$ denote the r-th elements of θ_0, θ_h, and $\hat{\theta}_n$, respectively. Let $\Sigma_{r,r}(\hat{\theta}_n)$ denote the (r,r)-th element of $\Sigma(\hat{\theta}_n)$.
(a) The t statistic for testing the null hypothesis $H_0 : \theta_{0,r} = \theta_{h,r}$ is

$$t_n(\theta_{h,r}) = \frac{\sqrt{n}(\hat{\theta}_{n,r} - \theta_{h,r})}{\Sigma_{r,r}^{1/2}(\hat{\theta}_n)}. \tag{2.1}$$

Let z_α denote the $1-\alpha$ quantile of the standard normal distribution.
(b) The *two-sided delta method* CI for $\theta_{0,r}$ with approximate confidence level $100(1-\alpha)\%$ based on the PML estimator $\hat{\theta}_n$ is

$$I_2(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - z_{\alpha/2}\frac{\Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \hat{\theta}_{n,r} + z_{\alpha/2}\frac{\Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}\right]. \tag{2.2}$$

(c) The *upper one-sided delta method* $100(1-\alpha)\%$ CI for $\theta_{0,r}$ is

$$I_{up}(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - z_\alpha\frac{\Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \infty\right). \tag{2.3}$$

(d) The *two-sided delta method t test* of $H_0 : \theta_{0,r} = \theta_{h,r}$ versus $H_1 : \theta_{0,r} \neq \theta_{h,r}$ with significance level α rejects H_0 if $|t_n(\theta_{h,r})| > z_{\alpha/2}$.
(e) The *one sided t test* of $H_0 : \theta_{0,r} \leq \theta_{h,r}$ versus $H_1 : \theta_{0,r} > \theta_{h,r}$ with significance level α rejects H_0 if $t_n(\theta_{h,r}) > z_\alpha$.

2.2 Background on Cumulants and Edgeworth Expansion

For a random variable U with a characteristic function $\varphi(t) = E(e^{itU})$, the jth cumulant, κ_j, of U is defined to be the coefficient of $\frac{1}{j!}(it)^j$ in a power series expansion of

$$\ln\varphi(t) = \sum_{j\geq 1}\frac{1}{j!}\kappa_j(it)^j. \tag{2.4}$$

Since

$$\varphi(t) = E(e^{itU}) = 1 + E(U)it + \frac{1}{2!}E(U^2)(it)^2 + ... + \frac{1}{j!}E(U^j)(it)^j + ..., \tag{2.5}$$

substituting (2.5) in (2.4) with $E(U^j)$ denoted by μ_j we obtain

$$\sum_{j\geq 1}\frac{1}{j!}\kappa_j(it)^j = \ln\left[1 + \sum_{j\geq 1}\frac{1}{j!}\mu_j(it)^j\right] \tag{2.6}$$

and using the Taylor series expansion identity $\ln(1+x) = \Sigma_{k\geq 1}(-1)^{k+1}\frac{x^k}{k}$, the right hand side of (2.6) above equals

$$\sum_{j\geq 1}(-1)^{k+1}\frac{1}{k}\left(\sum_{j\geq 1}\frac{1}{j!}\mu_j(it)^j\right)^k. \tag{2.7}$$

By equating the coefficients of $\frac{1}{j!}(it)^j$ from the left hand side of (2.6) with those in (2.7) we obtain $\kappa_1 = \mu_1, \kappa_2 = \mu_2 - \mu_1^2, \kappa_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3, \kappa_4 = \mu_4 - 4\mu_1\mu_3 - 3\mu_2^2 + 12\mu_1^2\mu_2 - 6\mu_1^4$, and so on, the formula for κ_{10} for example containing more than 41 such terms. (See Hall, (1992), pp. 41-46.)
Now, let $\hat{\theta}$ be an estimate of the parameter θ_0, constructed from a sample of size n. Under certain conditions $n^{1/2}(\hat{\theta}-\theta_0)$ is asymptotically normally distributed with zero mean and variance σ^2 and for many situations of practical interest the distribution function of $n^{1/2}(\hat{\theta}-\theta_0)$ is expanded as a power series in $n^{-1/2}$ as follows:

$$P(\frac{n^{1/2}(\hat{\theta}-\theta_0)}{\sigma} \leq u) = \Phi(u) + n^{-1/2}p_1(u)\phi(u) + ... + n^{-j/2}p_j(u)\phi(u) + ..., \tag{2.8}$$

where ϕ and Φ are the Standard Normal density and distribution function, respectively, and p_j is a polynomial in terms of cumulants and is of degree $3j-1$. The expansion on the right hand side of (2.8) is known as an Edgeworth expansion of the distribution function on the right. For example, if $U_1, U_2, ..., U_n$ are independent and identically distributed with mean $\mu = \theta_0$ and finite variance σ^2 and if $\hat{\theta}$ represent the sample mean, then p_1 and p_2 are of degrees 2 and 5, respectively, and are given by

$p_1(u) = -\frac{1}{6}\kappa_3(u^2-1)$, and
$p_2(u) = -\frac{1}{24}\kappa_4(u^3-3u) - \frac{1}{72}\kappa_3^2(u^5-10u^3+15u)$.

Details on more general cumulant and Edgeworth expansion can be found in Hall, (1992), and Barndorff et al. (1989).

2.3 Assumptions

Most of the assumptions stated below for the PML estimator are standard assumptions of long-memory processes and have appeared in numerous papers in the literature under different contexts including Fox et al. (1986), Dahlhaus (1989), Lieberman et al. (2003), Andrews et al. (2006), and Palma (2007) among others. Assumption A1 specifies the parameter space for which the results of this paper hold. Assumption A2 states that the sequence of estimators $\bar{\theta}_n$ for which we construct confidence intervals are consistent. Assumption A3 is essentially a modified version of Condition NS_s of Andrews et al. (2006) and Assumptions A4-A8 are also those of Andrews et al. (2006) restated here for convenience. This paper includes two additional assumptions, A9 and A10. Assumption A9 is included to impose a mild additional condition on the spectral density function $f_\theta(\lambda)$, requiring $f_\theta(\lambda)$ to be bounded away from zero. Assumption A10 puts a restriction on the design matrix **Z**. Assumptions A4-A8 below depend on a positive integer $s \geq 3$ that indexes the order of the PLL derivatives that are used in the Edgeworth expansions employed in the proofs of the CI coverage probability results.

A1. The parameter space Θ is a subset of $\mathbb{R}^r$ where $r = dim(\theta)$ with non-empty interior.

A2. For all $\varepsilon > 0$ and all compact subsets Θ_c of Θ, the sequence of PML estimators $\{\bar{\theta}_n : n \geq 1\}$ for which the results of this paper hold satisfy

$$\sup_{\theta_0 \in \Theta_c} P_{\theta_0}(||\bar{\theta}_n - \theta_0|| > n^{-1/2} ln(n)\varepsilon) = o(n^{1-s/2}) \;\; as \;\; n \to \infty$$

for some integer $s \geq 3$.

A3. The matrices $\mathcal{D}_n(\theta)$ and $\mathcal{D}(\theta)$ in (2.23) below are positive definite.

A4. For some integer $s \geq 3$, $f_\theta(\lambda)$ is $s + 1$ times continuously differentiable with respect to θ, and all of the derivatives are continuous in (λ, θ) for $\lambda \neq 0$. In addition, $f_\theta^{-1}(\lambda)$ is continuous in (λ, θ) for all $\lambda \in [0, \pi]$ and $\theta \in \Theta$.

A5. The derivatives $\frac{\partial}{\partial\lambda} f_\theta^{-1}(\lambda)$ and $\frac{\partial^2}{\partial\lambda^2} f_\theta^{-1}(\lambda)$ are continuous in (λ, θ) for $\lambda \neq 0$. In addition, there exists $c_1(\theta, \delta) < \infty$ such that $|\frac{\partial^k}{\partial\lambda^k} f_\theta^{-1}(\lambda)| \leq c_1(\theta, \delta)|\lambda|^{2d-k-\delta}$ for $k = 0, 1, 2$ and all $\delta > 0$, where $\theta = (d, \theta_1, ..., \theta_r)'$ and $d \in (0, 1/2)$.

A6. There exists $c_2(\theta, \delta) < \infty$ and $c_3(\theta, \delta) < \infty$ such that for all $\delta > 0$ and $\lambda \in [0, \pi]$:

(a) $|f_\theta(\lambda)| \leq c_2(\theta, \delta)|\lambda|^{-2d-\delta}$ and

(b) for all $(j_1, ..., j_k)$ with $k \leq s + 1$, with duplication among the j_i allowed, $|\frac{\partial^k}{\partial\theta_{j_1}, ..., \theta_{j_k}} f_\theta^{-1}(\lambda)| \leq c_1(\theta, \delta)|\lambda|^{2d-k-\delta}$.

A7. For any compact subset Θ_c of Θ, there exists a constant $C(\Theta_c, \delta) < \infty$ such that $c_1(\theta, \delta)$, $c_2(\theta, \delta)$, and $c_3(\theta, \delta)$ in a A3 and A4 are bounded by $C(\Theta_c, \delta)$ for all $\theta \in \Theta_c$.

A8. (a) There exists a function $\Omega(\lambda)$ that is integrable over $(0, \pi)$ and a constant $c_4(\theta) < \infty$ such that for all $(j_1, ..., j_k)$ with $k \leq s + 1$, with duplication among the j_i allowed, $|\frac{\partial^k}{\partial\theta_{j_1}, ..., \theta_{j_k}} f_\theta^{-1}(\lambda)| \leq c_4(\theta)\Omega(\lambda)$ for $\lambda \in (0, \pi)$. For any compact subset Θ_c of Θ, there exists a constant $\tilde{C}(\Theta_c) < \infty$ such that $c_4(\theta) \leq \tilde{C}(\Theta_c)$ for all $\theta \in \Theta_c$.

(b) When computing derivatives of the form $\frac{\partial^k}{\partial_{j_1}, ..., \partial_{j_k}} \gamma_\theta(u)$ for $k \leq s + 1$ and $u = 0, 1, 2, ...$, the derivatives may be taken inside the integral sign of (1.8), where $\gamma_\theta(u) = E_\theta(X_i - \mu_i)(X_i - \mu_{i+u})$ and E_θ denotes expectation when the true parameter is θ.

A9. The spectral density function is bounded away from zero on a compact subset of the parametric space.

A10. The design matrix Z is chosen in such a way that for the matrix

$$E = (e_{ij}), i = 1, ..., n, j = 1, ..., p \qquad (2.9),$$

defined in (1.10) there exists a constant $M_0 < \infty$ such that $|e_{ij}| \leq \frac{M_0}{\sqrt{n}}$ for $1 \leq i \leq n$, $1 \leq j \leq p$.

One drawback of the results of this paper is that the design matrix is required to satisfy assumption A10. A natural question would be whether one can find a design matrix of practical value that satisfies the condition imposed in this assumption. The next lemma provides such a matrix. The $n \times p$ matrix $\mathcal{Z}$ given in the lemma is a special case of the so called *Vandermonde matrix* which arises in many applications such as polynomial least squares fitting, Lagrange interpolating polynomials, and the reconstruction of a statistical distribution from the distribution's moments.

Lemma 2.1. There exists a design matrix Z of rank p which satisfies assumption A10.

Proof: Consider the Vandermonde Matrix

$$\mathcal{Z} = \begin{pmatrix} 1 & a_1 & a_1^2 & . & . & a_1^{p-1} \\ 1 & a_2 & a_2^2 & . & . & a_2^{p-1} \\ . & . & . & . & . & \\ 1 & a_n & a_n^2 & . & . & a_n^{p-1} \end{pmatrix}, \tag{2.10}$$

where $1 < a_1 < a_2 < ... < a_n$. We would like to form the matrix E of (1.10) from the design matrix $\mathcal{Z}$ above and show that this matrix satisfies assumption A10. Using the notation $\mathcal{U}$ for the analogue of matrix $\mathcal{V}$ defined in (1.7), we obtain

$$\mathcal{U} = \mathcal{Z}'\mathcal{Z} = \begin{pmatrix} n & \Sigma_{i=1}^n a_i & . & . & \Sigma_{i=1}^n a_i^{p-1} \\ \Sigma_{i=1}^n a_i & \Sigma_{i=1}^n a_i^2 & . & . & \Sigma_{i=1}^n a_i^p \\ . & . & . & . & \\ \Sigma_{i=1}^n a_i^{p-1} & \Sigma_{i=1}^n a_i^p & . & . & \Sigma_{i=1}^n a_i^{2p-2} \end{pmatrix} = \begin{pmatrix} O(n) & O(na_n) & . & . & O(na_n^{p-1}) \\ O(na_n) & O(na_n^2) & . & . & O(na_n^p) \\ . & . & . & . & \\ O(na_n^{p-1}) & O(na_n^p) & . & . & O(na_n^{2p-2}) \end{pmatrix}. \tag{2.11}$$

Looking at the orders of the entries of the matrix $\mathcal{U}$ in (2.11) above we see that its inverse should be of the form:

$$\mathcal{U}^{-1} = \begin{pmatrix} O(\frac{1}{n}) & O(\frac{1}{na_n}) & . & . & O(\frac{1}{a_n^{p-1}}) \\ O(\frac{1}{na_n}) & O(\frac{1}{na_n^2}) & . & . & O(\frac{1}{na_n^p}) \\ . & . & . & . & \\ O(\frac{1}{na_n^{p-1}}) & O(\frac{1}{na_n^p}) & . & . & O(\frac{1}{na_n^{2p-2}}) \end{pmatrix}. \tag{2.12}$$

Let $\lambda_k, k = 1, ..., p$ be the distinct eigenvalues and $u_k = (u_{1k}, u_{2k}, ..., u_{pk})'$ be the corresponding eigenvectors of $\mathcal{U}^{-1}$. λ_k are the roots of the symmetric characteristic polynomial $|\lambda I - \mathcal{U}^{-1}| = 0$, which may be written as

$$(\lambda - O(\frac{1}{n}))(\lambda - O(\frac{1}{na_n^2}))...(\lambda - O(\frac{1}{na_n^{2p-2}})) + O(\frac{1}{na_n^4}) + O(\frac{1}{na_n^5}) + ... + O(\frac{1}{na_n^{p+2}}) = 0. \tag{2.13}$$

From this we see that the vanishing of the characteristic polynomial $|\lambda I - \mathcal{U}^{-1}|$ is determined by the the first term

$$(\lambda - O(\frac{1}{n}))(\lambda - O(\frac{1}{na_n^2}))...(\lambda - O(\frac{1}{na_n^{2p-2}})) \tag{2.14}$$

of (2.13) because all the remaining terms are of order $\leq O(\frac{1}{na_n^4})$). But (2.14) vanishes for $\lambda = O(\frac{1}{n}), O(\frac{1}{na_n^2}), ..., O(\frac{1}{na_n^{2p-2}})$. Thus $O(\lambda_k) \leq O(\frac{1}{n})$ for $k = 1, 2, ..., p$.

Now let's use the notation $\mathcal{U}^{-1} = (v_{ij}), i, j = 1, ..., p$ for the matrix $\mathcal{U}^{-1}$ in (2.12). For each $k = 1, ..., p$, we find the eigenvectors u_k by solving the equations:

$$(v_{i1} - \lambda_k)u_{1k} + v_{i2}u_{2k} + ... + v_{ip}u_{pk} = 0 \tag{2.15}$$

for $i = 1, 2, ..., p$. Let $u_{1k} = 1$. From (2.12) we observe that for $i, k = 1, 2, ..., p$, $v_{ik} = O(\frac{1}{na_n^{k+i-1}})$ and therefore we may replace each v_{ik} by $\frac{N_{k+i}}{na_n^{k+i-1}}$ and each λ_k by $\frac{N}{n}$, where N, N_{k+i} are constants. Replacing these values in (2.15) above the first equation in (2.15) becomes: $(\frac{N_1}{n} - \frac{N}{n}) + \frac{N_2}{na_n^2}u_{2k} + ... + \frac{N_p}{na_n^p}u_{pk} = 0$ and the second equation in (2.15) becomes: $(\frac{N_2}{na_n^2} + (\frac{N_3}{na_n^3} - \frac{N}{n})u_{2k} + ... + \frac{N_{p+1}}{na_n^{p+1}}u_{pk} = 0$. Solving for u_{pk} in the first of the above equations and substituting in the second we obtain: $-na_n^pNu_{2k} + na_n^{p-1}(N_2 - \frac{N_{p+1}N_1}{N_p} - \frac{N_{p+1}N}{N_p}) + O(na_n^{p-2}) + ... + O(n) = 0$. Looking at the first two terms of the above equation we see that u_{2k} must have order $O(\frac{1}{na_n})$. Similarly, $u_{3k} = O(\frac{1}{na_n^2})$ obtained from the first and the 3rd equations of (2.12), and in general $u_{ik} = O(\frac{1}{na_n^{i-1}})$ for $i = 2, ..., p$. Now, the matrix E of (2.9) is given by $E = \mathcal{Z}\mathcal{B}$, where $\mathcal{B} = \mathcal{P}\mathcal{D}\mathcal{P}'$, $\mathcal{P} = (\frac{u_{ij}}{\|u_j\|}), i, j = 1, ..., p$ and $u_{1j} = 1$, and $\mathcal{D}$ is the matrix whose diagonal entries are $\sqrt{\lambda_k}$, for $k = 1, ..., p$ and zero elsewhere. Then we have

$$\mathcal{B} = \begin{pmatrix} \Sigma_{i=1}^p \frac{\sqrt{\lambda_i}}{\|u_i\|^2} & \Sigma_{i=1}^p \frac{u_{2i}\sqrt{\lambda_i}}{\|u_i\|^2} & . & . & \Sigma_{i=1}^p \frac{u_{pi}\sqrt{\lambda_i}}{\|u_i\|^2} \\ \Sigma_{i=1}^p \frac{u_{2i}\sqrt{\lambda_i}}{\|u_i\|^2} & \Sigma_{i=1}^p \frac{u_{2i}^2\sqrt{\lambda_i}}{\|u_i\|^2} & . & . & \Sigma_{i=1}^p \frac{u_{2i}u_{pi}\sqrt{\lambda_i}}{\|u_i\|^2} \\ . & . & . & . & \\ \Sigma_{i=1}^p \frac{u_{pi}\sqrt{\lambda_i}}{\|u_i\|^2} & \Sigma_{i=1}^p \frac{u_{2i}u_{pi}\sqrt{\lambda_i}}{\|u_i\|^2} & . & . & \Sigma_{i=1}^p \frac{u_{pi}^2\sqrt{\lambda_i}}{\|u_i\|^2} \end{pmatrix}. \tag{2.16}$$

Using the above notation of matrix $\mathcal{B}$, we can now form the matrix E as $E = \mathcal{Z}\mathcal{B} = (e_{ij})$ for $i = 1, ..., n$, $j = 1, ..., p$. To show that E satisfies assumption A10, it suffices to show that $O(e_{ij}) \leq O(\frac{1}{\sqrt{n}})$ for $i = 1, 2, ..., n$ and $j = 1, ..., p$. To this end it suffices to show that $O(e_{nj}) \leq \frac{1}{\sqrt{n}}$ for $j = 1, ..., p$ (because $e_{nj} \geq e_{ij}$ for $i = 1, 2, ..., n$ and $j = 1, ..., p$.) Now, for $j = 1$ we have

$$\begin{aligned} O(e_{n1}) &= O\left(\Sigma_{i=1}^{p} \frac{\sqrt{\lambda_i}}{\|u_i\|^2} + (a_n \Sigma_{i=1}^{p} \frac{u_{2i}\sqrt{\lambda_i}}{\|u_i\|^2} + .. + a_n^{p-1}\Sigma_{i=1}^{p} \frac{u_{pi}\sqrt{\lambda_i}}{\|u_i\|^2})\right) \\ &\leq O\left(\frac{p}{\sqrt{n}} + a_n(\frac{1}{na_n}\frac{p}{\sqrt{n}} + .. + a_n^{p-1}\frac{1}{na_n^{p-1}}\frac{p}{\sqrt{n}}\right) = O(\frac{1}{\sqrt{n}}) \end{aligned}. \tag{2.17}$$

A similar calculation shows that $O(e_{nj}) \leq O(\frac{1}{\sqrt{n}})$ for $j = 2, ..., p$ and this completes the proof of the lemma. □

Another drawback of the results of this paper is that the PML estimators are required to satisfy assumption A2, which implies that these estimators are consistent. The same drawback occurs in Andrews et al. (2006), Bhattacharya et al. (1978), and Aga et al. (2007). While Andrews et al. (2006) Lemma 1 provides a sequence of estimators that satisfies this condition, it is generally unknown to date whether or not the result of this paper and those of others in the literature are valid without it.

2.4 Parameter Values and Log-likelihood Derivatives

We begin by specifying the parameter values θ for which we establish delta method confidence intervals. To this end we introduce some additional notations.

Let $\upsilon = (r_1, r_2, ..., r_q)'$ denote a q-vector of positive integers each less than or equal to $r = dim(\theta)$ for $q \leq s$ where s is as given in assumptions A2 and A4. We write the real valued q-th order partial dervative of the PLL function indexed by υas

$$\Lambda_{n,\upsilon} = D_\upsilon \Lambda_n(\theta, \hat{\mu}) = \frac{\partial^q}{\partial\theta_{r_1} \dots \partial\theta_{r_q}} \Lambda_n(\theta, \hat{\mu}) = F_{n,\upsilon}(\theta) + Y' M_n B_{n,\upsilon}(\theta) M_n Y \tag{2.18}$$

where $F_{n,\upsilon}(\theta)$ and $B_{n,\upsilon}(\theta)$ are given by

$$F_{n,\upsilon}(\theta) = -\frac{1}{2} D_\upsilon \ln(\det(\Gamma_n(f_\theta))) = \sum_{k=1}^{b} a_k tr\left(\prod_{j=1}^{p_k} \Gamma_n^{-1}(f_\theta)\Gamma_n(g_{\theta,k,j})\right) \tag{2.19}$$

$$B_{n,\upsilon}(\theta) = -\frac{1}{2} D_\upsilon \Gamma_n^{-1}(f_\theta) = \sum_{k=1}^{b} a_k \left(\prod_{j=1}^{p_k} \Gamma_n^{-1}(f_\theta)\Gamma_n(g_{\theta,k,j})\right)\Gamma_n^{-1}(f_\theta) \tag{2.20}$$

for some constants b, a_k, and p_k that depend on υ and with $g_{\theta,k,j}$ being certain partial derivatives of the spectral density with respect to the components of θ of order q or less. Note that, when computing the partial derivatives of $\Gamma_n(f_\theta)_{j,k}$, we have used assumption A8 (b) to take the derivative in side the integral sign of (1.8) and assumption A4 to compute the derivatives of f_θ.

Delta method confidence intervals of the PML estimator are based on the Edgeworth expansion of a vector of centered and normalized log-likelihood derivatives (LLDs). To specify such a vector, let

$$Z_n(\theta) = (\Lambda_{n,\upsilon(1)}(\theta), \dots, \Lambda_{n,\upsilon(m)}(\theta)), \tag{2.21}$$

where each vector $\upsilon(j)$ is of the same form as υ defined in (2.18)-(2.20) above for $m = dim(Z_n(\theta))$ and $j = 1, 2, ..., m$. Let

$$W_n(\theta) = n^{-1/2}(Z_n(\theta) - E_\theta Z_n(\theta)). \tag{2.22}$$

Without loss of generality we may assume that $E_\theta Z_n(\theta) = 0$. Let

$$D_n(\theta) = E[W_n(\theta)W_n(\theta)'] \tag{2.23}$$

and let $D(\theta) = \lim_{n\to\infty} D_n(\theta)$.

Because $W_n(\theta)$ is a vector of centeral quadratic forms in Gaussian variables plus a vector of nonrandom quantities we have

$$D_n(\theta)_{i,j} = tr\left(B_{n,\nu_i} T_n(f_\theta) B_{n,\nu_j} T_n(f_\theta)\right) \tag{2.24}$$

and the (i, j) element of $D(\theta)$ is given by

$$D(\theta)_{i,j} = \frac{1}{\pi}\int_{-\pi}^{\pi} \{D_{\upsilon_i} f_\theta^{-1}(\lambda)\}\{D_{\upsilon_j} f_\theta^{-1}\} f_\theta^2(\lambda) d\lambda. \tag{2.25}$$

(See Anderson (1984) for details of equation (2.24) and Lemma 8 of Andrews et al. (2006) for that of (2.25).)

3. Delta Method Confidence Intervals

For $u \in \mathbb{R}^m$ let $h_n(u,\theta)$ be the density of $W_n(\theta)$ when the true parameter is θ, where m is as given in (2.22). Let $\tilde{h}_n^{\tau-2}(u,\theta)$ be the formal Edgeworth expansion of $h_n(u,\theta)$ of order $\tau-2$. Let $\xi_n(x,\theta) = E_\theta \exp(ix'Z_n(\theta))$ denote the characteristic function of $Z_n(\theta)$ when θ is the true value, and $x \in \mathbb{R}^m$. Let $\eta = (\eta_1, ..., \eta_q)$ be a q-vector of none-negative integers each of which is less than or equal to m. Define $D_{x,\eta} = \frac{\partial^q}{\partial x_{\eta_1}...\partial x_{\eta_q}}$. Let $\kappa_{n,s}(\theta)_\eta$ denote the η cumulants of $Z_n(\theta)$. By definition, $\kappa_{n,s}(\theta)_\eta = i^{-q} D_{x,\eta} \ln(\xi_n(x,\theta))|_{x=0}$, where $i = \sqrt{-1}$. The vector $\kappa_{n,s}(\theta)$ is composed of elements $\kappa_{n,s}(\theta)_\eta$ for vectors η of dimension $q \le s$. The following lemma gives the expilicit form of the Edgeworth expansion of the density function $h_n^{\tau-2}(u,\theta)$ of $W_n(\theta)$.

Lemma 3.1 Suppose assumptions A1-A10 hold and let $W_n(\theta)$ be as defined in (2.14). Then, the formal Edgeworth expansion $\tilde{h}_n^{\tau-2}(u,\theta)$ of the density of $W_n(\theta)$ is given by

$$\tilde{h}_n^{\tau-2}(u,\theta) = \frac{1}{\sqrt{|D_n(\theta)|(2\pi)^m}} \exp\{-\frac{1}{2}u'D_n^{-1}(\theta)u\}\left[1 + \sum_{j=3}^{\tau} n^{-(j-2)/2} P_{nj}(x,\theta)\right] \tag{3.1}$$

where the polynomial P_{nj} is as defined in (3.10) below.

Proof.

Let

$$\zeta_n(x,\theta) = \frac{1}{n}\ln \xi_n(x,\theta) + \frac{1}{2}x'D_n(\theta)x, \tag{3.2}$$

and denote by $\zeta_{n\tau}(x,\theta)$ the Taylor series approximation without remainder term of $\zeta_n(x,\theta)$ in powers of the components of x from 3 to τ, that is

$$\zeta_{n\tau}(x,\theta) = \sum_{j=3}^{\tau} \Sigma_j \left[\frac{\partial^j \zeta_n(x,\theta)}{\partial x_1^{j_1}...\partial x_m^{j_m}}\right]_{x=0} \frac{x_1^{j_1}...x_m^{j_m}}{j_1!...j_m!}, \tag{3.3}$$

where Σ_j denotes summation over integers $j_1, ..., j_m \ge 0$ such that $\Sigma_k j_k = j$, and m is as in (2.22). Thus, combining this with (3.2) and neglecting the reminder term of the Taylor expansion (3.3), we obtain

$$\zeta_{n\tau}(\frac{x}{\sqrt{n}},\theta) = \frac{1}{n}\ln\{\xi_n(\frac{x}{\sqrt{n}},\theta)\exp\{\frac{1}{2}x'D_n(\theta)x\}\}, \tag{3.4}$$

which is equivalent to

$$\exp\{n\zeta_{n\tau}(\frac{x}{\sqrt{n}},\theta)\} = \xi_n(\frac{x}{\sqrt{n}},\theta)\exp\{\frac{1}{2}x'D_n(\theta)x\}. \tag{3.5}$$

Expanding the left hand side of (3.5) we get

$$1 + \sum_{j=1}^{\infty} \frac{1}{j!}\{n\zeta_{n\tau}(\frac{x}{\sqrt{n}},\theta)\}^j = \xi_n(\frac{x}{\sqrt{n}},\theta)\exp\{\frac{1}{2}x'D_n(\theta)x\}. \tag{3.6}$$

We let

$$\sum_{j=1}^{\tau-2} \frac{1}{j!}\{n\zeta_{n\tau}(\frac{x}{\sqrt{n}},\theta)\}^j = \sum_{j=3}^{\tau} n^{-(j-2)/2}\pi_{nj}(x,\theta) + \alpha_{nr}(x,\theta), \tag{3.7}$$

be the expansion in powers of $n^{-1/2}$, where π_{nj} is a polynomial in powers of the components of x from 3 to j and α_{nr} regarded as a polynomial in powers of $n^{-1/2}$, is of degree greater than $\tau-2$ with coefficients which are polynomials in the components of x. Thus α_{nr} contributes an amount of order $n^{-(\tau-1)/2}$ which, in fact, we shall neglect.

Neglecting α_{nr} and combining (3.6) with (3.7) we obtain

$$\exp\{-\frac{1}{2}x'D_n(\theta)x\} + \exp\{-\frac{1}{2}x'D_n(\theta)x\}\sum_{j=3}^{\tau} n^{-(j-2)/2}\pi_{nj}(x,\theta) = \xi_n(\frac{x}{\sqrt{n}},\theta). \tag{3.8}$$

We note that the left hand side of (3.8) is only an approximate of the right hand side. Now, let $\tilde{h}_n^{\tau-2}(u,\theta)$ be the Fourier transform of the left hand side of (3.8). Then we have

$$\begin{aligned}\tilde{h}_n^{\tau-2}(u,\theta) &= \frac{1}{\sqrt{|D_n(\theta)|(2\pi)^m}}\exp\{-\tfrac{1}{2}u'D_n^{-1}(\theta)u\} \\ &+ \sum_{j=3}^{\tau} n^{-(j-2)/2}\frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}\exp\{-ix'u - \tfrac{1}{2}x'D_n(\theta)x\}\pi_{nj}(x,\theta)dx.\end{aligned} \tag{3.9}$$

Now, we define the polynomial P_{nj} by

$$\frac{1}{\sqrt{|D_n(\theta)|(2\pi)^m}}\exp\{-\frac{1}{2}u'D_n^{-1}(\theta)u\}P_{nj}(u,\theta)=\frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}\exp\{-ix'u-\frac{1}{2}x'D_n(\theta)x\}\pi_{nj}(x,\theta)dx. \quad (3.10)$$

As shown by Hall (1992) page 244, the polynomial P_{nj} is a polynomial in terms of $-D$ and $\kappa_{n,s}(\theta)$, where $D=\left(\frac{\partial}{\partial x^{(1)}},...,\frac{\partial}{\partial x^{(m)}}\right)'$ and $\kappa_{n,s}(\theta)$ is the vector cumulant defined in the paragraph preceding equation (3.1) above. Combining (3.9) and (3.10) above we see that the explicit form of the formal Edgeworth expansion $\tilde{h}_n^{\tau-2}(u,\theta)$ of the density of $W_n(\theta)$ is given by

$$\tilde{h}_n^{\tau-2}(u,\theta)=\frac{1}{\sqrt{|D_n(\theta)|(2\pi)^m}}\exp\{-\frac{1}{2}u'D_n^{-1}(\theta)u\}\left[1+\sum_{j=3}^{\tau}n^{-(j-2)/2}P_{nj}(x,\theta)\right]. \quad (3.11)$$

The following lemma is essentially Theorem 3.2 of Aga (2011) and is an important ingredient in the proof of the main results of this paper.

Lemma 3.2. Suppose assumptions A1-A10 hold. Then for all compact sets $\Theta_c\subset\Theta$ and all $\tau\geq 3$,

(a) $\sup_{\theta_0\in\Theta_c}\sup_{u\in\mathbb{R}^{d_s}}|h_n(u,\theta_0)-\tilde{h}_n^{\tau-2}(u,\theta_0)|=o(n^{-(\tau-2)/2})$ and

(b) $P_{\theta_0}(W_n(\theta_0)\in C)=\int_C\tilde{h}_n^{\tau-2}(u,\theta_0)du+o(n^{-(\tau-2)/2})$, uniformly over all Borel sets C and all $\theta_0\in\Theta_c$.

Our first main result determines the magnitude of the probability error of a two-sided delta method confidence interval and is stated in part (b) of the next theorem. We introduce some additional notations. Let $\Phi(\cdot)$ denote the distribution function of the standard normal random variable. Let $\bar{\kappa}_{n,s}(\theta)=\frac{\kappa_{n,s}(\theta)}{n}$. By Lemma 4.5(c) of Aga et al. (2007), the elements of $\bar{\kappa}_{n,s}(\theta)$ are $O(1)$. Let $p_i(\delta,\bar{\kappa}_{n,s}(\theta))$ be a polynomial in $\delta=\partial/\partial z$ whose coefficients are polynomials in the elements of $\bar{\kappa}_{n,s}(\theta)$ and for which $p_i(\delta,\bar{\kappa}_{n,s}(\theta))\Phi(z)$ is an even function of z when i is odd and an odd function of z when i is even for $i=1,2,...,s-2$. The Edgeworth expansion of the delta method t-statistic $t_n(\theta_{0,r})$ given in (2.1) depends on $p_i(\delta,\bar{\kappa}_{n,s}(\theta_0))$.

Theorem 3.3. Suppose assumption A1-A10 hold and let $s\geq 3$ be as given in assumption A2. Then, for all $\varepsilon>0$,

(a) $sup_{\theta_0\in\Theta_c}\sup_{z\in\mathbb{R}}|P_{\theta_0}(t_n(\theta_{0,r})\leq z)-[1+\sum_{i=1}^{s-2}n^{-i/2}p_i(\delta,\bar{\kappa}_{n,s}(\theta_0))]\Phi(z)|=o(n^{-(s-2)/2})$.
(b) $\sup_{\theta_0\in\Theta_c}P_{\theta_0}(\theta_0\in I_2(\bar{\theta}_n))=(1-\alpha)+O(n^{-1})\ \ for\ \ s=4$.

Proof

(a) Let $n^{-1}Z_n^+(\theta_0)$ denote the vector $n^{-1}Z_n(\theta_0)$ of normalized LLDs augmented to include the vector of expected values of all partial derivatives with respect to θ of order s of $n^{-1}\Lambda_n(\theta_0)$. By Theorem 3(b) of Bhattacharya and Ghosh (1978) the normalized PML estimator and the t statistic $t_n(\theta_{0,r})$ can be approximated by smooth functions of $n^{-1}Z_n^+(\theta_0)$. Specifically, there is an infinitely differentiable function $G(.)$ that does not depend on θ_0 that satisfies $G(n^{-1}E_{\theta_0}Z_n^+(\theta))=0$ for all n large and all $\theta_0\in\Theta_c$ and

$$\sup_{\theta_0\in\Theta_c}\sup_{B\in\mathcal{B}_d}|P_{\theta_0}(t_n(\theta_0)\in B)-P_{\theta_0}(n^{1/2}G(n^{-1}Z_n^+(\theta_0))\in B)|=o(n^{-(s-2)/2}). \quad (3.12)$$

Now, let $\mathcal{P}_s(\theta_0)=1+\sum_{i=1}^{s-2}n^{-i/2}p_i(\delta,\bar{\kappa}_{n,s}(\theta_0))$. Then,

$$\begin{aligned}|P_{\theta_0}(t_n(\theta_{0,r})\leq z)-\mathcal{P}_s(\theta_0)\Phi(z)| &\leq |P_{\theta_0}(t_n(\theta_{0,r})\leq z)-P_{\theta_0}(n^{1/2}G(n^{-1}Z_n^+(\theta_0))\leq z)|\\ &+ |P_{\theta_0}(n^{1/2}G(n^{-1}Z_n^+(\theta_0))\leq z)-\mathcal{P}_s(\theta_0)\Phi(z)|.\end{aligned} \quad (3.13)$$

The first term of the right hand side of (3.13) is equal to $o(n^{-(s-2)/2})$ by Lemma 10 of Andrews *et al.* [2006]. Thus, we only need to show that

$$|P_{\theta_0}(n^{1/2}G(n^{-1}Z_n^+(\theta_0))\leq z)-\mathcal{P}_s(\theta_0)\Phi(z)|=o(n^{-(s-2)/2}) \quad (3.14)$$

uniformly over $\theta_0\in\Theta_c$. That is, it suffices to show that $n^{1/2}G(n^{-1}Z_n^+(\theta_0))$ possesses an Edgeworth expansion given in the present Lemma with remainder $o(n^{-(s-2)/2})$ uniformly over $\theta_0\in\Theta_c$. We note that an Edgeworth expansion of $W_n(\theta_0)=n^{-1/2}(Z_n(\theta_0)-E_\theta Z_n(\theta_0))$ is established by lemma 3.2 above for each $\theta_0\in\Theta_c$. Now, an Edgeworth expansion for $n^{1/2}G(n^{-1}Z_n^+(\theta_0))$ is obtained from that of $n^{-1/2}(Z_n(\theta_0)-E_\theta Z_n(\theta_0))$ by the argument of Battacharya and Ghosh (1978), Theorem 2, p. 436 using the smoothness of G(.), $G(n^{-1}E_{\theta_0}Z_n^+(\theta_0))=0$ for all $n\geq 1$ and all $\theta_0\in\Theta_c$, and assumption A3.
(b) From the definition of $I_2(\bar{\theta}_n)$ given in section 2.1 we observe that

$$\begin{aligned}P_{\theta_0}(\theta_0\in I_2(\bar{\theta}_n)) &= P_{\theta_0}(\bar{\theta}_{n,r}-z_{\alpha/2}\frac{\sqrt{\Sigma_{r,r}}}{\sqrt{n}}\leq\theta_{0,r}\leq\bar{\theta}_{n,r}-z_{\alpha/2}\frac{\sqrt{\Sigma_{r,r}}}{\sqrt{n}})\\ &= P_{\theta_0}(-z_{\alpha/2}\leq\frac{\sqrt{n}(\theta_{0,r}-\bar{\theta}_{n,r})}{\sqrt{\Sigma_{r,r}(\bar{\theta}_n)}}\leq z_{\alpha/2})\\ &= P_{\theta_0}(|t_n(\theta_{0,r})|\leq z_{\alpha/2})\end{aligned} \quad (3.15)$$

and that

$$P_{\theta_0}(|t_n(\theta_{0,r})| \le z_{\alpha/2}) = P_{\theta_0}(t_n(\theta_{0,r}) \le z_{\alpha/2}) - P_{\theta_0}(t_n(\theta_{0,r}) \le -z_{\alpha/2}). \tag{3.16}$$

Denoting $1 + n^{-1/2}p_1(\delta, \bar{\kappa}_{n,4}(\theta_0)) + n^{-1}p_2(\delta, \bar{\kappa}_{n,4}(\theta_0))$ by $\mathcal{P}_4(\theta_0)$ as in (3.13) and substituting z by $z_{\alpha/2}$ and s by 4 in part (a) of this theorem we obtain

$$|P_{\theta_0}(|t_n(\theta_{0,r})| \le z_{\alpha/2}) - \mathcal{P}_4(\theta_0)(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}))| = o(n^{-1}). \tag{3.17}$$

We observe that $p_1(\delta, \bar{\kappa}_{n,4}(\theta_0))\Phi(z)$ is an even function in z which leads to $p_1(\delta, \bar{\kappa}_{n,4}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = 0$. It follows that

$$\sup_{\theta_0 \in \Theta_c} |P_{\theta_0}(|t_n(\theta_{0,r})| \le z_{\alpha/2}) - [1 + n^{-1}p_2(\delta, \bar{\kappa}_{n,4}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}))]| = o(n^{-1}). \tag{3.18}$$

Because $\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) = 1 - \alpha$ and $n^{-1}p_2(\delta, \bar{\kappa}_{n,4}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = o(n^{-1})$, this establishes part (b).□

Theorem 3.4. Suppose assumptions A1-A10 hold and let $s = 3$. Consider $\{\bar{\theta}_n : n \ge 1\}$ as given in assumption A2 and let Θ_c be any compact subset of Θ, the parameter space. Then,

(a) $\sup_{\theta_0 \in \Theta_c} P_{\theta_0}(\theta_0 \in I_2(\bar{\theta}_n)) = (1 - \alpha) + O(n^{-1/2})$.

(b) $\sup_{\theta_0 \in \Theta_c} P_{\theta_0}(\theta_0 \in I_{up}(\bar{\theta}_n)) = (1 - \alpha) + O(n^{-1/2})$.

Proof.

(a) We apply Theorem 3.3 (a) with $s = 3$ to obtain (3.18) above with $n^{-1}p_2(\delta, \bar{\kappa}_{n,4}(\theta_0))$ deleted and $o(n^{-1})$ replaced by $o(n^{-1/2})$. This yields part (a).
(b) We have $P_{\theta_0}(\theta_0 \in I_{up}(\bar{\theta}_n)) = P_{\theta_0}(t_n(\theta_{0,r}) \le z_\alpha)$. Again by Theorem 3.3 (a) with $s = 3$, and using the fact that $\Phi(z_\alpha) = 1 - \alpha$ we have

$$\begin{aligned}
\sup_{\theta_0 \in \Theta_c} |P_{\theta_0}(\theta_0 \in I_{up}(\bar{\theta}_n)) - (1 - \alpha)| &= \sup_{\theta_0 \in \Theta_c} |P_{\theta_0}(t_n(\theta_{0,r}) \le z_\alpha) - (1 - \alpha)| \\
&\le \sup_{\theta_0 \in \Theta_c} (|P_{\theta_0}(t_n(\theta_{0,r}) \le z_\alpha) - [1 + n^{-1/2}p_1(\delta, \bar{\kappa}_{n,3}(\theta_0))]\Phi(z_\alpha)| \\
&\quad + |[1 + n^{-1/2}p_1(\delta, \bar{\kappa}_{n,3}(\theta_0))]\Phi(z_\alpha) - (1 - \alpha)|) \\
&\le \sup_{\theta_0 \in \Theta_c} |(1 - \alpha)n^{-1/2}p_1(\delta, \bar{\kappa}_{n,3}(\theta_0))| + o(n^{-1/2}) = O(n^{-1/2}).
\end{aligned}$$

This proves part (b) of the Theorem. □

References

Aga, M., & Sun, T. (2007). Parametric Bootstrap Confidence Intervals for Linear Regression Processes with long-memory Errors. *Sankhya: The Indian Journal of Statistics, 69*(4) 615-634. https://doi.org/10.1080/03610920903447873

Aga, M. (2011). Edgeworth Expansion of Linear Regression Processes with Long-Memory Errors. *Communications in Statistics - Theory and Methods, 40*(4), 663-673.

Anderson, T. W. (1984). An Introduction to Multivariate Statistical Analysis (2nd ed.). New York: Wiley.

Andrews, D., Lieberman O., & Marmer, V. (2006). Higher-order Improvements of the Parametric Bootstrap for Long-memory Gaussian Processes. *Journal of Econometrics, 133*, 2, 673-702. https://doi.org/10.1016/j.jeconom.2005.06.010

Baillie, R. T. (1996). Long-memory Processes and Fractional Integration in Econometrics. *Journal of Econometrics, 73*, 5-59. https://doi.org/10.1016/0304-4076(95)01732-1

Barndorff, E., & Cox, R. (1989). Asymptotic Techniques for Use in Statistics. New York: Chapman and Hall. https://doi.org/10.1007/978-1-4899-3424-6

Battacharya, R. N., & Ghosh, J. K. (1978). On the Validity of the Formal Edgeworth Expansion. *Annals of Statistics, 6*, 434-451. https://doi.org/10.1214/aos/1176344134

Beran, R. (1992). Statistical Methods for Data with Long-range Dependence. *Statistical Sciences, 7*, 404-427. https://doi.org/10.1214/ss/1177011122

Beran, R. (1994). *Statistics for Long-Memory Processes.* New York: Chapman and Hall.

Dahlhaus R. (1989). Efficient Parameter Estimation for Self-similar Processes. *Annals of Statistics, 17*(4), 1749-1766. https://doi.org/10.1214/aos/1176347393

Fox, R., & Taqqu, M. S. (1986). Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series. *Annals of Statistics, 14*, 517-532. https://doi.org/10.1214/aos/1176349936

Granger, W. J., & Joyeux, R. (1980). An Introduction to Long-memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis, 1*, 15-29. https://doi.org/10.1111/j.1467-9892.1980.tb00297.x

Hall, P. (1992). The Bootstrap and Edgeworth Expansion. New York: Springer. https://doi.org/10.1007/978-1-4612-4384-7

Hosking, J. R. M. (1981). Fractional Differencing. *Biometricka, 68*, 165-176. https://doi.org/10.1093/biomet/68.1.165

Lieberman O., Rousseau J., & Zucker M. (2003). Valid Edgeworth Expansions For the Maximum Likelihood Estimator of the Parameter of a Stationary, Gaussian, Strongly Dependent Processes. *Annals of Statistics, 31*, 2, 586-612.

Palma, W. (2007). *Long-Memory Time Series - Theory and Methods.* New York: Wiley. https://doi.org/10.1002/9780470131466

Robinson, P. M. (1994). Time Series with Strong Dependence. *Advances in Econometrics, Sixth World Congress, 1*, 47-95. Cambridge University Press, New York. https://doi.org/10.1017/CCOL0521444594.002

Robinson, P. M. (1995). Log-periodogram Regression of Time Series with Long-range Dependence. *Annals of Statistics, 23*, 1048-1072. https://doi.org/10.1214/aos/1176324636

Taqqu, M. S. (1986). A bibliographical Guide to Self-similar processes and Long-Range Dependence. *Dependence in Probability and Statistics*, Boston: Birkhauser, pp. 137-162. https://doi.org/10.1007/978-1-4615-8162-8_7