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 introduced by Hosking (1980) and Grangr
et al. (1981) and defined by

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

where $$\Phi(B) = 1 + \Phi_1 B + \ldots + \Phi_p B^p$$ and $$\psi(B) = 1 + \Psi_1 B + \ldots + \Psi_q B^q$$ 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 $$1 - B)^d = \sum_{j=0}^{\infty} \theta_j B^j$$, where

$$\theta_j = \frac{\Gamma(j + d)}{\Gamma(j + 1) \Gamma(d)},$$ \hspace{1cm} (1.3)

and $$\Gamma$$ is the gamma function, and $$\epsilon_t$$ 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

$$f_\theta(\lambda) \sim |\lambda|^{-\alpha(\theta)} A_\theta(\lambda)$$ \hspace{1cm} (1.4)

as $$\lambda \rightarrow \infty$$, 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 $$\lambda^d A_\theta(\lambda)$$ is bounded for every $$\delta$$, 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:

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi} |2 \sin \frac{\lambda}{2} \Psi(e^{-i\lambda})|^2$$ \hspace{1cm} (1.5)

and it can be shown that

$$f_\theta(\lambda) \sim \frac{\sigma^2 |\Psi(1)|^2}{2\pi |\Phi(1)|^2}$$ \hspace{1cm} (1.6)
for $|t| \rightarrow 0$ (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 $\{\omega_i\}$ 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-memory process $\{\omega_i\}$ 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 $\beta_0 = (\beta_1, \beta_2, \ldots, \beta_p)'$ be a vector of deterministic, but unknown real numbers. Let $Y_i, i \geq 1$ be a process such that: $Y_i = Z_i \theta_0 + \omega_i$, where $Z_i = (z_{i1}, z_{i2}, \ldots, z_{ip})$ are non-stochastic regressors for $i = 1, 2, \ldots, n$ and $\omega_i$ are the stationary Gaussian processes mentioned above. With out loss of generality we can assume that the mean of the $\omega_i$ is zero.

Now, the process $\{Y_i, i \geq 1\}$ is non-stationary because the mean of $Y_i$ is $\mu = E Y_i = Z_0 \theta_0$, which varies with $i$. However, if $Y = (Y_1, Y_2, \ldots, Y_n)'$ is an observed sample of size $n$ and $W = (\omega_1, \omega_2, \ldots, \omega_n)'$, then clearly the covariance matrix of $Y$ is the same as that of $W$.

Let $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ be the true mean of $Y$. Then, the least square estimate (LSE) $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p)'$ of $\beta_0$ is given by $\hat{\beta} = V^{-1} Z' Y$, where $Z = (z_{ij})$ for $i = 1, \ldots, n$ and $j = 1, \ldots, p$ denote the design matrix of our regression model and

$$V = Z'Z$$

(1.7)

is a $p \times p$ matrix. Thus, the estimator of $\mu$ is $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_n)$, where $\hat{\mu}_i = Z_i \hat{\beta}$, $t = 1, 2, \ldots, n$. We shall assume the rank of $Z$ is $p$. Then, the matrix $V$ (and hence $V^{-1}$) is symmetric and positive definite.

The $n \times n$ (Toeplitz) covariance matrix corresponding to $f_0(\lambda)$ is denoted by $\Gamma_n(f_0)$ and has $(j, k)$ element defined by:

$$\Gamma_n(f_0)_{jk} = \int_{-\pi}^{\pi} e^{i(j-k)\lambda} f_0(\lambda)d\lambda.$$

(1.8)

The log-likelihood function is

$$\Lambda_n(\theta, \mu) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln(\det(\Gamma_n(f_0))) - \frac{1}{2} (Y - \mu)' \Gamma_n^{-1}(f_0)(Y - \mu).$$

(1.9)

We refer to $\Lambda_n(\theta, \hat{\mu})$, where $\hat{\mu}$ is replaced for $\mu$ in (1.9) above, as the plug-in log-likelihood (PLL) function. Let $A_n = Z'V^{-1}Z'$ 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

$$A_n = EE'.$$

(1.10)

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

$$\Lambda_n(\theta, \hat{\mu}) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln(\det(\Gamma_n(f_0))) - \frac{1}{2} Y'K_n \Gamma_n^{-1}(f_0)K_n Y.$$

(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 $\hat{\theta}_n$ an element of $\hat{\Theta}_n$. We call $\hat{\theta}_n$ a plug-in maximum likelihood (PML) estimator of the true parameter $\theta_0$. By Theorem 2.1 of Dahlhaus (1989), the asymptotic covariance matrix of a consistent PML estimator $\hat{\theta}_n$ is $\Sigma(\hat{\theta}_n)$, where

$$\Sigma(\theta) = \left[ \int_{-\pi}^{\pi} \frac{\partial}{\partial \theta} \ln(f_0(\lambda)) \frac{\partial}{\partial \theta} f_0(\lambda) \right]^{-1}.$$ A consistent estimator of $\Sigma(\theta_0)$ is $\Sigma(\hat{\theta}_n)$, provided that $f_0(\lambda)$ is smooth with respect to $\theta$. Our goal is to formulate the structure of delta method test statistic and confidence intervals of the true parameter $\theta_0$ 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 $\theta_0$ denote some element of $\Theta$, the parameter space. Let $\theta_{0,r}$, $\theta_{0,r}$, and $\hat{\theta}_{n,r}$ denote the $r$-th elements of $\theta_0$, $\theta_h$, 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 $r$ statistic for testing the null hypothesis $H_0 : \theta_{0,r} = \theta_{h,r}$ is

$$t_n(\theta_{0,r}) = \frac{\sqrt{n}(\hat{\theta}_{n,r} - \theta_{0,r})}{\Sigma_{r,r}^{1/2}(\hat{\theta}_n)}.$$  

(2.1)
Let $z_{\alpha}$ 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

\[
I_{2}(\hat{\theta}_n) = \left[ \hat{\theta}_{n,r} - z_{\alpha/2} \frac{\sum_{j=1}^{12} \kappa_j \hat{\theta}_j}{\sqrt{n}}, \hat{\theta}_{n,r} + z_{\alpha/2} \frac{\sum_{j=1}^{12} \kappa_j \hat{\theta}_j}{\sqrt{n}} \right].
\]

(2.2)

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

\[
I_{up}(\hat{\theta}_n) = \left[ \hat{\theta}_{n,r} - z_{\alpha} \frac{\sum_{j=1}^{12} \kappa_j \hat{\theta}_j}{\sqrt{n}}, \infty \right).
\]

(2.3)

(d) The two-sided delta method $t$ test of $H_0 : \theta_{0,r} = \theta_{0,r}$ versus $H_1 : \theta_{0,r} \neq \theta_{0,r}$ with significance level $\alpha$ rejects $H_0$ if $|t_n(\hat{\theta}_{n,r})| > z_{\alpha/2}$.

(e) The one sided $t$ test of $H_0 : \theta_{0,r} \leq \theta_{0,r}$ versus $H_1 : \theta_{0,r} > \theta_{0,r}$ with significance level $\alpha$ rejects $H_0$ if $t_n(\hat{\theta}_{n,r}) > z_{\alpha}.$

2.2 Background on Cumulants and Edgeworth Expansion

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

\[
\ln \varphi(t) = \sum_{j=1}^{\infty} \frac{1}{j!} \kappa_j (it)^j.
\]

(2.4)

Since

\[
\varphi(t) = E(e^{itU}) = 1 + E(U)it + \frac{1}{2!} E(U^2)(it)^2 + \ldots + \frac{1}{j!} E(U^j)(it)^j + \ldots,
\]

(2.5)

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

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

(2.6)

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

\[
\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \left( \sum_{j=1}^{\infty} \frac{1}{j!} \mu_j (it)^j \right)^k.
\]

(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 $\kappa_1 = \mu_1, \kappa_2 = \mu_2 - \mu_1^2, \kappa_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3, \kappa_4 = \mu_4 - 4\mu_1\mu_3 - 3\mu_1^2\mu_2 - 6\mu_1^2\mu_1^2$ and so on, the formula for $\kappa_{10}$ for example containing more than 41 such terms. (See Hall, (1992), pp. 41-46.)

Now, let $\hat{\theta}$ be an estimate of the parameter $\theta_0$, 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 $\sigma^2$ 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:

\[
P\left( \frac{n^{1/2}(\hat{\theta} - \theta_0)}{\sigma} \leq u \right) = \Phi(u) + n^{-1/2}p_1(u)\Phi(u) + \ldots + n^{-1/2}p_j(u)\Phi(u) + \ldots,
\]

(2.8)

where $\phi$ and $\Phi$ 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 $U_1, U_2, \ldots, U_j$ are independent and identically distributed with mean $\mu = \theta_0$ and finite variance $\sigma^2$ 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

\[
p_1(u) = -\frac{1}{6} \kappa_3(u^3 - 1), \quad \text{and} \quad p_2(u) = -\frac{1}{12} \kappa_4(u^3 - 3u) - \frac{1}{72} \kappa_5^2(u^5 - 10u^3 + 15u).
\]

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 \( \hat{\theta}_n \) for which we construct confidence intervals are consistent. Assumption A3 is essentially a modified version of Condition N5 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 \( f_0(\lambda) \), requiring \( f_0(\lambda) \) 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 \( \Theta \) is a subset of \( \mathbb{R}^r \) with non-empty interior.

A2. For all \( \varepsilon > 0 \) and all compact subsets \( \Theta_k \) of \( \Theta \), the sequence of PML estimators \( \{\tilde{\theta}_n : n \geq 1\} \) for which the results of this paper hold satisfy

\[
\sup_{\theta_0 \in \Theta_k} P_{\theta_0}(||\tilde{\theta}_n - \theta_0|| > n^{-1/2}\ln(n)\varepsilon) = o(n^{1-s/2}) \quad \text{as} \quad n \to \infty
\]

for some integer \( s \geq 3 \).

A3. The matrices \( D_n(\theta) \) and \( D(\theta) \) in (2.23) below are positive definite.

A4. For some integer \( s \geq 3 \), \( f_0(\lambda) \) is \( s + 1 \) times continuously differentiable with respect to \( \theta \), and all of the derivatives are continuous in \( (\lambda, \theta) \) for \( \lambda \neq 0 \). In addition, \( f_0^{-1}(\lambda) \) is continuous in \( (\lambda, \theta) \) for all \( \lambda \in [0, \pi] \) and \( \theta \in \Theta \).

A5. The derivatives \( \frac{\partial}{\partial \lambda} f_0^{-1}(\lambda) \) and \( \frac{\partial}{\partial \theta} f_0^{-1}(\lambda) \) are continuous in \( (\lambda, \theta) \) for \( \lambda \neq 0 \). In addition, there exists \( c_1(\theta, \delta) < \infty \) such that \( |\frac{\partial}{\partial \lambda} f_0^{-1}(\lambda)| \leq c_1(\theta, \delta)|\lambda|^{2d-k-\delta} \) for \( k = 0, 1, 2 \) and all \( \delta > 0 \), where \( \theta = (\theta_1, \theta_2, ..., \theta_d)' \) 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) \( |f_0(\lambda)| \leq c_2(\theta, \delta)|\lambda|^{2d-\delta} \)

(b) for all \( (j_1, ..., j_k) \) with \( k \leq s + 1 \), with duplication among the \( j_i \) allowed, \( |\frac{\partial}{\partial \lambda^{j_1}...\partial \lambda^{j_k}} f_0^{-1}(\lambda)| \leq c_3(\theta, \delta)|\lambda|^{2d-k-\delta} \).

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

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, \( |\frac{\partial}{\partial \lambda^{j_1}...\partial \lambda^{j_k}} f_0^{-1}(\lambda)| \leq c_4(\theta)|\lambda|^{2d-k-\delta} \) for all \( \lambda \in (0, \pi). \) For any compact subset \( \Theta_k \) of \( \Theta \), there exists a constant \( \tilde{C}(\Theta_k) < \infty \) such that \( c_4(\theta) \leq \tilde{C}(\Theta_k) \) for all \( \theta \in \Theta_k \).

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

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

\[
E = (e_{ij}), i = 1, ..., n, j = 1, ..., p
\]

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 \( Z \) 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

\[ Z = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{p-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{p-1} \end{bmatrix} \]  

(2.10)

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

\[ U = Z'Z = \begin{bmatrix} \sum_{i=1}^{n} a_i^{p-1} & \sum_{i=1}^{n} a_i^{p-2} & \cdots & \sum_{i=1}^{n} a_i^0 \\ \sum_{i=1}^{n} a_i^{p-1} & \sum_{i=1}^{n} a_i^{p-2} & \cdots & \sum_{i=1}^{n} a_i^1 \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_i^{p-1} & \sum_{i=1}^{n} a_i^{p-2} & \cdots & \sum_{i=1}^{n} a_i^{p-1} \end{bmatrix} = \begin{bmatrix} O(na_1^{-1}) & O(na_2^{-1}) & \cdots & O(na_n^{-1}) \\ O(na_1) & O(na_2^2) & \cdots & O(na_n^{p-1}) \\ \vdots & \vdots & \ddots & \vdots \\ O(na_1^{p-1}) & O(na_2^{p-2}) & \cdots & O(na_n^{p-2}) \end{bmatrix}. \]

(2.11)

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

\[ U^{-1} = \begin{bmatrix} O(\frac{1}{a_1}) & O(\frac{1}{na_1}) & \cdots & O(\frac{1}{na_1^{p-1}}) \\ O(\frac{1}{na_2}) & O(\frac{1}{na_2}) & \cdots & O(\frac{1}{na_2^{p-1}}) \\ \vdots & \vdots & \ddots & \vdots \\ O(\frac{1}{na_n}) & O(\frac{1}{na_n}) & \cdots & O(\frac{1}{na_n^{p-1}}) \end{bmatrix}. \]

(2.12)

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

\[ (\lambda - O(\frac{1}{n}))(\lambda - O(\frac{1}{na_1}))\cdots(\lambda - O(\frac{1}{na_n^{p-2}}))(\lambda - O(\lambda) + O(\frac{1}{na_1}) + O(\frac{1}{na_2}) + \ldots + O(\frac{1}{na_n^{p-2}}) = 0. \]

(2.13)

From this we see that the vanishing of the characteristic polynomial \( |U - \lambda I| \) is determined by the the first term

\[ (\lambda - O(\frac{1}{n}))(\lambda - O(\frac{1}{na_1}))\cdots(\lambda - O(\frac{1}{na_1^{p-2}})) \]

(2.14)

of (2.13) because all the remaining terms are of order \( O(\frac{1}{na_n^{p-1}}) \). But (2.14) vanishes for \( \lambda = O(\frac{1}{n}), O(\frac{1}{na_1}), \ldots, O(\frac{1}{na_n^{p-1}}) \). Thus \( O(\lambda_k) \leq O(\frac{1}{n}) \) for \( k = 1, 2, \ldots, p \).

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

\[ (v_{ij} - \lambda_k)u_{ik} + v_{i2}u_{2k} + \ldots + v_{ip}u_{pk} = 0 \]

(2.15)

for \( i = 1, 2, \ldots, p \), \( u_{i1} = 1 \). From (2.12) we observe that for \( i,k = 1, 2, \ldots, p \), \( v_{ik} = O(\frac{1}{na_i^{p-1}}) \) and therefore we may replace each \( v_{ik} \) by \( \frac{N_{ik}}{na_i^{p-1}} \) and each \( \lambda_k \) by \( \frac{N_k}{n} \), where \( N, N_{k+i} \) are constants. Replacing these values in (2.15) above the first equation in (2.15) becomes:

\[ \left( \frac{N_k}{n} - \frac{N_i}{n} \right)u_{ik} + \left( \frac{N_{k+i}}{n} \right)u_{2k} + \ldots + \left( \frac{N_{k+p}}{n} \right)u_{pk} = 0 \]

and the second equation in (2.15) becomes:

\[ \left( \frac{N_k}{na_i^{p-1}} \right)u_{ik} + \left( \frac{N_{k+i}}{na_i^{p-1}} \right)u_{2k} + \ldots + \left( \frac{N_{k+p}}{na_i^{p-1}} \right)u_{pk} = 0 \]

Solving for \( u_{pk} \) in the first of the above equations and substituting in the second we obtain:

\[ -nd_k u_{2k} + nd_k^{p-1}(N_k - N_{k+i}N_i - N_{k+i}N_k) + O(na_i^{p-2}) + \ldots + O(n) = 0. \]

Looking at the first two terms of the above equation we see that \( u_{2k} \) must have order \( O(\frac{1}{na_i^{p-1}}) \). Similarly, \( u_{3k} \) is \( O(\frac{1}{na_i^{p-2}}) \) obtained from the first and the 3rd equations of (2.12), and in general \( u_{ik} = O(\frac{1}{na_i^{p-1}}) \) for \( i = 2, \ldots, p \). Now, the matrix \( E \) of (2.9) is given by \( E = Z'B \), where \( B = PDP' \), \( P = \left( \frac{u_{ij}}{||u||^2} \right), i, j = 1, \ldots, p \) and \( u_{ij} = 1 \), and \( D \) is the matrix whose diagonal entries are \( \sqrt{\lambda_i} \), for \( k = 1, \ldots, p \) and zero elsewhere. Then we have

\[ B = \begin{bmatrix} \sum_{i=1}^{p} u_{ii} \sqrt{\lambda_i} \\ \sum_{i=1}^{p} u_{i2} \sqrt{\lambda_i} \\ \sum_{i=1}^{p} u_{i3} \sqrt{\lambda_i} \\ \vdots \\ \sum_{i=1}^{p} u_{ip} \sqrt{\lambda_i} \end{bmatrix}. \]

(2.16)
Using the above notation of matrix $B$, we can now form the matrix $E$ as $E = \mathcal{Z}B = (e_{ij})$ for $i = 1, \ldots, n$, $j = 1, \ldots, p$. To show that $E$ satisfies assumption A10, it suffices to show that $O(e_{ij}) \leq O(\frac{1}{\sqrt{n}})$ for $i = 1, 2, \ldots, n$ and $j = 1, \ldots, p$. To this end it suffices to show that $O(e_{nj}) \leq \frac{1}{\sqrt{n}}$ for $j = 1, \ldots, p$ (because $e_{nj} \geq e_{ij}$ for $i = 1, 2, \ldots, n$ and $j = 1, \ldots, p$). Now, for $j = 1$ we have

$$O(e_{nj}) = O\left(\sum_{k=1}^{p} \frac{Y_k}{|B_k|} + (a_{nj} \sum_{k=1}^{p} \frac{Y_k}{|B_k|} + \ldots + a_{p,1} \frac{Y_p}{|B_p|})\right) \leq O\left(\frac{1}{\sqrt{n}} + a_{nj} + \ldots + a_{p,1} \frac{1}{\sqrt{n}}\right) = O(\frac{1}{\sqrt{n}}).$$  

(2.17)

A similar calculation shows that $O(e_{nj}) \leq O(\frac{1}{\sqrt{n}})$ for $j = 2, \ldots, p$ 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 $\theta$ for which we establish delta method confidence intervals. To this end we introduce some additional notations.

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

$$\Lambda_{q,\nu}(\theta, \hat{\mu}) = \frac{\partial^q}{\partial \theta_{r_1} \ldots \partial \theta_{r_q}} \Lambda_{0}(\theta, \hat{\mu}) = F_{n,\nu}(\theta) + Y' B_{n,\nu}(\theta) M_n Y$$

(2.18)

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

$$F_{n,\nu}(\theta) = -\frac{1}{2} D_\nu \ln(\det(\Gamma_n(f_0))) = \sum_{k=1}^{b} a_k tr \left( \prod_{j=1}^{p} \Gamma_{n}^{-1}(f_0) \Gamma_n (g_{\theta,k,j}) \right)$$

(2.19)

$$B_{n,\nu}(\theta) = -\frac{1}{2} D_\nu \Gamma_n^{-1}(f_0) = \sum_{k=1}^{b} a_k \left( \prod_{j=1}^{p} \Gamma_n^{-1}(f_0) \Gamma_n (g_{\theta,k,j}) \right) \Gamma_n^{-1}(f_0)$$

(2.20)

for some constants $b$, $a_k$, and $p_k$ that depend on $\nu$ and with $g_{\theta,k,j}$ being certain partial derivatives of the spectral density with respect to the components of $\theta$ of order $q$ or less. Note that, when computing the partial derivatives of $\Gamma_n(f_0)_{k,j}$, 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 $f_0$.

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

$$Z_n(\theta) = (\Lambda_{m,1}(\theta), \ldots, \Lambda_{m,m}(\theta)),$$

(2.21)

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

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

(2.22)

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

$$D_n(\theta) = E[W_n(\theta) W_n(\theta)']$$

(2.23)

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

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

$$D_n(\theta)_{i,j} = \text{tr} \left( B_{m,\nu} T_{nm}(f_0) B_{m,\nu} T_{nm}(f_0) \right)$$

(2.24)

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

$$D(\theta)_{i,j} = \frac{1}{\pi} \int_{-\infty}^{\infty} \{D_{t_0} f_0^{-1}(\lambda)\} \{D_{t_0} f_0^{-1}(\lambda)\}^2 d\lambda.$$
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 \( \theta \), where \( m \) is as given in (2.22). Let \( \tilde{h}_n^{-2}(u, \theta) \) be the formal Edgeworth expansion of \( h_n(u, \theta) \) of order \( \tau - 2 \). Let \( \xi_n(x, \theta) = E_0 \exp(i x' Z_\eta(\theta)) \) denote the characteristic function of \( Z_\eta(\theta) \) when \( \theta \) 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 \( D_{x, \eta} = \frac{\partial}{\partial x_{\eta_1}} \cdots \frac{\partial}{\partial x_{\eta_q}} \). Let \( \kappa_{n, \eta}(\theta) \) denote the \( \eta \) cumulants of \( Z_n(\theta) \). By definition, \( \kappa_{n, \eta}(\theta) = i^q D_{x, \eta} \ln(\xi_n(x, \theta)) \)|\( x = 0 \), where \( i = \sqrt{-1} \). The vector \( \kappa_{n, \eta}(\theta) \) is composed of elements \( \kappa_{n, j}(\theta) \eta \) for vectors \( \eta \) of dimension \( q \leq s \). The following lemma gives the explicit form of the Edgeworth expansion of the density function \( h_n^{-2}(u, \theta) \) 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 \( \tilde{h}_n^{-2}(u, \theta) \) of the density of \( W_n(\theta) \) is given by

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

(3.1)

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

**Proof.**

Let

\[
\xi_n(x, \theta) = \frac{1}{n} \ln(\xi_n(x, \theta) + \frac{1}{2} x' D_n(\theta)x),
\]

(3.2)

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

\[
\xi_{n, \tau}(x, \theta) = \sum_{j=3}^{\tau} \sum_{j_1, ..., j_m \geq 0} \left( \frac{\partial^j}{\partial x_{j_1} \cdots \partial x_{j_m}} \right) x_{j_1}^{j_1} \cdots x_{j_m}^{j_m} \bigg|_{x=0} \frac{1}{j!} \sum_{j=3}^{\tau} n^{-j/2} \xi_{n, j}(x, \theta),
\]

where \( \Sigma_j \) denotes summation over integers \( j_1, ..., j_m \geq 0 \) such that \( j = j_1 \cdots j_m \) 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

\[
\xi_{n, \tau}(x, \theta) = \frac{1}{n} \ln(\xi_n(x, \theta) \exp(\frac{1}{2} x' D_n(\theta)x)),
\]

(3.4)

which is equivalent to

\[
\exp(n \xi_{n, \tau}(x, \theta)) = \xi_n(x, \theta) \exp(\frac{1}{2} x' D_n(\theta)x).
\]

(3.5)

Expanding the left hand side of (3.5) we get

\[
1 + \sum_{j=1}^{\infty} \frac{1}{j!} \left( n \xi_{n, \tau}(x, \theta) \right)^j = \xi_n(x, \theta) \exp(\frac{1}{2} x' D_n(\theta)x).
\]

(3.6)

We let

\[
\sum_{j=1}^{\infty} \frac{1}{j!} \left( n \xi_{n, \tau}(x, \theta) \right)^j = \sum_{j=3}^{\tau} n^{-j/2} \pi_{nj}(x, \theta) + \alpha_{n, \tau}(x, \theta),
\]

(3.7)

be the expansion in powers of \( n^{-1/2} \), where \( \pi_{nj} \) is a polynomial in powers of the components of \( x \) from 3 to \( j \) and \( \alpha_{n, \tau} \) 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 \( \alpha_{n, \tau} \) contributes an amount of order \( n^{-(\tau - 1)/2} \) which, in fact, we shall neglect.

Neglecting \( \alpha_{n, \tau} \) and combining (3.6) with (3.7) we obtain

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

(3.8)

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

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

(3.9)
Now, we define the polynomial $P_{n,j}$ by

$$
\frac{1}{\sqrt{|D_n(\theta)|(2\pi)^m}} \exp\left(-\frac{1}{2} u' D_n^{-1}(\theta) u\right) P_{n,j}(u, \theta) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} \exp\left[-\frac{i}{2} x'D_n(\theta) x\right] \pi_{n,j}(x, \theta) dx.
$$

(3.10)

As shown by Hall (1992) page 244, the polynomial $P_{n,j}$ is a polynomial in terms of $-D$ and $\kappa_{n,\theta}(\theta)$, where $D = \left(\frac{\partial}{\partial \theta_1}, \ldots, \frac{\partial}{\partial \theta_m}\right)'$ and $\kappa_{n,\theta}(\theta)$ 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 $\tilde{h}^r_n(u, \theta)$ of the density of $W_n(\theta)$ is given by

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

(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 $\Theta_r \subset \Theta$ and all $r \geq 3$,

(a) $\sup_{0 \leq \theta \leq \Theta_r} \sup_{u \in \mathbb{R}^m} |h_n(u, \theta) - \tilde{h}^r_n(u, \theta)| = o(n^{-(r-2)/2})$

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

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 additional notations. Let $\Phi(\cdot)$ denote the distribution function of the standard normal random variable. Let $\kappa_{n,\theta}(\theta) = \frac{\kappa_{n,\theta}(\theta)}{n}$ By Lemma 4.5(c) of Aga et al. (2007), the elements of $\kappa_{n,\theta}(\theta)$ are $O(1)$. Let $p_i(\delta, \kappa_{n,\theta}(\theta))$ be a polynomial in $\delta = \frac{\delta}{\Delta}$ whose coefficients are polynomials in the elements of $\kappa_{n,\theta}(\theta)$ and for which $p_i(\delta, \kappa_{n,\theta}(\theta)) \Phi(z)$ is an even function of $z$ when $i$ is odd and an odd function of $z$ when $i$ is even for $i = 1, 2, \ldots, s - 2$. The Edgeworth expansion of the delta method $t$-statistic $t_n(\theta_{0,r})$ given in (2.1) depends on $p_i(\delta, \kappa_{n,\theta}(\theta))$.

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

(a) $\sup_{0 \leq \theta \leq \Theta_r} \sup_{|u| \leq n^{1/4}} |P_{\Theta_r}(t_n(\theta_{0,r}) \leq z) - \left[1 + \sum_{i=1}^{s-2} n^{-i/2} p_i(\delta, \kappa_{n,\theta}(\theta))\right] \Phi(z)| = o(n^{-(s-2)/2})$.

(b) $\sup_{0 \leq \theta \leq \Theta_r} P_{\Theta_r}(\Delta(\theta_{0,r}) = (1 - \alpha) + O(n^{-1})$ 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 $\theta$ of order $s$ of $n^{-1}A_n(\theta_{0})$. 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 $G(\cdot)$ that does not depend on $\theta_0$ that satisfies $G(n^{-1}E_{\theta_0}Z_n^*(\theta_0)) = 0$ for all large and all $\theta_0 \in \Theta_r$ and

$$
\sup_{\theta_0 \in \Theta_r} \sup_{B \in \mathcal{B}_B} \left|P_{\theta_0}(t_n(\theta_{0,r}) \leq z) - P_{\theta_0}(n^{1/2}G(n^{-1}Z^*_n(\theta_0)) \leq z)\right| = o(n^{-(s-2)/2}).
$$

(3.12)

Now, let $P_{\theta_0}(\theta_0) = 1 + \sum_{i=1}^{s-2} n^{-i/2} p_i(\delta, \kappa_{n,\theta}(\theta_{0}))$. Then,

$$
\left|P_{\theta_0}(t_n(\theta_{0,r}) \leq z) - P_{\theta_0}(\Delta(\theta_{0,r}) \leq z) - P_{\theta_0}(n^{1/2}G(n^{-1}Z^*_n(\theta_0)) \leq z)\right| = o(n^{-(s-2)/2}).
$$

(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

$$
\left|P_{\theta_0}(n^{1/2}G(n^{-1}Z^*_n(\theta_{0})) = 0 - P_{\theta_0}(\Delta(\theta_{0,r}) \leq z)ight| = o(n^{-(s-2)/2})
$$

(3.14)

uniformly over $\theta_0 \in \Theta_r$. 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 $\theta_0 \in \Theta_r$. We note that an Edgeworth expansion of $W_n(\theta) = n^{-1/2}(Z_n(\theta) - E_nZ_n(\theta_{0}))$ is established by lemma 3.2 above for each $\theta_0 \in \Theta_r$. Now, an Edgeworth expansion for $n^{1/2}G(n^{-1}Z^*_n(\theta_{0}))$ is obtained from that of $n^{-1/2}(Z_n(\theta) - E_nZ_n(\theta_{0}))$ by the argument of Bhattacharya and Ghosh (1978), Theorem 2, p. 436 using the smoothness of $G(\cdot)$, $G(n^{-1}E_{\theta_0}Z^*_n(\theta_{0})) = 0$ for all $n \geq 1$ and all $\theta_0 \in \Theta_r$, and assumption A3.

(b) From the definition of $I_2(\theta_{0,r})$ given in section 2.1 we observe that

$$
P_{\theta_0}(I_2(\tilde{\theta}_{n})) = P_{\theta_0}(\tilde{\theta}_{n,r} - \bar{\theta}_{a,r} \leq 0.5 \sqrt{n^{-1/2} - \tilde{\theta}_{a,r} - \bar{\theta}_{a,r} \leq \sqrt{n^{-1/2}}})
$$

(3.15)
and that
\[ P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha/2}) = P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha/2}) - P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq -z_{\alpha/2}). \] (3.16)

Denoting 1 + n^{-1/2}p_{1}(\delta_{n},\theta_{0}\delta_{n},\theta_{0}) + n^{-1}p_{2}(\delta_{n},\theta_{0}\delta_{n},\theta_{0}) by \( \mathcal{P}(\theta_{0}) \) as in (3.13) and substituting \( z \) by \( z_{\alpha/2} \) and \( s \) by 4 in part (a) of this theorem we obtain
\[ |P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha/2}) - \mathcal{P}(\theta_{0})(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}))| = o(n^{-1}). \] (3.17)

We observe that \( p_{1}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})\Phi(z) \) is an even function in \( z \) which leads to \( p_{1}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = 0 \). It follows that
\[ \sup_{\theta_0,\theta_0} |P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha/2}) - [1 + n^{-1/2}p_{2}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}))]| = o(n^{-1}). \] (3.18)

Because \( \Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) = 1 - \alpha \) and \( n^{-1/2}p_{2}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = o(n^{-1}) \), this establishes part (b). □

**Theorem 3.4.** Suppose assumptions A1-A10 hold and let \( s = 3 \). Consider \( \tilde{\theta}_{n} : n \geq 1 \) as given in assumption A2 and let \( \Theta_{k} \) be any compact subset of \( \Theta \), the parameter space. Then,
(a) \[ \sup_{\theta_0,\theta_0} P_{\theta_0}(\theta_{0} \in I_{L}(\tilde{\theta}_{n})) = (1 - \alpha) + O(n^{-1/2}). \]
(b) \[ \sup_{\theta_0,\theta_0} P_{\theta_0}(\theta_{0} \in I_{L}(\bar{\theta}_{n})) = (1 - \alpha) + O(n^{-1/2}). \]

**Proof.**
(a) We apply Theorem 3.3 (a) with \( s = 3 \) to obtain (3.18) above with \( n^{-1/2}p_{2}(\delta_{n},\theta_{0}\delta_{n},\theta_{0}) \) deleted and \( o(n^{-1}) \) replaced by \( o(n^{-1/2}) \). This yields part (a).
(b) We have \( P_{\theta_0}(\theta_{0} \in I_{L}(\tilde{\theta}_{n})) = P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha} \). Again by Theorem 3.3 (a) with \( s = 3 \), and using the fact that \( \Phi(z_{\alpha}) = 1 - \alpha \) we have
\[ \sup_{\theta_0,\theta_0} |P_{\theta_0}(\theta_{0} \in I_{L}(\bar{\theta}_{n})) - (1 - \alpha)| = \sup_{\theta_0,\theta_0} |P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha}) - (1 - \alpha)| \]
\[ \leq \sup_{\theta_0,\theta_0} |P_{\theta_0}(t_{n}\theta_{0}\sigma) \leq z_{\alpha}) - [1 + n^{-1/2}p_{1}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})]\Phi(z_{\alpha})| \]
\[ + |[1 + n^{-1/2}p_{1}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})]\Phi(z_{\alpha}) - (1 - \alpha)| \]
\[ \leq \sup_{\theta_0,\theta_0} |(1 - \alpha)n^{-1/2}p_{1}(\delta_{n},\theta_{0}\delta_{n},\theta_{0})| + o(n^{-1/2}) = O(n^{-1/2}). \]

This proves part (b) of the Theorem. □

**References**


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).