Statistical Higher Order Asymptotic Theory and Its Applications to Analysis of Financial Time Series

統計的高次漸近理論とその金融時系列解析への応用

Kenichiro TAMAKI

February, 2006

Statistical Higher Order Asymptotic Theory and Its Applications to Analysis of Financial Time Series

統計的高次漸近理論とその金融時系列解析への応用

2006 年 2 月

早稲田大学大学院理工学研究科

数理科学専攻　数理統計学研究

玉置　健一郎

Contents

1. Introduction

This dissertation is concerned with statistical higher order asymptotic theory and its applications to analysis of financial time series. In statistical analysis, because it is difficult to use the exact distribution theory, the discussion is based on the asymptotic theory. It is well known that we can construct infinitely many first order asymptotically efficient estimators for unknown parameters. Thus it is required to illuminate their distinction, by considering higher order terms in the expansions of their asymptotic distributions.

There has been much demand for statistical analysis of dependent observation in many fields, for example, economics, engineering and nature sciences. Financial engineering is the application of engineering methods to financial problems. Time series analysis enables financial engineers to measure and manage their financial risks and to design and analyze sophisticated financial contracts.

In this dissertation, using higher order approximations of the distribution of estimators and tests we elucidate their higher order asymptotic properties. One of the main topics in financial engineering is option pricing. Thus we discuss the option pricing problems using statistical series expansion for the price process of an underlying asset.

This dissertation is organized as follows. In Chapter 2, under the existence of nuisance parameters, we consider a class of tests $\mathcal{S}$ which contains the likelihood ratio, Wald and Rao's score tests as special cases. To investigate the influence of nuisance parameters, we derive the second order asymptotic expansion of the distribution of $T \in \mathcal{S}$ under a sequence of local alternatives. This result and concrete examples illuminate some interesting features of influences due to nuisance parameters. Optimum properties for a modified likelihood ratio test proposed in Mukerjee [32] are shown under the criteria of second order local maximinity.

Chapter 3 discusses the option pricing problems using statistical series expansion for the price process of an underlying asset. We derive the Edgeworth expansion for the stock log return via extracting dynamics structure of time series. Using this result, we investigate influences of the non-Gaussianity and the dependency of log return processes for option pricing. Numerical studies show some interesting features of them.

In Chapter 4, we consider the second order asymptotic properties of an efficient frequency domain regression coefficient estimator $\hat{\beta}$ proposed by Hannan [18]. This estimator is a semiparametric estimator based on nonparametric spectral estimators. We derive the second order Edgeworth expansion of the distribution of $\hat{\beta}$. Then it is shown that the second order asymptotic properties are independent of the bandwidth choice for residual spectral estimator, which implies that $\hat{\beta}$ has the same rate of convergence as in regular parametric estimation. This is a sharp contrast with the general semiparametric estimation theory. We also examine the second order Gaussian efficiency of $\hat{\beta}$. Numerical studies are given to confirm the theoretical results.

In Chapter 5, we investigate an optimal property of the maximum likelihood estimator of Gaussian locally stationary processes by the second order approximation. In the case where the model is correctly specified, it is shown that appropriate modifications of the maximum likelihood estimator for Gaussian locally stationary processes is second order asymptotically efficient. We also discuss second order robustness properties.

Finally, in Chapter 6, we place the proofs of the theorems and lemmas.

2. Second Order Asymptotic Properties of a Class of Test Statistics under the Existence of Nuisance Parameters

2.1. Introduction

In multivariate analysis, the second order asymptotic powers of various test statistics have been investigated by Hayakawa [21], and Harris and Peers [20]. Under the absence of nuisance parameters, results on optimality are now known for the likelihood ratio (LR) test in terms of second order local maximinity and Rao's score (R) test in terms of third order local average power (Mukerjee [33]). Under the existence of nuisance parameters, Eguchi [17] studied the effect of the composite null hypothesis from a geometric point of view. Mukerjee [32] suggested a test that is superior to the usual LR test with regard to second order local maximinity. The test proposed in Mukerjee [32] is motivated from the principle of conditional likelihood and also from that of adjusted likelihood.

In time series analysis, under a set-up involving an unknown scalar parameter, Taniguchi [44] considered the problem of second order comparison of tests. He worked with a large class of tests that contains LR, R and Wald's (Wesss as special cases. Taniguchi [45] showed that the local powers of all the modified tests which are second order asymptotically unbiased are identical up to $N^{-1/2}$. Also Taniguchi [46] considered the problem of third order comparison of tests, and suggested a Bartlett-type adjustment for the tests in the class and then, on the basis of such adjusted versions, explored the point-by-point maximization of third order power.

Bartlett's adjustment procedure has been elucidated in various directions. Cordeiro and Ferrari [8] gave a general formula of Bartlett-type adjustment to order N^{-1} for the test statistic whose asymptotic expansion is a finite linear combination of chi-squared distribution with suitable degrees of freedom. Kakizawa [25] considered the extension of Cordeiro and Ferrari's [8] adjustment to the case of order N^{-k}, where k is an integer $k \geq 1$. Rao and Mukerjee [34] compared various Bartlett-type adjustments for the R statistic. Rao and Mukerjee [35] addressed the problem of comparing the higher order power of tests in their original forms and not via their bias-corrected or Bartlett-type adjusted versions.

In this chapter, under the existence of nuisance parameters, we consider the second order properties of a class of tests $\mathcal{S}$ which contains LR, R and W tests as special cases. If nuisance parameters are present, sensitivity of test statistics to perturbation of the nuisance parameters becomes important. It is shown that the powers and sizes of $T \in \mathcal{S}$ are equally sensitive to perturbation of the nuisance parameter. In Section 2.3 we compare the second order local power. It is seen that the local average powers of all $T \in \mathcal{S}$ are identical. It is shown that optimality properties hold for a modified test of the LR test in terms of second order local maximinity. Section 2.4 provides a decomposition formula of local powers for LR, R and W test statistics under local orthogonality for parameters. The decomposition consists of the sum of the three parts; one is the local power for the case of known nuisance parameters, another represents sensitivity to perturbation of nuisance parameters and the other part can be interpreted as an effect of nuisance parameters in test statistics. In Section 2.5, we discuss the local unbiasedness of $T \in \mathcal{S}$. The results and their examples illuminate some interesting features of effects due to nuisance parameters. The proofs of theorems are relegated to Section 6.1.

2.2. Asymptotic expansion of a class of tests

Suppose that $\boldsymbol{X}_N = (X_1, \ldots, X_N)$ be a collection of m-dimensional random vectors forming a stochastic process. Let $p_N(\boldsymbol{x}_N; \theta)$, $\boldsymbol{x}_N \in \mathbf{R}^{mN}$, be the probability density function of $\boldsymbol{X}_N$, where $\theta = (\theta^1, \ldots, \theta^{p+q})' \in \Theta$ an open subset of $\mathbf{R}^{p+q}$. Let $\theta_1 = (\theta^1, \ldots, \theta^p)'$ be the p-dimensional parameter of interest and $\theta_2 = (\theta^{p+1}, \ldots, \theta^{p+q})'$ be the q-dimensional nuisance parameter. We consider the problem of testing the hypothesis $H : \theta_1 = \theta_{10}$, where $\theta_{10} = (\theta_0^1, \ldots, \theta_0^p)'$, against the alternative $A : \theta_1 \neq \theta_{10}$. For this problem we introduce a class of test $\mathcal{S}$ which contains LR, W and R tests as special cases. In the presence of nuisance parameters, the powers and sizes of $T \in \mathcal{S}$ are affected by the true but unknown nuisance parameter. Therefore we investigate the influence of perturbation by the sequence of local alternatives $\theta = \theta_0 + c_N^{-1}\varepsilon$ where $\theta_0' = (\theta_{10}', \theta_{20}')$, $\theta_{20} = (\theta_0^{p+1}, \ldots, \theta_0^{p+q})'$ and $\varepsilon = (\varepsilon^1, \ldots, \varepsilon^{p+q})'$. As in Li [28], we shall use Greek letters $\{\alpha, \beta, \gamma, \ldots\}$ as indices that run from 1 to $p+q$, the set of English letters $\{i, j, k, \ldots, q\}$ as indices that run from 1 to p, and the set of $\{r, s, t, \ldots, z\}$ as indices that run from $p+1$ to $p+q$. The indices i, r and α will serve two purposes, first to denote a typical term in a sum and second to indicate the range of a sum. For example, $a_\alpha X^\alpha = \sum_{\alpha=1}^{p+q} a_\alpha X^\alpha$, $a_i X^i = \sum_{i=1}^{p} a_i X^i$ and $a_r X^r = \sum_{r=p+1}^{p+q} a_r X^r$.

We make the following assumptions:

Assumption 2.1. (i) $l_N(\theta) = \log p_N(\boldsymbol{X}_N; \theta)$ is continuously four times differentiable with respect to θ.

(ii) The expectation E_θ with respect to $p_N(\boldsymbol{x}_N; \theta)$ and the partial derivative $\partial_\alpha = \partial/\partial\theta^\alpha$ are interchangeable.

(iii) For an appropriate sequence $\{c_N\}$ satisfying $c_N \to +\infty$ as $N \to +\infty$, the asymptotic moments (cumulants) of

$$Z_\alpha(\theta) = c_N^{-1}\partial_\alpha l_N(\theta),$$
$$Z_{\alpha\beta}(\theta) = c_N^{-1}[\partial_\alpha\partial_\beta l_N(\theta) - \mathrm{E}_\theta\{\partial_\alpha\partial_\beta l_N(\theta)\}],$$

possess the following asymptotic expansions

$$\mathrm{E}_\theta\{Z_\alpha(\theta)Z_\beta(\theta)\} = I_{(\alpha\beta)}(\theta) + O(c_N^{-2}),$$

$$\mathrm{E}_\theta\{Z_\alpha(\theta)Z_{\beta\gamma}(\theta)\} = J_{\alpha,\beta\gamma}(\theta) + O(c_N^{-2}),$$
$$\mathrm{E}_\theta\{Z_\alpha(\theta)Z_\beta(\theta)Z_\gamma(\theta)\} = c_N^{-1}K_{\alpha\beta\gamma}(\theta) + O(c_N^{-3}),$$

and J-th order ($J \geq 2$) cumulants of $Z_\alpha(\theta)$ and $Z_{\alpha\beta}(\theta)$ are all $O(c_N^{-J+2})$.

(iv) (iv1) $I_{(\alpha\beta)}(\theta)$ is continuously two times differentiable with respect to θ.

(iv2) $J_{\alpha,\beta\gamma}(\theta)$ and $K_{\alpha\beta\gamma}(\theta)$ are continuously differentiable functions.

(v) (v1) $I(\theta) = \{I_{(\alpha\beta)}(\theta)\}$ is positive definite for all $\theta \in \Theta$.

(v2) $L(\theta) = \{-c_N^{-2}\partial_\alpha\partial_\beta l_N(\theta)\}$ is positive definite almost surely for all $\theta \in \Theta$.

Let $\hat{\theta} = (\hat{\theta}^1, \ldots, \hat{\theta}^{p+q})'$ be the global maximum likelihood estimator of θ, and let $\tilde{\theta}_2 = (\tilde{\theta}^{p+1}, \ldots, \tilde{\theta}^{p+q})'$ be the restricted maximum likelihood estimator of θ_2 given $\theta_1 = \theta_{10}$. The partition $\theta' = (\theta_1', \theta_2')$ induces the following corresponding partitions

$$\hat{\theta} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix}, \qquad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix},$$
$$I(\theta) = \begin{bmatrix} I_{11}(\theta) & I_{12}(\theta) \\ I_{21}(\theta) & I_{22}(\theta) \end{bmatrix}, \qquad L(\theta) = \begin{bmatrix} L_{11}(\theta) & L_{12}(\theta) \\ L_{21}(\theta) & L_{22}(\theta) \end{bmatrix}.$$

Let

$$g(\theta) = \{g_{\alpha\beta}(\theta)\} = \begin{bmatrix} I_{11\cdot 2}(\theta) & I_{12}(\theta) \\ 0 & I_{22}(\theta) \end{bmatrix},$$

where $I_{11\cdot 2}(\theta) = I_{11}(\theta) - I_{12}(\theta)\{I_{22}(\theta)\}^{-1}I_{21}(\theta)$.

We consider the transformation

$$W_i(\theta) = Z_i(\theta) - I_{(ir)}(\theta)g^{rs}(\theta)Z_s(\theta), \quad W_r(\theta) = Z_r(\theta),$$
$$W_{\alpha\beta}(\theta) = Z_{\alpha\beta}(\theta) - J_{\gamma,\alpha\beta}(\theta)I^{\gamma\delta}(\theta)Z_\delta(\theta),$$

where $I^{\alpha\beta}(\theta)$ and $g^{\alpha\beta}(\theta)$ are the (α,β) component of the inverse matrix of $I(\theta)$ and $g(\theta)$, respectively. Henceforth we use the simpler notations Z_α, W_α, $I_{(\alpha\beta)}$, $K_{\alpha\beta\gamma}$, etc. if $Z_\alpha(\theta)$, $W_\alpha(\theta)$, $I_{(\alpha\beta)}(\theta)$, $K_{\alpha\beta\gamma}(\theta)$, etc. are evaluated at $\theta = \theta_0$. Any function evaluated at the point $\theta = \hat{\theta}$ will be distinguished by the addition of a circumflex. Similarly any function evaluated at the point $\theta_1 = \theta_{10}$, $\theta_2 = \tilde{\theta}_2$ will be distinguished by the addition of a tilde. For the testing problem $H : \theta_1 = \theta_{10}$ against the alternative $A : \theta_1 \neq \theta_{10}$, we introduce the following class of tests:

$$\begin{aligned} \mathcal{S} = \{T \mid T = {} & g^{ij}W_iW_j + c_N^{-1}a_1 g^{i\alpha}g^{j\beta}W_{\alpha\beta}W_iW_j + 2c_N^{-1}g^{i\alpha}g^{rs}W_{\alpha r}W_iW_s \\ & + c_N^{-1}a_2^{ijk}W_iW_jW_k - c_N^{-1}g^{i\alpha}g^{j\beta}g^{rs}K_{\alpha,\beta,r}W_iW_jW_s \\ & - c_N^{-1}g^{i\alpha}g^{rt}g^{su}(K_{\alpha rs} + J_{\alpha,rs})W_iW_tW_u + c_N^{-1}a_3^iW_i \\ & + o_p(c_N^{-1}), \\ & \text{under } H, \text{ where } a_1, a_2^{ijk} \text{ and } a_3^i \text{ are nonrandom constants}\}. \end{aligned} \tag{2.1}$$

This class $\mathcal{S}$ is a very natural one. We can show that famous tests based on the maximum likelihood estimator belong to $\mathcal{S}$.

EXAMPLE 2.1. (i) The likelihood ratio test LR $= 2(\hat{l}_N - \tilde{l}_N)$ belongs to $\mathcal{S}$. In fact, from Bickel and Ghosh [3], the expansion for the r-th component of $c_N^{-1}(\hat{\theta}_2 - \tilde{\theta}_2)$ is given by

$$\begin{aligned} c_N^{-1}(\hat{\theta}^r - \tilde{\theta}^r) &= \eta^r + c_N^{-1}\hat{g}^{rs}\hat{Z}_{s\alpha}\eta^\alpha \\ &\quad + \frac{1}{2}c_N^{-1}\hat{g}^{rs}(\hat{K}_{s\alpha\beta} + \hat{J}_{s,\alpha\beta}[3])\eta^\alpha\eta^\beta + o_p(c_N^{-1}), \end{aligned} \tag{2.2}$$

where $\eta^i = c_N^{-1}(\hat{\theta}^i - \theta_0^i)$, $\eta^r = -\hat{g}^{rs}\hat{I}_{(si)}\eta^i$ and $\hat{J}_{\alpha,\beta\gamma}[3] = \hat{J}_{\alpha,\beta\gamma} + \hat{J}_{\beta,\gamma\alpha} + \hat{J}_{\gamma,\alpha\beta}$. Expanding LR in a Taylor series at $\theta = \hat{\theta}$ and noting (2.2), we obtain

$$\begin{aligned} 2(\hat{l}_N - \tilde{l}_N) &= \hat{g}_{ij}\eta^i\eta^j - c_N^{-1}\hat{Z}_{\alpha\beta}\eta^\alpha\eta^\beta \\ &\quad - c_N^{-1}\left(\frac{1}{3}\hat{K}_{\alpha\beta\gamma} + \hat{J}_{\alpha,\beta\gamma}\right)\eta^\alpha\eta^\beta\eta^\gamma + o_p(c_N^{-1}). \end{aligned} \tag{2.3}$$

By Taylor expansion around θ_0,

$$\begin{aligned} \hat{g}_{ij} &= g_{ij} + g_{ik}g^{k\alpha}g_{jl}g^{l\beta}(K_{\alpha\beta\gamma} + J_{\alpha,\beta\gamma} + J_{\beta,\gamma\alpha})(\hat{\theta}^\gamma - \theta_0^\gamma) \\ &\quad + o_p(c_N^{-1}). \end{aligned} \tag{2.4}$$

Furthermore, the stochastic expansion of $c_N^{-1}(\hat{\theta}^\alpha - \theta_0^\alpha)$ is given by

$$\begin{aligned} c_N^{-1}(\hat{\theta}^\alpha - \theta_0^\alpha) &= g^{\beta\alpha}W_\beta + c_N^{-1}I^{\alpha\beta}g^{\delta\gamma}W_{\beta\gamma}W_\delta \\ &\quad - \frac{1}{2}c_N^{-1}I^{\alpha\alpha'}g^{\beta\beta'}g^{\gamma\gamma'}(K_{\alpha'\beta'\gamma'} + J_{\alpha',\beta'\gamma'})W_\beta W_\gamma \\ &\quad + o_p(c_N^{-1}). \end{aligned} \tag{2.5}$$

Inserting (2.4) and (2.5) in (2.3) and noting $\hat{Z}_{\alpha\beta} = W_{\alpha\beta} + o_p(1)$, we have

$$\begin{aligned} 2(\hat{l}_N - \tilde{l}_N) &= g^{ij}W_iW_j + c_N^{-1}g^{i\alpha}g^{j\beta}W_{\alpha\beta}W_iW_j + 2c_N^{-1}g^{i\alpha}g^{rs}W_{\alpha r}W_iW_s \\ &\quad - \frac{1}{3}c_N^{-1}g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha\beta\gamma}W_iW_jW_k \\ &\quad - c_N^{-1}g^{i\alpha}g^{j\beta}g^{rs}K_{\alpha\beta r}W_iW_jW_s \\ &\quad - c_N^{-1}g^{i\alpha}g^{rt}g^{su}(K_{\alpha rs} + J_{\alpha,rs})W_iW_tW_u + o_p(c_N^{-1}). \end{aligned}$$

Hence, LR $= 2(\hat{l}_N - \tilde{l}_N)$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha\beta\gamma}/3$ and $a_3^i = 0$.

Similarly, we can get results (ii)–(v):

(ii) Wald's test $\mathrm{W}_1 = \hat{g}_{ij}\eta^i\eta^j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 2$, $a_2^{ijk} = g^{i\alpha}g^{j\beta}g^{k\gamma}J_{\alpha,\beta\gamma}$ and $a_3^i = 0$.

(iii) A modified Wald's test $\mathrm{W}_2 = \tilde{g}_{ij}\eta^i\eta^j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 2$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}(K_{\alpha\beta\gamma} + J_{\alpha,\beta\gamma})$ and $a_3^i = 0$.

(iv) Rao's score test $\mathrm{R}_1 = \hat{g}^{ij}\tilde{Z}_i\tilde{Z}_j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 0$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}(K_{\alpha\beta\gamma} + 2J_{\alpha,\beta\gamma})$ and $a_3^i = 0$.

(v) A modified version of Rao's score test $\mathrm{R}_2 = \tilde{g}^{ij}\tilde{Z}_i\tilde{Z}_j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 0$, $a_2^{ijk} = 0$ and $a_3^i = 0$.

Furthermore, it is shown that modified versions of the four tests W_1, W_2, R_1 and R_2 which are based on the observed information belong to $\mathcal{S}$. Let $\{l_{ij}(\theta)\} = L_{11\cdot 2}(\theta) = L_{11}(\theta) - L_{12}(\theta)\{L_{22}(\theta)\}^{-1}L_{21}(\theta)$ and $\{l^{ij}(\theta)\}$ be the (i, j) component of the inverse matrix of $L_{11\cdot 2}(\theta)$.

(vi) A modified version of Wald's test $\mathrm{W}_3 = \hat{l}_{ij}\eta^i\eta^j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = g^{i\alpha}g^{j\beta}g^{k\gamma}J_{\alpha,\beta\gamma}$ and $a_3^i = 0$.

(vii) A modified version of Wald's test $\mathrm{W}_4 = \tilde{l}_{ij}\eta^i\eta^j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}(K_{\alpha\beta\gamma} + 2J_{\alpha,\beta\gamma})$ and $a_3^i = 0$.

(viii) A modified version of Rao's score test $\mathrm{R}_3 = \hat{l}^{ij}\tilde{Z}_i\tilde{Z}_j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}(K_{\alpha\beta\gamma} + 2J_{\alpha,\beta\gamma})$ and $a_3^i = 0$.

(ix) A modified version of Rao's score test $\mathrm{R}_4 = \tilde{l}^{ij}\tilde{Z}_i\tilde{Z}_j$ belongs to $\mathcal{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = g^{i\alpha}g^{j\beta}g^{k\gamma}J_{\alpha,\beta\gamma}$ and $a_3^i = 0$.

(x) The test $\mathrm{LR}^* = \mathrm{LR} + c_N^{-1}\tilde{g}^{i\alpha}\tilde{g}^{rs}(\tilde{K}_{\alpha rs} + \tilde{J}_{\alpha,rs})\tilde{Z}_i$ proposed in Mukerjee [32] belongs to $\mathcal{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha\beta\gamma}/3$ and $a_3^i = g^{i\alpha}g^{rs}(K_{\alpha rs} + J_{\alpha,rs})$.

Li [28] compared the sensitivities of LR, W_2 and R_2 statistics to nuisance parameters. In the one-parameter case, Taniguchi [46] discussed the third order asymptotic properties of a class of tests $\mathcal{S}_1$. Rao and Mukerjee [35] studied a wider class $\mathcal{S}_2(\supset \mathcal{S}_1)$ which enables us to compare the various Bartlett-type adjustments available for the members of $\mathcal{S}_1$. Our class $\mathcal{S}$ contains $\mathcal{S}_1$ and $\mathcal{S}_2$, hence the class $\mathcal{S}$ is sufficiently rich.

REMARK 2.1. Test statistics in Example 2.1 are based on the maximum likelihood estimator. From (2.2) and (2.3), these statistics can be written as

$$\begin{aligned} T = {} & \hat{g}_{ij}\eta^i\eta^j + c_N^{-1}b_1\hat{Z}_{\alpha\beta}\eta^\alpha\eta^\beta + c_N^{-1}(b_2\hat{K}_{\alpha\beta\gamma} + b_3\hat{J}_{\alpha,\beta\gamma})\eta^\alpha\eta^\beta\eta^\gamma \\ & + c_N^{-1}b_4\hat{g}^{rs}(\hat{K}_{\alpha rs} + \hat{J}_{\alpha,rs})\eta^\alpha + o_p(c_N^{-1}), \end{aligned} \tag{2.6}$$

where the coefficient $(b_1, b_2, b_3, b_4) \in \mathbf{R}^4$. For these statistics,

$$
\begin{array}{lllll}
b_1 = -1, & b_2 = -1/3, & b_3 = -1, & b_4 = 0, & \text{for LR}, \\
b_1 = -1, & b_2 = -1/3, & b_3 = -1, & b_4 = 1, & \text{for LR}^*, \\
b_1 = 0, & b_2 = 0, & b_3 = 0, & b_4 = 0, & \text{for W}_1, \\
b_1 = 0, & b_2 = -1, & b_3 = -2, & b_4 = 0, & \text{for W}_2, \\
b_1 = -1, & b_2 = 0, & b_3 = 0, & b_4 = 0, & \text{for W}_3, \\
b_1 = -1, & b_2 = -1, & b_3 = -3, & b_4 = 0, & \text{for W}_4, \\
b_1 = -2, & b_2 = -1, & b_3 = -3, & b_4 = 0, & \text{for R}_1, \\
b_1 = -2, & b_2 = 0, & b_3 = -1, & b_4 = 0, & \text{for R}_2, \\
b_1 = -1, & b_2 = -1, & b_3 = -3, & b_4 = 0, & \text{for R}_3, \\
b_1 = -1, & b_2 = 0, & b_3 = 0, & b_4 = 0, & \text{for R}_4.
\end{array}
\tag{2.7}
$$

Inserting (2.4) and (2.5) in (2.6), we obtain

$$
\begin{aligned}
T = {} & g^{ij} W_i W_j + c_N^{-1}(b_1 + 2) g^{i\alpha} g^{j\beta} W_{\alpha\beta} W_i W_j + 2c_N^{-1} g^{i\alpha} g^{rs} W_{\alpha r} W_i W_s \\
& + c_N^{-1} g^{i\alpha} g^{j\beta} g^{k\gamma} \{b_2 K_{\alpha\beta\gamma} + (b_3 + 1) J_{\alpha,\beta\gamma}\} W_i W_j W_k \\
& - c_N^{-1} g^{i\alpha} g^{j\beta} g^{rs} K_{\alpha\beta r} W_i W_j W_s \\
& - c_N^{-1} g^{i\alpha} g^{rt} g^{su} (K_{\alpha rs} + J_{\alpha,rs}) W_i W_t W_u \\
& + c_N^{-1} b_4 g^{i\alpha} g^{rs} (K_{\alpha rs} + J_{\alpha,rs}) W_i + o_p(c_N^{-1}).
\end{aligned}
\tag{2.8}
$$

The class $\mathcal{S}$ in (2.1) is motivated from (2.8).

First, we give the second-order asymptotic expansion of the distribution function of $T \in \mathcal{S}$ under a sequence of local alternatives. This result can be applied to the i.i.d. case, multivariate analysis and time series analysis. Let $G_{\mu,\nu}(z)$ is the distribution function for a non-central chi-square variate with degree of freedom μ and non-centrality parameter ν.

THEOREM 2.1. *The distribution function of $T \in \mathcal{S}$ under a sequence of local alternatives $\theta = \theta_0 + c_N^{-1}\varepsilon$ has the asymptotic expansion*

$$
P_{\theta_0 + c_N^{-1}\varepsilon}[T < z] = G_{p,\Delta}(z) + c_N^{-1} \sum_{j=0}^{3} m_j G_{p+2j,\Delta}(z) + o(c_N^{-1}),
$$

where

$$
\begin{aligned}
m_3 &= \frac{1}{6} K_{\alpha\beta\gamma} d^\alpha d^\beta d^\gamma + \frac{1}{2} a_2^{ijk} g_{ii'} g_{jj'} g_{kk'} d^{i'} d^{j'} d^{k'}, \\
m_2 &= -\frac{1}{2} a_2^{ijk} g_{ii'} g_{jj'} g_{kk'} d^{i'} d^{j'} d^{k'} + \frac{1}{2} B^{\alpha\beta} K_{\alpha\beta\gamma} d^\gamma + \frac{1}{2} a_2^{ijk}[3] g_{il} g_{jk} d^l, \\
m_1 &= \frac{1}{2} J_{\alpha,\beta\gamma} d^\alpha d^\beta d^\gamma - \frac{1}{2}(K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) d^\alpha d^\beta (d^r - \varepsilon^r) \\
&\quad - \frac{1}{2} B^{\alpha\beta} K_{\alpha\beta\gamma} d^\gamma - \frac{1}{2} a_2^{ijk}[3] g_{il} g_{jk} d^l
\end{aligned}
$$

$$
\begin{aligned}
&- \frac{1}{2} g^{rs} (K_{\alpha rs} + J_{\alpha,rs}) d^\alpha + \frac{1}{2} a_3^i g_{ij} d^j, \\
m_0 &= -\frac{1}{6} (K_{\alpha\beta\gamma} + 3 J_{\alpha,\beta\gamma}) d^\alpha d^\beta d^\gamma \\
&+ \frac{1}{2} (K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) d^\alpha d^\beta (d^r - \varepsilon^r) \\
&+ \frac{1}{2} g^{rs} (K_{\alpha rs} + J_{\alpha,rs}) d^\alpha - \frac{1}{2} a_3^i g_{ij} d^j,
\end{aligned}
$$

$\Delta = g_{ij} \varepsilon^i \varepsilon^j$, $d^\alpha = g_{ij} g^{j\alpha} \varepsilon^i$, $a_2^{ijk}[3] = a_2^{ijk} + a_2^{jki} + a_2^{kij}$ *and*

$$
\{B^{\alpha\beta}\} = \{I^{\alpha\beta}\} - \begin{bmatrix} 0 & 0 \\ 0 & (I_{22})^{-1} \end{bmatrix}.
$$

Second, we consider the sensitivity of $T \in \mathcal{S}$ to the change ε_2 in the nuisance parameter. Test statistics that are less sensitive to such changes are generally more desirable because their sizes and powers are less affected by the estimation of the nuisance parameter. Then we have

THEOREM 2.2. (i) *For $T \in \mathcal{S}$, the sensitivity of the distribution function of T to nuisance parameters is given by*

$$
\begin{aligned}
&P_{\theta_0 + c_N^{-1}\varepsilon}[T < z] - P_{\theta_{10} + c_N^{-1}\varepsilon_1, \theta_{20}}[T < z] \\
&\quad = \frac{1}{2} c_N^{-1} (K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) d^\alpha d^\beta \varepsilon^r \{G_{p+2,\Delta}(z) - G_{p,\Delta}(z)\} + o(c_N^{-1}).
\end{aligned}
$$

(ii) *If*

$$
g_{ii'} g^{i'\alpha} g_{jj'} g^{j'\alpha} (K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) = 0, \tag{2.9}
$$

is satisfied, then the distribution function of $T \in \mathcal{S}$ is asymptotically independent of ε_2 with an error $o(c_N^{-1})$.

REMARK 2.2. Note that

$$
\partial_r g_{ij}(\theta) = g_{ii'}(\theta) g^{i'\alpha}(\theta) g_{jj'}(\theta) g^{j'\alpha}(\theta) \{K_{\alpha\beta r}(\theta) + J_{\alpha,\beta r}(\theta) + J_{\beta,\alpha r}(\theta)\}.
$$

If $g_{ij}(\theta)$ is independent of θ_2, then the condition (2.9) holds.

REMARK 2.3. In the case of i.i.d. observations, Li [28] gave factorizations of LR, W_2 and R_2 test statistics as quadratic forms and compared density functions of these factors. Then he showed that the powers and sizes of these statistics are equally sensitive to nuisance parameters. Form (i) in Theorem 2.2, we can see that the powers and sizes of all $T \in \mathcal{S}$ are equally sensitive to nuisance parameters. Hence, our results agree with that of Li [28].

EXAMPLE 2.2. Suppose that X_i, $i = 1, \ldots, N$ are i.i.d. random variables distributed as $N_1(\mu, \sigma^2)$.

(i) If $\theta_1 = \sigma^2$ and $\theta_2 = \mu$, then $g_{11}(\sigma^2, \mu) = (2\sigma^4)^{-1}$. Hence, the condition (2.9) holds.

(ii) If $\theta_1 = \mu$ and $\theta_2 = \sigma^2$, then $g_{11}(\mu, \sigma^2) = (\sigma^2)^{-1}$. Hence, the condition (2.9) does not hold.

EXAMPLE 2.3. Consider the nonlinear regression model

$$X_t = \alpha + \beta \cos(t-1)\lambda + u_t, \qquad t = 1, \ldots, N, \tag{2.10}$$

where $\theta_1 = \beta$, $\theta_2 = (\alpha, \lambda)$, $\lambda = 2\pi l/N$ (l an integer), $\{u_t\}$ is a sequence of i.i.d. $N(0, \sigma^2)$ random variables. Then it follows that

$$I(\theta) = \begin{bmatrix} 1/(2\sigma^2) & 0 & \beta/(4l\sigma^2) \\ 0 & 1/\sigma^2 & \beta/(l\sigma^2) \\ \beta/(4l\sigma^2) & \beta/(l\sigma^2) & \beta^2(8\pi^2 l^2 - 3)/(12 l^2 \sigma^2) \end{bmatrix}. \tag{2.11}$$

For our model (2.10) we calculate $g_{11}(\theta)$. From (2.11)

$$g_{11}(\theta) = \frac{1}{2\sigma^2} - \frac{3}{4\sigma^2(8\pi^2 l^2 - 15)}$$

which implies that the condition (2.9) does not hold.

2.3. Comparison of power

Taking $\varepsilon_1 = \mathbf{0}$ in Theorem 2.1, it can be seen that all $T \in \mathcal{S}$ have sizes $\alpha + o(c_N^{-1})$. Hence, it would be meaningful to compare $T \in \mathcal{S}$ in terms of power up to $o(c_N^{-1})$. From Theorem 2.1, we can see that there is no test which is second order uniformly most powerful in $\mathcal{S}$. Thus we attempt to compare the tests in $\mathcal{S}$ on the basis of their second order power. First, we derive the explicit formula to compare the local power of $T \in \mathcal{S}$. Note that the first order powers of all $T \in \mathcal{S}$ are identical and independent of ε_2. Write the power function of $T \in \mathcal{S}$ under $\theta_0 + c_N^{-1}\varepsilon$ as $P^T(\varepsilon) = P_1(\varepsilon_1) + c_N^{-1} P_2^T(\varepsilon) + o(c_N^{-1})$. From Theorem 2.1, we can state

THEOREM 2.3. *For $T_l \in \mathcal{S}$ with the coefficient $(a_{1l}, a_{2l}^{ijk}, a_{3l}^i)$ $(l = 1, 2)$, respectively,*

$$P_2^{T_1}(\varepsilon) - P_2^{T_2}(\varepsilon) = \sum_{j=0}^{2} m_j' \{G_{p+2j,\Delta}(z) - G_{p+2j+2,\Delta}(z)\},$$

where

$$m_2' = \frac{1}{2}(a_{21}^{ijk} - a_{22}^{ijk}) g_{ii'} g_{jj'} g_{kk'} d^{i'} d^{j'} d^{k'},$$

$$
\begin{aligned}
m_1' &= \frac{1}{2}(a_{21}^{ijk}[3] - a_{22}^{ijk}[3])g_{il}g_{jk}d^l, \\
m_0' &= \frac{1}{2}(a_{31}^i - a_{32}^i)g_{ij}d^j.
\end{aligned}
$$

Note that m_2', m_1' and m_0' are independent of ε_2. From Theorem 2.3 we have

COROLLARY 2.1. *For $T_l \in \mathcal{S}$ with the coefficient $(a_{1l}, a_{2l}^{ijk}, a_{3l}^i)$ $(l = 1, 2)$, respectively,*

$$
P_2^{T_1}(\varepsilon_1, 0) - P_2^{T_2}(\varepsilon_1, 0) = \sum_{j=1}^{2} m_j' \{G_{p+2j,\Delta}(z) - G_{p+2j+2,\Delta}(z)\},
$$

where m_2', m_1' and m_0' are the same as Theorem 2.3.

EXAMPLE 2.4. Suppose that X_i, $i = 1, \ldots, N$ are i.i.d. random variables distributed as

$$
N_2\left(\mu, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).
$$

Then parametric orthogonality holds. If $\theta_1 = \rho$ and $\theta_2 = \mu$, then

$$
\begin{aligned}
g_{11}(\rho, \mu) &= \frac{1 + \rho^2}{(1 - \rho^2)^2}, \quad K_{111}(\rho, \mu) = -\frac{6\rho + 2\rho^3}{(1 - \rho^2)^3} = -J_{1,11}(\rho, \mu), \\
g_{22}(\rho, \mu) &= \frac{2}{1 + \rho}, \qquad K_{122}(\rho, \mu) = \frac{2}{(1 + \rho)^2},
\end{aligned}
\tag{2.12}
$$

and $J_{1,22}(\rho, \mu) = 0$.

For test statistics T_1 and T_2 in (2.7) with the coefficient $(b_{11}, b_{21}, b_{31}, b_{41})$ and $(b_{12}, b_{22}, b_{32}, b_{42})$, respectively,

$$
\begin{aligned}
m_2' &= -\frac{3\rho + \rho^3}{(1 - \rho^2)^3}\{(b_{21} - b_{22}) - (b_{31} - b_{32})\}(\varepsilon_1)^3, \\
m_1' &= -3\frac{3\rho + \rho^3}{(1 - \rho^2)(1 + \rho^2)}\{(b_{21} - b_{22}) - (b_{31} - b_{32})\}\varepsilon_1, \\
m_0' &= \frac{1}{2(1 + \rho)}(b_{41} - b_{42})\varepsilon_1.
\end{aligned}
$$

Based on the above we can compare the second order power among W_i, R_i $(i = 1, 2, 3, 4)$, LR and LR*.

(i) (i1) If $\rho > 0$ and $\varepsilon_1 > 0$, then

$$
\begin{aligned}
&P_2^{\mathrm{W}_4}(\varepsilon) = P_2^{\mathrm{R}_1}(\varepsilon) = P_2^{\mathrm{R}_3}(\varepsilon) < P_2^{\mathrm{W}_2}(\varepsilon) = P_2^{\mathrm{R}_2}(\varepsilon) < P_2^{\mathrm{LR}}(\varepsilon) \\
&\quad < P_2^{\mathrm{W}_1}(\varepsilon) = P_2^{\mathrm{W}_3}(\varepsilon) = P_2^{\mathrm{R}_4}(\varepsilon).
\end{aligned}
$$

(i2) If $\rho < 0$ and $\varepsilon_1 > 0$, then

$$P_2^{\mathrm{W}_1}(\varepsilon) = P_2^{\mathrm{W}_3}(\varepsilon) = P_2^{\mathrm{R}_4}(\varepsilon) < P_2^{\mathrm{LR}}(\varepsilon) < P_2^{\mathrm{W}_2}(\varepsilon) = P_2^{\mathrm{R}_2}(\varepsilon)$$
$$< P_2^{\mathrm{W}_4}(\varepsilon) = P_2^{\mathrm{R}_1}(\varepsilon) = P_2^{\mathrm{R}_3}(\varepsilon).$$

(ii) LR versus LR*,

$$P_2^{\mathrm{LR}}(\varepsilon) - P_2^{\mathrm{LR}^*}(\varepsilon) = -\frac{1}{2(1+\rho)}\varepsilon_1\{G_{p,\Delta}(z) - G_{p+2,\Delta}(z)\}$$

implies, for $\varepsilon_1 > 0$, $P_2^{\mathrm{LR}}(\varepsilon) < P_2^{\mathrm{LR}^*}(\varepsilon)$ and $P_2^{\mathrm{LR}^*}(-\varepsilon_1, \varepsilon_2) < P_2^{\mathrm{LR}}(-\varepsilon_1, \varepsilon_2)$ unless $\rho = -1$.

From (2.12) in Example 2.4, it is seen that cumulants g_{11}, K_{111} and $J_{1,11}$ tend to ∞ as $\rho \to \pm 1$, and g_{22} and K_{122} tend to ∞ as $\rho \to -1$. Hence, we need to inspect second order power functions if ρ is close to ± 1. Note the relation

$$G_{p,\Delta}(z) - G_{p+2,\Delta}(z) = 2f_{p+2,\Delta}(z), \tag{2.13}$$

where $f_{p,\Delta}(z)$ is the probability density function of non-central chi-square variate with p degree of freedom and non-centrality parameter Δ. From (2.12) and (2.13) it follows that second order powers of all test statistics in Example 2.1 converge to 0 as $\rho \to \pm 1$ at each fixed ε_1.

In Figure 2.1, we plotted P_2^{LR} (solid line), $P_2^{\mathrm{LR}^*}$ (dotted line), $P_2^{\mathrm{R}_1}$ (dashed line) and $P_2^{\mathrm{W}_1}$ (dash-dotted line) of Example 2.4 with $\alpha = 0.05$, $\varepsilon_1 = 1$ and $-1 < \rho < 1$. Figure 2.1 illustrates that second order powers of these statistics converge to 0 as $\rho \to \pm 1$.

In Figure 2.2, we plotted P_2^{LR} (solid line), $P_2^{\mathrm{LR}^*}$ (dotted line), $P_2^{\mathrm{R}_1}$ (dashed line) and $P_2^{\mathrm{W}_1}$ (dash-dotted line) of Example 2.4 with $\alpha = 0.05$, $\varepsilon_1 = 0.1$ and $-1 < \rho < 1$. We can see that the extreme points is close to ± 1 in comparison with Figure 2.1.

Figures 2.1 and 2.2 are about here.

EXAMPLE 2.5. Let $\{X_t\}$ be a Gaussian $MA(1)$ process with the spectral density

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}|1 - \psi e^{i\lambda}|^2.$$

If $\theta_1 = \psi$ and $\theta_2 = \sigma^2$, then,

$$g_{11}(\psi, \sigma^2) = \frac{1}{1-\psi^2}, \quad K_{111}(\psi, \sigma^2) = -\frac{6\psi}{(1-\psi^2)^2}, \quad J_{1,11}(\psi, \sigma^2) = \frac{4\psi}{(1-\psi^2)^2},$$
$$g_{22}(\psi, \sigma^2) = \frac{1}{2\sigma^4}, \qquad K_{122}(\psi, \sigma^2) = J_{1,22}(\psi, \sigma^2) = 0. \tag{2.14}$$

Note that $g^{22}(K_{122} + J_{1,22}) = 0$. For test statistics T_1 and T_2 in (2.7) with the coefficient $(b_{11}, b_{21}, b_{31}, b_{41})$ and $(b_{12}, b_{22}, b_{32}, b_{42})$, respectively,

$$\begin{aligned} m_2' &= -\frac{\psi}{(1-\psi^2)^2}\{3(b_{21}-b_{22}) - 2(b_{31}-b_{32})\}(\varepsilon_1)^3, \\ m_1' &= -\frac{3\psi}{1-\psi^2}\{3(b_{21}-b_{22}) - 2(b_{31}-b_{32})\}\varepsilon_1, \\ m_0' &= 0. \end{aligned}$$

Based on the above we can compare the second order power among W_i, R_i $(i = 1, 2, 3, 4)$, LR and LR*.

(i) If $\psi > 0$ and $\varepsilon_1 > 0$, then

$$\begin{aligned} P_2^{\mathrm{W}_4}(\varepsilon) = P_2^{\mathrm{R}_1}(\varepsilon) = P_2^{\mathrm{R}_3}(\varepsilon) < P_2^{\mathrm{R}_2}(\varepsilon) < P_2^{\mathrm{LR}}(\varepsilon) = P_2^{\mathrm{LR}^*}(\varepsilon) = P_2^{\mathrm{W}_2}(\varepsilon) \\ < P_2^{\mathrm{W}_1}(\varepsilon) = P_2^{\mathrm{W}_3}(\varepsilon) = P_2^{\mathrm{R}_4}(\varepsilon). \end{aligned}$$

(ii) If $\psi < 0$ and $\varepsilon_1 > 0$, then

$$\begin{aligned} P_2^{\mathrm{W}_1}(\varepsilon) = P_2^{\mathrm{W}_3}(\varepsilon) = P_2^{\mathrm{R}_4}(\varepsilon) < P_2^{\mathrm{LR}}(\varepsilon) = P_2^{\mathrm{LR}^*}(\varepsilon) = P_2^{\mathrm{W}_2}(\varepsilon) < P_2^{\mathrm{R}_2}(\varepsilon) \\ < P_2^{\mathrm{W}_4}(\varepsilon) = P_2^{\mathrm{R}_1}(\varepsilon) = P_2^{\mathrm{R}_3}(\varepsilon). \end{aligned}$$

From (2.14) in Example 2.5, it is seen that cumulants g_{11}, K_{111} and $J_{1,11}$ tend to ∞ as $\psi \to \pm 1$. Hence, we need to examine second order powers if ψ is close to ± 1. From (2.13) and (2.14) it follows that second order powers of all test statistics in Example 2.1 converge to 0 as $\psi \to \pm 1$ at each fixed ε_1.

In Figure 2.3, we plotted P_2^{LR} (solid line), $P_2^{\mathrm{R}_1}$ (dashed line) and $P_2^{\mathrm{W}_1}$ (dotted line) of Example 2.5 with $\alpha = 0.01$, $\varepsilon_1 = 6.5$ and $-1 < \psi < 1$. From Figure 2.3 we observe that second order powers of these statistics converge to 0 as $\psi \to \pm 1$.

In Figure 2.4, we plotted P_2^{LR} (solid line), $P_2^{\mathrm{R}_1}$ (dashed line) and $P_2^{\mathrm{W}_1}$ (dotted line) of Example 2.5 with $\alpha = 0.01$, $\varepsilon_1 = 0.65$ and $-1 < \psi < 1$. We can see that the extreme points is close to ± 1 in comparison with Figure 2.3.

Figures 2.3 and 2.4 are about here.

EXAMPLE 2.6. Let $\{X_t\}$ be a Gaussian $AR(1)$ process with the spectral density

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{1}{|1-\rho e^{i\lambda}|^2}.$$

If $\theta_1 = \rho$ and $\theta_2 = \sigma^2$, then

$$\begin{aligned} &g_{11}(\rho,\sigma^2) = \frac{1}{1-\rho^2}, \quad K_{111}(\rho,\sigma^2) = \frac{6\rho}{(1-\rho^2)^2}, \quad J_{1,11}(\rho,\sigma^2) = -\frac{2\rho}{(1-\rho^2)^2}, \\ &g_{22}(\rho,\sigma^2) = \frac{1}{2\sigma^4}, \qquad K_{122}(\rho,\sigma^2) = J_{1,22}(\rho,\sigma^2) = 0. \end{aligned} \tag{2.15}$$

Note that $g^{22}(K_{122} + J_{1,22}) = 0$. For test statistics T_1 and T_2 in (2.7) with the coefficient $(b_{11}, b_{21}, b_{31}, b_{41})$ and $(b_{12}, b_{22}, b_{32}, b_{42})$, respectively,

$$\begin{aligned}
m_2' &= -\frac{\rho}{(1-\rho^2)^2}\{3(b_{21}-b_{22}) - (b_{31}-b_{32})\}(\varepsilon_1)^3,\\
m_1' &= -\frac{3\rho}{1-\rho^2}\{3(b_{21}-b_{22}) - (b_{31}-b_{32})\}\varepsilon_1,\\
m_0' &= 0.
\end{aligned}$$

Based on the above we can compare the second order power among W_i, R_i $(i = 1, 2, 3, 4)$, LR and LR*.

(i) If $\rho > 0$ and $\varepsilon_1 > 0$, then

$$\begin{aligned}
P_2^{\mathrm{W}_2}(\varepsilon) < P_2^{\mathrm{LR}}(\varepsilon) &= P_2^{\mathrm{LR}^*}(\varepsilon) = P_2^{\mathrm{W}_1}(\varepsilon) = P_2^{\mathrm{W}_3}(\varepsilon) = P_2^{\mathrm{W}_4}(\varepsilon)\\
&= P_2^{\mathrm{R}_1}(\varepsilon) = P_2^{\mathrm{R}_3}(\varepsilon) = P_2^{\mathrm{R}_4}(\varepsilon) < P_2^{\mathrm{R}_2}(\varepsilon).
\end{aligned}$$

(ii) If $\rho < 0$ and $\varepsilon_1 > 0$, then

$$\begin{aligned}
P_2^{\mathrm{R}_2}(\varepsilon) < P_2^{\mathrm{LR}}(\varepsilon) &= P_2^{\mathrm{LR}^*}(\varepsilon) = P_2^{\mathrm{W}_1}(\varepsilon) = P_2^{\mathrm{W}_3}(\varepsilon) = P_2^{\mathrm{W}_4}(\varepsilon)\\
&= P_2^{\mathrm{R}_1}(\varepsilon) = P_2^{\mathrm{R}_3}(\varepsilon) = P_2^{\mathrm{R}_4}(\varepsilon) < P_2^{\mathrm{W}_2}(\varepsilon).
\end{aligned}$$

From (2.15) in Example 2.6, it is seen that cumulants g_{11}, K_{111} and $J_{1,11}$ tend to ∞ as $\rho \to \pm 1$. Hence, we need to examine second order powers if ρ is close to ± 1. From (2.13) and (2.15) it follows that second order powers of all test statistics in Example 2.1 converge to 0 as $\rho \to \pm 1$ at each fixed ε_1.

In Figure 2.5, we plotted P_2^{LR} (solid line), $P_2^{\mathrm{R}_2}$ (dashed line) and $P_2^{\mathrm{W}_2}$ (dotted line) of Example 2.6 with $\alpha = 0.01$, $\varepsilon_1 = 3$ and $-1 < \rho < 1$. From Figure 2.5 it is seen that second order powers of these statistics converge to 0 as $\rho \to \pm 1$.

In Figure 2.6, we plotted P_2^{LR} (solid line), $P_2^{\mathrm{R}_2}$ (dashed line) and $P_2^{\mathrm{W}_2}$ (dotted line) of Example 2.6 with $\alpha = 0.01$, $\varepsilon_1 = 0.8$ and $-1 < \rho < 1$. We can see that the extreme points is close to ± 1 in comparison with Figure 2.5.

Figures 2.5 and 2.6 are about here.

Next we consider the criterion of average power $P_2^T(\varepsilon_1, \varepsilon_2) + P_2^T(-\varepsilon_1, \varepsilon_2)$. Then from Theorem 2.1 it is easily seen that for each $T \in \mathcal{S}$,

$$\begin{aligned}
&P_2^T(\varepsilon_1, \varepsilon_2) + P_2^T(-\varepsilon_1, \varepsilon_2)\\
&\qquad = (K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) d^\alpha d^\beta \varepsilon^r \{G_{p,\Delta}(z) - G_{p+2,\Delta}(z)\}.
\end{aligned}$$

It is, therefore, clear that the average powers of all $T \in \mathcal{S}$ are identical up to c_N^{-1}. However, even in this situation, with a more detailed analysis it is possible to compare tests in $\mathcal{S}$ in a meaningful way under suitable choice of criterion. Under the absence of nuisance parameters, Mukerjee [33] showed that LR statistic is optimal in terms of second-order local maximinity. However, in the presence of nuisance parameters, optimality properties do not generally hold for LR test in terms of second-order local maximinity. We can see the optimality of LR* statistic in terms of second-order local maximinity. For each fixed Δ, let

$$P_{\varepsilon_2}^{T}(\Delta) = \min P_2^{T}(\varepsilon), \quad P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) = \min P_2^{\mathrm{LR}^*}(\varepsilon),$$

where the minimum is taken over ε_1 such that $g_{ij}\varepsilon^i\varepsilon^j = \Delta$. Then we can get the following result.

THEOREM 2.4. *For $T \in \mathcal{S}$ whose coefficients do not satisfy*

$$z(a_2^{ijk}[3]g_{jk} + g^{i\alpha}B^{\beta\gamma}K_{\alpha\beta\gamma}) + (p+2)\{a_3^i - g^{i\alpha}g^{rs}(K_{\alpha rs} + J_{\alpha,rs})\} = 0,$$

(*the coefficients of LR* satisfy* $a_2^{ijk}[3]g_{jk} = -g^{i\alpha}B^{\beta\gamma}K_{\alpha\beta\gamma}$ *and* $a_3^i = g^{i\alpha}g^{rs}(K_{\alpha rs} + J_{\alpha,rs})$), *there exists a positive* Δ_0 *such that*

$$P_{\varepsilon_2}^{T}(\Delta) < P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta),$$

whenever $0 < \Delta < \Delta_0$.

EXAMPLE 2.7. (i) In Example 2.4, W_1, W_3 and R_4 are most powerful in Example 2.1 except LR* at each fixed $\varepsilon_1 > 0$ and $\rho > 0$ with an error $o(c_N^{-1})$. Hence, we compare W_1 and LR^N tests in terms of second-order local maximinity. Note that the condition (2.9) holds. From Theorem 2.1 and Example 2.4,

$$\begin{aligned}
P_2^{\mathrm{LR}^*}(\varepsilon) &= \frac{3\rho+\rho^3}{(1-\rho^2)^3}(\varepsilon_1)^3\left\{\frac{1}{3}G_{5,\Delta}(z) - G_{3,\Delta}(z) + \frac{2}{3}G_{1,\Delta}(z)\right\},\\
P_2^{\mathrm{W}_1}(\varepsilon) &= \frac{3\rho+\rho^3}{(1-\rho^2)^3}(\varepsilon_1)^3\left\{-\frac{2}{3}G_{7,\Delta}(z) + G_{5,\Delta}(z) - G_{3,\Delta}(z) + \frac{2}{3}G_{1,\Delta}(z)\right\}\\
&\quad + \frac{2(3\rho+\rho^3)}{(1-\rho^2)(1+\rho^2)}\varepsilon_1\{-G_{5,\Delta}(z) + G_{3,\Delta}(z)\}\\
&\quad + \frac{1}{2(1+\rho)}\varepsilon_1\{G_{3,\Delta}(z) - G_{1,\Delta}(z)\},
\end{aligned}$$

where $\Delta = (\varepsilon_1)^2(1+\rho^2)/(1-\rho^2)^2$. If $\rho = 1/2$, $\Delta \le 1$ and $\alpha = 0.05$, then

$$\begin{aligned}
P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) &= P_2^{\mathrm{LR}^*}\left\{-\frac{(1-\rho^2)\Delta^{1/2}}{(1+\rho^2)^{1/2}}, \varepsilon_2\right\},\\
P_{\varepsilon_2}^{\mathrm{W}_1}(\Delta) &= P_2^{\mathrm{W}_1}\left\{-\frac{(1-\rho^2)\Delta^{1/2}}{(1+\rho^2)^{1/2}}, \varepsilon_2\right\}.
\end{aligned}$$

Thus we can see

$$P_{\varepsilon_2}^{\mathrm{W}_1}(\Delta) < P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta),$$

whenever $0 < \Delta \le 1$.

(ii) If $\rho < 0$ and $\varepsilon_1 > 0$, then R_1, R_3 and W_4 are most powerful in Example 2.1 except LR* with an error $o(c_N^{-1})$. Hence, we compare R_1 and LR* tests in terms of second-order local maximinity. Then

$$
\begin{aligned}
P_2^{\mathrm{R}_1}(\varepsilon) = & \frac{3\rho+\rho^3}{(1-\rho^2)^3}(\varepsilon_1)^3\left\{\frac{4}{3}G_{7,\Delta}(z) - G_{5,\Delta}(z) - G_{3,\Delta}(z) + \frac{2}{3}G_{1,\Delta}(z)\right\} \\
& + \frac{4(3\rho+\rho^3)}{(1-\rho^2)(1+\rho^2)}\varepsilon_1\{G_{5,\Delta}(z) - G_{3,\Delta}(z)\} \\
& + \frac{1}{2(1+\rho)}\varepsilon_1\{G_{3,\Delta}(z) - G_{1,\Delta}(z)\}.
\end{aligned}
$$

If $\rho = -1/2$, $\Delta \le 1$ and $\alpha = 0.05$, then

$$
\begin{aligned}
P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) &= P_2^{\mathrm{LR}^*}\left\{\frac{(1-\rho^2)\Delta^{1/2}}{(1+\rho^2)^{1/2}}, \varepsilon_2\right\}, \\
P_{\varepsilon_2}^{\mathrm{R}_1}(\Delta) &= P_2^{\mathrm{R}_1}\left\{-\frac{(1-\rho^2)\Delta^{1/2}}{(1+\rho^2)^{1/2}}, \varepsilon_2\right\}.
\end{aligned}
$$

Thus we can see

$$
P_{\varepsilon_2}^{\mathrm{R}_1}(\Delta) < P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta),
$$

whenever $0 < \Delta \le 1$.

Example 2.8. (i) In Example 2.5, W_1, W_3 and R_4 are most powerful in Example 2.1 at each fixed $\varepsilon_1 > 0$ and $\psi > 0$ with an error $o(c_N^{-1})$. Hence, we compare W_1 and LR* test in terms of second-order local maximinity. For $MA(1)$ model in Example 2.5, the condition (2.9) holds. From Theorem 2.1, we obtain

$$
\begin{aligned}
P_2^{\mathrm{LR}^*}(\varepsilon) = & \frac{\psi}{(1-\psi^2)^2}(\varepsilon_1)^3\{G_{5,\Delta}(z) - 2G_{3,\Delta}(z) + G_{1,\Delta}(z)\}, \\
P_2^{\mathrm{W}_1}(\varepsilon) = & \frac{\psi}{(1-\psi^2)^2}(\varepsilon_1)^3\{-G_{7,\Delta}(z) + 2G_{5,\Delta}(z) - 2G_{3,\Delta}(z) + G_{1,\Delta}(z)\} \\
& + \frac{3\psi}{1-\psi^2}\varepsilon_1\{-G_{5,\Delta}(z) + G_{3,\Delta}(z)\},
\end{aligned}
$$

where $\Delta = (\varepsilon_1)^2/(1-\psi^2)$. If $\psi = 1/2$, $\Delta \le 1$ and $\alpha = 0.01$, then we have

$$
\begin{aligned}
P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) &= P_2^{\mathrm{LR}^*}\{(1-\psi^2)^{1/2}\Delta^{1/2}, \varepsilon_2\}, \\
P_{\varepsilon_2}^{\mathrm{W}_1}(\Delta) &= P_2^{\mathrm{W}_1}\{-(1-\psi^2)^{1/2}\Delta^{1/2}, \varepsilon_2\}.
\end{aligned}
$$

Hence,

$$
P_{\varepsilon_2}^{\mathrm{W}_1}(\Delta) < P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta),
$$

whenever $0 < \Delta \le 1$.

(ii) If $\psi < 0$ and $\varepsilon_1 > 0$, then R_1, R_3 and W_4 are most powerful in Example 2.1 at each fixed $\varepsilon_1 > 0$ and $\psi > 0$ with an error $o(c_N^{-1})$. Hence, we compare R_1 and LR^* test in terms of second-order local maximinity. From Theorem 2.1, we get

$$\begin{aligned} P_2^{\mathrm{R}_1}(\varepsilon) = {} & \frac{\psi}{(1-\beta^2)^2}(\varepsilon_1)^3\{2G_{7,\Delta}(z) - G_{5,\Delta}(z) - 2G_{3,\Delta}(z) + G_{1,\Delta}(z)\} \\ & + \frac{6\psi}{1-\psi^2}\varepsilon_1\{G_{5,\Delta}(z) - G_{3,\Delta}(z)\}. \end{aligned}$$

If $\psi = -1/2$, $\Delta \le 1$ and $\alpha = 0.01$, then

$$\begin{aligned} P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) &= P_2^{\mathrm{LR}^*}\big\{-(1-\psi^2)^{1/2}\Delta^{1/2}, \varepsilon_2\big\}, \\ P_{\varepsilon_2}^{\mathrm{R}_1}(\Delta) &= P_2^{\mathrm{R}_1}\big\{-(1-\psi^2)^{1/2}\Delta^{1/2}, \varepsilon_2\big\}. \end{aligned}$$

Hence,

$$P_{\varepsilon_2}^{\mathrm{R}_1}(\Delta) < P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta),$$

whenever $0 < \Delta \le 1$.

2.4. Effect of nuisance parameters

In this section, we consider the case where the nuisance parameter $\theta_2 = \theta_{20}$ is known. Let $\bar{\theta}_1 = (\bar{\theta}^1, \ldots, \bar{\theta}^p)'$ be the maximum likelihood estimator of θ_1 under $\theta_2 = \theta_{20}$. Any function evaluated at the point $\theta_1 = \bar{\theta}_1$, $\theta_2 = \theta_{20}$ will be distinguished by the addition of a horizontal bar. Then the corresponding statistics with that in Example 2.1 are given by

$$\begin{aligned} & \mathrm{LR}_0 = \mathrm{LR}_0^* = 2(\bar{l}_N - l_N), \\ & \mathrm{W}_{10} = \bar{I}_{(ij)}\tau^i\tau^j, \quad \mathrm{W}_{20} = I_{(ij)}\tau^i\tau^j, \quad \mathrm{W}_{30} = \bar{L}_{(ij)}\tau^i\tau^j, \quad \mathrm{W}_{40} = L_{(ij)}\tau^i\tau^j, \\ & \mathrm{R}_{10} = \bar{I}_0^{ij}Z_iZ_j, \quad \mathrm{R}_{20} = I_0^{ij}Z_iZ_j, \quad \mathrm{R}_{30} = \bar{L}_0^{ij}Z_iZ_j, \quad \mathrm{R}_{40} = L_0^{ij}Z_iZ_j, \end{aligned} \tag{2.16}$$

where $\tau^i = c_N^{-1}(\bar{\theta}^i - \theta_0^i)$, $\{L_{(ij)}(\theta)\} = L_{11}(\theta)$, and $I_0^{ij}(\theta)$ and $L_0^{ij}(\theta)$ are the (i,j) component of the inverse matrix of $I_{11}(\theta)$ and $L_{11}(\theta)$, respectively.

The stochastic expansions of test statistics in (2.16) are given by

$$\begin{aligned} T_0 = {} & I_0^{ij}Z_iZ_j + c_N^{-1}(b_1+2)I_0^{ik}I_0^{jl}W'_{kl}Z_iZ_j \\ & + c_N^{-1}I_0^{ii'}I_0^{jj'}I_0^{kk'}\{b_2K_{i'j'k'} + (b_3+1)J_{i',j'k'}\}Z_iZ_jZ_k + o_p(c_N^{-1}), \end{aligned}$$

where the coefficient (b_1, b_2, b_3) is the same as in (2.7) and $W'_{ij} = Z_{ij} - J_{k,ij}I_0^{kl}Z_l$. Hence, we consider the following class of tests:

$$\mathcal{S}_0 = \{T_0 \mid T_0 = I_0^{ij}Z_iZ_j + c_N^{-1}a_1I_0^{ik}I_0^{jl}W'_{kl}Z_iZ_j$$

$$+ c_N^{-1} a_2^{ijk} Z_i Z_j Z_k + o_p(c_N^{-1}),$$
$$\text{under } H, \text{ where } a_1 \text{ and } a_2^{ijk} \text{ are nonrandom constants}\}.$$

For simplicity we assume the local parametric orthogonality at $\theta = \theta_0$, namely

ASSUMPTION 2.2. $I_{ir} = 0 \;\; i = 1, \ldots, p, r = p+1, \ldots, p+q.$

Then the class $\mathcal{S}$ can be written as

$$\begin{aligned}
\mathcal{S} = \{T \mid T = {} & I_0^{ij} Z_i Z_j + c_N^{-1} a_1 I_0^{ik} I_0^{jl} W_{kl} Z_i Z_j + 2c_N^{-1} I_0^{ij} g^{rs} W_{jr} Z_i Z_s \\
& + c_N^{-1} a_2^{ijk} Z_i Z_j Z_k - c_N^{-1} I_0^{ik} I_0^{jl} g^{rs} K_{klr} Z_i Z_j Z_s \\
& - c_N^{-1} I_0^{ij} g^{rt} g^{su} (K_{jrs} + J_{j,rs}) Z_i Z_t Z_u + c_N^{-1} a_3^i Z_i + o_p(c_N^{-1}), \\
& \text{under } H, \text{ where } a_1, a_2^{ijk} \text{ and } a_3^i \text{ are nonrandom constants}\}.
\end{aligned}$$

Thus the comparison between T and T_0 with the same coefficient will illustrate what influence nuisance parameters exert on the performance of test statistics. Then we have the following theorem.

THEOREM 2.5. (i) *Under* Assumption 2.2, *for $T \in \mathcal{S}$ and $T_0 \in \mathcal{S}_0$ with the same coefficient, the distribution functions of T are decomposed into*

$$\begin{aligned}
& P_{\theta_0 + c_N^{-1}\varepsilon}[T < z] \\
& \quad = P_{\theta_{10} + c_N^{-1}\varepsilon_1, \theta_{20}}[T_0 < z] \\
& \qquad + \frac{1}{2} c_N^{-1} (K_{ijr} + J_{i,jr} + J_{j,ir}) \varepsilon^i \varepsilon^j \varepsilon^r \{G_{p+2,\Delta}(z) - G_{p,\Delta}(z)\} \\
& \qquad + \frac{1}{2} c_N^{-1} \{I_{(ij)} a_3^j - g^{rs}(K_{irs} + J_{i,rs})\} \varepsilon^i \{G_{p+2,\Delta}(z) - G_{p,\Delta}(z)\} + o(c_N^{-1}).
\end{aligned} \tag{2.17}$$

(ii) *If*

$$K_{ijr} + J_{i,jr} + J_{j,ir} = 0, \tag{2.18}$$
$$g^{rs}(K_{irs} + J_{i,rs}) = 0, \tag{2.19}$$

are satisfied, then the distribution function of $T \in \mathcal{S}$ with $a_3^i = 0$ is equal to that of $T_0 \in \mathcal{S}_0$ with the same coefficient as T up to order c_N^{-1}.

REMARK 2.4. The condition (2.18) agrees with (2.9) in Theorem 2.2 under Assumption 2.2. If the condition (2.19) holds, then LR test is second order asymptotically unbiased. Therefore, the third term of the right hand in (2.17) can be interpreted as second order local bias in the usual likelihood ratio test (see Mukerjee [32]). In Section 5, we will observe that this term can also be interpreted as an effect of nuisance parameters in test statistics. Thus, we provide a decomposition formula of local powers for test statistics under local orthogonality for parameters.

EXAMPLE 2.9. This example relates to the ratio of independent exponential means. Let

$$p(x_1, x_2; \mu_1, \mu_2) = (\mu_1 \mu_2)^{-1} \exp\{-(\mu_1^{-1} x_1 + \mu_2^{-1} x_2)\}, \quad x_1, x_2 > 0.$$

(i) If $\theta_1 = \mu_1/\mu_2$ and $\theta_2 = (\mu_1\mu_2)^{1/2}$, then parametric orthogonality holds and $g_{11}(\theta) = (\theta_1)^{-2}/2$ and $g^{22}(K_{122} + J_{1,22}) = 0$. Hence, the conditions (2.18) and (2.19) hold.

(ii) If $\theta_1 = (\mu_1\mu_2)^{1/2}$ and $\theta_2 = \mu_1/\mu_2$, then parametric orthogonality holds and $g_{11}(\theta) = 2(\theta_1)^{-2}$ and $g^{22}(K_{122} + J_{1,22}) = (\theta_1)^{-1}$. Hence, the condition (2.18) holds and (2.19) does not hold.

EXAMPLE 2.10. Let $\{X_t\}$ be a Gaussian $ARMA(1,1)$ process with the spectral density

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi} \frac{|1 - \psi e^{i\lambda}|^2}{|1 - \rho e^{i\lambda}|^2}.$$

(i) If $\theta_1 = \sigma^2$ and $\theta_2 = (\rho, \psi)'$, then parameter orthogonality holds,

$$g_{11}(\sigma^2, \rho, \psi) = (2\sigma^4)^{-1}, \quad I_{22}(\sigma^2, \rho, \psi) = \begin{bmatrix} (1-\rho^2)^{-1} & -(1-\rho\psi)^{-1} \\ -(1-\rho\psi)^{-1} & (1-\psi^2)^{-1} \end{bmatrix},$$

$$K_{122}(\sigma^2\rho\psi) = \frac{2\sigma^{-2}}{1-\rho^2}, \quad J_{1,22}(\sigma^2, \rho, \psi) = -\frac{\sigma^{-2}}{1-\rho^2},$$

$$K_{133}(\sigma^2\rho\psi) = \frac{2\sigma^{-2}}{1-\psi^2}, \quad J_{1,33}(\sigma^2, \rho, \psi) = -\frac{\sigma^{-2}}{1-\psi^2},$$

$$K_{123}(\sigma^2\rho\psi) = -\frac{2\sigma^{-2}}{1-\rho\psi}, \quad J_{1,23}(\sigma^2, \rho, \psi) = \frac{\sigma^{-2}}{1-\rho\psi}.$$

Hence, the condition (2.18) hold, and $g^{rs}(K_{1rs} + J_{1,rs}) = 2\sigma^{-2}$ shows that the condition (2.19) does not hold.

(ii) If $\theta_1 = (\rho, \psi)\prime$ and $\theta_2 = \sigma^2$, then parameter orthogonality holds,

$$I_{11\cdot 2}(\rho, \psi, \sigma^2) = \begin{bmatrix} (1-\rho^2)^{-1} & -(1-\rho\psi)^{-1} \\ -(1-\rho\psi)^{-1} & (1-\psi^2)^{-1} \end{bmatrix},$$

and $g^{33}(K_{i33} + J_{i,33}) = 0$. Hence, the conditions (2.18) and (2.19) hold.

2.5. Unbiased test

We discuss the local unbiasedness of $T \in \mathcal{S}$. Under the absence of nuisance parameters, LR test is locally unbiased. However, under the existence of nuisance parameters, LR test is not generally locally unbiased. From Theorem 2.1, among the test statistics in Example 2.1, LR* test is the only one which is second order asymptotically unbiased unless $g_{ij}g^{j\alpha}g^{rs}(K_{\alpha rs}+J_{\alpha,rs})=0$. If $g_{ij}g^{j\alpha}g^{rs}(K_{\alpha rs}+J_{\alpha,rs})=0$, then $\mathrm{LR}=\mathrm{LR}^*+o_p(c_N^{-1})$. Hence, LR test is locally unbiased. Since $T\in\mathcal{S}$ is not generally unbiased, we consider modification of $T\in\mathcal{S}$ to $T^*=h(\hat{\theta}_1)T+c_N^{-1}A^i\tilde{Z}_i$ so that T^* is second order asymptotically unbiased, where $h(\theta_1)$ is a smooth function and A^i is a nonrandom constant. The following theorem asserts that this is accomplished by choosing A^i and $h_i(\theta_1)=\partial_i h(\theta_1)$ satisfy appropriate conditions.

THEOREM 2.6. *Suppose that $h(\theta_1)$ is a continuously two times differentiable function with $h(\theta_{10})=1$ and A^i is a nonrandom constant. Then, for $T\in\mathcal{S}$, the modified test $T^*=h(\hat{\theta}_1)T+c_N^{-1}A^i\tilde{Z}_i$ is second order asymptotically unbiased if $h_i=h_i(\theta_{10})$ and A^i satisfy*

(i) $h_i=-\frac{1}{p+2}(g_{ij}g^{j\alpha}B^{\beta\gamma}K_{\alpha\beta\gamma}+g_{ij}g_{kl}a_2^{jkl}[3])$,

(ii) $A^i=g^{i\alpha}g^{rs}(K_{\alpha rs}+J_{\alpha,rs})-a_3^i$.

For $h(\theta_1)$ and A^i satisfying (i) and (ii), respectively, in Theorem 2.6, from Theorem 2.1, we can get the asymptotic expansion of the distribution function of T^*.

THEOREM 2.7. *Suppose that $h(\theta_1)$ and A^i satisfy* (i) *and* (ii), *respectively, in Theorem 2.6. Then, for $T\in\mathcal{S}$, the distribution function of the modified test $T^*=h(\hat{\theta}_1)T+c_N^{-1}A^i\tilde{Z}_i$ under a sequence of local alternatives $\theta=\theta_0+c_N^{-1}\varepsilon$ has the second order asymptotic expansion*

$$P_{\theta_0+c_N^{-1}\varepsilon}[T^*<z]=G_{p,\Delta}(z)+c_*^{-1}\sum_{j=0}^{3}m_j^*G_{p+2j,\Delta}(z)+o(c_N^{-1}),$$

where

$$\begin{aligned}
m_3^* &= \frac{1}{6}K_{\alpha\beta\gamma}d^\alpha d^\beta d^\gamma-\frac{1}{2(p+2)}g_{ij}B^{\alpha\beta}K_{\alpha\beta\gamma}d^\gamma d^i d^j\\
&\quad+\frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'}-\frac{1}{2(p+2)}a_2^{ijk}[3]g_{ii'}g_{jk}g_{j'k'}d^{i'}d^{j'}d^{k'},\\
m_2^* &= \frac{1}{2(p+2)}g_{ij}B^{\alpha\beta}K_{\alpha\beta\gamma}d^\gamma d^i d^j\\
&\quad-\frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'}+\frac{1}{2(p+2)}a_2^{ijk}[3]g_{ii'}g_{jk}g_{j'k'}d^{i'}d^{j'}d^{k'},\\
m_1^* &= \frac{1}{2}J_{\alpha,\beta\gamma}d^\alpha d^\beta d^\gamma-\frac{1}{2}(K_{\alpha\beta r}+J_{\alpha,\beta r}+J_{\beta,\alpha r})d^\alpha d^\beta(d^r-\varepsilon^r),\\
m_0^* &= -\frac{1}{6}(K_{\alpha\beta\gamma}+3J_{\alpha,\beta\gamma})d^\alpha d^\beta d^\gamma\\
&\quad+\frac{1}{2}(K_{\alpha\beta r}+J_{\alpha,\beta r}+J_{\beta,\alpha r})d^\alpha d^\beta(d^r-\varepsilon^r).
\end{aligned} \tag{2.20}$$

If $p = 1$, then

$$a_2^{ijk} g_{ii'} g_{jj'} g_{kk'} - \frac{1}{(p+2)} a_2^{ijk}[3] g_{ii'} g_{jk} g_{j'k'} = 0. \tag{2.21}$$

In this case we observe that the coefficients m_3^*, m_2^*, m_1^* and m_0^* in (2.20) are independent of $T \in \mathcal{S}$, and hence all the powers of the modified tests T^* are identical up to second order. On the other hand, if $p \geq 2$, then uniform results are not available.

EXAMPLE 2.11. Consider the $ARMA(1,1)$ model in Example 2.10 (ii). For test statistics in (2.7),

$$\begin{aligned} a_2^{ijk} g_{i1} g_{j1} g_{k1} &= b_2 K_{111} + (b_3 + 1) J_{1,11} \\ &= \frac{2\rho}{(1-\rho^2)^2}(3b_2 - b_3 - 1), \end{aligned} \tag{2.22}$$

and

$$\begin{aligned} \frac{1}{4} a_2^{ijk}[3] g_{i1} g_{jk} g_{11} &= \frac{3}{4} b_2 K_{1ij} g^{ij} g_{11} + \frac{1}{4}(b_3+1) J_{1,ij}[3] g^{ij} g_{11} \\ &= \frac{3}{4} b_2 \left\{ \frac{6\rho}{(1-\rho^2)^2} + \frac{4\psi}{(1-\rho^2)(1-\rho\psi)} \right\} \\ &\quad + \frac{1}{4}(b_3+1) \frac{12\rho^3\psi - 10\rho^2\psi^2 - 8\rho^2 + 4\psi^2 + 2}{(1-\rho^2)^2(1-\rho\psi)(\rho-\psi)}. \end{aligned} \tag{2.23}$$

(2.22) and (2.23) show that (2.21) does not holds.

We give factorizations of $T \in \mathcal{S}$ as quadratic forms. By direct computation, $T \in \mathcal{S}$ can be factorized as

$$T = g^{ij} T_i T_j + o_p(c_N^{-1}),$$

where

$$\begin{aligned} T_i = W_i &+ \frac{1}{2} c_N^{-1} a_1 g_{ij} g^{j\alpha} g^{k\beta} W_{\alpha\beta} W_k + c_N^{-1} g_{ij} g^{j\alpha} g^{rs} W_{\alpha r} W_s \\ &+ \frac{1}{2} c_N^{-1} g_{ij} a_2^{jkl} W_k W_l - \frac{1}{2} c_N^{-1} g_{ij} g^{j\alpha} g^{k\beta} g^{rs} K_{\alpha\beta r} W_k W_s \\ &- \frac{1}{2} c_N^{-1} g_{ij} g^{j\alpha} g^{rt} g^{su} (K_{\alpha rs} + J_{\alpha,rs}) W_t W_u + \frac{1}{2} c_N^{-1} g_{ij} a_3^j + o_p(c_N^{-1}). \end{aligned}$$

Then the asymptotic mean of T_i under $\theta = \theta_0$ is given by

$$\begin{aligned} E_{\theta_0}[T_i] = &\frac{1}{2} c_N^{-1} g_{ij} g_{kl} a_2^{jkl} - \frac{1}{2} c_N^{-1} g_{ij} g^{j\alpha} g^{rs} (K_{\alpha rs} + J_{\alpha,rs}) \\ &+ \frac{1}{2} c_N^{-1} g_{ij} a_3^j + o(c_N^{-1}). \end{aligned}$$

Similarly, we consider factorizations of $T_0 \in \mathcal{S}_0$ as quadratic forms. The asymptotic mean of T_{i0} under $\theta = \theta_0$, where $T_0 = I_0^{ij} T_{i0} T_{j0} + o_p(c_N^{-1})$, is given by

$$E_{\theta_0}[T_{i0}] = \frac{1}{2} c_N^{-1} I_{(ij)} I_{(kl)} a_2^{jkl} + o(c_N^{-1}).$$

Under Assumption 2.2, A^i in Theorem 2.6 can be written as

$$c_N^{-1} A^i = 2 I_0^{ij} (E_{\theta_0}[T_{j0}] - E_{\theta_0}[T_j]) + o(c_N^{-1}).$$

Note that the third term of the right hand in (2.17) is given by $E_{\theta_0}[T_i] - E_{\theta_0}[T_{i0}]$. Therefore, this term (and hence A^i) can be interpreted as a effect of nuisance parameters in $T \in \mathcal{S}$.

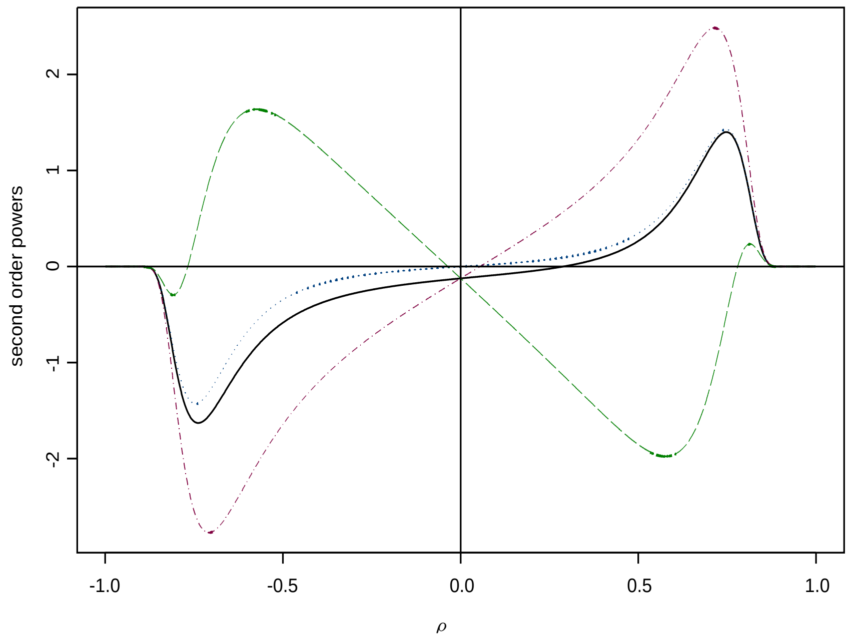

Figure 2.1: For the bivariate normal model with correlation coefficient $\theta_1 = \rho$, both means $\theta_2 = \mu$ and both variances 1 in Example 2.4, second order powers of LR, LR*, R_1 and W_1 statistics are plotted. $P_2^{\mathrm{LR}}(\varepsilon)$ (solid line), $P_2^{\mathrm{LR}^*}(\varepsilon)$ (dotted line), $P_2^{\mathrm{R}_1}(\varepsilon)$ (dashed line) and $P_2^{\mathrm{W}_1}(\varepsilon)$ (dash-dotted line) with $\alpha = 0.05$ and $\varepsilon_1 = 1$.

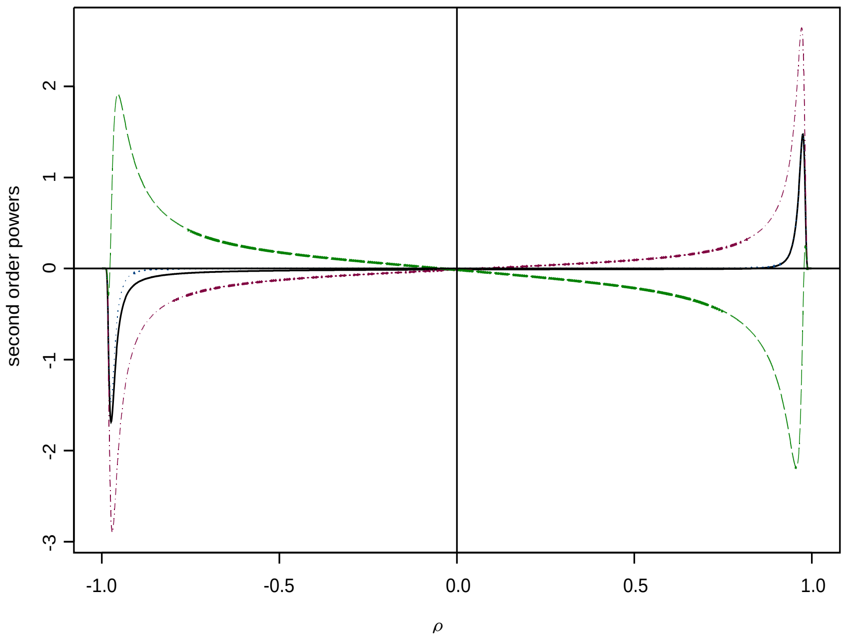

Figure 2.2: For the bivariate normal model with correlation coefficient $\theta_1 = \rho$, both means $\theta_2 = \mu$ and both variances 1 in Example 2.4, second order powers of LR, LR*, R_1 and W_1 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{LR}^*}$ (dotted line), $P_2^{\mathrm{R}_1}$ (dashed line) and $P_2^{\mathrm{W}_1}$ (dash-dotted line) with $\alpha = 0.05$ and $\varepsilon_1 = 0.1$.

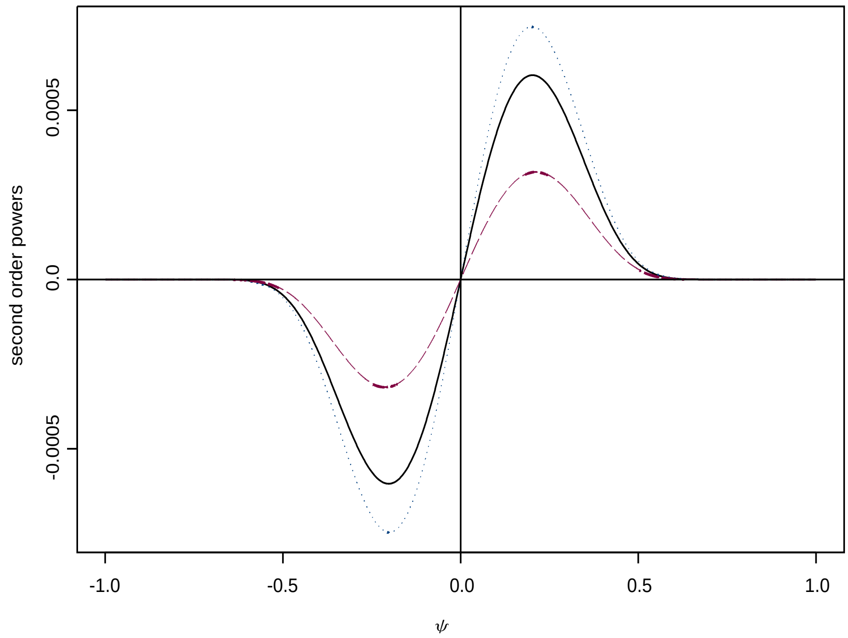

Figure 2.3: For $MA(1)$ model in Example 2.5, second order powers of LR, W_1 and R_1 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_1}$ (dotted line) and $P_2^{\mathrm{R}_1}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 6.5$.

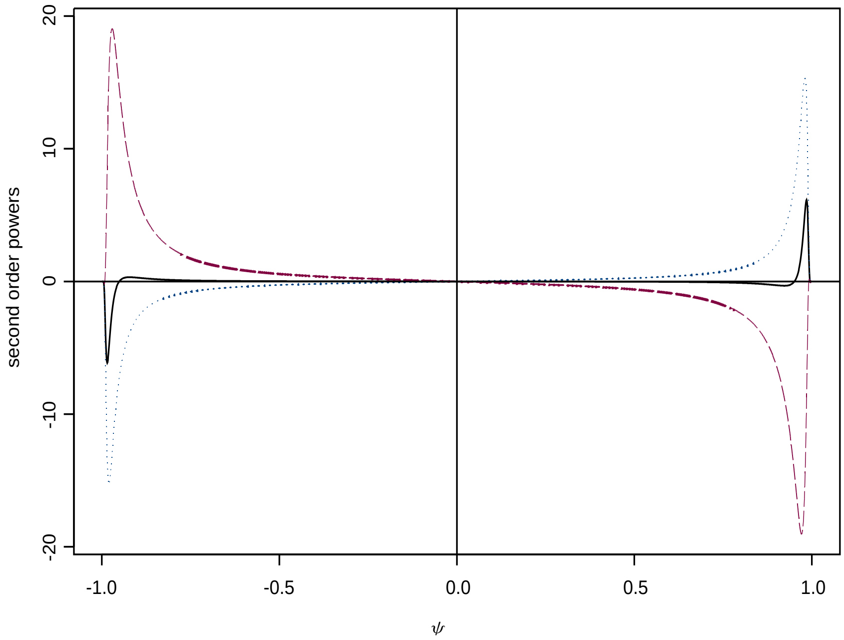

Figure 2.4: For $MA(1)$ model in Example 2.5, second order powers of LR, W_1 and R_1 statistics are plotted. P_2^{LR} (solid line), $P_2^{\text{W}_1}$ (dotted line) and $P_2^{\text{R}_1}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 0.65$.

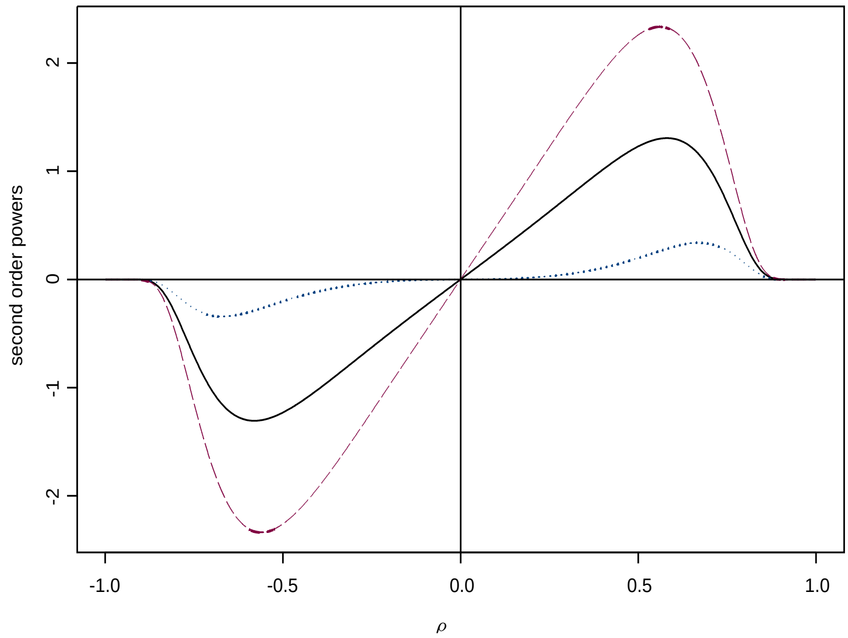

Figure 2.5: For $AR(1)$ model in Example 2.6, second order powers of LR, W_2 and R_2 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_2}$ (dotted line) and $P_2^{\mathrm{R}_2}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 3$.

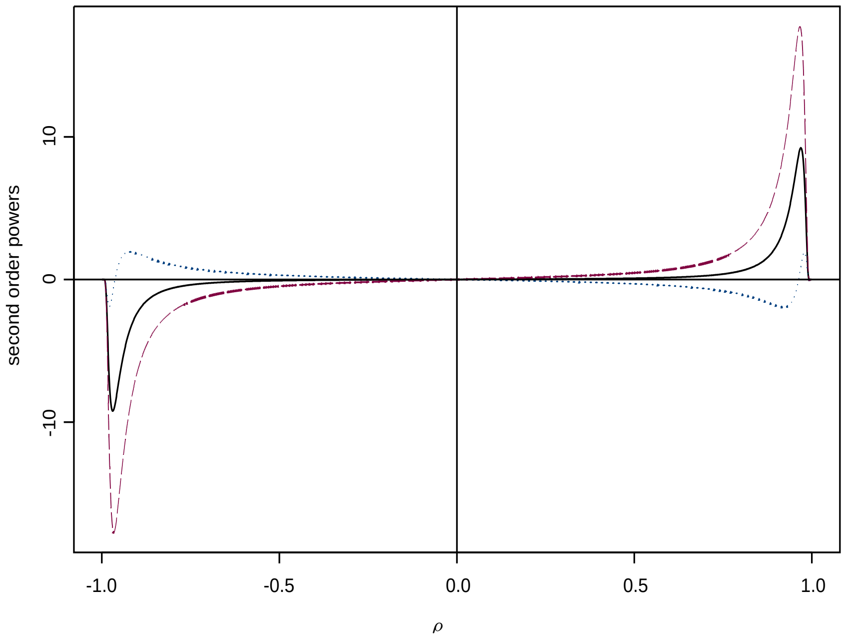

Figure 2.6: For $AR(1)$ model in Example 2.6, second order powers of LR, W_2 and R_2 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_2}$ (dotted line) and $P_2^{\mathrm{R}_2}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 0.8$.

3. Higher Order Asymptotic Option Valuation for Non-Gaussian Dependent Returns

3.1. Introduction

There has been much demand for statistical analysis of dependent observation in many fields, for example, economics, engineering and nature sciences. Financial engineering is the application of engineering methods to financial problems. Time series analysis enables financial engineers to measure and manage their financial risks and to design and analyze sophisticated financial contracts.

One of the main topics in financial engineering is option pricing. Black and Scholes [4] provided the foundation of modern option pricing theory. Despite its usefulness, however, the Black and Scholes theory entails some inconsistencies. It is well known that the model frequently misprices deep in-the-money and deep out-of-the-money options. This result is generally attributed to the unrealistic assumptions used to derive the model. In particular, the Black and Scholes model assumes that stock prices follow a geometric Brownian motion with a constant volatility under an equivalent martingale measure.

In order to avoid this drawback, Jarrow and Rudd [23] proposed a semiparametric option pricing model to account for non-normal skewness and kurtosis in stock returns. This approach aims to approximate the risk-neutral density by a statistical series expansion. Jarrow and Rudd [23] approximated the density of the state price by an Edgeworth series expansion involving the log-normal density. Corrado and Su [9] implemented Jarrow and Rudd's formula to price options. Corrado and Su [10, 11] considered Gram-Charlier expansions for the stock log return rather than the stock price itself. Rubinstein [38] used the Edgeworth expansion for the stock log return. Jurczenko et al. [24] compared these different multi-moment approximate option pricing models. Also they investigated in particular the conditions that ensure the martingale restriction.

As in Kariya [26] and Kariya and Liu [27], the time series structure of return series does not always admit a measure which makes the discounted process a martingale. Hence, we will not able to develop an arbitrage pricing theory by forming an equivalent portfolio. In such a case, we often regard the expected value of the present value of a contingent claim as a proxy for pricing maybe with help of a risk neutrality argument. In view of this, Kariya [26] considered pricing problems with no martingale property and approximated the density of the state price by the Gram-Charlier expansion for the stock log return.

In this chapter, we consider option pricing problems by using Kariya's approach. In Section 3.2, we derive the Edgeworth expansion for the stock log return via extracting dynamics structure of time series. Using this result, we investigate influences of the non-Gaussianity and the dependency of log return processes for option pricing. Numerical studies illuminate some interesting features of the influences. In Section 3.3, we give option prices based on the risk neutrality argument. In Section 3.4, we discuss a consistent estimator of the quantities in our results. Section 3.5 concludes. The proofs of theorems are relegated to Section 6.2.

3.2. Edgeworth expansion of log return

Let $\{S_t; t \geq 0\}$ be the price process of an underlying security at trading time t. The j-th period log return X_j is defined as

$$\log S_{T_0+j\Delta} - \log S_{T_0+(j-1)\Delta} = \Delta\mu + \Delta^{1/2} X_j, \quad j = 1, 2, \ldots, N, \tag{3.1}$$

where T_0 is present time, $N = \tau/\Delta$ is the number of unit time intervals of length Δ during a period $\tau = T - T_0$ and T is the maturity date. Then the terminal price S_T of the underlying security is given by

$$S_T = S_{T_0} \exp\left\{\tau\mu + \left(\frac{\tau}{N}\right)^{1/2} \sum_{j=1}^{N} X_j\right\}. \tag{3.2}$$

REMARK 3.1. In the Black and Scholes option theory the price process is assumed to be a geometric Brownian motion

$$S_T = S_{T_0} \exp(\tau\mu + \sigma W_\tau), \tag{3.3}$$

where the process $\{W_t; t \in \mathbf{R}\}$ is a Wiener process with drift 0 and variance t. From (3.3), the log return at discreterized time point can be written as

$$\log S_{t+j\Delta} - \log S_{t+(j-1)\Delta} = \Delta\mu + \Delta^{1/2}\sigma\nu_j, \quad \nu_j \sim \text{iid } N(0, 1). \tag{3.4}$$

The expression of (3.1) is motivated from (3.4).

First, we derive an analytical expression for the density function of S_T. Since from (3.2) the distribution of S_T depends on that of $Z_N = N^{-1/2} \sum_{j=1}^{N} X_j$, we consider the Edgeworth expansion of the density function of Z_N. If we assume that X_j are independently and identically distributed random variables with mean zero and finite variance, it is easy to give the Edgeworth expansion for Z_N (the classical Edgeworth expansion).

However, a lot of financial empirical studies show that X_j's are not independent. Thus we suppose that $\{X_j\}$ is a dependent process which satisfies the following assumption.

ASSUMPTION 3.1. (i) $\{X_t; t \in \mathbf{Z}\}$ is fourth order stationary in the sense that

(i1) $\mathrm{E}(X_t) = 0$,

(i2) $\mathrm{cum}(X_t, X_{t+u}) = c_{X,2}(u)$,

(i3) $\mathrm{cum}(X_t, X_{t+u_1}, X_{t+u_2}) = c_{X,3}(u_1, u_2)$, (i4) $\mathrm{cum}(X_t, X_{t+u_1}, X_{t+u_2}, X_{t+u_3}) = c_{X,4}(u_1, u_2, u_3)$.

(ii) The cumulants $c_{X,k}(u_1, \ldots, u_{k-1})$, $k = 2, 3, 4$, satisfy

$$\sum_{u_1,\ldots,u_{k-1}=-\infty}^{\infty} \Big(1 + |u_j|^{2-k/2}\Big)|c_{X,k}(u_1, \ldots, u_{k-1})| < \infty$$

for $j = 1, \ldots, k-1$.

(iii) J-th order ($J \geq 5$) cumulants of Z_N are all $O(N^{-J/2+1})$.

Under Assumption 3.1 (ii), $\{X_t; t \in \mathbf{Z}\}$ has the k-th order cumulant spectral density. Let $f_{X,k}$ be the k-th order cumulant spectral density evaluated at frequency $\mathbf{0}$

$$f_{X,k} = (2\pi)^{-(k-1)} \sum_{u_1,\ldots,u_{k-1}=-\infty}^{\infty} c_{X,k}(u_1, \ldots, u_{k-1})$$

for $k = 2, 3, 4$.

First, we state the following result.

THEOREM 3.1. *Suppose that Assumption 3.1* (i)-(iii) *hold. The third order Edgeworth expansion of the density function of $Z = (2\pi f_{X,2})^{-1/2} Z_N$ is given by*

$$\begin{aligned} g(z) = \phi(z)\Bigg\{1 &+ \frac{(2\pi)^{1/2}}{6} N^{-1/2} \frac{f_{X,3}}{(f_{X,2})^{3/2}} H_3(z) - \frac{1}{4\pi} N^{-1} \frac{f'_{X,2}}{f_{X,2}} H_2(z) \\ &+ \frac{\pi}{12} N^{-1} \frac{f_{X,4}}{(f_{X,2})^2} H_4(z) + \frac{\pi}{36} N^{-1} \frac{(f_{X,3})^2}{(f_{X,2})^3} H_6(z)\Bigg\} + o(N^{-1}), \end{aligned} \tag{3.5}$$

where $\phi(\cdot)$ is the standard normal density function, $H_k(\cdot)$ is the k-th order Hermite polynomial and

$$f'_{X,2} = \sum_{u=-\infty}^{\infty} |u| c_{X,2}(u).$$

Many authors have proposed to use different statistical series expansion to price options (see Jarrow and Rudd [23], Corrado and Su [9, 10, 11], Rubinstein [38] and Kariya [26]). Here we give the Edgeworth expansion for the stock log return in powers of $N^{-1/2}$.

A European call option can be viewed as a security which pays at time T its holder the amount

$$X_T^* = \max(S_T - K, 0),$$

where K is the exercise or strike price. As in Kariya [26], we price X_T^* by its discounted expected value;

$$C = \exp(-r\tau)\mathrm{E}_{T_0}(X_T^*), \tag{3.6}$$

where r is the interest rate which is regarded as a constant for the remaining period τ and $\mathrm{E}_{T_0}(\cdot)$ is evaluated at T_0. Evaluate (3.6) based on the density in (2.5). Then writing

$$
\begin{aligned}
d_1 &= (\log S_{T_0}/K + \tau\mu + 2\pi\tau f_{X,2})/(2\pi\tau f_{X,2})^{1/2},\\
d_2 &= d_1 - (2\pi\tau f_{X,2})^{1/2},
\end{aligned}
$$

we obtain the following theorem

THEOREM 3.2. *Let $a_1 = \exp(-r\tau)$ and $a_2 = \exp(\tau\mu + \pi\tau f_{X,2})$. Then*

$$
\begin{aligned}
C = G_0 &+ \frac{(2\pi)^{1/2}}{6} N^{-1/2} \frac{f_{X,3}}{(f_{X,2})^{3/2}} G_3 - \frac{1}{4\pi} N^{-1} \frac{f'_{X,2}}{f_{X,2}} G_2 \\
&+ \frac{\pi}{12} N^{-1} \frac{f_{X,4}}{(f_{X,2})^2} G_4 + \frac{\pi}{36} N^{-1} \frac{(f_{X,3})^2}{(f_{X,2})^3} G_6 + o(N^{-1}),
\end{aligned}
\tag{3.7}
$$

where

$$
\begin{aligned}
G_0 &= a_1\{a_2 S_{T_0}\Phi(d_1) - K\Phi(d_2)\},\\
G_k &= a_1 a_2 S_{T_0} \left\{ \sum_{j=1}^{k-1} (2\pi\tau f_{X,2})^{j/2} H_{k-j-1}(-d_2)\phi(d_1) + (2\pi\tau f_{X,2})^{k/2}\Phi(d_1) \right\},
\end{aligned}
$$

for $k = 2, 3, 4, 6$, where $\Phi(\cdot)$ is the standard normal distribution function.

From (3.7) it is seen that the asymptotic expansion of the option price depends on $f_{X,2}$, $f'_{X,2}$, $f_{X,3}$ and $f_{X,4}$. Hence, we can see influences of the non-Gaussianity and the dependency of the log return processes for the higher order option valuation.

COROLLARY 3.1. *Write*

$$
C = G_0 + N^{-1/2} C_{\mathrm{G},2} + N^{-1} C_{\mathrm{G},3} + N^{-1} C_{\mathrm{D},3} + o(N^{-1}),
$$

where

$$
\begin{aligned}
C_{\mathrm{G},2} &= \frac{(2\pi)^{1/2}}{6} \frac{f_{X,3}}{(f_{X,2})^{3/2}} G_3,\\
C_{\mathrm{G},3} &= \frac{\pi}{12} \frac{f_{X,4}}{(f_{X,2})^2} G_4 + \frac{\pi}{36} \frac{(f_{X,3})^2}{(f_{X,2})^3} G_6,\\
C_{\mathrm{D},3} &= -\frac{1}{4\pi} \frac{f'_{X,2}}{f_{X,2}} G_2.
\end{aligned}
$$

If $\{X_t; t \in \mathbf{Z}\}$ is independent, then $C_{\mathrm{D},3} = 0$. If $\{X_t; t \in \mathbf{Z}\}$ is a Gaussian process, then $C_{\mathrm{G},2} = C_{\mathrm{G},3} = 0$.

Example 3.1. Suppose that X_j, $j = 1, \dots, N$, are independently and identically distributed random variables. Let $c_{X,k} = c_{X,k}(\mathbf{0})$, $k = 2, 3, 4$. Note that $f'_{X,2} = 0$ and $f_{X,k} = (2\pi)^{-(k-1)} c_{X,k}$, $k = 2, 3, 4$. The price of a European call option C_{IID} is given by

$$\begin{aligned} C_{\text{IID}} = G_0 &+ \frac{1}{6} N^{-1/2} \frac{c_{X,3}}{(c_{X,2})^{3/2}} G_3 + \frac{1}{24} N^{-1} \frac{c_{X,4}}{(c_{X,2})^2} G_4 \\ &+ \frac{1}{72} N^{-1} \frac{(c_{X,3})^2}{(c_{X,2})^3} G_6 + o(N^{-1}), \end{aligned}$$

where G_k, $k = 0, 3, 4, 6$, are defined in Theorem 3.2 with $f_{X,2} = (2\pi)^{-1} c_{X,2}$.

If $\mu = r - c_{X,2}/2$, then $a_1 a_2 = 1$ so that G_0 equals the Black and Scholes formula.

Example 3.2. In Example 3.1, suppose that X_j, $j = 1, \dots, N$, are distributed as t-distribution with ν degrees of freedom. Then, for $\nu > 4$

$$C_t = G_{t,0} + N^{-1} G_{t,3} + o(N^{-1}),$$

where

$$\begin{aligned} G_{t,0} &= a_1 \{ a_2 S_{T_0} \Phi(d_1) - K \Phi(d_2) \}, \\ a_2 &= \exp\left\{ \tau\mu + \frac{\tau\nu}{2(\nu-2)} \right\}, \\ d_1 &= \left(\log S_{T_0}/K + \tau\mu + \frac{\tau\nu}{\nu-2} \right) \Big/ \left(\frac{\tau\nu}{\nu-2} \right)^{1/2}, \\ d_2 &= d_1 - \left(\frac{\tau\nu}{\nu-2} \right)^{1/2}, \\ G_{t,3} &= \frac{a_1 a_2 S_{T_0}}{4(\nu-4)} \left\{ \sum_{j=1}^{3} \left(\frac{\tau\nu}{\nu-2} \right)^{j/2} H_{3-j}(-d_2)\phi(d_1) + \left(\frac{\tau\nu}{\nu-2} \right)^2 \Phi(d_1) \right\}. \end{aligned}$$

In order to show influences of higher order terms, in Figure 3.1, we plotted $C_{t,1} = G_{t,0}$ (dotted line) and $C_{t,3} = G_{t,0} + N^{-1} G_{t,3}$ (solid line) of Example 3.2 with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$ ($\Delta = 1/365$), $r = \mu = 0.05$ and $4 < \nu < 9$. From this, we observe that $C_{t,3}$ diverges as $\nu \to 4$.

Figure 3.1 is about here.

Example 3.3. Let $\{X_t : t \in \mathbf{Z}\}$ be the ARCH(1) process

$$X_t = h_t^{1/2} \eta_t \quad \text{and} \quad h_t = \psi_0 + \psi_1 X_{t-1}^2,$$

where $\psi_0 > 0$, $\psi_1 \geq 0$, $\{\eta_t : t \in \mathbf{Z}\}$ is a sequence of independently and identically distributed random variables with

$$E(\eta_t) = 0, \qquad E(\eta_t^2) = 1,$$

$$E(\eta_t^3) = 0, \qquad\qquad E(\eta_t^4) = m, \quad m > 1,$$

and η_t is independent of X_{t-s}, $s > 0$. Then

$$f_{X,2} = \frac{1}{2\pi}\frac{\psi_0}{1-\psi_1}, \qquad f_{X,3} = 0, \qquad f'_{X,2} = 0,$$
$$f_{X,4} = \frac{1}{(2\pi)^3}\frac{\psi_0^2(m - 3 + 5m\psi_1 - 3\psi_1 + 2m\psi_1^2 - 2m\psi_1^3)}{(1-\psi_1)^3(1-m\psi_1^2)},$$

for $ma_1^2 < 1$. Hence,

$$C_{\text{ARCH}(1)} = G_{\text{ARCH}(1),0} + N^{-1}G_{\text{ARCH}(1),3} + o(N^{-1}),$$

where

$$G_{\text{ARCH}(1),0} = a_1 a_2 S_{T_0}\Phi(d_1) - a_1 K\Phi(d_2),$$
$$a_2 = \exp\left\{\tau\mu + \frac{\tau\psi_0}{2(1-\psi_1)}\right\},$$
$$d_1 = \left(\log S_{T_0}/K + \tau\mu + \frac{\tau\psi_0}{1-\psi_1}\right) \bigg/ \left(\frac{\tau\psi_0}{1-\psi_1}\right)^{1/2},$$
$$d_2 = d_1 - \left(\frac{\tau\psi_0}{1-\psi_1}\right)^{1/2},$$
$$G_{\text{ARCH}(1),3} = \frac{a_1 a_2 S_{T_0}}{24}\frac{m - 3 + 5m\psi_1 - 3\psi_1 + 2m\psi_1^2 - 2m\psi_1^3}{(1-\psi_1)(1-m\psi_1^2)}$$
$$\times \left\{\sum_{j=1}^{3}\left(\frac{\tau\psi_0}{1-\psi_1}\right)^{j/2} H_{3-j}(-d_2)\phi(d_1) + \left(\frac{\tau\psi_0}{1-\psi_1}\right)^2 \Phi(d_1)\right\}.$$

In Figure 3.2, we plotted $C_{\text{ARCH}(1),1} = G_{\text{ARCH}(1),0}$ (dotted line) and $C_{\text{ARCH}(1),3} = G_{\text{ARCH}(1),0} + N^{-1}G_{\text{ARCH}(1),3}$ (solid line) of Example 3.3 with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$ ($\Delta = 1/365$), $r = \mu = 0.05$, $m = 3$, $\psi_0 = 0.5$ and $-1/\sqrt{3} < \psi_1 < 1/\sqrt{3}$. Figure 3.2 illuminates influences of higher order terms under Gaussian innovations. From this, we can see that $C_{\text{ARCH}(1),3}$ diverges as $\psi_1 \to \pm 1/\sqrt{3}$.

In Figure 3.3, we plotted $C_{\text{ARCH}(1),1}$ (dotted line) and $C_{\text{ARCH}(1),3}$ (solid line) of Example 3.3 with $S_{T_0} = 100$, $K = 95$, $\tau = 30/365$, $N = 30$, $r = \mu = 0.05$, $\psi_0 = 0.5$, $\psi_1 = 0.3$ and $1 < m < 9$. Figure 3.3 illuminates influences of non-Gaussian innovations. From this, we observe that $C_{\text{ARCH}(1),3}$ decreases as $m \to 9$. The first order term $C_{\text{ARCH}(1),1}$ is a constant because of independence from m.

Figures 3.2 and 3.3 are about here.

Next we consider option pricing problems for a class of processes generated by uncorrelated random variables, which includes the linear process and an important class in time series analysis. Here we are concerned with the following process

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad t \in \mathbf{Z}, \tag{3.8}$$

where $\{\varepsilon_t; t \in \mathbf{Z}\}$ is a sequence of uncorrelated random variables. Instead of (i) and (ii) in Assumption 3.1 we make the following assumption.

ASSUMPTION 3.2. (i′) $\{\varepsilon_t; t \in \mathbf{Z}\}$ is fourth order stationary in the sense that

(i′1) $\mathrm{E}(\varepsilon_t) = 0$,

(i′2) $\mathrm{Var}(\varepsilon_t) = \sigma^2$,

(i′3) $\mathrm{cum}(\varepsilon_t, \varepsilon_{t+u_1}, \varepsilon_{t+u_2}) = c_{\varepsilon,3}(u_1, u_2)$,

(i′4) $\mathrm{cum}(\varepsilon_t, \varepsilon_{t+u_1}, \varepsilon_{t+u_2}, \varepsilon_{t+u_3}) = c_{\varepsilon,4}(u_1, u_2, u_3)$.

(ii′) The cumulants $c_{\varepsilon,k}(u_1, \ldots, u_{k-1})$, $k = 3, 4$, satisfy

$$\sum_{u_1,\ldots,u_{k-1}=-\infty}^{\infty} \left(1 + |u_j|^{2-k/2}\right) |c_{\varepsilon,k}(u_1, \ldots, u_{k-1})| < \infty,$$

for $j = 1, \ldots, k-1$.

(iii′) $\{a_j; j \in \mathbf{Z}\}$ satisfies

$$\sum_{j=0}^{\infty} (1 + |j|)|a_j| < \infty.$$

Under (ii′) in Assumption 3.2, $\{\varepsilon_t; t \in \mathbf{Z}\}$ has the k-th order cumulant spectral density. Let $f_{\varepsilon,k}$ be the k-th order cumulant spectral density evaluated at frequency $\mathbf{0}$

$$f_{\varepsilon,k} = (2\pi)^{-(k-1)} \sum_{u_1,\ldots,u_{k-1}=-\infty}^{\infty} c_{\varepsilon,k}(u_1, \ldots, u_{k-1})$$

for $k = 2, 3, 4$. The response function of $\{a_j; j \in \mathbf{Z}\}$ is defined by

$$A(\lambda) = \sum_{j=0}^{\infty} a_j e^{-ij\lambda}$$

for $-\pi \le \lambda \le \pi$.

Under (i′)-(iii′) in Assumption 3.2, (i) and (ii) in Assumption 3.1 hold. Hence, from Theorem 3.1, we have

COROLLARY 3.2. Suppose that (i′)-(iii′) in Assumption 3.2 and (iii) in Assumption 3.1 hold. Let $a_1 = \exp(-r\tau)$, $a_2 = \exp\bigl(\tau\mu + \frac{1}{2}\tau\sigma^2 A^2\bigr)$ and $A = A(0)$. Then

$$
\begin{aligned}
C = G_0 &+ \frac{2\pi^2 A^3}{3\sigma^3|A|^3} N^{-1/2} f_{\varepsilon,3} G_3 - \frac{1}{2A^2} N^{-1} f'_{\varepsilon,2} G_2 \\
&+ \frac{\pi^3}{3\sigma^4} N^{-1} f_{\varepsilon,4} G_4 + \frac{2\pi^4}{9\sigma^6} N^{-1} f_{\varepsilon,3}^2 G_6 + o(N^{-1}),
\end{aligned}
$$

where

$$
f'_{\varepsilon,2} = 2 \sum_{j_1,j_2=0}^{\infty} |j_2| a_{j_1} a_{j_1+j_2},
$$

G_k, $k = 0, 2, 3, 4, 6$, are given in Theorem 3.2 with

$$
f_{X,2} = \frac{\sigma^2}{2\pi} A^2.
$$

EXAMPLE 3.4. Let $\{X_t; t \in \mathbf{Z}\}$ be AR(1) process

$$
X_t = \rho X_{t-1} + \varepsilon_t, \quad |\rho| < 1.
$$

Note that

$$
A = \frac{1}{1-\rho}, \qquad f'_{\varepsilon,2} = \frac{2\rho}{(1+\rho)(1-\rho)^3}.
$$

The price of a European call option $C_{\mathrm{AR}(1)}$ is given by

$$
C_{\mathrm{AR}(1)} = G_{\mathrm{AR}(1),0} + N^{-1/2} G_{\mathrm{AR}(1),2} + N^{-1} G_{\mathrm{AR}(1),3} + o(N^{-1}),
$$

where

$$
\begin{aligned}
G_{\mathrm{AR}(1),0} &= a_1\{a_2 S_{T_0} \Phi(d_1) - K\Phi(d_2)\}, \\
G_{\mathrm{AR}(1),2} &= \frac{2\pi^2}{3\sigma^3} f_{\varepsilon,3} G_3, \\
G_{\mathrm{AR}(1),3} &= -\frac{\rho}{1-\rho^2} G_2 + \frac{\pi^3}{3\sigma^4} f_{\varepsilon,4} G_4 + \frac{2\pi^4}{9\sigma^6} (f_{\varepsilon,3})^2 G_6, \\
a_2 &= \exp\left\{\tau\mu + \frac{\tau\sigma^2}{2(1-\rho)^2}\right\}, \\
d_1 &= \left\{\log S_{T_0}/K + \tau\mu + \frac{\tau\sigma^2}{(1-\rho)^2}\right\} \Bigg/ \left(\frac{\tau^{1/2}\sigma}{1-\rho}\right), \\
d_2 &= d_1 - \left(\frac{\tau^{1/2}\sigma}{1-\rho}\right), \\
G_k &= a_1 a_2 S_{T_0} \left\{\sum_{j=1}^{k-1} \left(\frac{\tau^{1/2}\sigma}{1-\rho}\right)^j H_{k-j-1}(-d_2)\phi(d_1) + \left(\frac{\tau^{1/2}\sigma}{1-\rho}\right)^k \Phi(d_1)\right\},
\end{aligned}
$$

for $k = 2, 3, 4, 6$.

In order to show influences of higher order terms, in Figure 3.4, $C_{\mathrm{AR}(1),1} = G_{\mathrm{AR}(1),0}$ (dotted line), $C_{\mathrm{AR}(1),2} = G_{\mathrm{AR}(1),0} + N^{-1/2}G_{\mathrm{AR}(1),2}$ (dashed line) and $C_{\mathrm{AR}(1),3} = G_{\mathrm{AR}(1),0} + N^{-1/2}G_{\mathrm{AR}(1),2} + N^{-1}G_{\mathrm{AR}(1),3}$ (solid line) of Example 3.4 are plotted with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$ $(\Delta = 1/365)$, $r = \mu = 0.05$, $\sigma = 1$, $f_{X,3} = -0.1$, $f_{X,4} = 0.2$ and $-1 < \rho < 0.75$. From this, we observe that $C_{\mathrm{AR}(1),k}$, $k = 1, 2, 3$ diverges as $\rho \to 1$.

Figure 3.4 is about here.

In Examples 3.2 and 3.3, although the third order terms diverge, the first order terms do not diverge. On the other hand, in Example 3.4, even the first order term does not converge as $\rho \to 1$. This fact is attributed to finiteness of the variances.

3.3. Martingale restriction

In the previous section, we considered pricing problems with no martingale property. Now we recall that the theoretical price of a option is based on the risk neutrality argument. In this section, to investigate influences of the martingale restriction. we derive the option price based on the risk neutrality argument (see Cox and Ross [12] and Longstaff [31]).

Let

$$d_1^* = (\log S_{T_0}/K + r\tau + \pi\tau f_{X,2})/(2\pi\tau f_{X,2})^{1/2},$$
$$d_2^* = d_1^* - (2\pi\tau f_{X,2})^{1/2}.$$

Then we have

THEOREM 3.3. The fair price C^* of a European call option is given by

$$\begin{aligned} C^* = G_0^* &+ \frac{(2\pi)^{1/2}}{6} N^{-1/2} \frac{f_{X,3}}{(f_{X,2})^{3/2}} G_3^* - \frac{1}{4\pi} N^{-1} \frac{f'_{X,2}}{f_{X,2}} G_2^* \\ &+ \frac{\pi}{12} N^{-1} \frac{f_{X,4}}{(f_{X,2})^2} G_4^* + \frac{\pi}{36} N^{-1} \frac{(f_{X,3})^2}{(f_{X,2})^3} G_6^* + o(N^{-1}), \end{aligned} \tag{3.9}$$

where

$$G_0^* = S_{T_0}\Phi(d_1^*) - e^{-r\tau}K\Phi(d_2^*),$$
$$G_k^* = S_{T_0}\sum_{j=1}^{k-1}(2\pi\tau f_{X,2})^{j/2} H_{k-j-1}(-d_2^*)\phi(d_1^*),$$

for $k = 2, 3, 4$ and

$$G_6^* = S_{T_0}\left[\sum_{j=1}^{2}(2\pi\tau f_{X,2})^{j/2}\left\{H_{5-j}(-d_2^*) - 2\pi\tau f_{X,2}H_{3-j}(-d_2^*)\right\}\right]\phi(d_1^*).$$

EXAMPLE 3.5. Suppose that $\{X_t; t \in \mathbf{Z}\}$ is AR(1) process in Example 3.4. Then the fair price of a European call option $C^*_{\mathrm{AR}(1)}$ is given by

$$C^*_{\mathrm{AR}(1)} = G^*_{\mathrm{AR}(1),0} + N^{-1/2}G^*_{\mathrm{AR}(1),2} + N^{-1}G^*_{\mathrm{AR}(1),3} + o(N^{-1}),$$

where

$$\begin{aligned}
G^*_{\mathrm{AR}(1),0} &= S_{T_0}\Phi(d_1^*) - e^{-r\tau}K\Phi(d_2^*),\\
G^*_{\mathrm{AR}(1),2} &= \frac{2\pi^2}{3\sigma^3} f_{\varepsilon,3}G_3^*,\\
G^*_{\mathrm{AR}(1),3} &= -\frac{\rho}{1-\rho^2}G_2^* + \frac{\pi^3}{3\sigma^4} f_{\varepsilon,4}G_4 + \frac{2\pi^4}{9\sigma^6}(f_{\varepsilon,3})^2 G_6^*,\\
d_1^* &= \left\{\log S_{T_0}/K + r\tau + \frac{\tau\sigma^2}{2(1-\rho)^2}\right\} \bigg/ \left(\frac{\tau^{1/2}\sigma}{1-\rho}\right),\\
d_2^* &= d_1^* - \left(\frac{\tau^{1/2}\sigma}{1-\rho}\right),\\
G_k^* &= S_{T_0}\left\{\sum_{j=1}^{k-1}\left(\frac{\tau^{1/2}\sigma}{1-\rho}\right)^j H_{k-j-1}(-d_2^*)\phi(d_1^*)\right\},
\end{aligned}$$

for $k = 2, 3, 4$ and

$$G_6^* = S_{T_0}\left[\sum_{j=1}^{2}\left(\frac{\tau^{1/2}\sigma}{1-\rho}\right)^j \left\{H_{5-j}(-d_2^*) - \frac{\tau\sigma^2}{(1-\rho)^2}H_{3-j}(-d_2^*)\right\}\right]\phi(d_1^*).$$

In Figure 3.5, we plotted $C^*_{\mathrm{AR}(1),1} = G^*_{\mathrm{AR}(1),0}$ (dotted line), $C^*_{\mathrm{AR}(1),2} = G^*_{\mathrm{AR}(1),0} + N^{-1/2}G^*_{\mathrm{AR}(1),2}$ (dashed line) and $C^*_{\mathrm{AR}(1),3} = G^*_{\mathrm{AR}(1),0} + N^{-1/2}G^*_{\mathrm{AR}(1),2} + N^{-1}G^*_{\mathrm{AR}(1),3}$ (solid line) of Example 3.5 with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$ ($\Delta = 1/365$), $r = 0.05$, $\sigma = 1$, $f_{X,3} = -0.1$, $f_{X,4} = 0.2$ and $-1 < \rho < 1$. Unlike Example 3.4, we observe that $C_{\mathrm{AR}(1),k}$, $k = 1, 2, 3$ converge to $S_{T_0}(= 100)$ as $\rho \to 1$.

3.4. Estimation

From (3.1), X_{j-N_0}, $j = 1, \ldots, N_0$, are available, where $N_0 = T_0/\Delta$. Therefore, in this section we consider to estimate μ, $f_{X,2}$, $f'_{X,2}$, $f_{X,3}$ and $f_{X,4}$ in Theorems 3.1 and 3.2 consistently based on the past observations. From (i) in Assumption 3.1, $\Delta\mu$ is the mean of stock log returns. Hence, a natural unbiased estimator of μ is the sample mean

$$\hat{\mu} = \frac{1}{\Delta N_0}\sum_{j=1}^{N_0}\{\log S_{j\Delta} - \log S_{(j-1)\Delta}\}, \tag{3.10}$$

The variance of $\hat{\mu}$ is given by

$$\mathrm{Var}(\hat{\mu}) = \frac{1}{\Delta N_0} \sum_{u=-(N_0-1)}^{N_0-1} \left(1 - \frac{|u|}{N_0}\right) c_{X,2}(u).$$

Hence, under (ii) in Assumption 3.1, $\hat{\mu}$ given in (3.10) is consistent estimator of μ.

Moreover in order to construct consistent estimator of $f'_{X,2}$, we define the lag window function $w(\cdot)$ which is an even and piecewise continuous function satisfying the conditions,

$$\begin{aligned} & w(0) = 1, \\ & |w(x)| \le 1, \qquad \text{for all } x, \\ & w(x) = 0, \qquad \text{for } |x| > 1. \end{aligned} \tag{3.11}$$

Let

$$\hat{f}'_{X,2} = \sum_{u=-(N_0-1)}^{N_0-1} |u| \hat{c}_{X,2}(u) w(B_{N_0} u),$$

where $\hat{c}_{X,2}(u)$ is the sample autocovariance function at lag u

$$\begin{aligned} \hat{c}_{X,2}(u) = \frac{1}{\Delta N_0} \sum_{j=1}^{N_0-|u|} & \{\log S_{(j+|u|)\Delta} - \log S_{(j+|u|-1)\Delta} - \Delta\hat{\mu}\} \\ & \times \{\log S_{j\Delta} - \log S_{(j-1)\Delta} - \Delta\hat{\mu}\}, \end{aligned} \tag{3.12}$$

and $B_{N_0} \to 0$ as $N_0 \to \infty$, but $(B_{N_0})^3 N_0 \to \infty$. Then we can easily see that under (ii) in Assumption 3.1, $\hat{f}'_{X,2}$ given in (3.12) is a consistent estimator of $f'_{X,2}$.

Since $f_{X,k}$, $k = 2, 3, 4$, are the k-th order cumulant spectral density evaluated at frequency $\mathbf{0}$, using Brillinger and Rosenblatt [6, 7] formula, we construct consistent estimators $\hat{f}_{X,k}$ of $f_{X,k}$ ($k = 2, 3, 4$). See also Brillinger [5]. Thus we can consistently estimate all the quantities in Theorems 3.1 and 3.2 (e.g., G_j, $j = 0, 2, 3, 4, 6.$) by the corresponding quantities replacing μ, $f'_{X,2}$ and $f_{X,k}$ by $\hat{\mu}$, $\hat{f}'_{X,2}$ and $\hat{f}_{X,k}$ ($k = 2, 3, 4$).

For example, we discuss a consistent estimator for New York stock exchange data. The data are daily returns of AMOCO, FORD HP, IBM and MERCK companies. The individual time series are the last 1024 data points from stocks, representing the daily returns for the five companies from February 2, 1984, to December 31, 1991. We used the window functions

$$W(u_1, \ldots, u_{k-1}) = \begin{cases} 2^{-(k-1)} & \text{If } |u_1|, \ldots, |u_{k-1}| \le 1, \\ 0 & \text{otherwise} \end{cases}$$

for $\hat{f}_{X,k}$ ($k = 2, 3, 4$) and Let $w(u) = 1$ for $|u| \le 1$, where $w(u)$ is defined in (3.11). Also we used the bandwidth in frequency direction with $B_{N_0} = 1/50$ for $\hat{f}_{X,2}$, $B_{N_0} = 1/30$ for $\hat{f}_{X,3}$ and $B_{N_0} = 1/10$ for $\hat{f}_{X,4}$ and $\hat{f}'_{X,2}$ (see Brillinger and Rosenblatt [6, 7], and Brillinger [5]).

Table 3.1: Values of Consistent estimators

	AMOCO	FORD	HP	IBM	MERCK
$\hat{\mu}$	0.235103	0.045337	0.133815	0.017165	0.481340
$\hat{f}_{X,2}$	0.002937	0.016006	0.016202	0.003085	0.004534
$\frac{\hat{f}_{X,3}}{(\hat{f}_{X,2})^{3/2}}$	-0.706149	-3.078889	8.501363	0.470144	2.419969
$\frac{\hat{f}_{X,4}}{(\hat{f}_{X,2})^{2}}$	2.278478	-0.280973	8.651378	15.0914	-2.249174
$\frac{\hat{f}'_{X,2}}{\hat{f}_{X,2}}$	-22.78799	-5.520428	0.169291	27.18047	-37.3221

Table 3.1 show these values of consistent estimators of μ, $f'_{X,2}$ and $f_{X,k}$ $(k = 2, 3, 4)$ for the five companies. From this result, we can see that the quantities involved in higher order terms is quite different from the Black and Scholes model. Therefore, in general the assumptions of the Gaussianity and the independency of stock log returns will not hold.

Table 3.2: Option prices

	AMOCO	FORD	HP	IBM	MERCK
C_1	2.776419	4.031663	4.472833	1.699889	4.689151
C_2	2.809884	3.979554	4.434833	1.700269	4.495491
C_3	2.881406	4.345765	6.392765	1.374588	4.650024

Table 3.2 show these values of the approximation up to the first C_1, second C_2 and third order C_3 of the option prices with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$, $r = 0.05$. From this result, we observe that option prices are strongly affected by third order terms except for AMOCO and MERCK.

Table 3.3: Fair prices

	AMOCO	FORD	HP	IBM	MERCK
C_1^*	1.764254	3.827175	3.849221	1.80241	2.138307
C_2^*	1.769475	3.784867	3.954549	1.798532	2.124842
C_3^*	1.83751	4.111153	6.09142	1.481998	2.459177

Table 3.3 show these values of the approximation up to the first C_1^*, second C_2^* and third order C_3^* of the fair prices with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$, $r = 0.05$. From this result, we observe that option prices are strongly affected by third order terms.

3.5. Concluding remark

The Black and Scholes model assumes the Gaussianity and the independency of stock log returns. Empirical studies, however, report that they are not Gaussian nor independent. In this chapter, dropping these two assumptions, we derive a European option pricing. Then, we observed that option prices are strongly affected by the non-Gaussianity and the dependency of stock log returns. Hence, it should be noted that we use option pricing models taking account of the non-Gaussianity and the dependency of stock log returns.

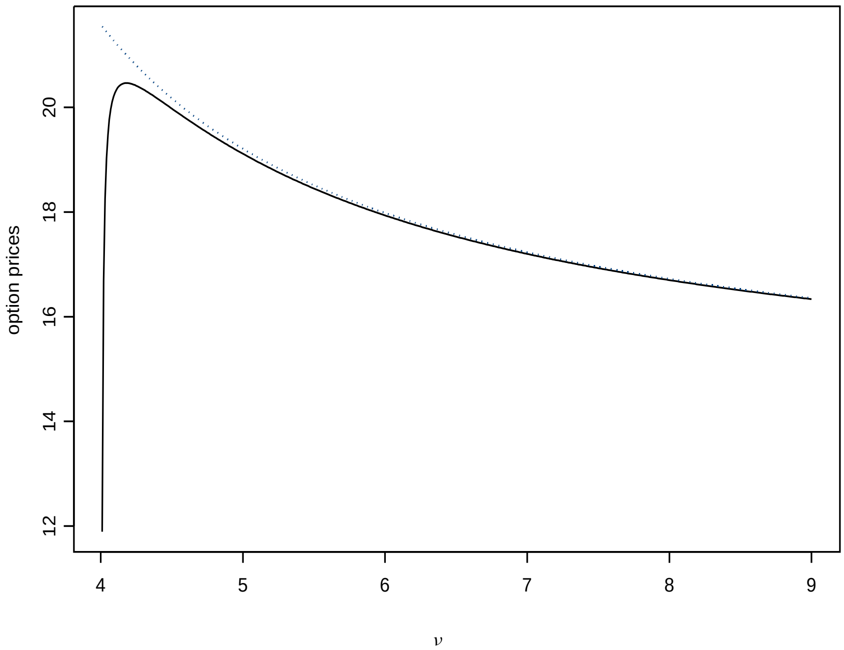

Figure 3.1: For t-distribution with ν degrees of freedom in Example 3.2, the approximation up to the first ($C_{t,1}$, dotted line) and third order ($C_{t,3}$, solid line) of the option price are plotted with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$, $r = \mu = 0.05$ and $4 < \nu < 9$.

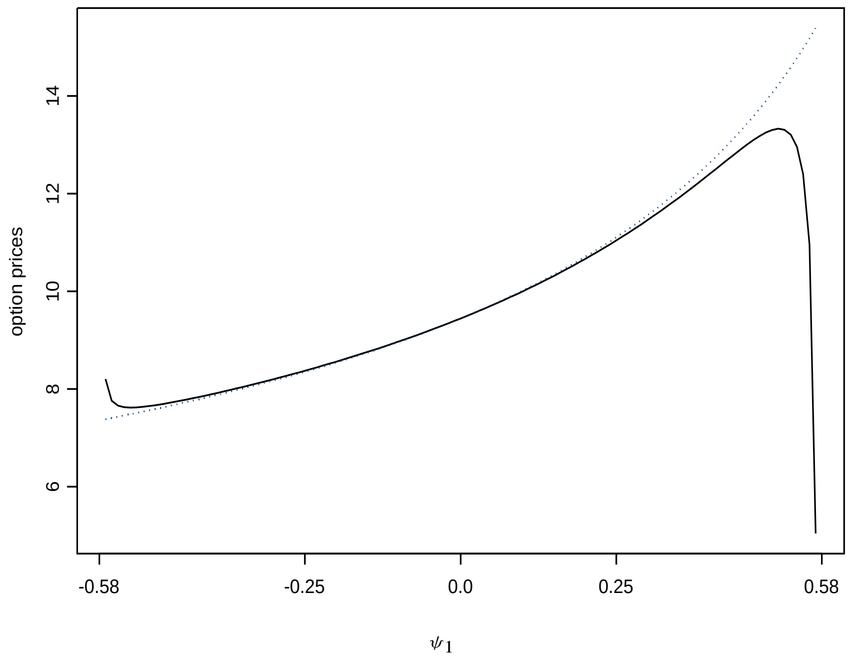

Figure 3.2: For ARCH(1) in Example 3.3, the approximation up to the first ($C_{\mathrm{ARCH}(1),1}$, dotted line) and third order ($C_{\mathrm{ARCH}(1),3}$, solid line) of the option price are plotted with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$, $r = \mu = 0.05$, $m = 3$, $\psi_0 = 0.5$ and $-1/\sqrt{3} < \psi_1 < 1/\sqrt{3}$.

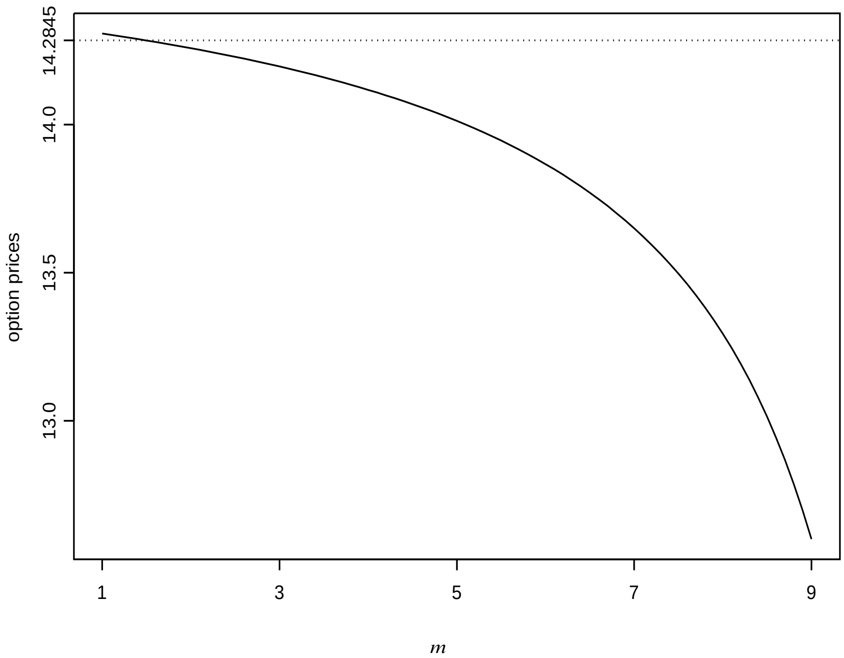

Figure 3.3: For ARCH(1) in Example 3.3, the approximation up to the first ($C_{\text{ARCH(1)},1}$, dotted line) and third order ($C_{\text{ARCH(1)},3}$, solid line) of the option price are plotted with $S_{T_0} = 100$, $K = 95$, $\tau = 30/365$, $N = 30$, $r = \mu = 0.05$, $\psi_0 = 0.5$, $\psi_1 = 0.3$ and $1 < m < 9$.

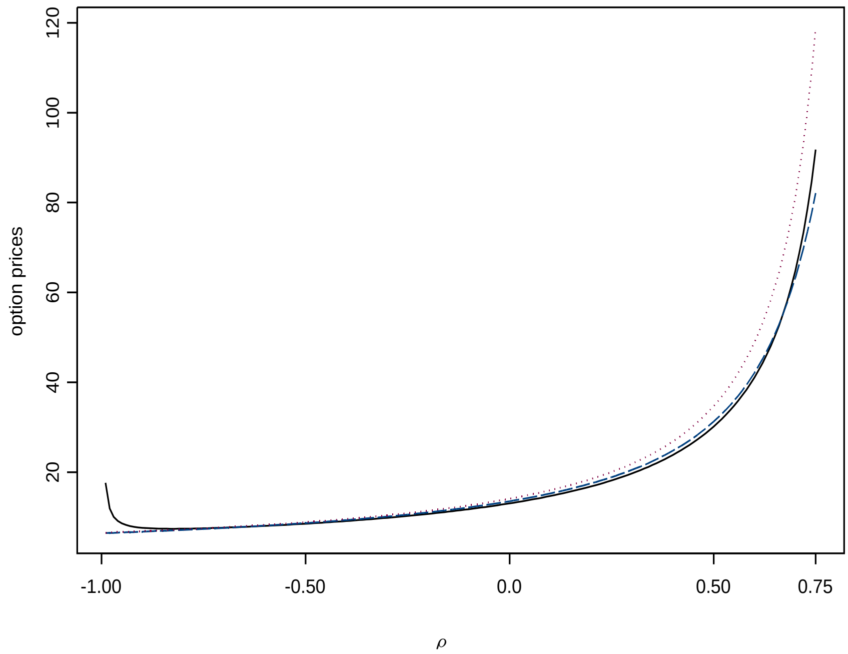

Figure 3.4: For AR(1) in Example 3.4, the approximation up to the first ($C_{\mathrm{AR}(1),1}$, dotted line), second ($C_{\mathrm{AR}(1),2}$, dashed line) and third order ($C_{\mathrm{AR}(1),3}$, solid line) of the option price are plotted with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$, $r = \mu = 0.05$, $\sigma = 1$, $f_{X,3} = -0.1$, $f_{X,4} = 0.2$ and $-1 < \rho < 0.75$.

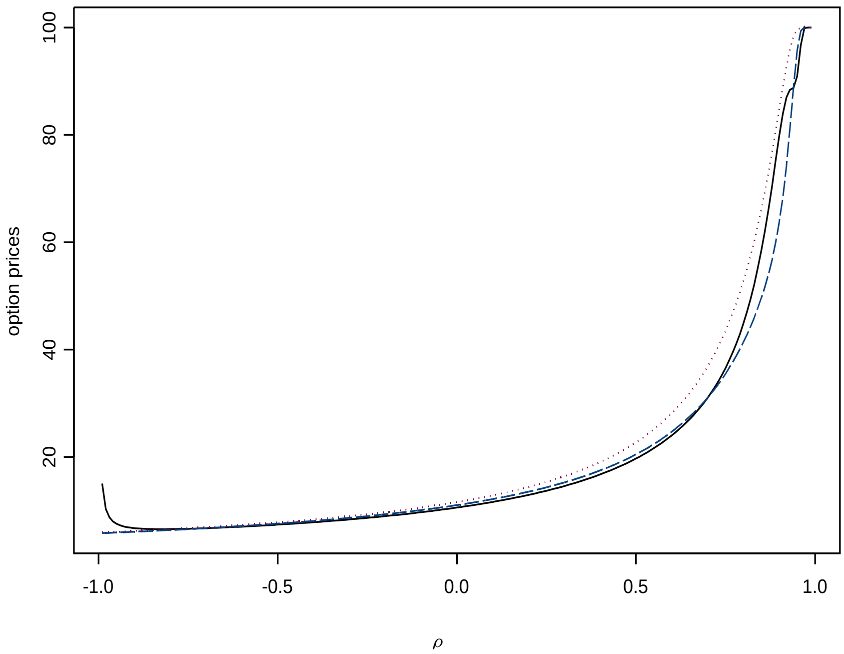

Figure 3.5: For AR(1) in Example 3.5, the approximation up to the first ($C^*_{\mathrm{AR}(1),1}$, dotted line), second ($C^*_{\mathrm{AR}(1),2}$, dashed line) and third order ($C^*_{\mathrm{AR}(1),3}$, solid line) of the option price are plotted with $S_{T_0} = K = 100$, $\tau = 30/365$, $N = 30$, $r = 0.05$, $\sigma = 1$, $f_{X,3} = -0.1$, $f_{X,4} = 0.2$ and $-1 < \rho < 1$.

4. Second Order Optimality for Estimators in Time Series Regression Models

4.1. Introduction

The problem of efficiently estimating the coefficients in a linear regression model has been investigated widely. When the error covariance matrix depends on unknown parameters, the regression coefficients are often estimated by generalized least squares (GLS), using appropriate consistent estimators of the parameters. It is well known that standardized GLS estimators have the same limiting distribution as the best linear unbiased estimator. Rothenberg [37] gave higher order approximations to the distribution of GLS estimators. Toyooka [50, 51] derived the asymptotic expansion of the mean squared errors (MSE). Since these methods are parametric, standard root N asymptotics hold for time domain GLS estimators, where N is the sample size.

If the autocorrelation structure of the unobservable residuals is not parameterized, we then construct efficient estimators by spectral methods. This technique is semiparametric since it relies on a nonparametric spectral estimator of the residuals.

The semiparametric method of a linear regression model was introduced by Hannan [18], who showed that a frequency domain GLS estimator achieves asymptotically the Gauss-Markov efficiency bound under smoothness and Grenander's conditions on the residual spectral density and the regressor sequence, respectively.

There are principal differences between parametric and nonparametric estimation technique that are often given in terms of consistency and rates of convergence. Velasco and Robinson [52] derived Edgeworth expansions for the distribution of nonparametric estimates. Taniguchi et al. [49] discussed higher order asymptotic theory for minimum contrast estimators of spectral parameters. They established that for semiparametric estimation it does not hold in general that first order efficiency implies second order efficiency.

The semiparametric estimation entails the problem of the bandwidth selection. Applications of higher order asymptotic expansions to this problem have been studied by many authors. Robinson [36] studied frequency domain inference on semiparametric and nonparametric models in the presence of a data dependent bandwidth. Linton [29] investigated the second order properties of various quantities in the partially linear model. Xiao and Phillips [54] gave higher order approximations of the MSE of the frequency domain GLS estimators. Linton and Xiao [30] derived asymptotic expansions for semiparametric adaptive regression estimators. They discussed the bandwidth selection based on minimizing the (integrated) MSE. Also Xiao and Phillips [55] discussed higher order approximations for Wald statistics in frequency domain regressions with integrated processes.

Taniguchi et al. [47] established the root N asymptotic theory for functionals of nonparametric spectral density estimators. This is due to the fact that integration of nonparametric spectral density estimators recovers root N consistency. Since the Hannan estimator is based on integral functionals of nonparametric estimators, it may be expected that the Hannan estimator has attractive properties in higher order asymptotic theory.

In this chapter, we will develop the second order asymptotic theory for the frequency domain GLS estimator proposed by Hannan [18]. First, we give the second order Edgeworth expansion of the distribution of the Hannan estimator. Next, we show that the bias-adjusted version of the Hannan estimator is not second order asymptotically Gaussian efficient in general. Of course, if the residual is Gaussian, it is second order asymptotically efficient. As in Xiao and Phillips [54], if the error is a Gaussian process, then it holds that first order efficiency implies second order efficiency.

An interesting result in this chapter is that the second order asymptotic properties are independent of the bandwidth choice for the residual spectral estimator. This implies that the Hannan estimator has the same rate of convergence as in regular parametric estimation. This is a sharp contrast with the general semiparametric estimation theory, where it is known that the second order asymptotic properties are strongly influenced by the bandwidth (e.g., Taniguchi et al. [49]).

This chapter is organized as follows. Section 4.2 gives the basic assumptions entertained in this chapter. Section 4.3 gives a number of preliminary results and the main results on the second order Edgeworth expansions. Section 4.4 reviews the concept of efficiency which is introduced by Akahira and Takeuchi [1]. Section 4.5 contains the discussion on Gaussian efficiency. Proofs are relegated to Section 6.3.

4.2. The model

We consider the following linear regression model

$$y(t) = B'x(t) + u(t), \quad t = 1, \dots, N, \tag{4.1}$$

where $x(t) = (x_1(t), \dots, x_q(t))'$ is a known vector and nonrandom design sequence, $B = \{\beta_{jk}\}$ is a $(q \times p)$-matrix of unknown regression parameters, and $u(t) = (u_1(t), \dots, u_p(t))'$ is an unobserved stationary residual.

The vector process $u(t)$ is supposed to satisfy the following assumption

ASSUMPTION 4.1. (i) $\{u(t)\}$ is a linear process generated by

$$u(t) = \sum_{s=-\infty}^{\infty} A(s)\varepsilon(t-s),$$

where $\varepsilon(t) = (\varepsilon_1(t), \dots, \varepsilon_r(t))'$ are independent identically distributed random vectors with $\mathrm{E}[\varepsilon(t)] = \mathbf{0}$, $\mathrm{E}[\varepsilon(t)\varepsilon(t)'] = G$ and finite absolute moments.

(ii) The $(p \times r)$-matrices $A(s), s = 0, \pm 1, \dots,$ satisfy

$$\sum_{s=-\infty}^{\infty} (1 + |s|^2)\|A(s)\| < \infty,$$

where $\|A\|$ is the square root of the greatest eigenvalue of A^*A and A^* is the conjugate transpose of a matrix A. Then $\{u(t)\}$ has the spectral density matrix

$$F(\lambda) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \Gamma(s) e^{-is\lambda},$$

where $\Gamma(s) = \mathrm{E}[u(t)u(t+s)']$.

(iii) There exists a positive constant γ_1 such that

$$\det\{F(\lambda)\} \geq \gamma_1 > 0$$

for $\lambda \in (-\pi, \pi]$.

REMARK 4.1. Assumption 4.1 (i) and (ii) are satisfied by a wide class of time series models which contains the usual VARMA processes. Under (i) and (ii) in Assumption 4.1, the joint k-th order cumulants of $u_{j_1}(s), u_{j_2}(s+s_1), \ldots, u_{j_k}(s+s_{k-1})$

$$\Gamma_{j_1 \ldots j_k}(s_1, \ldots, s_{k-1}) = \mathrm{cum}^{(k)}[u_{j_1}(s), u_{j_2}(s+s_1), \ldots, u_{j_k}(s+s_{k-1})]$$

exist and satisfy

$$\sum_{s_1, \ldots, s_{k-1}=-\infty}^{\infty} (1 + |s_l|^2) |\Gamma_{j_1 \ldots j_k}(s_1, \ldots, s_{k-1})| < \infty, \quad j_1, \ldots, j_k = 1, \ldots, p$$

for $l = 1, \ldots, k-1$. Then $\{u(t)\}$ has the k-th order cumulant spectral density

$$\begin{aligned} &F_{j_1 \ldots j_k}(\lambda_1, \ldots, \lambda_{k-1}) \\ &\quad = \left(\frac{1}{2\pi}\right)^{k-1} \sum_{s_1, \ldots, s_{k-1}=-\infty}^{\infty} \Gamma_{j_1 \ldots j_k}(s_1, \ldots, s_{k-1}) e^{-i(s_1\lambda_1 + \cdots + s_{k-1}\lambda_{k-1})}. \end{aligned}$$

Assumption 4.1 (i)-(iii) imply that $F(\lambda)^{-1}$ exists and has the Fourier series representation

$$F(\lambda)^{-1} = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \Delta(s) e^{is\lambda}, \qquad \sum_{s=-\infty}^{\infty} (1 + |s|^2) \|\Delta(s)\| < \infty.$$

This follows from an application of a famous theorem due to Wiener (see, for example, [53, Section 12]).

Let $d_j(N)$ be the positive square root of $\sum_{t=1}^{N} \{x_j(t)\}^2$ for $j = 1, \ldots, q$ and

$$D_N = \mathrm{diag}\{d_1(N), \ldots, d_q(N)\}.$$

We impose some assumptions on $\{x(t)\}$.

ASSUMPTION 4.2. (i) $\{x(t)\}$ is uniformly bounded; that is, there exists a positive constant γ_2 such that

$$\sup_{t\in\mathbf{Z}}|x_j(t)| < \gamma_2, \quad j = 1, \dots, q.$$

(ii) There exists $\gamma_3 > 0$ such that $\{d_j(N)\}^2 \geq \gamma_3 N$ for $j = 1, \dots, q$.

(iii) There exist η_j such that

$$\sum_{t=1}^{N} \frac{x_j(t)}{d_j(N)} = N^{1/2}\eta_j + O(N^{-1/2}), \quad j = 1, \dots, q.$$

(iv) There exist regression spectral measures $M_{j_1\dots j_k}(\lambda_1, \dots, \lambda_{k-1})$ such that

$$\begin{aligned}\sum_{t=1}^{N} &\frac{x_{j_1}(t)x_{j_2}(t+l_1)\cdots x_{j_k}(t+l_{k-1})}{d_{j_1}(N)\cdots d_{j_k}(N)} \\ &= N^{-k/2+1}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi} e^{i(l_1\lambda_1+\cdots+l_{k-1}\lambda_{k-1})}dM_{j_1\dots j_k}(\lambda_1, \dots, \lambda_{k-1}) \\ &\quad + O(N^{-k/2})\end{aligned}$$

for $k = 2, 3, \dots$.

(v) $R(0)$ is nonsingular. Here $R(0)$ is the $(q \times q)$-matrix given by

$$R(l) = \int_{-\pi}^{\pi} e^{il\lambda}dM(\lambda), \quad l = 0, \pm 1, \dots,$$

where $M(\lambda) = \{M_{jk}(\lambda)\}$.

REMARK 4.2. Assumption 4.2 is a higher order version for Grenander's conditions. For example, linear combinations of harmonic functions satisfy Assumption 4.2 (i)-(v). Let us consider a example of η_j and $M_{j_1\dots j_k}(\lambda_1, \dots, \lambda_{k-1})$.

EXAMPLE 4.1 (Harmonic trend)**.** Suppose $x_j(t) = \cos\nu_j t$, $j = 1, \dots, q$, where $0 < \nu_1 < \cdots < \nu_q < \pi$. From the relation

$$\sum_{t=1}^{N}\cos\nu t = \frac{1}{2}\left\{\frac{\sin(N+1/2)\nu}{\sin\nu/2} - 1\right\}, \quad \nu \neq 0, \pm 2\pi, \dots,$$

it is seen that

$$\sum_{t=1}^{N}\frac{x_j(t)}{d_j(N)} = \frac{1}{\sqrt{2}}N^{-1/2}\left\{\frac{\sin(N+1/2)\nu_j}{\sin\nu_j/2} - 1\right\} + O(N^{-3/2}),$$

which means $\eta_j = 0$.

It is well known that $M(\lambda)$ has a jump $\mathrm{diag}(0, \dots, 0, 1/2, 0, \dots, 0)$ (1/2 is in the j-th diagonal) at $\lambda = \pm\nu_j$.

To construct the Hannan estimator, we use the spectral window $W_N(\cdot)$ and the lag window $w(\cdot)$ which satisfy the following assumption

ASSUMPTION 4.3. (i) The function $W_N(\lambda)$ can be expanded as

$$W_N(\lambda) = \frac{1}{2\pi} \sum_{l=-M}^{M} w\left(\frac{l}{M}\right) e^{-il\lambda}.$$

(ii) $w(x)$ is a continuous, even function with $w(0) = 1$ and $w(x) = 0$ for $|x| \geq 1$, and satisfies

$$|w(x)| \leq 1,$$
$$\lim_{x \to 0} \frac{1 - w(x)}{|x|^2} < \infty.$$

(iii) $M = M(N)$ satisfies

$$M/N^{1/3} + N^{1/4}/M \to 0 \quad \text{as } N \to \infty.$$

REMARK 4.3. It is easy to see that the Tukey-Hanning window and Parzen window satisfy Assumption 4.3 (i) and (ii) (see Hannan [19, pp. 278–279]).

As in Hannan [18], we define for two sequences $y(t)$ and $x(t)$ of N scalars

$$\widehat{F}_{yx}(\lambda) = \frac{1}{2\pi N} \sum_{l=-M}^{M} w\left(\frac{l}{M}\right) \sum_{m=1+\underline{l}}^{N-\overline{l}} y(m)x(m+l)e^{-il\lambda},$$

where $\underline{l} = \max(0, -l)$ and $\overline{l} = \max(0, l)$ for $l \in \mathbf{Z}$.

This serves to define all such functions as

$$\hat{F}_{y_j y_k}(\lambda), \qquad \hat{F}_{x_j x_k}(\lambda), \qquad \hat{F}_{u_j u_k}(\lambda), \qquad \hat{F}_{y_j x_k}(\lambda), \qquad \hat{F}_{u_j x_k}(\lambda).$$

We also use the matrix notation

$$\hat{F}_{yy}(\lambda) = \{\hat{F}_{y_j y_k}(\lambda)\}, \quad \hat{F}_{xx}(\lambda) = \{\hat{F}_{x_j x_k}(\lambda)\}, \quad \hat{F}_{uu}(\lambda) = \{\hat{F}_{u_j u_k}(\lambda)\},$$
$$\hat{F}_{yx}(\lambda) = \{\hat{F}_{y_j x_k}(\lambda)\}, \quad \hat{F}_{ux}(\lambda) = \{\hat{F}_{x_j u_k}(\lambda)\}.$$

It is not assumed that all of them are estimates of well defined spectral density matrices. Indeed $\hat{F}_{uu}(\lambda)$ is constructed from the actual $u(t)$ and not estimates of them.

We consider a frequency domain version of (4.1), viz.

$$\hat{F}_{yx}(\lambda) = B' \hat{F}_{xx}(\lambda) + \hat{F}_{ux}(\lambda),$$

which we rewrite in the tensor notation

$$\hat{f}_{yx}(\lambda) = \{I_p \otimes \hat{F}_{xx}(\lambda)'\}\beta + \hat{f}_{ux}(\lambda),$$

where $\hat{f}_{yx}(\lambda) = \mathrm{vec}\big[\hat{F}_{yx}(\lambda)'\big]$, $\hat{f}_{ux}(\lambda) = \mathrm{vec}\big[\hat{F}_{ux}(\lambda)'\big]$, $\beta = \mathrm{vec}[B]$, and I_p is the $(p \times p)$ identity matrix.

The Hannan estimator of β in an integration version is given by

$$\hat{\beta} = \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{F}_{uu}(\lambda)^{-1} \otimes \hat{F}_{xx}(\lambda)' d\lambda\right]^{-1} \times \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \big\{\tilde{F}_{uu}(\lambda) \otimes I_q\big\}^{-1} \hat{f}_{yx}(\lambda) d\lambda\right]. \tag{4.2}$$

Since the actual $u(t)$ is unobservable, the quantity $\hat{F}_{uu}(\lambda)$ is infeasible. Therefore, we use $\tilde{F}_{uu}(\lambda)$ for the estimate of $F(\lambda)$ obtained from the residuals, $\tilde{u}(t) = y(t) - \hat{B}_{\mathrm{LS}}' x(t)$, from the least squares regression. Then $\tilde{F}_{uu}(\lambda)$ can be calculated directly as

$$\tilde{F}_{uu}(\lambda) = \hat{F}_{yy}(\lambda) - \hat{F}_{yx}(\lambda)\hat{B}_{\mathrm{LS}} - \hat{B}_{\mathrm{LS}}'\hat{F}_{xy}(\lambda) + \hat{B}_{\mathrm{LS}}'\hat{F}_{xx}(\lambda)\hat{B}_{\mathrm{LS}}.$$

Hannan [18] showed that under very general conditions, $\hat{\beta}$ is first order asymptotically Gaussian efficient; that is, the distribution of $(I_p \otimes D_N)(\hat{\beta} - \beta)$ converges as $N \to \infty$ to the multivariate normal distribution with zero mean vector and covariance matrix given by

$$\mathcal{I}^{-1} = \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} F(\lambda)^{-1} \otimes dM(\lambda)'\right]^{-1},$$

(see also Hannan [19]).

4.3. Second order asymptotic theory

It is well known that integration of nonparametric estimators recovers root N consistency (cf. Taniguchi et al. [47]). Since $\hat{\beta}$ in (4.2) is based on integral functionals of nonparametric estimators, it may be expected that $\hat{\beta}$ has attractive properties in higher order asymptotic theory. Thus we consider the second order asymptotic properties of the estimator $\hat{\beta}$. First, we give the following theorem.

THEOREM 4.1. *The stochastic expansion for $(I_p \otimes D_N)(\hat{\beta} - \beta)$ is given by*

$$(I_p \otimes D_N)(\hat{\beta} - \beta) = \mathcal{I}^{-1} Z_1 - N^{-1/2}\mathcal{I}^{-1}(Z_2 - \mathrm{E}[Z_2]) - N^{-1/2}\mathcal{I}^{-1}\mathrm{E}[Z_2] + N^{-1/2}\mathcal{I}^{-1} Z_3 \mathcal{I}^{-1} Z_1 + o_p(N^{-1/2}),$$

where

$$Z_1 = \frac{N}{2\pi}\int_{-\pi}^{\pi} \big\{F(\lambda)^{-1} \otimes D_N{}^{-1}\big\} \hat{f}_{ux}(\lambda) d\lambda,$$

$$Z_2 = \frac{N^{3/2}}{2\pi}\int_{-\pi}^{\pi} \big\{F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} \otimes D_N{}^{-1}\big\} \hat{f}_{ux}(\lambda) d\lambda,$$

$$
\begin{aligned}
Z_3 &= \frac{N^{3/2}}{2\pi} \int_{-\pi}^{\pi} \{F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\} \otimes \{D_N{}^{-1} \hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda, \\
V_1(\lambda) &= \hat{F}_{uu}(\lambda) - \mathrm{E}[\hat{F}_{uu}(\lambda)].
\end{aligned}
$$

Next, we evaluate the asymptotic cumulants of Z_j, $j = 1, 2, 3$ given in Theorem 4.1. Denote by $Z_1(jk)$ and $Z_2(jk)$ the $(j-1)q+k$-th component of the vectors Z_1 and Z_2, respectively. Similarly, denote by $Z_3(j_1k_1, j_2k_2)$ the $((j_1-1)q+k_1, (j_2-1)q+k_2)$-th element of the matrix Z_3. Then we have the following lemma.

LEMMA 4.1.

$$
\begin{aligned}
&\mathrm{E}[Z_1] = \mathbf{0}, \\
&\mathrm{E}[Z_2(jk)] = \sum_{j_1, j_2=1}^{p} \mathcal{K}_{jj_1j_2}(0,0) F_{j_1j_2}(0) \eta_k + \mathrm{o}(1), \\
&\mathrm{E}[Z_3] = \mathbf{0}, \\
&\mathrm{Cov}[Z_1] = \mathcal{I} + o(N^{-1/2}), \\
&\mathrm{Cov}[Z_1, Z_2] = O(M/N^{1/2}), \\
&\mathrm{Cov}[Z_1(j_1k_1), Z_3(j_2k_2, j_3k_3)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{K}_{j_1j_2j_3}(\lambda, -\lambda) \eta_{k_1} dM_{k_2k_3}(\lambda) + o(1), \\
&\mathrm{cum}[Z_1(j_1k_1), Z_1(j_2k_3), Z_1(j_3k_3)] \\
&\qquad = N^{-1/2} \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathcal{K}_{j_1j_2j_3}(\lambda_1, \lambda_2) dM_{k_1k_2k_3}(\lambda_1, \lambda_2) + o(N^{-1/2}),
\end{aligned}
$$

where

$$
\mathcal{K}_{jkl}(\lambda_1, \lambda_2) = F^{jj'}(-\lambda_1 - \lambda_2) F^{kk'}(\lambda_1) F^{ll'}(\lambda_2) F_{j'k'l'}(-\lambda_1, -\lambda_2),
$$

and $F^{jk}(\lambda)$ is the (j,k)-th element of the matrix $F(\lambda)^{-1}$. Here we use the Einstein summation convention.

Denote by $\mathcal{I}^{j_1k_1, j_2k_2}$ the $((j_1-1)q+k_1, (j_2-1)q+k_2)$-th element of the matrix $\mathcal{I}^{-1}$. From Theorem 4.1 and Lemma 4.1 the asymptotic cumulants of $(I_p \otimes D_N)(\hat{\beta} - \beta)_{jk} = d_k(N)(\hat{\beta}_{kj} - \beta_{kj})$ are evaluated as follows:

$$
\begin{aligned}
&\mathrm{E}[(I_p \otimes D_N)(\hat{\beta} - \beta)_{jk}] \\
&\quad = -N^{-1/2} \mathcal{I}^{jk, j_1k_1} \sum_{j_2, j_3=1}^{p} \mathcal{K}_{j_1j_2j_3}(0,0) F_{j_2j_3}(0) \eta_{k_1} \\
&\qquad + N^{-1/2} \frac{1}{2\pi} \mathcal{I}^{jk, j_1k_1} \mathcal{I}^{j_2k_2, j_3k_3} \int_{-\pi}^{\pi} \mathcal{K}_{j_3j_1j_2}(\lambda, -\lambda) \eta_{k_3} dM_{k_1k_2}(\lambda) \\
&\qquad + o(N^{-1/2})
\end{aligned}
$$

$$= N^{-1/2} C^{jk} + o(N^{-1/2}), \quad \text{(say)},$$

$$\mathrm{Cov}[(I_p \otimes D_N)(\hat{\beta} - \beta)_{j_1 k_1}, (I_p \otimes D_N)(\hat{\beta} - \beta)_{j_2 k_2}] = \mathcal{I}^{j_1 k_1, j_2 k_2} + o(N^{-1/2}),$$

$$\begin{aligned}
&\mathrm{cum}[(I_p \otimes D_N)(\hat{\beta} - \beta)_{j_1 k_1}, (I_p \otimes D_N)(\hat{\beta} - \beta)_{j_2 k_2}, (I_p \otimes D_N)(\hat{\beta} - \beta)_{j_3 k_3}] \\
&\quad = N^{-1/2} \frac{1}{2\pi} \mathcal{I}^{j_1 k_1, j_1' k_1'} \mathcal{I}^{j_2 k_2, j_2' k_2'} \mathcal{I}^{j_3 k_3, j_3' k_3'} \\
&\qquad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathcal{K}_{j_1' j_2' j_3'}(\lambda_1, \lambda_2) dM_{k_1' k_2' k_3'}(\lambda_1, \lambda_2) \\
&\qquad + o(N^{-1/2}), \\
&\quad = N^{-1/2} C^{j_1 k_1, j_2 k_2, j_3 k_3} + o(N^{-1/2}), \quad \text{(say)}.
\end{aligned}$$

The L-th order cumulants of $(I_p \otimes D_N)(\hat{\beta} - \beta)_{jk}$ satisfy

$$\mathrm{cum}^{(L)}[(I_p \otimes D_N)(\hat{\beta} - \beta)_{j_1 k_1}, \ldots, (I_p \otimes D_N)(\hat{\beta} - \beta)_{j_L k_L}] = O(N^{-L/2+1})$$

for each $L \geq 3$.

From the general Edgeworth expansion formula (e.g., Taniguchi and Kakizawa [48, pp. 169]) we get the following theorem.

THEOREM 4.2.

$$\begin{aligned}
P_\beta[(I_p \otimes D_N)(\hat{\beta} - \beta) \leq z] = \int_{-\infty}^{z} N(w : \mathcal{I}^{-1}) \Big[& 1 + N^{-1/2} C^{jk} H_{jk}(w, \mathcal{I}^{-1}) \\
& + \frac{1}{6} N^{-1/2} C^{j_1 k_1, j_2 k_2, j_3 k_3} H_{j_1 k_1, j_2 k_2, j_3 k_3}(w, \mathcal{I}^{-1}) \Big] dw \\
& + o(N^{-1/2}),
\end{aligned}$$

where z and w are the pq-vectors with z_{jk} and w_{jk} in $(j-1)q + k$-th place, respectively,

$$N(w : \mathcal{I}^{-1}) = (2\pi)^{-pq/2} |\mathcal{I}|^{1/2} \exp\left(-\frac{1}{2} w' \mathcal{I} w \right),$$

the multivariate normal distribution, and multivariate Hermite polynomials:

$$H_{j_1 k_1, \ldots, j_s k_s}(w, \mathcal{I}^{-1}) = \frac{(-1)^s}{N(w : \mathcal{I}^{-1})} \frac{\partial^s}{\partial w_{j_1 k_1} \ldots \partial w_{j_s k_s}} N(w : \mathcal{I}^{-1}).$$

The preceding results are unexpected.

REMARK 4.4. In the context of semiparametric estimation, it is known that root-N asymptotics in general do not hold (e.g., Taniguchi et al. [49]). However, our results claim that, in a linear regression model, standard root-N asymptotics hold up to second order. This means that the Hannan estimator has the same rate of convergence as regular parametric estimation. Moreover, it is seen that our Edgeworth expansion is independent of the bandwidth and the window type function for the residual spectra. This is in sharp contrast with the general semiparametric estimation theory.

We examine of the performance of the second order Edgeworth expansion given in Theorem 4.2. The model used for data generation is the following:

$$\begin{aligned} y(t) &= \beta x(t) + u(t), \quad (p = q = 1) \\ u(t) &= au(t-1) + \varepsilon(t), \end{aligned}$$

where $|a| < 1$, $\varepsilon(t)$'s are i.i.d. $Exp(0,1)$ random variables with probability density

$$p(z) = \exp\{-(z+1)\}, \quad z > -1.$$

In the following Figure 4.1-4.4, we plotted of the first (solid) and the second (dotted) order approximation, and empirical distribution (dashes) which is obtained by 10000 times replications. From Figure 4.1-4.4, we observed that the second order Edgeworth expansions are quite accurate in the neighborhood of $z = 0$.

Figures 4.1-4.4 are about here.

4.4. Second order efficiency

We consider the approach of Akahira and Takeuchi [1] whose argument proceeds as follows. Let $X_1, \ldots, X_N$ be a sequence of random variables forming a stochastic process, and possessing the probability measure $P_\theta^N[\cdot]$, where $\theta = (\theta^1, \ldots, \theta^p) \in \Theta$, a subset of $\mathbf{R}^p$. We assume that $\theta_2 = (\theta^2, \ldots, \theta^p)$ is a nuisance parameter (see, Section 1.2 and 4.4 in Akahira and Takeuchi [1]). If an estimator $\hat{\theta}^1$ of θ^1 satisfies the equation

$$\lim_{N\to\infty} \sqrt{N} \left| P_\theta^N[\sqrt{N}(\hat{\theta}^1 - \theta^1) \le 0] - 1/2 \right| = 0,$$

then $\hat{\theta}^1$ is called a second order asymptotically median unbiased (second order AMU) estimator. For this $\hat{\theta}^1$, the asymptotic distribution functions $F_\theta^+(x) + G_\theta^+(x)/\sqrt{N}$ and $F_\theta^-(x) + G_\theta^-(x)/\sqrt{N}$ are defined to be the second order asymptotically distribution of $\sqrt{N}(\hat{\theta}^1 - \theta^1)$ if

$$\lim_{N\to\infty} \sqrt{T} \left| P_\theta^N[\sqrt{N}(\hat{\theta}^1 - \theta^1) \le x^1] - F_\theta^+(x^1) - G_\theta^+(x^1)/\sqrt{N} \right| = 0$$

for all $x^1 \ge 0$,

$$\lim_{N\to\infty} \sqrt{N} \left| P_\theta^N[\sqrt{N}(\hat{\theta}^1 - \theta^1) \le x^1] - F_\theta^-(x^1) - G_\theta^-(x^1)/\sqrt{N} \right| = 0$$

for all $x^1 < 0$.

For $\theta_0 = (\theta_0^1, \dots, \theta_0^p) \in \Theta$, consider the problem of testing hypothesis $H : \theta^1 = \theta_0^1 + x^1/\sqrt{N}$ ($x^1 > 0$) against alternative $A : \theta = \theta_0$. We define $\beta_{\theta_0}^+(x^1)$ and $\gamma_{\theta_0}^+(x^1)$ as follows:

$$\sup_{\{\phi_N \in \Phi_{1/2}\}} \limsup_{N\to\infty} \sqrt{N}\big\{E_{\theta_0}^N[\phi_N] - \beta_{\theta_0}^+(x^1) - \gamma_{\theta_0}^+(x^1)/\sqrt{N}\big\} = 0, \tag{4.3}$$

where

$$\Phi_{1/2} = \big\{\phi_N : E^N_{\theta_0^1 + x^1/\sqrt{N}, \theta_2}[\phi_N] = 1/2 + o(1/\sqrt{N}),\ 0 \le \phi_N \le 1\big\}.$$

Then we have for $x^1 \ge 0$

$$F_{\theta_0}^+(x^1) \le \beta_{\theta_0}^+(x^1) \quad \text{and} \quad G_{\theta_0}^+(x^1) \le \gamma_{\theta_0}^+(x^1).$$

Also consider the problem of testing hypothesis $H : \theta^1 = \theta_0^1 + x^1/\sqrt{N}$ ($x^1 < 0$) against alternative $A : \theta = \theta_0$. We define $\beta_{\theta_0}^-(x^1)$ and $\gamma_{\theta_0}^-(x^1)$ as follows:

$$\inf_{\{\phi_N \in \Phi_{1/2}\}} \liminf_{N\to\infty} \sqrt{N}\big\{E_{\theta_0}^N[\phi_N] - \beta_{\theta_0}^-(x^1) - \gamma_{\theta_0}^-(x^1)/\sqrt{N}\big\} = 0.$$

In the same way as for the case $x^1 > 0$, we have for each $x^1 < 0$

$$F_{\theta_0}^-(x^1) \ge \beta_{\theta_0}^-(x^1) \quad \text{and} \quad G_{\theta_0}^-(x^1) \ge \gamma_{\theta_0}^-(x^1).$$

Thus we make the following definition.

DEFINITION 4.1 (Akahira and Takeuchi [1]). A second order AMU estimator $\hat{\theta}^1$ is called second order asymptotically efficient if for each $\theta \in \Theta$

$$P_\theta^N[\sqrt{N}(\hat{\theta}^1 - \theta^1) \le x^1] = \begin{cases} \beta_{\theta_0}^+(x^1) + \gamma_{\theta_0}^+(x^1)/\sqrt{N} + o(1/\sqrt{N}) & \text{for all } x^1 \ge 0 \\ \beta_{\theta_0}^-(x^1) + \gamma_{\theta_0}^-(x^1)/\sqrt{N} + o(1/\sqrt{N}) & \text{for all } x^1 < 0. \end{cases}$$

The above definition means that second order asymptotic efficiency implies highest probability concentration around the true value with respect to the second order asymptotic distribution.

4.5. Efficiency of Hannan's estimator

In this section we discuss higher order asymptotic efficiency of the Hannan estimator $\hat{\beta}$ defined by (4.2). To discuss higher order efficiency and establish unified higher order results we need to restrict the class of estimators to second order asymptotically median unbiased (AMU).

From theorem 4.2, it can be seen that $\hat{\beta}$ is not second order AMU. Thus we modify $\hat{\beta}$ as follows:

$$\begin{aligned}\hat{\beta}^{*jk} &= \hat{\beta}^{jk} - N^{-1/2}(I_p \otimes D_N)^{-1}\tilde{C}^{jk} \\ &\quad + \frac{1}{6}N^{-1/2}(I_p \otimes D_N)^{-1}(\tilde{\mathcal{I}}^{jk,jk})^{-1}\tilde{C}^{jk,jk,jk},\end{aligned}$$

where

$$\tilde{\mathcal{I}} = \frac{N}{2\pi}\int_{-\pi}^{\pi} \tilde{F}_{uu}(\lambda)^{-1} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda,$$

and, $\tilde{C}^{jk}$ and $\tilde{C}^{jk,jk,jk}$ are the quantities replacing the cumulant spectrum by the nonparametric spectral estimator in C^{jk} and $C^{jk,jk,jk}$, respectively.

Then we have the following theorem

THEOREM 4.3. (i) *The estimator $\hat{\beta}^{*jk}$ is second order AMU.*

(ii) *The second order asymptotic distribution of $\hat{\beta}^* = \{\hat{\beta}^{*jk}\}$ is*

$$\begin{aligned}&P_\beta[(I_p \otimes D_N)(\hat{\beta}^* - \beta) \le z] \\ &= \int_{-\infty}^{z} N(w : \mathcal{I}^{-1})\bigg[1 + \frac{1}{6}N^{-1/2}C^{jk,jk,jk}H_{jk}(w, \mathcal{I}^{-1}) \\ &\quad + \frac{1}{6}N^{-1/2}C^{j_1k_1,j_2k_2,j_3k_3}H_{j_1k_1,j_2k_2,j_3k_3}(w, \mathcal{I}^{-1})\bigg]dw + o(N^{-1/2}).\end{aligned}$$

Since $\hat{\beta}$ is first order asymptotically efficient under Gaussian errors, we concentrate our attention only the Gaussian efficiency. From Akahira and Takeuchi [1], the second order Gaussian efficient bound distribution of jk-component is given by

$$\Phi((\mathcal{I}_{\mathrm{B}}^{jk,jk})^{-1/2}z) + \mathrm{o}(N^{-1/2}),$$

where $\mathcal{I}_{\mathrm{B}}^{j_1k_1,j_2k_2}$ is (j_1k_1, j_2k_2)-component of the covariance matrix $\mathcal{I}_{\mathrm{B}}^{-1}$ of the best linear unbiased estimator. Hence, we have the following result.

THEOREM 4.4. *The bias-corrected estimator $\hat{\beta}^{*jk}$ is second order asymptotically Gaussian efficient, if and only if*

$$C^{jk,jk,jk} = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathcal{K}_{jjj}(\lambda_1, \lambda_2)dM_{kkk}(\lambda_1, \lambda_2) = 0. \tag{4.4}$$

REMARK 4.5. If the residual $\{u(t)\}$ is a Gaussian process, then (4.4) holds. However, in general, the bias-corrected estimator $\hat{\beta}^*$ is not second order asymptotically Gaussian efficient.

REMARK 4.6. Theorem 4.3 can be employed to check whether the Hannan estimator leads to a second order Gaussian efficient estimator. Since we do not assume the normality of the error process, in general we have $\mathcal{K}_{j_1 j_2 j_3}(\lambda_1, \lambda_2) \neq 0$. Here, we give four examples of the regressor $\{x(t)\}$ in the case where $p = q = 1$.

(i) $x_1(t) = 1$ for $t = 1, 2, \ldots$. Then $\eta_1 = 1$, $M_{11}(\lambda)$ has the jump 1 at $\lambda = 0$ and $M_{111}(\lambda_1, \lambda_2)$ has the jump 1 at $\lambda_1 = \lambda_2 = 0$. Hence, the Hannan estimator is second order Gaussian efficient if and only if $F_{111}(0) = 0$.

(ii) $x_1(t) = \cos \nu t$, $\nu \in (0, 2\pi/3)$ for $t = 1, 2, \ldots$. Then $M_{111}(\lambda_1, \lambda_2)$ has the jump $\mathrm{O_p}(N^{-3/2})$. Hence, the Hannan estimator is always second order Gaussian efficient.

(iii) $x_1(t) = 1 + \cos \nu t$ for $t = 1, 2, \ldots$. Then $\eta_1 = (2/3)^{1/2}$, $M_{11}(\lambda)$ has the jump $2/3$ and $1/6$ at $\lambda = 0$ and $\lambda = \pm\nu$, respectively, and $M_{111}(\lambda_1, \lambda_2)$ has the jump $(2/3)^{3/2}$ and $(2/3)^{3/2}/2$ at $\lambda_1 = \lambda_2 = 0$ and $(\lambda_1, \lambda_2) = (0, \pm\nu)$, $(\pm\nu, 0)$, $(\nu, -\nu)$, $(-\nu, \nu)$, respectively. Hence, the Hannan estimator is not second order Gaussian efficient.

(iv) $x_1(t) = t/N$ for $t = 1, 2, \ldots$. Then $\eta_1 = \sqrt{3}/2$, $M_{11}(\lambda)$ has the jump 1 at $\lambda = 0$ and $M_{111}(\lambda_1, \lambda_2)$ has the jump $3^{3/2}/4$ at $\lambda_1 = \lambda_2 = 0$ Hence, the Hannan estimator is second order Gaussian efficient if and only if $F_{111}(0) = 0$.

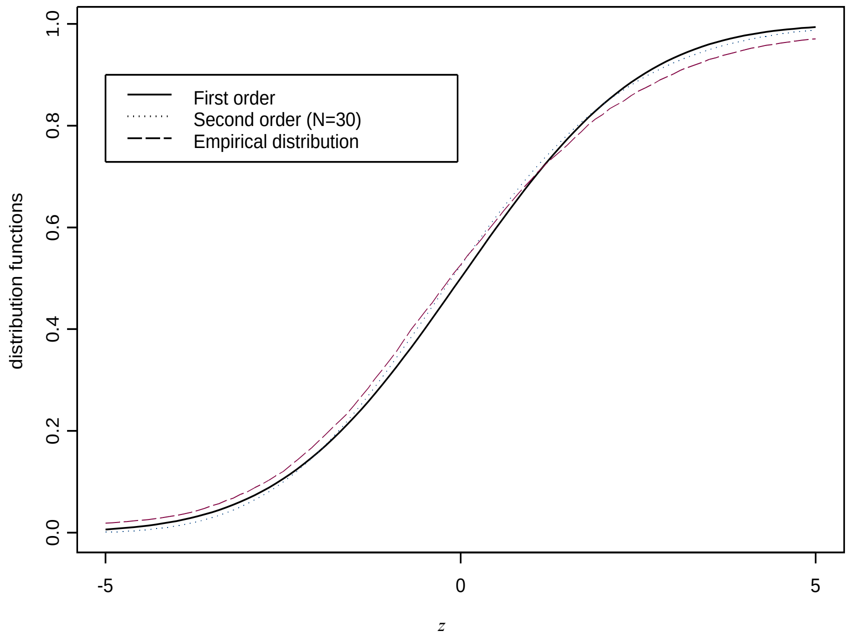

Figure 4.1: $a = 0.5$ and $x(t) = 1$.

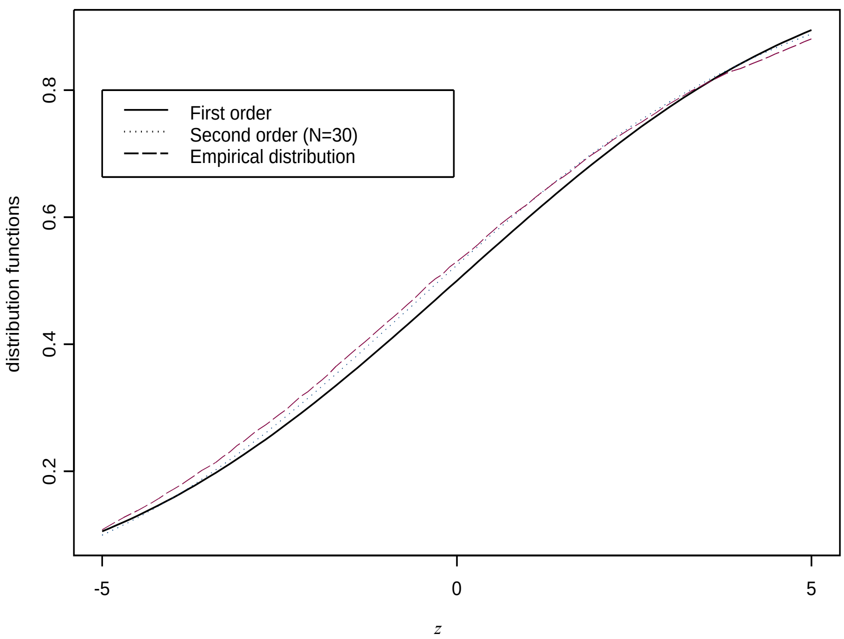

Figure 4.2: $a = 0.75$ and $x(t) = 1$.

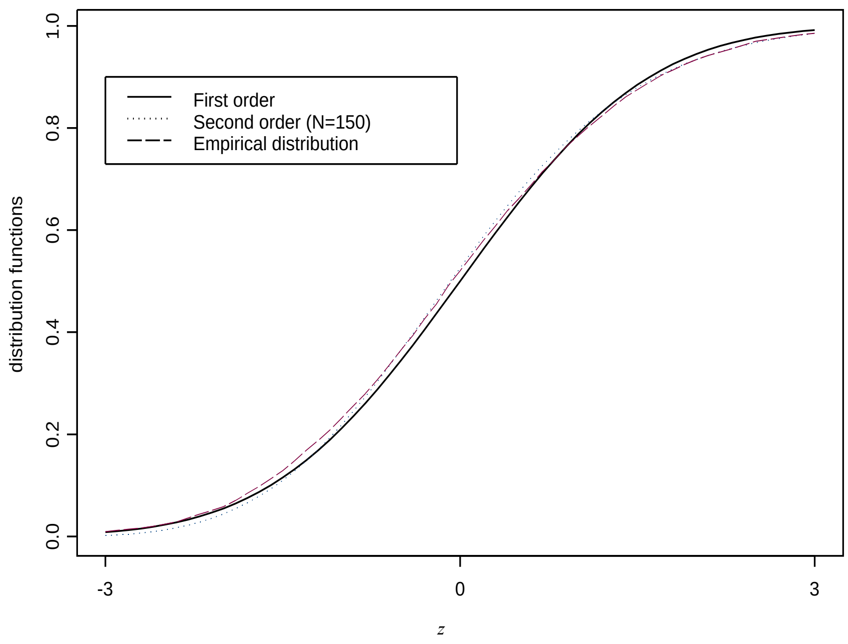

Figure 4.3: $a = 0.25$ and $x(t) = 1 + \cos t$.

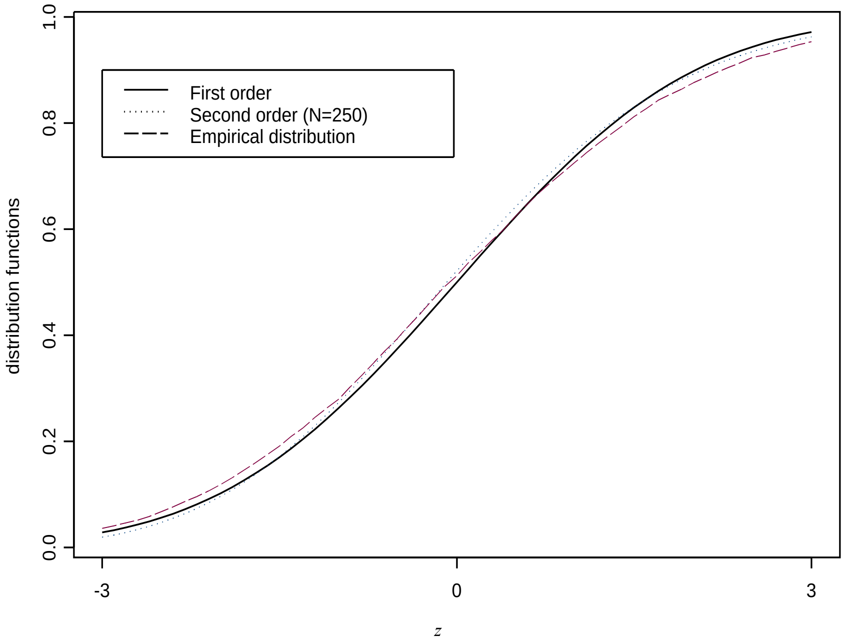

Figure 4.4: $a = 0.5$ and $x(t) = 1 + \cos t$.

5. Second Order Properties of Locally Stationary Processes

5.1. Introduction

There has been much discussion of the efficiency in estimation of stationary time series. Hosoya [22], Akahira and Takeuchi [1] and Taniguchi [43] deal with higher order efficiencies for time series analysis. Taniguchi [43] and Taniguchi and Kakizawa [48] showed that appropriately modified maximum likelihood and quasi maximum likelihood estimators of Gaussian autoregressive moving average processes is second order asymptotically efficient in the sense of degree of concentration of the sampling distribution up to second order. This concept of efficiency was introduced by Akahira and Takeuchi [1], and these results was reviewed in Section 4.4.

Although the analysis for stationary time series is well established, there are many cases where the stationary assumption seems to be restrictive. Because all the results above deal with stationary processes we are led to the problem of efficiently estimating parameters of non-stationary processes. Dahlhaus [13, 14, 15, 16] has introduced a class of locally stationary processes (non-stationary processes), and formulated in a rigorous asymptotic framework.

In this chapter, we investigate the problems of efficiently estimating parameters of multivariate Gaussian locally stationary processes in the sense of Akahira and Takeuchi [1]. In Section 5.2, we discuss second order robustness properties.

5.2. Second order efficiency of the maximum likelihood estimator in locally stationary processes

In this section we shall show that if we appropriately modify the maximum likelihood estimator in Gaussian locally stationary processes, then it is second order asymptotically efficient in the sense of Definition 4.1. First we give the precise definition of multivariate locally stationary processes which is due to Dahlhaus [16].

DEFINITION 5.1. A sequence of Gaussian multivariate stochastic processes $X_{t,T} = (X_{t,T}^{(1)}, \ldots, X_{t,T}^{(d)})'$ $(t = 1, \ldots, T)$ is called locally stationary with transfer function matrix A° and mean function vector μ if there exists a representation

$$X_{t,T} = \mu\left(\frac{t}{T}\right) + \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^\circ(\lambda) d\xi(\lambda)$$

with the following properties:

(i) $\xi(\lambda)$ is a complex valued Gaussian vector process on $[-\pi, \pi]$ with $\overline{\xi_a(\lambda)} = \xi_a(-\lambda)$, $\mathrm{E}\xi_a(\lambda) = 0$ and

$$\mathrm{E}\{d\xi_a(\lambda) d\xi_b(\mu)\} = \delta_{ab}\eta(\lambda + \mu) d\lambda d\mu,$$

where $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ is the period 2π extension of the Dirac delta function.

(ii) There exist 2π-periodic matrix valued functions $A : [0,1] \times \mathbf{R} \to \mathbf{C}^{d\times d}$ with $A(u, -\lambda) = \overline{A(u,\lambda)}$ and

$$\sup_{t,\lambda}\left| A^{\circ}_{t,T}(\lambda)_{ab} - A(t/T, \lambda)_{ab} \right| = O(T^{-1})$$

for all $a, b = 1, \ldots, d$ and $T \in \mathbf{N}$. $A(u,\lambda)$ and $\mu(u)$ are assumed to be continuous in u.

$f(u,\lambda) := A(u,\lambda)\overline{A(u,\lambda)}'$ is called the time varying spectral density of the process.

Throughout this section we assume $A^{\circ}_{t,T}(u,\lambda) = A_{\theta,t,T}(u,\lambda)$ and $\mu(u) = \mu_\theta(u)$, so that efficiency is discussed when the model is correctly specified.

We now set down the following assumptions.

ASSUMPTION 5.1. (i) There exist 2π-periodic matrix valued functions $A_\theta : [0,1]\times \mathbf{R} \to \mathbf{C}^{d\times d}$ with $A_\theta(u,-\lambda) = \overline{A_\theta(u,\lambda)}$ whose components are four times differentiable in θ and

$$\sup_{t,\lambda}\left| \partial^k_{j_1\ldots j_k}\left\{ A^{\circ}_{\theta,t,T}(\lambda)_{ab} - A_\theta(t/T,\lambda)_{ab}\right\} \right| = O(T^{-1}) \quad \text{for } k = 0,1,2,3,$$

where $\partial^k_{j_1\ldots j_k} = \partial^k/\partial\theta^{j_1}\ldots\partial\theta^{j_k}$. The components of $\partial^k_{j_1\ldots j_k} A_\theta(u,\lambda)$ are differentiable in u and λ with uniformly bounded derivatives.

(ii) All eigenvalues of $f_\theta(u,\lambda) = A_\theta(u,\lambda)\overline{A_\theta(u,\lambda)}'$ are bounded from below by some $C > 0$ uniformly in u and λ.

(iii) The components of $\mu_\theta(u)$ are four times differentiable in θ. The components of $\partial^k_{j_1\ldots j_k}\mu_\theta(u)$ are differentiable in u with uniformly bounded derivatives.

Second we give the bound distributions of $\beta^+_{\theta_0}(x^1) + \gamma^+_{\theta_0}(x^1)/\sqrt{T}$ and $\beta^+_{\theta_-}(x^1) + \gamma^+_{\theta_-}(x^1)/\sqrt{T}$ defined in Section 4.4. Using the fundamental lemma of Neyman and Pearson these are given by the likelihood ratio test. Thus we consider the problem of testing hypothesis $H : \theta = \theta_0 + x/\sqrt{T}$ against the alternative $A : \theta = \theta_0$, where $x = (x^1, \ldots, x^p)$ and $x_2 = (x^2, \ldots, x^p)$ is an arbitrary but fixed constant. Let $\underline{X} = (X'_{1,T}, \ldots, X'_{T,T})'$, $\underline{\mu} = (\mu(1/T)', \ldots, \mu(T/T)')'$ and $\Sigma_T(A,B)$ be $T \times T$ block matrix whose (r,s) block is

$$\left[\Sigma_T(A,B)\right]_{r,s} = \int_{-\pi}^{\pi} \exp\{i\lambda(r-s)\} A_{r,T}(\lambda) B_{s,T}(-\lambda)' d\lambda$$

$r, s = 1, \ldots, T$. The log likelihood function based on $\underline{X}$ is given by

$$L_T(\theta) = -\frac{d}{2}\log(2\pi) - \frac{1}{2T}\log\det\Sigma_\theta - \frac{1}{2T}(\underline{X} - \underline{\mu}_\theta)'\Sigma_\theta^{-1}(\underline{X} - \underline{\mu}_\theta), \tag{5.1}$$

where $\Sigma_\theta = \Sigma_T(A_\theta^\circ, A_\theta^\circ)$. Let $LR = L_T(\theta_0) - L_T(\theta_0 + x/\sqrt{T})$. Using Lemma A.8 in Dahlhaus [16], we can show that

$$E_{\theta_0}[LR] = \frac{1}{2} I_{ij} x^i x^j + \frac{1}{6\sqrt{T}} (3J_{ij,k} + K_{ijk}) x^i x^j x^k + o(T^{-1}),$$
$$\mathrm{cum}_{\theta_0}[LR, LR] = I_{ij} x^i x^j + \frac{1}{\sqrt{T}} J_{ij,k} x^i x^j x^k + o(T^{-1}),$$
$$\mathrm{cum}_{\theta_0}[LR, LR, LR] = -\frac{1}{\sqrt{T}} K_{ijk} x^i x^j x^k + o(T^{-1}),$$
$$E_{\theta_0 + x/\sqrt{T}}[LR] = -\frac{1}{2} I_{ij} x^i x^j - \frac{1}{6\sqrt{T}} (3J_{ij,k} + 2K_{ijk}) x^i x^j x^k + o(T^{-1}),$$
$$\mathrm{cum}_{\theta_0 + x/\sqrt{T}}[LR, LR] = I_{ij} x^i x^j + \frac{1}{\sqrt{T}} (J_{ij,k} + K_{ijk}) x^i x^j x^k + o(T^{-1}),$$
$$\mathrm{cum}_{\theta_0 + x/\sqrt{T}}[LR, LR, LR] = -\frac{1}{\sqrt{T}} K_{ijk} x^i x^j x^k + o(T^{-1}),$$

where

$$\begin{aligned}
I_{ij}(\theta) = &-\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[(\partial_i^1 f_\theta)(\partial_j^1 f_\theta^{-1})] d\lambda du \\
&+ \frac{1}{2\pi} \int_0^1 \{\partial_i^1 \mu_\theta(u)\}' f_\theta(u, 0)^{-1} \{\partial_j^1 \mu_\theta(u)\} du, \\
J_{ij,k}(\theta) = &-\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[(\partial_{ij}^2 f_\theta^{-1})(\partial_k^1 f_\theta)] d\lambda du \\
&+ \frac{1}{2\pi} \int_0^1 \{\partial_{ij}^2 \mu_\theta(u)\}' f_\theta(u, 0)^{-1} \{\partial_k^1 \mu_\theta(u)\} du \\
&+ \frac{1}{2\pi} \int_0^1 \{\partial_i^1 \mu_\theta(u)\}' \{\partial_j^1 f_\theta(u, 0)^{-1}\} \{\partial_k^1 \mu_\theta(u)\} du[2, ij], \\
K_{ijk}(\theta) = &\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[f_{1,\theta}^{-1} (\partial_i^1 f_\theta) f_\theta^{-1} (\partial_j^1 f_\theta) f_{1,\theta}^{-1} (\partial_k^1 f_\theta)] d\lambda du[2] \\
&- \frac{1}{2\pi} \int_0^1 \{\partial_i^1 \mu_\theta(u)\}' \{\partial_j^1 f_\theta(u, 0)^{-1}\} \{\partial_k^1 \mu_\theta(u)\} du[3].
\end{aligned}$$

Here we use the Einstein summation convention and the simpler notations I_{ij}, $J_{ij,k}$, K_{ijk} etc. are evaluated at $\theta = \theta_0$. By (4.3) and the fundamental lemma of Neyman and Perason, the asymptotic power of the most powerful test LR is given by

$$\Phi(\sigma) + \frac{1}{6\sqrt{T}\sigma} \phi(\sigma)(3J_{ij,k} + K_{ijk}) x^i x^j x^k + o(T^{-1/2}),$$

where $\Phi(z) = \int_{-\infty}^{z} \phi(u) du$, $\phi(u) = (2\pi)^{-1/2} \exp(-u^2/2)$, $\sigma = (I_{ij} x^i x^j)^{1/2}$.

Denote by I^{ij} the (i,j)-th element of the inverse matrix of $I = \{I_{ij}\}$. The partition $x = (x^1, x_2)$ induces the following corresponding partition

$$I(\theta) = \begin{bmatrix} I_{(11)}(\theta) & I_{(12)}(\theta) \\ I_{(21)}(\theta) & I_{(22)}(\theta) \end{bmatrix}.$$

Since x_2 can take arbitrary values, then the power function of the tests is not larger than the infimum of (5.1) with respect to x_2. A x_2 minimizing σ is given by $x_2 = (I_{(22)})^{-1} I_{(21)} x^1$, then $\sigma^2 = (I^{11})^{-1}(x^1)^2$. Thus we have the following:

THEOREM 5.1. If $\hat{\theta}^1$ is second order AMU and

$$\begin{aligned} &P^T_{\theta_0}[\sqrt{T}(\hat{\theta}^1 - \theta^1_0) \leq x^1] \\ &\quad = \Phi(x^1(I^{11})^{-1/2}) \\ &\qquad + \frac{(x^1)^2}{6(I^{11})^{5/2}\sqrt{T}} \phi(x^1(I^{11})^{-1/2}) I^{1i} I^{1j} I^{1k} (3J_{ij,k} + K_{ijk}) \\ &\qquad + o(T^{-1/2}) \end{aligned}$$

is satisfied, then $\hat{\theta}^1$ is second order asymptotically efficient estimator.

Let $\hat{\theta}_{ML} = (\hat{\theta}^1_{ML}, \dots, \hat{\theta}^p_{ML})$ be maximum likelihood estimator which is defined by a value of θ that satisfies the equation

$$0 = \partial^1_i L_T(\theta).$$

Write

$$\begin{aligned} &U^i = \sqrt{T}(\hat{\theta}^i_{ML} - \theta^i_0), \quad Z_i(\theta) = \sqrt{T}\big[\partial^1_i L_T(\theta) - E\{\partial^1_i L_T(\theta)\}\big], \\ &Z_{ij}(\theta) = \sqrt{T}\big[\partial^2_{ij} L_T(\theta) - E\{\partial^2_{ij} L_T(\theta)\}\big]. \end{aligned}$$

Then we can show the following.

LEMMA 5.1.

$$\begin{aligned} U^i = I^{ij} Z_j &+ \frac{1}{\sqrt{T}} I^{ij} I^{kl} Z_{jk} Z_l - \frac{1}{2\sqrt{T}} I^{ii'} I^{jj'} I^{kk'} (J_{i'j,k}[3] + K_{i'jk}) Z_{j'} Z_{k'} \\ &+ o_p(T^{-1/2}). \end{aligned}$$

It is seen that

$$\begin{aligned} &\mathrm{E}_{\theta_0}[U^i] = -\frac{1}{2\sqrt{T}} I^{ij} I^{kl} (J_{kl,j} + K_{jkl}) + o(T^{-1/2}), \\ &\mathrm{cum}_{\theta_0}[U^i, U^j] = I^{ij} + o(T^{-1/2}), \\ &\mathrm{cum}_{\theta_0}[U^i, U^j, U^k] = -T^{-1/2} I^{ii'} I^{jj'} I^{kk'} (J_{i'j',k'}[3] + 2K_{i'j'k'}) + o(T^{-1/2}), \\ &\mathrm{cum}^J_{\theta_0}[U^{i_1}, \dots, U^{i_J}] = O(T^{-J/2+1}) \qquad \text{for } J \geq 3. \end{aligned}$$

Applying a general Edgeworth expansion formula (e.g., Taniguchi and Kakizawa, [48], p.168-170), we have the following theorem

THEOREM 5.2.

$$\begin{aligned} P^T_{\theta_0}&[\sqrt{T}(\hat{\theta}_{ML}-\theta_0)\le x] \\ &= \int_{-\infty}^{x} \phi(z, I^{-1})\bigg[1-\frac{1}{2\sqrt{T}}I^{ij}I^{kl}(J_{kl,j}+K_{jkl})H_i(z, I^{-1}) \\ &\quad -\frac{1}{6\sqrt{T}}I^{ii'}I^{jj'}I^{kk'}(J_{i'j',k'}[3]+2K_{i'j'k'})H_{ijk}(z, I^{-1})\bigg]dz \\ &\quad + o(T^{-1/2}), \end{aligned}$$

where $z=(z^1,\dots,z^p)'$,

$$\phi(z,\Omega)=(2\pi)^{-p/2}|\Omega|^{-1/2}\exp\left(-\frac{1}{2}z'\Omega^{-1}z\right),$$

the multivariate normal distribution, and multivariate Hermite polynomials:

$$H_{j_1\dots j_s}(z,\Omega)=\frac{(-1)^s}{\phi(z,\Omega)}\frac{\partial^s}{\partial x_{j_1}\dots x_{j_s}}\phi(z,\Omega).$$

From Theorem 5.2, it can be seen that $\hat{\theta}^1_{ML}$ is not second order AMU. Thus we modify $\hat{\theta}^1_{ML}$ as follows:

$$\begin{aligned} \hat{\theta}^{1*}_{ML} &= \hat{\theta}^1_{ML} + \frac{1}{2T}I^{1i}(\hat{\theta}_{ML})I^{jk}(\hat{\theta}_{ML})\{J_{jk,i}(\hat{\theta}_{ML})+K_{ijk}(\hat{\theta}_{ML})\} \\ &\quad -\frac{1}{6I^{11}T}I^{1i}(\hat{\theta}_{ML})I^{1j}(\hat{\theta}_{ML})I^{1k}(\hat{\theta}_{ML})\{3J_{ij,k}(\hat{\theta}_{ML})+2K_{ijk}(\hat{\theta}_{ML})\}. \end{aligned} \tag{5.2}$$

Then we obtain

$$\begin{aligned} P^T_{\theta_0}&[\sqrt{T}(\hat{\theta}^{1*}_{ML}-\theta^1_0)\le x^1] \\ &= \Phi(x^1(I^{11})^{-1/2}) \\ &\quad + \frac{(x^1)^2}{6(I^{11})^{5/2}\sqrt{T}}\phi(x^1(I^{11})^{-1/2})I^{1i}I^{1j}I^{1k}(3J_{ij,k}+K_{ijk}) \\ &\quad + o(T^{-1/2}). \end{aligned}$$

Remembering Theorem 5.1, we can see that (5.2) coincides with the bound distribution. Thus we have

THEOREM 5.3. The modified MLE $\hat{\theta}^{1*}_{ML}$ is second order asymptotically efficient.

5.3. Higher order robustness

In this section, we discuss second order misspecified and time varying robustness of the maximum likelihood estimator. To discuss the problem of higher order asymptotic estimation for parameters of locally stationary processes, the following assumptions are imposed

ASSUMPTION 5.2. (i)

$$A^\circ_{t,T}(\lambda) = A^\circ_{1,\theta,t,T}(\lambda) + \frac{1}{\sqrt{T}} A^\circ_{2,\theta,t,T}(\lambda) + \frac{1}{T} A^\circ_{3,\theta,t,T}(\lambda),$$
$$\mu(u) = \mu_\theta(u).$$

(ii) There exist 2π-periodic matrix valued functions $A_{i,\theta} : [0,1] \times \mathbf{R} \to \mathbf{C}^{d\times d}$ with $A_{i,\theta}(u,-\lambda) = \overline{A_{i,\theta}(u,\lambda)}$ whose components are four times differentiable in θ and

$$\sup_{t,\lambda} \left| \partial^k_{j_1\dots j_k} \{ A^\circ_{i,\theta,t,T}(\lambda)_{ab} - A_{i,\theta}(t/T,\lambda)_{ab} \} \right| = o(T^{-1})$$

for $k = 0,1,2,3$ and $i = 1,2,3$. The components of $\partial^k_{j_1\dots j_k} A_{i,\theta}(u,\lambda)$ ($i = 1,2,3$) are differentiable in u and λ with uniformly bounded derivatives.

(iii) Let

$$f_\theta(u,\lambda) = f_{1,\theta}(u,\lambda) + \frac{1}{\sqrt{T}} f_{2,\theta}(u,\lambda) + \frac{1}{T} f_{3,\theta}(u,\lambda) + o(T^{-1}).$$

Then, $f_{i,\theta}(u,\lambda)$ ($i = 1,2,3$) fulfill Assumption 5.1 (ii).

(iv) $\mu_\theta(u)$ fulfills Assumption 5.1. (iii).

We define the MLE $\tilde{\theta}_{ML}$ in the misspecified case by a solution of equation

$$0 = \partial^1_i \tilde{L}_T(\theta), \quad i = 1, \dots p,$$

where

$$\tilde{L}_T(\theta) = -\frac{d}{2}\log(2\pi) - \frac{1}{2T}\log\det\Sigma_{1,\theta} - \frac{1}{2T}(\underline{X} - \underline{\mu}_\theta)'\Sigma^{-1}_{1,\theta}(\underline{X} - \underline{\mu}_\theta),$$

and $\Sigma_{1,\theta} = \Sigma_T(A^\circ_{1,\theta}, A^\circ_{1,\theta})$.

Write

$$\tilde{U}^i = \sqrt{T}(\tilde{\theta}^i_{ML} - \theta^i), \quad \tilde{Z}_i(\theta) = -\sqrt{T}\big[\partial_i \tilde{L}_T(\theta) - E_\theta\{\partial_i \tilde{L}_T(\theta)\}\big],$$
$$\tilde{Z}_{ij}(\theta) = -\sqrt{T}\big[\partial^2_{ij} \tilde{L}_T(\theta) - E_\theta\{\partial^2_{ij} \tilde{L}_T(\theta)\}\big],$$

where E_θ denotes the expectation under the true model.

In the same way as the previous calculations, it follows that

$$\begin{aligned}\tilde{U}^i = {} & I^{ij}(\tilde{Z}_j - \Gamma^{(1)}_j) - \frac{1}{\sqrt{T}} I^{ij}\Gamma^{(2)}_j - \frac{1}{\sqrt{T}} I^{ij} I^{kl}\Delta_{jk}(\tilde{Z}_l - \Gamma^{(1)}_l) \\ & + \frac{1}{\sqrt{T}} I^{ij} I^{kl} \tilde{Z}_{jk}(\tilde{Z}_l - \Gamma^{(1)}_l) \\ & - \frac{1}{2\sqrt{T}} I^{ii'} I^{jj'} I^{kk'} (J_{i'j,k}[3] + K_{i'jk})(\tilde{Z}_{j'} - \Gamma^{(1)}_{j'})(\tilde{Z}_{k'} - \Gamma^{(1)}_{k'}) \\ & + o_p(T^{-1/2}).\end{aligned} \tag{5.3}$$

where

$$\Gamma_i^{(1)}(\theta) = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} \mathrm{tr}[(\partial_i f_{1,\theta}^{-1}) f_{2,\theta}] d\lambda du,$$
$$\Gamma_i^{(2)}(\theta) = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} \mathrm{tr}[(\partial_i f_{1,\theta}^{-1}) f_{3,\theta}] d\lambda du,$$
$$\Delta_{ij}(\theta) = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} \mathrm{tr}[(\partial_{ij}^2 f_{1,\theta}^{-1}) f_{2,\theta}] d\lambda du.$$

From direct verification, we can show that

$$\begin{aligned}
&E_{\theta_0}[\tilde{Z}_i\tilde{Z}_j] \\
&\quad = I_{ij} + \frac{1}{4\pi\sqrt{T}}\int_0^1\int_{-\pi}^{\pi} \mathrm{tr}[f_{2,\theta_0}(\partial_i f_{1,\theta_0}^{-1}) f_{1,\theta_0}(\partial_j f_{1,\theta_0}^{-1})] d\lambda du[2] \\
&\qquad + \frac{1}{2\pi\sqrt{T}}\int_0^1 \{\partial_i \mu_{\theta_0}(u)\}' f_{1,\theta_0}(u,0)^{-1} f_{2,\theta_0}(u,0) f_{1,\theta_0}(u,0)^{-1}\{\partial_j \mu_{\theta_0}(u)\} du \\
&\qquad + o(T^{-1/2}) \\
&\quad = I_{ij} + \frac{1}{\sqrt{T}}\Delta_{1,ij} + o(T^{-1/2}) \qquad \text{(say)}, \\
&E_{\theta_0}[\tilde{Z}_{ij}\tilde{Z}_k] = J_{ij,k} + O(T^{-1/2}), \\
&E_{\theta_0}[\tilde{Z}_i\tilde{Z}_j\tilde{Z}_k] = T^{-1/2}K_{ijk} + O(T^{-1}),
\end{aligned} \tag{5.4}$$

and the Jth ($J \geq 3$) order cumulant of $\tilde{Z}_{i_1}(\theta), \ldots, \tilde{Z}_{i_{J_1}}(\theta), \tilde{Z}_{j_1k_1}(\theta), \ldots, \tilde{Z}_{j_{J_2}k_{J_2}}(\theta)$ ($J_1 + J_2 = J$) satisfies

$$\mathrm{cum}^{(J)}[\tilde{Z}_{i_1}(\theta), \ldots, \tilde{Z}_{i_{J_1}}(\theta), \tilde{Z}_{j_1k_1}(\theta), \ldots, \tilde{Z}_{j_{J_2}k_{J_2}}(\theta)] = O(T^{-J/2+1}). \tag{5.5}$$

From (5.3)-(5.5), it is seen that

$$\begin{aligned}
E_{\theta_0}[\tilde{U}^i] &= -I^{ij}\Gamma_j^{(1)} - \frac{1}{\sqrt{T}}I^{ij}\Gamma_j^{(2)} + \frac{1}{\sqrt{T}}I^{ij}I^{kl}\Delta_{jk}\Gamma_l^{(1)} \\
&\quad - \frac{1}{2\sqrt{T}}I^{ii'}I^{jj'}I^{kk'}(J_{i'j',k'}[3] + K_{i'j'k'})\Gamma_j^{(1)}\Gamma_k^{(1)} \\
&\quad - \frac{1}{2\sqrt{T}}I^{ij}I^{kl}(J_{kl,j} + K_{jkl}) + o(T^{-1/2}), \\
\mathrm{cum}_{\theta_0}[\tilde{U}^i, \tilde{U}^j] &= I^{ij} + \frac{1}{\sqrt{T}}I^{ik}I^{jl}(\Delta_{1,kl} - 2\Delta_{kl}) \\
&\quad + \frac{1}{\sqrt{T}}I^{ii'}I^{jj'}I^{kk'}\Gamma_k^{(1)}(J_{i'j',k'}[3] + J_{i'j',k'} + 2K_{i'j'k'}) + o(T^{-1/2}), \\
\mathrm{cum}_{\theta_0}[\tilde{U}^i, \tilde{U}^j, \tilde{U}^k] &= -T^{-1/2}I^{ii'}I^{jj'}I^{kk'}(J_{i'j',k'}[3] + 2K_{i'j'k'}) + O(T^{-1/2}),
\end{aligned}$$

$$\mathrm{cum}^J_{\theta_0}[\tilde{U}^{i_1}, \ldots, \tilde{U}^{i_J}] = O(T^{-J/2+1}) \qquad \text{for } J \geq 4.$$

Applying a general formula (e.g., Taniguchi and Kakizawa, [48], p.168-170), we have

THEOREM 5.4. If $\Gamma_i^{(1)} = 0$ ($i = 1, \ldots, p$), then the Edgeworth expansion of the distribution function of $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ is given by

$$\begin{aligned}
&P^T_{\theta_0}[\sqrt{T}(\tilde{\theta}_T - \theta_0) \leq z] \\
&= \int_{-\infty}^{z} \phi(\boldsymbol{x}, I^{-1})\Big[1 - \frac{1}{2\sqrt{T}} I^{ij}\{\Gamma_j^{(2)} + I^{kl}(J_{kl,j} + K_{jkl})\} H_i(x, I^{-1}) \\
&\quad + \frac{1}{2\sqrt{T}} I^{ik} I^{jl} (\Delta_{1,kl} - 2\Delta_{kl}) H_{ij}(x, I^{-1}) \\
&\quad - \frac{1}{6\sqrt{T}} I^{ii'} I^{jj'} I^{kk'} (J_{i'j',k'}[3] + 2K_{i'j'k'}) H_{ijk}(x, I^{-1})\Big] dx + o(T^{-1/2}).
\end{aligned}$$

REMARK 5.1. The condition $\Gamma_i^{(1)} = 0$ ensures that the distribution of $\sqrt{T}(\tilde{\theta}_{ML} - \theta)$ converges to the multivariate normal distribution with zero mean vector. If $\Gamma_i^{(2)} = 0$ is satisfied, then the bias of $\tilde{\theta}_{ML}$ is equal to that of $\hat{\theta}_{ML}$ up to second order.

From

$$\begin{aligned}
I_{ij}(\theta) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[(\partial_i f_\theta) f_\theta^{-1} (\partial_j f_\theta) f_\theta^{-1}] d\lambda du \\
&\quad + \frac{1}{2\pi} \int_0^1 \{\partial_i \mu_\theta(u)\}' f_\theta(u, 0)^{-1} \{\partial_j \mu_\theta(u)\} du \\
&= I_{ij}(\theta) + \frac{1}{\sqrt{T}} \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[-(\partial_i f_{1,\theta}) f_{1,\theta}^{-1} (\partial_j f_{1,\theta}) f_{1,\theta}^{-1} f_{2,\theta} f_{1,\theta}^{-1}[2]] d\lambda du \\
&\quad + \frac{1}{\sqrt{T}} \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[(\partial_i f_{2,\theta}) f_{1,\theta}^{-1} (\partial_j f_{1,\theta}) f_{1,\theta}^{-1}[2]] d\lambda du \\
&\quad - \frac{1}{\sqrt{T}} \frac{1}{2\pi} \int_0^1 \{\partial_i \mu_\theta(u)\}' f_{1,\theta}(u, 0)^{-1} f_{2,\theta}(u, 0) f_{1,\theta}(u, 0)^{-1} \{\partial_j \mu_\theta(u)\} du \\
&\quad + o(T^{-1/2}) \\
&= I_{1,ij}(\theta) + \frac{1}{\sqrt{T}} \Delta_{2,ij}(\theta) + o(T^{-1/2}) \quad \text{(say)},
\end{aligned}$$

we have

$$I^{ij} = I_1^{ij} - \frac{1}{\sqrt{T}} I^{ik} I^{jl} \Delta_{2,kl} + o(T^{-1/2}).$$

It is easy to see that $\Gamma_i^{(1)} = 0$ implies $2\Delta_{ij} - \Delta_{1,ij} - \Delta_{2,ij} = 0$. Therefore, if we put

$$\tilde{\theta}_{ML}^{*i} = \tilde{\theta}_{ML}^{i} + \frac{1}{T} I^{ij}(\tilde{\theta}_{ML}) \Gamma_j^{(2)}(\tilde{\theta}_{ML}),$$

then we obtain

$$P[\sqrt{T}(\tilde{\theta}^{*}_{ML} - \theta_0) \leq z] = P[\sqrt{T}(\hat{\theta}_{ML} - \theta_0) \leq z] + o(T^{-1/2}).$$

Thus we have

COROLLARY 5.1. If $\Gamma_j^{(1)} = 0$ is satisfied, then the distribution function of the modified maximum likelihood estimator $\tilde{\theta}^{i}_{ML}$ is equal to that of the $\hat{\theta}_{ML}$ with an error $o(T^{-1/2})$.

If

$$P[\sqrt{T}(\tilde{\theta}_{ML} - \theta_0) \leq z] = \Pr[\sqrt{T}(\hat{\theta}_{ML} - \theta_0) \leq z] + o(T^{-1/2})$$

is satisfied, then we say that the estimator $\tilde{\theta}_{ML}$ of θ is asymptotically misspecified robustness with an error $o(T^{-1/2})$.

COROLLARY 5.2. If $\Gamma_j^{(1)} = \Gamma_j^{(2)} = 0$ is satisfied, then $\tilde{\theta}_T$ is asymptotically misspecified robustness with an error $o(T^{-1/2})$.

If

$$\int_0^1 \int_{-\pi}^{\pi} \mathrm{tr}[(\partial_i^1 f_\theta^{-1}) f_\theta] d\lambda du = 0$$

is satisfied, then we say that the parameter θ is innovation-free w.r.t. f_θ.

REMARK 5.2. From (5.3), if the parameter θ is innovation-free w.r.t. $f_{1,\theta}$, $f_{2,\theta} = af_{1,\theta}$ and $f_{3,\theta} = bf_{1,\theta}$ $a, b \in R$, then $\Gamma_j^{(1)} = \Gamma_j^{(2)} = 0$ holds.

We consider the situation where all of the quantities appearing in second order Edgeworth expansion for an estimator have the form

$$\int_0^1 \int_{-\pi}^{\pi} g_1(\lambda, u) d\lambda du + \int_0^1 g_2(u) du.$$

If $g_1(\lambda, u)$ and $g_2(u)$ are independent of u, then we say that the estimator is time varying robustness up to second order.

COROLLARY 5.3. If

$$\begin{aligned} A_\theta(u, \lambda) &= B(u) C_\theta(\lambda), \\ \mu_\theta(u) &= B(u) \nu_\theta, \end{aligned} \tag{5.6}$$

are satisfied, then $\hat{\theta}_{ML}$, $\tilde{\theta}_{ML}$ are time varying robustness.

REMARK 5.3. If the condition (5.6) holds, then locally stationary processes $X_{t,T}$ can be written as

$$
\begin{aligned}
X_{t,T} &= B\left(\frac{t}{T}\right)\left\{\nu_\theta + \int_{-\pi}^{\pi} \exp(i\lambda t) C_\theta^\circ(\lambda) d\xi(\lambda)\right\}, \\
&= B\left(\frac{t}{T}\right) \times \{\text{ stationary process }\}.
\end{aligned}
$$

EXAMPLE 5.1. To observe the non-stationary effect, we consider the following model:

$$
X_{t,T} + b_{\theta^2}\left(\frac{t}{T}\right) X_{t-1,T} = a_{\theta^1}\left(\frac{t}{T}\right)\varepsilon_t, \quad t = 1, \dots, T,
$$

where $a_\theta(u) = a\exp\{-(u-\theta)^2/2\}$, $b_\theta(u) = u\theta$, $|a| < 1$, $\theta^1 < 0$, $1 < \theta^1$, $|\theta^2| < 1$ and ε_t's are i.i.d. $(0, 1)$ random variables. Then the time varying spectral density is given by

$$
f_\theta(u, \lambda) = \frac{1}{2\pi}\left|\frac{a_{\theta^1}(u)}{1 + b_{\theta^2}(u)e^{-i\lambda}}\right|^2, \quad \theta = (\theta^1, \theta^2).
$$

By the residue theorem, it is shown that

$$
\begin{aligned}
I_{11} &= 2\int_0^1 \left\{\frac{\partial_1 a_{\theta^1}(u)}{a_{\theta^1}(u)}\right\}^2 du, & I_{12} &= 0, \\
I_{21} &= 0, & I_{22} &= \int_0^1 \frac{\{\partial_2 b_{\theta^2}(u)\}^2}{1 - \{b_{\theta^2}(u)\}^2} du,
\end{aligned}
$$

$$
\begin{aligned}
J_{11,1} &= 2\int_0^1 \frac{\partial_1 a_{\theta^1}(u)}{a_{\theta^1}(u)} \frac{\partial_1^2 a_{\theta^1}(u)}{a_{\theta^1}(u)} du - \frac{3}{4}K_{111}, \quad J_{11,2} = J_{12,1} = 0, \\
J_{12,2} &= J_{22,1} = -\frac{1}{2}K_{122}, \quad J_{22,2} = -\frac{1}{3}K_{222} + \int_0^1 \frac{\partial_2 b_{\theta^2}(u)\partial_2^2 b_{\theta^2}(u)}{1 - \{b_{\theta^2}(u)\}^2} du,
\end{aligned}
$$

and

$$
\begin{aligned}
K_{111} &= 8\int_0^1 \left\{\frac{\partial_1 a_{\theta^1}(u)}{a_{\theta^1}(u)}\right\}^3 du, & K_{112} &= 0, \\
K_{122} &= 4\int_0^1 \frac{\partial_1 a_{\theta^1}(u)}{a_{\theta^1}(u)} \frac{\{\partial_2 b_{\theta^2}(u)\}^2}{1 - \{b_{\theta^2}(u)\}^2} du, & K_{222} &= 6\int_0^1 \frac{\{\partial_2 b_{\theta^2}(u)\}^3 b_{\theta^2}(u)}{[1 - \{b_{\theta^2}(u)\}^2]^2} du,
\end{aligned}
$$

Let $\Delta^S(u)$ be Δ^{LS} in stationary case (i.e., u is treated as a known parameter). We introduce the criterion

$$
D(\theta) = \int_0^1 \left\{\Delta^{LS} - \Delta^S(u)\right\}^2 du,
$$

which measures the time varying effect in efficient estimation.

(i) Suppose that θ^1 is unknown, and that θ^2 is known. Then it is easy to show

$$\Delta^{LS} = \frac{3}{4}\frac{(1-\theta^1)^4 - (\theta^1)^4}{(1-\theta^1)^3 + (\theta^1)^3}, \qquad \Delta^S(u) = \frac{1}{3(u-\theta^1)}.$$

In Figure 5.1, we plotted $D(\theta^1)$ with $-2 < \theta^1 < 0$ and $1 < \theta^1 < 3$. From the figure we observe that the time varying effect becomes large as $\theta^1 \nearrow 0$ or $\theta^1 \searrow 1$.

Figures 5.1 is about here.

(ii) Suppose that θ^2 is unknown, and that θ^1 is known. Then it is easy to show

$$\begin{aligned}\Delta^{LS} &= \frac{1}{6}\left\{-\frac{1}{(\theta^2)^2} + \frac{1}{2(\theta^2)^3}\log\frac{1+\theta}{1-\theta}\right\}^{-2} \\ &\quad\times\left[\frac{3\{3-2(\theta^2)^2\}}{(\theta^2)^3\{1-(\theta^2)^2\}} - \frac{9}{2(\theta^2)^4}\log\frac{1+\theta}{1-\theta}\right], \\ \Delta^S(u) &= \theta^2.\end{aligned}$$

In Figure 5.2, we plotted $D(\theta^2)$ with $-1 < \theta^2 < 1$. From the figure we observe that the time varying effect becomes large as $|\theta^2| \nearrow 1$.

Figures 5.2 is about here.

(iii) Suppose that θ^1 is a parameter of interest, and that θ^2 is a nuisance parameter. Then it is easy to show

$$\begin{aligned}\Delta^{LS} &= \frac{3}{4}\frac{(1-\theta^1)^4 - (\theta^1)^4}{(1-\theta^1)^3 + (\theta^1)^3} + \frac{3}{4}\frac{1}{(1-\theta^1)^3 + (\theta^1)^3} \\ &\quad\times\left\{-\frac{1}{(\theta^2)^2} + \frac{1}{2(\theta^2)^3}\log\frac{1+\theta^2}{1-\theta^2}\right\} \\ &\quad\times\left[-\frac{1}{(\theta^2)^2} - \frac{1}{2(\theta^2)^4}\log\{1-(\theta^2)^2\} + 2\frac{\theta^1}{(\theta^2)^2} + \frac{\theta^1}{2(\theta^2)^3}\log\frac{1+\theta^2}{1-\theta^2}\right], \\ \Delta^S(u) &= \frac{5}{6(u-\theta^1)}.\end{aligned}$$

(iv) Suppose that θ^2 is a parameter of interest, and that θ^1 is a nuisance parameter. From $J_{11,2} = K_{112} = 0$, it is seen that the modification term is not affected by the nuisance parameter. Hence, Δ^{LS} and $\Delta^S(u)$ are the same as the case (ii).

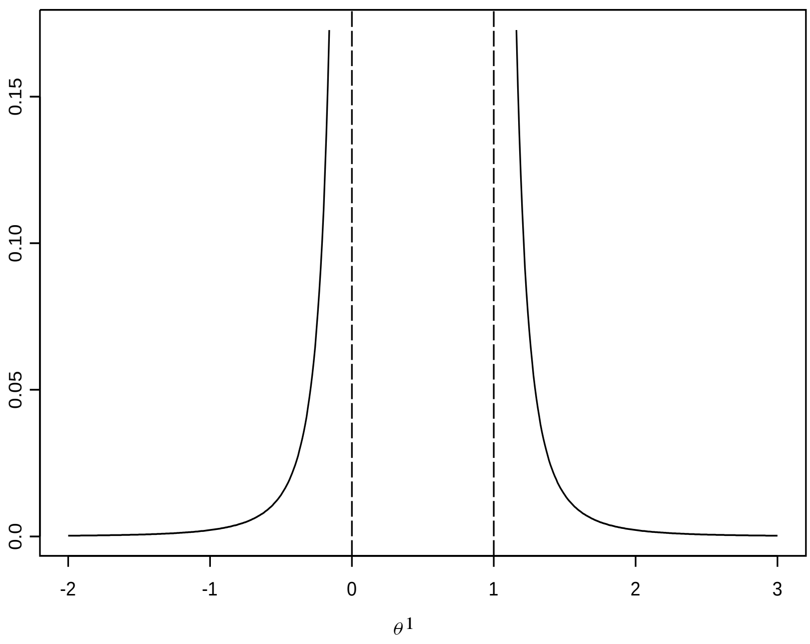

Figure 5.1: In Example 5.1 (i), $D(\theta^1)$ is plotted with $-2 < \theta^1 < 0$ and $1 < \theta^1 < 3$.

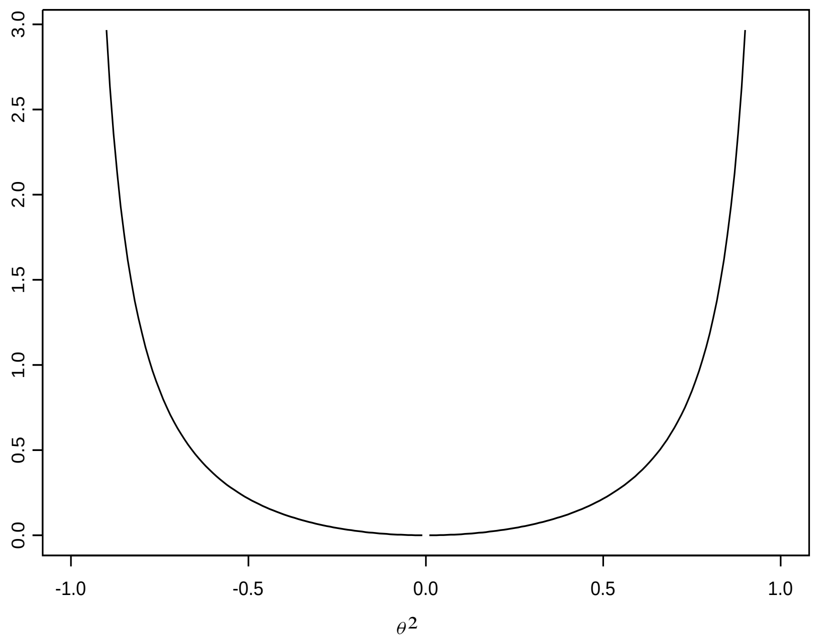

Figure 5.2: In Example 5.1 (ii), $D(\theta^2)$ is plotted with $-1 < \theta^2 < 1$.

6. Proofs

6.1. Proofs of Chapter 2

In this section, we give the proofs of theorems in Chapter 2.

PROOF OF THEOREM 2.1. Since the actual calculation procedure is formidable, we give a sketch of the derivation. First, we evaluate the characteristic function of T,

$$\psi_N(\xi,\varepsilon) = E_{\theta_0+c_N^{-1}\varepsilon}[\exp(tT)], \quad T \in \mathcal{S},$$

where $t = (-1)^{1/2}\xi$. Let $D(\theta) = \{D_{\alpha\beta}(\theta)\}$ be the unique lower triangular matrix with positive diagonal such that

$$D(\theta)D'(\theta) = \begin{pmatrix} I_{11\cdot 2}(\theta) & 0 \\ 0 & I_{22}(\theta) \end{pmatrix}.$$

We consider the transformation

$$Y^\alpha = D^{\alpha\beta}W_\beta,$$

where $D^{\alpha\beta}(\theta)$ is the (α,β) component of the inverse matrix of $D(\theta)$.

Denoting $L_N(\boldsymbol{x}_N) = p_N(\boldsymbol{x}_N;\theta_0 + c_N^{-1}\varepsilon)/p_N(\boldsymbol{x}_N;\theta_0)$, we have

$$\begin{aligned} \psi_N(\xi,\varepsilon) &= \int \exp\{tT(\boldsymbol{x}_N)\}L_N(\boldsymbol{x}_N)p_N(\boldsymbol{x}_N;\theta_0)d\boldsymbol{x}_N \\ &= E_{\theta_0}[\exp\{tT + \log L_N(\boldsymbol{x}_N)\}]. \end{aligned} \tag{6.1}$$

We expand $\log L_N(\boldsymbol{x}_N)$ in a Taylor series in $c_N^{-1}\varepsilon$, leading to

$$\begin{aligned} \log L_N(\boldsymbol{x}_N) &= W_i\varepsilon^i + g_{\alpha r}g^{rs}W_s\varepsilon^\alpha - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta + \frac{1}{2}c_N^{-1}W_{\alpha\beta}\varepsilon^\alpha\varepsilon^\beta \\ &\quad + \frac{1}{2}c_N^{-1}J_{\gamma,\alpha\beta}g^{\delta\gamma}W_\delta\varepsilon^\alpha\varepsilon^\beta - \frac{1}{6}c_N^{-1}(K_{\alpha\beta\gamma} + J_{\alpha,\beta\gamma}[3])\varepsilon^\alpha\varepsilon^\beta\varepsilon^\gamma \\ &\quad + o_p(c_N^{-1}) \\ &= D_{ij}\varepsilon^iY^j + g_{\alpha r}g^{rs}D_{st}\varepsilon^\alpha Y^t - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta + \frac{1}{2}c_N^{-1}W_{\alpha\beta}\varepsilon^\alpha\varepsilon^\beta \\ &\quad + \frac{1}{2}c_N^{-1}J_{\gamma,\alpha\beta}g^{\delta\gamma}D_{\delta\zeta}\varepsilon^\alpha\varepsilon^\beta Y^\zeta - \frac{1}{6}c_N^{-1}(K_{\alpha\beta\gamma} + J_{\alpha,\beta\gamma}[3])\varepsilon^\alpha\varepsilon^\beta\varepsilon^\gamma \\ &\quad + o_p(c_N^{-1}). \end{aligned} \tag{6.2}$$

Inserting (6.2) in $\exp\{tT + \log L_N(\boldsymbol{x}_N)\}$ we obtain, after further expansion and collection of terms,

$$\begin{aligned} \exp\{tT + \log L_N(\boldsymbol{x}_N)\} = \exp\Bigg\{ & t\sum_{i=1}^p (Y^i)^2 + D_{ij}\varepsilon^iY^j + g_{\alpha r}g^{rs}D_{st}\varepsilon^\alpha Y^t \\ & - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta\Bigg\}\{1 + c_N^{-1}q_1(Y^\alpha, W_{\beta\gamma})\} + o_p(c_N^{-1}), \end{aligned} \tag{6.3}$$

where $q_1(\cdot, \cdot)$ is a polynomial. In view of Assumption 2.1 (iii) we can easily evaluate the asymptotic cumulants of $(Y^\alpha, W_{\beta\gamma})$. Since $E_\theta\{Y^\alpha(\theta)W_{\beta\gamma}(\theta)\} = o(c_N^{-1})$, we derive the second order Edgeworth expansion of the distribution of Y^α. Thus the second order Edgeworth expansion of the distribution of Y^α is given by

$$\begin{aligned}
P_{\theta_0}(Y^\alpha < y^\alpha) &= \int_{-\infty}^{y^\alpha} f(y^\alpha)\left\{1 + \frac{1}{6}c_N^{-1} \sum_{\beta,\gamma,\delta=1}^{p+q} C_{\beta\gamma\delta} H_{\beta\gamma\delta}(y^\alpha)\right\} dy^\alpha + o(c_N^{-1}) \\
&= \int_{-\infty}^{y^\alpha} q(y^\alpha) dy^\alpha + o(c_N^{-1}),
\end{aligned} \tag{6.4}$$

where

$$\begin{aligned}
f(y^\alpha) &= \frac{1}{(2\pi)^{(p+q)/2}} \exp\left\{-\frac{1}{2}\sum_{\alpha=1}^{p+q}(y^\alpha)^2\right\}, \\
C_{\alpha\beta\gamma} &= D_{\alpha_1\alpha} g^{\alpha_1\alpha_2} D_{\beta_1\beta} g^{\beta_1\beta_2} D_{\gamma_1\gamma} g^{\gamma_1\gamma_2} K_{\alpha_2\beta_2\gamma_2},
\end{aligned}$$

and $H_{\beta\gamma\delta}(y^\alpha)$ are the Hermite polynomials. Note that

$$\begin{aligned}
t\sum_{i=1}^{p}(y^i)^2 &+ D_{ij}\varepsilon^i y^j + g_{\alpha r} g^{rs} D_{st}\varepsilon^\alpha y^t - \frac{1}{2} I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta - \frac{1}{2}\sum_{\alpha=1}^{p+q}(y^\alpha)^2 \\
&= \frac{t g_{ij}\varepsilon^i\varepsilon^j}{1-2t} - \frac{1}{2}\sum_{i=1}^{p}\{(1-2t)^{1/2} y^i - (1-2t)^{-1/2} D_{ji}\varepsilon^j\}^2 \\
&\quad - \frac{1}{2}\sum_{r=p+1}^{p+q}\{y^r - g_{\alpha s} g^{st} D_{tr}\varepsilon^\alpha\}^2.
\end{aligned}$$

From (6.1), (6.3) and (6.4) it follows that

$$\begin{aligned}
\psi_N(\xi, \varepsilon) &= \int \exp\left\{t\sum_{i=1}^{p}(y^i)^2 + D_{ij}\varepsilon^i y^j + g_{\alpha r} g^{rs} D_{st}\varepsilon^\alpha y^t - \frac{1}{2} I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta\right\} \\
&\quad \times \{1 + c_N^{-1} q_1(y^\gamma, 0)\} q(y^\delta) dy^\zeta + o(c_N^{-1}) \\
&= \exp\left(\frac{t g_{ij}\varepsilon^i\varepsilon^j}{1-2t}\right)(1-2t)^{-p/2}\left\{1 + c_N^{-1}\sum_{j=0}^{3} m_j (1-2t)^{-j}\right\} \\
&\quad + o(c_N^{-1}).
\end{aligned} \tag{6.5}$$

Inverting (6.5) by Fourier inverse transform we can prove Theorem 2.1. □

PROOF OF THEOREM 2.2. (i) Note that d^α is independent of ε_2. From Theorem 2.1, for $T \in \mathcal{S}$ we have

$$P_{\theta_0 + c_N^{-1}\varepsilon}[T < z] - P_{\theta_{10} + c_N^{-1}\varepsilon_1, \theta_{20}}[T < z]$$

$$= \frac{1}{2} c_N^{-1} (K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) d^\alpha d^\beta \varepsilon^r \{G_{p+2,\lambda}(z) - G_{p,\lambda}(z)\} + o(c_N^{-1}),$$

which leads to (i).

(ii) From $d^\alpha = g_{ij} g^{j\alpha} \varepsilon^i$, clearly

$$(K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) d^\alpha d^\beta \varepsilon^r = g_{ii'} g^{i'\alpha} g_{jj'} g^{j'\alpha} (K_{\alpha\beta r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) \varepsilon^i \varepsilon^j \varepsilon^r.$$

Hence, we get (ii) in Theorem 2.2.

□

PROOF OF THEOREM 2.3 AND COROLLARY 2.1. From Theorem 2.1 we can see that

$$\begin{aligned} m_3 &= \frac{1}{2} a_2^{ijk} g_{ii'} g_{jj'} g_{kk'} d^{i'} d^{j'} d^{k'} + C_3, \\ m_2 &= -\frac{1}{2} a_2^{ijk} g_{ii'} g_{jj'} g_{kk'} d^{i'} d^{j'} d^{k'} + \frac{1}{2} a_2^{ijk} [3] g_{il} g_{jk} d^l + C_2, \\ m_1 &= -\frac{1}{2} a_2^{ijk} [3] g_{il} g_{jk} d^l + \frac{1}{2} a_3^i g_{ij} d^j + C_1, \\ m_0 &= -\frac{1}{2} a_3^i g_{ij} d^j + C_0, \end{aligned} \tag{6.6}$$

where C_0, C_1, C_2 and C_3 are independent of a_1, a_2^{ijk} and a_3^i and hence are the same for all test statistics in $\mathcal{S}$. Theorem 2.1 and Corollary 2.1 follow from (6.6). □

PROOF OF THEOREM 2.4. Let a_2^{ijk} and a_3^i be the coefficients of $T \in \mathcal{S}$. Then, we can rewrite

$$\begin{aligned} P_2^T(\varepsilon) = {} & Q_{1,i'j'k'}(a_2^{ijk}) \varepsilon^{i'} \varepsilon^{j'} \varepsilon^{k'} + Q_{2,ijr} \varepsilon^i \varepsilon^j \varepsilon^r \\ & + \frac{1}{2} g_{li} (g^{i\alpha} B^{\beta\gamma} K_{\alpha\beta\gamma} + a_2^{ijk} [3] g_{jk}) \varepsilon^l \{G_{p+2,\Delta}(z) - G_{p+4,\Delta}(z)\} \\ & + \frac{1}{2} g_{ij} \{a_3^i - g^{i\alpha} g^{rs} (K_{\alpha rs} + J_{\alpha,rs})\} \varepsilon^j \{G_{p,\Delta}(z) - G_{p+2,\Delta}(z)\}. \end{aligned} \tag{6.7}$$

Note that $|\varepsilon^i| \le (\Delta/\lambda)^{1/2}$, where λ is the smallest eigenvalue of $I_{11\cdot 2}$. By (6.7)

$$\begin{aligned} P_2^T(\varepsilon) \le {} & \Psi_1(\Delta, a_2^{ijk}) \Delta^{3/2} + \Psi_{2r}(\Delta) \Delta |\varepsilon^r| \\ & + \frac{1}{2} g_{li} (g^{i\alpha} B^{\beta\gamma} K_{\alpha\beta\gamma} + a_2^{ijk} [3] g_{jk}) \varepsilon^l \{G_{p+2,\Delta}(z) - G_{p+4,\Delta}(z)\} \\ & + \frac{1}{2} g_{ij} \{a_3^i - g^{i\alpha} g^{rs} (K_{\alpha rs} + J_{\alpha,rs})\} \varepsilon^j \{G_{p,\Delta}(z) - G_{p+2,\Delta}(z)\}, \end{aligned}$$

where

$$\Psi_1(\Delta, a_2^{ijk}) = \sum_{i',j',k'=1}^{p} \left| Q_{1,i'j'k'}(a_2^{ijk}) \right| \lambda^{-3/2}, \qquad \Psi_{2r}(\Delta) = \sum_{i,j=1}^{p} \left| Q_{2,ijr} \right| \lambda^{-1}.$$

Hence, we obtain

$$P_{\varepsilon_2}^{T}(\Delta) \le \Psi_1(\Delta, a_2^{ijk})\Delta^{3/2} + \Psi_{2r}(\Delta)\Delta|\varepsilon^r| + M(\Delta), \tag{6.8}$$

where

$$\begin{aligned} M(\Delta) = \min_{g_{ij}\varepsilon^i\varepsilon^j=\Delta}\Big[&\frac{1}{2}g_{li}(g^{i\alpha}B^{\beta\gamma}K_{\alpha\beta\gamma} + a_2^{ijk}[3]g_{jk})\varepsilon^l\{G_{p+2,\Delta}(z) - G_{p+4,\Delta}(z)\} \\ &+ \frac{1}{2}g_{ij}\{a_3^i - g^{i\alpha}g^{rs}(K_{\alpha rs} + J_{\alpha,rs})\}\varepsilon^j\{G_{p,\Delta}(z) - G_{p+2,\Delta}(z)\}\Big]. \end{aligned}$$

Similarly, we have

$$P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) \ge -\Psi_1(\Delta, -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha\beta\gamma}/3)\Delta^{3/2} - \Psi_{2r}(\Delta)\Delta|\varepsilon^r|. \tag{6.9}$$

From (6.8) and (6.9),

$$\begin{aligned} P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) - P_{\varepsilon_2}^{T}(\Delta) \ge -\big\{&\Psi_1(\Delta, a_2^{ijk}) + \Psi_1(\Delta, -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha\beta\gamma}/3)\big\}\Delta^{3/2} \\ &- 2\Psi_{2r}(\Delta)\Delta|\varepsilon^r| - M(\Delta). \end{aligned}$$

Hence, for $T \in \mathcal{S}$ whose coefficients do not satisfy $z(a_2^{ijk}[3]g_{jk} + g^{i\alpha}B^{\beta\gamma}K_{\alpha\beta\gamma}) + (p+2)\{a_3^i - g^{i\alpha}g^{rs}(K_{\alpha rs} + J_{\alpha,rs})\} = 0$, there exists a positive Δ_0 such that

$$P_{\varepsilon_2}^{\mathrm{LR}^*}(\Delta) - P_{\varepsilon_2}^{T}(\Delta) > 0,$$

whenever $0 < \Delta < \Delta_0$. □

PROOF OF THEOREM 2.5. The distribution function of $T_0 \in \mathcal{S}_0$ under a sequence of local alternatives $\theta_1 = \theta_{10} + c_N^{-1}\varepsilon_1$ has the asymptotic expansion

$$P_{\theta_{10}+c_N^{-1}\varepsilon_1,\theta_{20}}[T_0 < z] = G_{p,\Delta}(z) + c_N^{-1}\sum_{j=0}^{3} m_{j0}G_{p+2j,\Delta}(z) + o(c_N^{-1}),$$

where

$$\begin{aligned} m_{30} &= \left(\frac{1}{6}K_{ijk} + \frac{1}{2}a_2^{i'j'k'}I_{(i'i)}I_{(j'j)}I_{(k'k)}\right)\varepsilon^i\varepsilon^j\varepsilon^k, \\ m_{20} &= -\frac{1}{2}a_2^{i'j'k'}I_{(i'i)}I_{(j'j)}I_{(k'k)}\varepsilon^i\varepsilon^j\varepsilon^k + \frac{1}{2}I_0^{ij}K_{ijk}\varepsilon^k + \frac{1}{2}a_2^{ijk}[3]I_{(il)}I_{(jk)}\varepsilon^l, \\ m_{10} &= \frac{1}{2}J_{i,jk}\varepsilon^i\varepsilon^j\varepsilon^k - \frac{1}{2}I_0^{ij}K_{ijk}\varepsilon^k - \frac{1}{2}a_2^{ijk}[3]I_{(il)}I_{(jk)}\varepsilon^l, \\ m_{00} &= -\frac{1}{6}(K_{ijk} + 3J_{i,jk})\varepsilon^i\varepsilon^j\varepsilon^k. \end{aligned} \tag{6.10}$$

Note that, under Assumption 2.2, $d^r = 0$ and

$$\{B^{\alpha\beta}\} = \begin{pmatrix} (I_{11})^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Then the coefficients m_3, m_2, m_1 and m_0 in Theorem 2.1 can be written as

$$\begin{aligned} m_3 &= m_{30}, \\ m_2 &= m_{20}, \\ m_1 &= m_{10} + \frac{1}{2}(K_{ijr} + J_{i,jr} + J_{j,ir})\varepsilon^i\varepsilon^j\varepsilon^r \\ &\quad + \frac{1}{2}\{a_3^i I_{(ij)} - g^{rs}(K_{jrs} + J_{j,rs})\}\varepsilon^j, \\ m_0 &= m_{00} - \frac{1}{2}(K_{ijr} + J_{i,jr} + J_{j,ir})\varepsilon^i\varepsilon^j\varepsilon^r \\ &\quad - \frac{1}{2}\{a_3^i I_{(ij)} - g^{rs}(K_{jrs} + J_{j,rs})\}\varepsilon^j. \end{aligned} \tag{6.11}$$

The comparison of (6.10) and (6.11) leads to Theorem 2.5. □

PROOF OF THEOREM 2.6 AND 2.7. Note that $\tilde{Z}_i = W_i + o_p(1)$. Expand T^* as

$$\begin{aligned} T^* &= h(\hat{\theta}_1)T + c_N^{-1}A^i\tilde{Z}_i \\ &= (1 + c_N^{-1}h_i\eta^i)T + c_N^{-1}A^iW_i + o_p(c_N^{-1}). \end{aligned} \tag{6.12}$$

Inserting (2.5) in (6.12) we obtain

$$\begin{aligned} T^* = g^{ij}W_iW_j &+ c_N^{-1}a_1g^{i\alpha}g^{j\beta}W_{\alpha\beta}W_iW_j + 2c_N^{-1}g^{i\alpha}g^{rs}W_{\alpha r}W_iW_s \\ &+ c_N^{-1}a_2^{*ijk}W_iW_jW_k - c_N^{-1}g^{i\alpha}g^{j\beta}g^{rs}K_{\alpha\beta r}W_iW_jW_s \\ &- c_N^{-1}g^{i\alpha}g^{rt}g^{su}(K_{\alpha rs} + J_{\alpha,rs})W_iW_tW_u + c_N^{-1}a_3^{*i}W_i + o_p(c_N^{-1}), \end{aligned}$$

where

$$\begin{aligned} a_2^{*ijk} &= a_2^{ijk} + h_lg^{li}g^{jk}, \\ a_3^{*i} &= a_3^i + A^i. \end{aligned} \tag{6.13}$$

This implies $T^* \in \mathcal{S}$, and hence a necessary and sufficient condition for its locally unbiasedness is that the coefficients in (6.13) satisfy

$$a_2^{*ijk}[3]g_{il}g_{jk} + g_{li}g^{i\alpha}B^{\beta\gamma}K_{\alpha\beta\gamma} = 0, \tag{6.14}$$

$$a_3^{*i}g_{ij} - g_{ji}g^{i\alpha}g^{rs}(K_{\alpha rs} + J_{\alpha,rs}) = 0. \tag{6.15}$$

Note that

$$\begin{aligned} a_2^{*ijk}[3]g_{il}g_{jk} &= a_2^{ijk}[3]g_{il}g_{jk} + (h_{l'}g^{l'i}g^{jk} + h_{l'}g^{l'j}g^{ki} + h_{l'}g^{l'k}g^{ij})g_{il}g_{jk} \\ &= a_2^{ijk}[3]g_{il}g_{jk} + (p+2)h_l. \end{aligned}$$

Solving (6.14) and (6.15) with respect to h_i and A^i, we obtain the relations in Theorem 2.6. Theorem 2.7 follows from the above argument and Theorem 2.1. □

6.2. Proofs of Chapter 3

PROOF OF THEOREM 3.1. First, we evaluate the asymptotic cumulants of Z_N. From (i) and (ii) in Assumption 3.1, $E(Z_N) = 0$,

$$\begin{aligned} \operatorname{cum}(Z_N, Z_N) &= N^{-1} \sum_{j=-(N-1)}^{N-1} (N - |j|) c_{X,2}(j) \\ &= 2\pi f_{X,2} - N^{-1} f'_{X,2} + o(N^{-1}), \\ \operatorname{cum}(Z_N, Z_N, Z_N) &= N^{-3/2} \sum_{j_1, j_2=-(N-1)}^{N-1} (N - S_{j_1 j_2}) c_{X,3}(j_1, j_2) \\ &= N^{-1/2} (2\pi)^2 f_{X,3} + o(N^{-1}), \end{aligned}$$

where

$$S_{j_1 j_2} = \begin{cases} \max(|j_1|, |j_2|) & \text{if } \operatorname{sign}(j_1) = \operatorname{sign}(j_2), \\ \min(|j_1| + |j_2|, N) & \text{if } \operatorname{sign}(j_1) = -\operatorname{sign}(j_2), \end{cases}$$

and

$$\begin{aligned} \operatorname{cum}(Z_N, Z_N, Z_N, Z_N) &= N^{-2} \sum_{j_1, j_2, j_3=-(N-1)}^{N-1} (N - S_{j_1 j_2 j_3}) c_{X,4}(j_1, j_2, j_3) \\ &= N^{-1} (2\pi)^3 f_{X,4} + o(N^{-1}), \end{aligned}$$

where

$$S_{j_1 j_2 j_3} = \begin{cases} \max(|j_1|, |j_2|, |j_3|) & \text{if } \operatorname{sign}(j_1) = \operatorname{sign}(j_2) = \operatorname{sign}(j_2), \\ \min\{\max(|j_1| + |j_2|) + |j_3|, N\} & \text{if } \operatorname{sign}(j_1) = \operatorname{sign}(j_2) = -\operatorname{sign}(j_3). \end{cases}$$

Applying the general formula for the Edgeworth expansion (e.g., Taniguchi and Kakizawa [48, p168–170]), we obtain (3.5). □

PROOF OF THEOREM 3.2. From Theorem 3.1 and (3.6),

$$C = e^{-r\tau} \int_{-d_2}^{\infty} \left[S_{T_0} \exp\left\{ \mu\tau + (2\pi\tau f_{X,2})^{1/2} z \right\} - K \right] g(z) dz. \tag{6.16}$$

Integrating by parts and using the following equality

$$\exp\{-(2\pi\tau f_{X,2})^{1/2} d_2\} \phi(-d_2) = \exp(\pi\tau f_{X,2}) \phi(d_1),$$

yield

$$\begin{aligned} &\int_{-d_2}^{\infty} \left[S_{T_0} \exp\left\{ \mu\tau + (2\pi\tau f_{X,2})^{1/2} z \right\} - K \right] H_k(z) \phi(z) dz \\ &= a_2 S_{T_0} \left\{ \sum_{j=1}^{k-1} (2\pi\tau f_{X,2})^{j/2} H_{k-j-1}(-d_2) \phi(d_1) + (2\pi\tau f_{X,2})^{k/2} \Phi(d_1) \right\} \end{aligned} \tag{6.17}$$

for $k = 2, 3, 4, 6$. Inserting (6.17) in (6.16), we obtain (3.7). □

PROOF OF COROLLARY 3.1. If $\{X_t; t \in \mathbf{Z}\}$ is independent, then $f'_{X,2} = 0$. If $\{X_t; t \in \mathbf{Z}\}$ is a Gaussian process, then $f_{X,3} = f_{X,4} = 0$. Hence, Corollary 3.1 follows. □

PROOF OF COROLLARY 3.2. From (i)-(iii) in Assumption 3.2 and (3.8),

$$f_{X,k} = A^k f_{\varepsilon,k}, \quad k = 2, 3, 4,$$

and (ii) in Assumption 3.1 holds. Note that

$$\begin{aligned} c_{X,2}(u) &= Var\left(\sum_{j_1=0}^{\infty} a_{j_1}\varepsilon_{t-j_1}, \sum_{j_2=0}^{\infty} a_{j_2}\varepsilon_{t+u-j_2}\right) \\ &= \sigma^2 \sum_{j=0}^{\infty} a_j a_{|u|+j}. \end{aligned}$$

We can see $f'_{X,2} = \sigma^2 f'_{\varepsilon,2}$. From above arguments Corollary 3.2 follows. □

PROOF OF THEOREM 3.3. From the martingale restriction,

$$\begin{aligned} S_{T_0} &= e^{-r\tau} E_{T_0}[S_T], \\ &= e^{-r\tau} \int_{-\infty}^{\infty} S_{T_0} \exp\left\{\tau\mu + (2\pi\tau f_{X,2})^{1/2} z\right\} g(z)dz. \end{aligned} \tag{6.18}$$

Note that

$$\int_{-\infty}^{\infty} \exp\left\{(2\pi\tau f_{X,2})^{1/2} z\right\} H_k(z)\phi(z)dz = (2\pi\tau f_{X,2})^{k/2} \exp(\pi\tau f_{X,2})$$

for $k = 2, 3, 4, 6$. The equation (6.18) implies that

$$\begin{aligned} 1 = \exp(-r\tau + \tau\mu + \pi\tau f_{X,2}) &\left\{ 1 + \frac{2}{3}\pi^2\tau^{3/2}N^{-1/2} f_{X,3} \right. \\ &\left. - \frac{1}{2}\tau N^{-1} f'_{X,2} + \frac{1}{3}\pi^3\tau^2 N^{-1} f_{X,4} + \frac{2}{9}\pi^4\tau^3 N^{-1}(f_{X,3})^2 \right\} + o(N^{-1}). \end{aligned} \tag{6.19}$$

Taking the logarithm of the equation (6.19) and using Taylor expansion, yield

$$\begin{aligned} \mu = r - \pi f_{X,2} &- \frac{2}{3}\pi^2\tau^{1/2}N^{-1/2} f_{X,3} \\ &+ \frac{1}{2}N^{-1} f'_{X,2} - \frac{1}{3}\pi^3\tau N^{-1} f_{X,4} + o(N^{-1}). \end{aligned} \tag{6.20}$$

Substituting (6.20) into $G_k,\ k = 0, 2, 3, 4, 6$ in Theorem 3.2, further expansion and collection of terms, we obtain

$$\begin{aligned} G_0 = G_0^* &- \frac{2}{3}\pi^2\tau^{3/2} S_{T_0} N^{-1/2} f_{X,3}\Phi(d_1^*) \\ &+ S_{T_0} N^{-1}\left\{\frac{1}{2}\tau f'_{X,2} - \frac{1}{3}\pi^3\tau^2 f_{X,4} + \frac{2}{9}\pi^4\tau^3 (f_{X,3})^2\right\}\Phi(d_1^*) \\ &+ \frac{\pi}{36} S_{T_0} N^{-1}\frac{(f_{X,3})^2}{(f_{X,2})^3}(2\pi\tau f_{X,2})^{5/2}\phi(d_1^*) + o(N^{-1}), \end{aligned} \tag{6.21}$$

$$\begin{aligned} G_3 = G_3^* &+ S_{T_0}(2\pi\tau f_{X,2})^{3/2}\Phi(d_1^*) \\ &- \frac{(2\pi)^{7/2}}{6}\tau^3 S_{T_0} N^{-1/2} f_{X,3}(f_{X,2})^{3/2}\Phi(d_1^*) \\ &- \frac{(2\pi)^{1/2}}{6} S_{T_0} N^{-1/2}\frac{f_{X,3}}{(f_{X,2})^{3/2}}\sum_{j=1}^{3}(2\pi\tau f_{X,2})^{j/2+1} H_{3-j}(-d_2^*)\phi(d_1^*) \\ &+ o(N^{-1/2}), \end{aligned} \tag{6.22}$$

and

$$G_k = S_{T_0}\left\{\sum_{j=1}^{k-1}(2\pi\tau f_{X,2})^{j/2} H_{k-j-1}(-d_2^*)\phi(d_1^*) + (2\pi\tau f_{X,2})^{k/2}\Phi(d_1^*)\right\} + o(1) \tag{6.23}$$

for $k = 2, 4, 6$. From (6.21)-(6.23), Theorem 3.3 follows. □

6.3. Proofs of Chapter 4

In this section we give the proofs of lemmas and theorems and state some lemmas related to the results in Chapter 4.

PROOF OF THEOREM 4.1. We decompose $\tilde{F}_{uu}(\lambda)$ as follow:

$$\tilde{F}_{uu}(\lambda) = F(\lambda) + \sum_{j=1}^{4} V_j(\lambda), \tag{6.24}$$

where

$$\begin{aligned} V_2(\lambda) &= \int_{-\pi}^{\pi} W_N(\lambda - \mu) F(\mu) d\mu - F(\lambda), \\ V_3(\lambda) &= \tilde{F}_{uu}(\lambda) - \hat{F}_{uu}(\lambda), \end{aligned}$$

$$V_4(\lambda) = \mathrm{E}\big[\hat{F}_{uu}(\lambda)\big] - \int_{-\pi}^{\pi} W_N(\lambda - \mu) F(\mu) d\mu.$$

The order of magnitude for each of these terms in our decomposition (6.24) is given by the standard texts (e.g., [2, 5, 19]) and stated in the following lemma for convenience.

LEMMA 6.1. *$V_1(\lambda) = O_p((M/N)^{1/2})$, $V_2(\lambda) = O(M^{-2})$, $V_3(\lambda) = O_p(M/N)$, and $V_4(\lambda) = O(N^{-1})$.*

Expanding $\tilde{F}_{uu}(\lambda)^{-1}$ about $F(\lambda)^{-1}$, we obtain, after application of Lemma 6.1,

$$\begin{aligned}\tilde{F}_{uu}(\lambda)^{-1} = {} & F(\lambda)^{-1} - F(\lambda)^{-1} \sum_{j=1}^{3} V_j(\lambda) F(\lambda)^{-1} \\ & + F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} + O_p(M^{-3/2} N^{-1/2}).\end{aligned} \tag{6.25}$$

Let

$$\tilde{Z} = \frac{N}{2\pi} \int_{-\pi}^{\pi} \big\{\tilde{F}_{uu}(\lambda)^{-1} \otimes D_N{}^{-1}\big\} \hat{f}_{ux}(\lambda) d\lambda. \tag{6.26}$$

We then have

$$(I_p \otimes D_N)(\hat{\beta} - \beta) = \tilde{\mathcal{I}}^{-1} \tilde{Z}.$$

Inserting (6.25) into (6.26) we have

$$\begin{aligned}\tilde{Z} = {} & Z_1 - N^{-1/2} Z_2 - \frac{N}{2\pi} \int_{-\pi}^{\pi} \big[\big\{F(\lambda)^{-1} V_2(\lambda) F(\lambda)^{-1} \\ & + F(\lambda)^{-1} V_3(\lambda) F(\lambda)^{-1} \\ & - F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\big\} \otimes D_N{}^{-1}\big] \hat{f}_{ux}(\lambda) d\lambda \\ & \quad + o_p(N^{-1/2}),\end{aligned} \tag{6.27}$$

where we used the fact that $(I_p \otimes D_N{}^{-1}) \hat{f}_{ux}(\lambda) = O_p(M/N)$.

The order of magnitude for each of these terms in (6.27) is given in the next lemma.

LEMMA 6.2.

$$\begin{aligned}& Z_1 = O_p(1), \quad Z_2 = O_p(M^{1/2}), \\ & \frac{N}{2\pi} \int_{-\pi}^{\pi} \big[\big\{F(\lambda)^{-1} V_2(\lambda) F(\lambda)^{-1}\big\} \otimes D_N{}^{-1}\big] \hat{f}_{ux}(\lambda) d\lambda = O_p(M^{-2}), \\ & \frac{N}{2\pi} \int_{-\pi}^{\pi} \big[\big\{F(\lambda)^{-1} V_3(\lambda) F(\lambda)^{-1}\big\} \otimes D_N{}^{-1}\big] \hat{f}_{ux}(\lambda) d\lambda = O_p(M/N), \\ & \frac{N}{2\pi} \int_{-\pi}^{\pi} \big[\big\{F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\big\} \otimes D_N{}^{-1}\big] \hat{f}_{ux}(\lambda) d\lambda = o_p(N^{-1/2}).\end{aligned}$$

Inserting (6.25) into $\tilde{\mathcal{I}}$ we have

$$\begin{aligned}\tilde{\mathcal{I}} = \frac{N}{2\pi}\int_{-\pi}^{\pi} F(\lambda)^{-1} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda - N^{-1/2} Z_3 \\ - \frac{N}{2\pi}\int_{-\pi}^{\pi} \{F(\lambda)^{-1} V_2(\lambda) F(\lambda)^{-1} + F(\lambda)^{-1} V_3(\lambda) F(\lambda)^{-1} \\ - F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda \\ + o_p(N^{-1/2}).\end{aligned} \tag{6.28}$$

where we used the fact that $D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1} = O_p(M/N)$.

The order of magnitude for each of these terms in (6.28) is given in the next lemma.

LEMMA 6.3.

$$\begin{aligned}&\frac{N}{2\pi}\int_{-\pi}^{\pi} F(\lambda)^{-1} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda = \mathcal{I} + o_p(N^{-1/2}),\\ &Z_3 = O_p(M^{1/2}),\\ &\frac{N}{2\pi}\int_{-\pi}^{\pi} \{F(\lambda)^{-1} V_2(\lambda) F(\lambda)^{-1}\} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda = O_p(M^{-2}),\\ &\frac{N}{2\pi}\int_{-\pi}^{\pi} \{F(\lambda)^{-1} V_3(\lambda) F(\lambda)^{-1}\} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda = O_p(M/N),\\ &\frac{N}{2\pi}\int_{-\pi}^{\pi} \{F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda \\ &\qquad = o_p(N^{-1/2}).\end{aligned}$$

Theorem 4.1 follows from Lemma 6.2 and Lemma 6.3. □

PROOF OF LEMMA 6.2. The proofs of the first four equalities follow directly by evaluating the absolute moments. Hence, we only give the proofs of the last equality.

Note that

$$\mathrm{E}\left[\int_{-\pi}^{\pi} \|(I_p \otimes D_N{}^{-1})\hat{f}_{ux}(\lambda)\|^2 d\lambda\right] = O(M/N^2)$$

and

$$\mathrm{E}\big[\|V_1(\lambda)\|^4\big] = \mathrm{E}\Big[\big\|\hat{F}_{uu}(\lambda) - \mathrm{E}\big[\hat{F}_{uu}(\lambda)\big]\big\|^4\Big] = O((M/N)^2),$$

(see the proof of Theorem 7.4.4 in Brillinger [5]). We have

$$\begin{aligned}&\left\|\frac{N}{2\pi}\int_{-\pi}^{\pi} \big[\{F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\} \otimes D_N{}^{-1}\big]\hat{f}_{ux}(\lambda) d\lambda\right\| \\ &\qquad \le \frac{N}{2\pi}\int_{-\pi}^{\pi} \|F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} \otimes I_q\|\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad \times \|(I_p \otimes D_N{}^{-1})\hat{f}_{ux}(\lambda)\| d\lambda \\
&\le \frac{N}{2\pi}\left\{\int_{-\pi}^{\pi} \|F(\lambda)^{-1}\|^6 \|V_1(\lambda)\|^4 d\lambda\right\}^{1/2} \left\{\int_{-\pi}^{\pi} \|(I_p \otimes D_N{}^{-1})\hat{f}_{ux}(\lambda)\|^2 d\lambda\right\}^{1/2} \\
&= N \times O_p(M/N) \times O_p(M^{1/2}/N) \\
&= o_p(N^{-1/2}).
\end{aligned}
$$

□

PROOF OF LEMMA 6.3. Similarly to Lemma 6.2, we give the proofs of the first and last equalities. The first one is evaluated as follows:

$$
\begin{aligned}
&\frac{N}{2\pi}\int_{-\pi}^{\pi} F(\lambda)^{-1} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda \\
&\quad = \left(\frac{1}{2\pi}\right)^2 \sum_{l=-M}^{M} \Delta(l) w\left(\frac{l}{M}\right) \otimes \left\{R(l)' + O\left(\frac{1+|l|}{N}\right)\right\} \\
&\quad = \left(\frac{1}{2\pi}\right)^2 \sum_{l=-\infty}^{\infty} \Delta(l) \otimes R(l)' + O(M^{-2}) + O(N^{-1}) \\
&\quad = \mathcal{I} + o(N^{-1/2}).
\end{aligned}
$$

From

$$
\int_{-\pi}^{\pi} \|D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\|^2 d\lambda = O(M/N^2),
$$

We have

$$
\begin{aligned}
&\left\| \frac{N}{2\pi}\int_{-\pi}^{\pi} \{F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1} V_1(\lambda) F(\lambda)^{-1}\} \otimes \{D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\} d\lambda \right\| \\
&\quad \le \frac{N}{2\pi}\int_{-\pi}^{\pi} \|F(\lambda)^{-1}\|^3 \|V_1(\lambda)\|^2 \|D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\| d\lambda \\
&\quad \le \frac{N}{2\pi}\left\{\int_{-\pi}^{\pi} \|F(\lambda)^{-1}\|^6 \|V_1(\lambda)\|^4 d\lambda\right\}^{1/2} \left\{\int_{-\pi}^{\pi} \|D_N{}^{-1}\hat{F}_{xx}(\lambda)' D_N{}^{-1}\|^2 d\lambda\right\}^{1/2} \\
&\quad = o_p(N^{-1/2}).
\end{aligned}
$$

Thus we complete the proofs of Lemma 6.3. □

PROOF OF LEMMA 4.1 AND THEOREM 4.2. From direct verifications, it is seen that

$$
\begin{aligned}
\mathrm{E}[Z_2(jk)] = &\left(\frac{1}{2\pi}\right)^4 N^{-1/2} \sum_{s_1,s_2=-\infty}^{\infty} \Delta^{jj_1}(s_1)\Delta^{j_2j_3}(s_2) \sum_{l_1=-M}^{M} w\left(\frac{l_1}{M}\right) w\left(\frac{l_2}{M}\right) \\
&\times \sum_{m_2=1+\underline{l_2}}^{N-\overline{l_2}} \frac{x_k(m_2+l_2)}{d_k(N)} \sum_{m_1=1+\underline{l_1}}^{N-\overline{l_1}} \Gamma_{j_1j_2j_3}(l_1, m_2-m_1)
\end{aligned}
$$

$$
\begin{aligned}
&\quad (l_2 = s_1 + s_2 - l_1) \\
&= \left(\frac{1}{2\pi}\right)^4 \eta_k \sum_{s_1,s_2=-\infty}^{\infty} \Delta^{jj_1}(s_1)\Delta^{j_2j_3}(s_2) \sum_{l_1=-M}^{M} w\left(\frac{l_1}{M}\right) \\
&\quad \times w\left(\frac{s_1+s_2-l_1}{M}\right) \sum_{m=-\infty}^{\infty} \Gamma_{j_1j_2j_3}(l_1, m) + \mathrm{o}(1) \\
&= \left(\frac{1}{2\pi}\right)^4 \eta_k \sum_{s_1=-\infty}^{\infty} \Delta^{jj_1}(s_1) \sum_{l_1,l_2=-M}^{M} w\left(\frac{l_1}{M}\right) w\left(\frac{l_2}{M}\right) \\
&\quad \times \int_{-\pi}^{\pi} F^{j_2j_3}(\lambda) e^{-i(l_1+l_2-s_1)\lambda} d\lambda \sum_{m=-\infty}^{\infty} \Gamma_{j_1j_2j_3}(l_1, m) + o(1) \\
&= F^{jj_1}(0) F_{j_1j_2j_3}(0,0) F^{j_2j_3}(0) \eta_k + \mathrm{o}(1) \\
&= \sum_{j_1,j_2=1}^{p} \mathcal{K}_{jj_1j_2}(0,0) F_{j_1j_2}(0) \eta_k + o(1),
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{E}[Z_1(j_1k_1)Z_2(j_2k_2)] \\
&\quad = \left(\frac{1}{2\pi}\right)^6 N^{-1/2} \sum_{s_1,s_2,s_3=-\infty}^{\infty} \Delta^{jj_1}(s_1)\Delta^{kk_1}(s_2)\Delta^{k_2k_3}(s_3) \\
&\qquad \times \sum_{l_2=-M}^{M} w\left(\frac{l_1}{M}\right) w\left(\frac{l_2}{M}\right) w\left(\frac{l_3}{M}\right) \sum_{m_1=1+\underline{l_1}}^{N-\overline{l_1}} \sum_{m_2=1+\underline{l_2}}^{N-\overline{l_2}} \sum_{m_3=1+\underline{l_3}}^{N-\overline{l_3}} \\
&\qquad \times \frac{x_{j'}(m_1+l_1)x_{k'}(m_3+l_3)}{d_{j'}(N)d_{k'}(N)} \\
&\qquad \times E[u_{j_1}(m_1)\{u_{k_1}(m_2)u_{k_2}(m_2+l_2) - \Gamma_{k_1k_2}(l_2)\}u_{k_3}(m_3)] \\
&\qquad\quad (l_1 = s_1, \quad l_3 = s_2 + s_3 - l_2) \\
&\quad = \left(\frac{1}{2\pi}\right)^6 N^{-1/2} \sum_{s_1,s_2,s_3=-\infty}^{\infty} \Delta^{jj_1}(s_1)\Delta^{kk_1}(s_2)\Delta^{k_2k_3}(s_3) \\
&\qquad \times \sum_{l_2=-M}^{M} w\left(\frac{l_1}{M}\right) w\left(\frac{l_2}{M}\right) w\left(\frac{l_3}{M}\right) \sum_{m_1=1+\underline{l_1}}^{N-\overline{l_1}} \sum_{m_2=1+\underline{l_2}}^{N-\overline{l_2}} \sum_{m_3=1+\underline{l_3}}^{N-\overline{l_3}} \\
&\qquad \times \frac{x_{j'}(m_1+l_1)x_{k'}(m_3+l_3)}{d_{j'}(N)d_{k'}(N)} \\
&\qquad \times \{\Gamma_{j_1k_1k_2j_3}(m_2-m_1, m_2-m_1+l_2, m_3-m_1) \\
&\qquad \times \Gamma_{j_1k_1}(m_2-m_1)\Gamma_{k_2j_3}(m_3-m_2-l_2) \\
&\qquad + \Gamma_{j_1k_2}(m_2-m_1+l_2)\Gamma_{k_1j_3}(m_3-m_2)\} \\
&\quad = O(M/N^{1/2}),
\end{aligned}
$$

$$
\begin{aligned}
&E[Z_1(j_1k_1)Z_3(j_2k_2,\, j_3k_3)] \\
&\quad = \left(\frac{1}{2\pi}\right)^6 N^{-1/2} \sum_{s_1,s_2,s_3=-\infty}^{\infty} \Delta^{j_1a_1}(s_1)\Delta^{j_2a_2}(s_2)\Delta^{a_3j_3}(s_3) \\
&\qquad \times \sum_{l_2=-M}^{M} w\left(\frac{l_1}{M}\right) w\left(\frac{l_2}{M}\right) w\left(\frac{l_3}{M}\right) \\
&\qquad \times \sum_{m_1=1+\underline{l_1}}^{N-\overline{l_1}} \sum_{m_2=1+\underline{l_2}}^{N-\overline{l_2}} \sum_{m_3=1+\underline{l_3}}^{N-\overline{l_3}} E[u_{a_1}(m_1)\{u_{a_2}(m_2)u_{a_3}(m_2+l_2) - \Gamma_{a_2a_3}(l_2)\} \\
&\qquad \times \frac{x_{k_1}(m_1+l_1)x_{k_2}(m_3)x_{k_3}(m_3+l_3)}{d_{k_1}(N)d_{k_2}(N)d_{k_3}(N)} \\
&\quad\; (l_1 = s_1, \quad l_3 = s_2 + s_3 - l_2) \\
&\quad = \left(\frac{1}{2\pi}\right)^6 \eta_{k_1} \sum_{s_1,s_2,s_3=-\infty}^{\infty} \Delta^{j_1a_1}(s_1)\Delta^{j_2a_2}(s_2)\Delta^{a_3j_3}(s_3) \\
&\qquad \times \sum_{l_2=-M}^{M} w\left(\frac{l_1}{M}\right) w\left(\frac{l_2}{M}\right) w\left(\frac{l_3}{M}\right) \\
&\qquad \times \sum_{m=-\infty}^{\infty} \Gamma_{a_1a_2a_3}(m, m+l_2) R_{k_1k_2}(l_3) + o(1) \\
&\quad = \frac{1}{2\pi} N^{-1/2} \eta_k F^{jj_1}(0) \int_{-\pi}^{\pi} F^{aa_1}(\lambda) F^{a_2b}(\lambda) F_{j_1a_1a_2}(-\lambda,\lambda)\, dM_{cd}(\lambda) \\
&\qquad + o(N^{-1/2}), \\
&\quad = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{K}_{j_1j_2j_3}(\lambda,-\lambda)\eta_{k_1}\, dM_{k_2k_3}(\lambda) + o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathrm{cum}[Z_1(j_1k_1), Z_1(j_2k_2), Z_1(j_3k_3)] \\
&\quad = \left(\frac{1}{2\pi}\right)^6 \sum_{s_1,s_2,s_3=-M}^{M} \Delta^{j_1j_1'}(s_1)\Delta^{j_2j_2'}(s_2)\Delta^{j_3j_3'}(s_3) w\left(\frac{s_1}{M}\right) w\left(\frac{s_2}{M}\right) w\left(\frac{s_3}{M}\right) \\
&\qquad \times \sum_{m_1=1+\underline{s_1}}^{N-\overline{s_1}} \sum_{m_2=1+\underline{s_2}}^{N-\overline{s_2}} \sum_{m_3=1+\underline{s_3}}^{N-\overline{s_3}} \frac{x_{k_1}(m_1+s_1)x_{k_2}(m_2+s_2)x_{k_3}(m_3+s_3)}{d_{k_1}(N)d_{k_2}(N)d_{k_3}(N)} \\
&\qquad \times \Gamma_{j_1'j_2'j_3'}(m_2-m_1, m_3-m_1) \\
&\quad = \left(\frac{1}{2\pi}\right)^6 \sum_{s_1,s_2,s_3=-M}^{M} \Delta^{j_1j_1'}(s_1)\Delta^{j_2j_2'}(s_2)\Delta^{j_2j_2'}(s_2) w\left(\frac{s_1}{M}\right) w\left(\frac{s_2}{M}\right) w\left(\frac{s_3}{M}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\times \sum_{m_1=1}^{N} \sum_{m_2=-\infty}^{\infty} \sum_{m_3=-\infty}^{\infty} \Gamma_{j_1' j_2' j_3'}(m_2, m_3) \\
&\times \frac{x_{k_1}(m_1) x_{k_2}(m_1 + m_2 + s_2 - s_1) x_{k_3}(m_1 + m_3 + s_3 - s_1)}{d_{k_1}(N) d_{k_2}(N) d_{k_2}(N)} + O(N^{-3/2}) \\
= &\frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathcal{K}_{j_1 j_2 j_3}(\lambda_1, \lambda_2) dM_{k_1 k_2 k_3}(\lambda_1, \lambda_2) + o(N^{-1/2}).
\end{aligned}
$$

□

PROOF OF THEOREM 4.3 AND THEOREM 4.4. It is sufficient to show that $\mathcal{I}_{\mathrm{B}} = \mathcal{I} + o(N^{-1/2})$. The proof is substantially a modification of that of Theorem 5 in Hannan [19, pp. 427], see also Theorem 10.2.7 in Anderson [2, pp. 575].

From Assumption 4.1 (i)-(iii), we can find spectral matrices $F_1(\lambda)^{-1}$ and $F_2(\lambda)^{-1}$ of moving average processes of order M such that

$$
\begin{aligned}
&0 < F_2(\lambda)^{-1} \leq F(\lambda)^{-1} \leq F_1(\lambda)^{-1}, \\
&F_1(\lambda)^{-1} - F_2(\lambda)^{-1} < \delta I_p,
\end{aligned} \tag{6.29}
$$

where $\delta = O(M^{-2})$. Here these inequalities are to be interpreted in the usual way as between Hermitian matrices. In fact, let

$$
\begin{aligned}
F_1(\lambda)^{-1} &= \frac{1}{2\pi} \sum_{s=-M}^{M} \Delta(s) e^{is\lambda} + \frac{K_1}{M^2} I_p, \\
F_2(\lambda)^{-1} &= \frac{1}{2\pi} \sum_{s=-M}^{M} \Delta(s) e^{is\lambda} - \frac{K_1}{M^2} I_p,
\end{aligned}
$$

then we can choose a constant $K_1 > 0$ such that (6.29) holds. Thus we have approximated $F(\lambda)$ by autoregressive processes of order M. Let $\{u(t)\}$ satisfy the equation

$$
\sum_{s=0}^{M} C_1(s) u(t-s) = \eta(t),
$$

where $C_1(s)$ are the autoregressive matrices corresponding to $F_1(\lambda)$ and the $\eta(t)$ are independent and identically distributed random vectors with mean zero and covariance matrix unity. Let $\tilde{u}$ have $u_k(t)$ in the $(t-1)p+k$-th place and $\Gamma^{(1)} = \mathrm{Cov}[\tilde{u}'\tilde{u}]$. Then, we obtain

$$
\begin{aligned}
&(D_N{}^{-1} \otimes I_p)(X' \otimes I_p)\Gamma^{(1)-1}(X \otimes I_p)(D_N{}^{-1} \otimes I_p) \\
&\quad = \sum_{t=M+1}^{N} \sum_{j_1, j_2=1}^{M} \{D_N{}^{-1} x(t-j_1) x(t-j_2)' D_N{}^{-1}\} \otimes C_1(j_1)' C_1(j_2) + R_N,
\end{aligned}
$$

$$= \sum_{j_1, j_2=1}^{M} \{R(j_1 - j_2) + O(M/N)\} \otimes C_1(j_1)' C_1(j_2) + R_N,$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} dM(\lambda) \otimes F_1(-\lambda)^{-1} + R_N + O(M/N),$$

where

$$\|R_N\| = O(1/N) \sum_{j_1, j_2=1}^{M} \|\Gamma(j_1 - j_2)\| = O(M/N).$$

Reversing the order of the indices of the tensors we obtain the result. □

Acknowledgements

The author would like to express his deepest gratitude to his supervisor Professor Masanobu Taniguchi of Waseda University for his constant encouragement, advice and guidance, and leading him to study of time series analysis. He also has no words with which to thank Professor Takeru Suzuki of Waseda University who supervised him during his bachelor, master course and doctoral study, for leading him to study of fundamental statistical inference.

The author would like to express his thanks to Associate Professor Kiyoshi Inoue of Waseda University for his helpful comments and useful critical readings of the manuscript as the members of the doctoral committee.

The author would like to express sincere thanks to the students of Professor Taniguchi and Professor Suzuki laboratories who worked with him, for their continuing discussions, comments and helps in studies. Moreover, thanks are extended to all the other staffs of the Department of Mathematical Science, Waseda University for their various helps in all respects.

Finally, the author thanks with his whole heart to his parents and sister for their successive and cordial supports

Bibliography

[1] AKAHIRA, M. and TAKEUCHI, K. (1981). *Asymptotic Efficiency of Statistical Estimators, Concepts and Higher Order Asymptotic Efficiency*. Lecture Notes in Statistics, 7. Springer-Verlag, New York.

[2] ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series.* John Wiley, New York.

[3] BICKEL, P. J. and GHOSH, J. K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction—a Bayesian argument. *Ann. Statist.* **18** 1070–1090.

[4] BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *J. Political Economy* **81** 637–654.

[5] BRILLINGER, D. R. (1981). *Time Series, Data Analysis and Theory.* Holden-Day, San Francisco.

[6] BRILLINGER, D. R. and ROSENBLATT, M. (1967). Asymptotic theory of estimates of k-th order spectra. *Spectral Analysis Time Series (Proc. Advanced Sem., Madison, Wis., 1966)* pp. 153–188 John Wiley, New York.

[7] BRILLINGER, D. R. and ROSENBLATT, M. (1967). Computation and interpretation of k-th order spectra. *Spectral Analysis Time Series (Proc. Advanced Sem., Madison, Wis., 1966)* pp. 189–232 John Wiley, New York.

[8] CORDEIRO, G. M. and FERRARI, S. L. P. (1991). A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika* **78** 573–582.

[9] CORRADO, C. J. and SU, T. (1996). S&P 500 index option tests of Jarrow and Rudd's approximate option valuation formula. *J. Futures Markets* **16** 611–629.

[10] CORRADO, C. J. and SU, T. (1996). Skewness and kurtosis in S&P 500 index returns implied by option prices. *J. Financial Research* **19** 175–192.

[11] CORRADO, C. J. and SU, T. (1997). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *European J. Finance* **3** 73–85.

[12] COX, J. C. and ROSS, S. A. (1976). The valuation of options for alternative stochastic processes. *J. Financial Economics* **3** 145–166.

[13] DAHLHAUS, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.* **62** 139–168.

[14] DAHLHAUS, R. (1996). Maximum likelihood estimation and model selection for locally stationary processes. *J. Nonparametr. Statist.* **6** 171–191.

[15] DAHLHAUS, R. (1996). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. *Athens Conference on Applied Probability and Time Series Analysis, Vol. II (1995)* pp. 145–159, Lecture Notes in Statist., 115, Springer, New York.

[16] DAHLHAUS, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28** 1762–1794.

[17] EGUCHI, S. (1991). A geometric look at nuisance parameter effect of local powers in testing hypothesis. *Ann. Inst. Statist. Math.* **43** 245–260.

[18] HANNAN, E. J. (1963). Regression for time series. *Proc. Sympos. Time Series Analysis* pp. 17–37 John Wiley, New York.

[19] HANNAN, E. J. (1970). *Multiple Time Series.* John Wiley, New York.

[20] HARRIS, P. and PEERS, H. W. (1980). The local power of the efficient scores test statistic. *Biometrika* **67** 525–529.

[21] HAYAKAWA, T. (1975). The likelihood ratio criterion for a composite hypothesis under a local alternative. *Biometrika* **62** 451–460.

[22] HOSOYA, Y. (1979). High-order efficiency in the estimation of linear processes. *Ann. Statist.* **7** 516–530.

[23] JARROW, R. and RUDD, A. (1982). Approximate option valuation for arbitrary stochastic processes. *J. Financial Economics* **10** 347–369.

[24] JURCZENKO, E., MAILLET, B. and NEGREA, B. (2002). Multi-moment approximate option pricing models: a general comparison (Part 1). University of Paris I Panthéon-Sorbonne.

[25] KAKIZAWA, Y. (1997). Higher-order Bartlett-type adjustment. *J. Statist. Plann. Inference* **65** 269–280.

[26] KARIYA, T. (1993). *Quantitative Methods for Portfolio Analysis.* Kluwer Academic, Dordrecht.

[27] KARIYA, T. and LIU, R. Y. (2003). *Asset Pricing, Discrete Time Approach.* Kluwer Academic, Boston.

[28] LI, B. (2001). Sensitivity of Rao's score test, the Wald test and the likelihood ratio test to nuisance parameters. *J. Statist. Plann. Inference* **97** 57–66.

[29] LINTON, O. (1995). Second order approximation in the partially linear regression model. *Econometrica* **63** 1079–1112.

[30] LINTON, O. and XIAO, Z. (2001). Second-order approximation for adaptive regression estimators. *Econometric Theory* **17** 984–1024.

[31] LONGSTAFF, F. A. (1995). Option pricing and the martingale restriction. *Review of Financial Studies* **8** 1091–1124.

[32] MUKERJEE, R. (1993). An extension of the conditional likelihood ratio test to the general multiparameter case. *Ann. Inst. Statist. Math.* **45** 759–771.

[33] MUKERJEE, R. (1994). Comparison of tests in their original forms. *Sankhyā Ser. A* **56** 118–127.

[34] RAO, C. R. and MUKERJEE, R. (1995). Comparison of Bartlett-type adjustments for the efficient score statistic. *J. Statist. Plann. Inference* **46** 137–146.

[35] RAO, C. R. and MUKERJEE, R. (1997). Comparison of LR, score and Wald tests in a non-IID setting. *J. Multivariate Anal.* **60** 99–110.

[36] ROBINSON, P. M. (1991). Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* **59** 1329–1363.

[37] ROTHENBERG, T. J. (1984). Approximate normality of generalized least squares estimates. *Econometrica* **52** 811–825.

[38] RUBINSTEIN, M. (1998). Edgeworth binomial trees. *J. Derivatives* **5** 20–27.

[39] TAMAKI, K. (2005). Second Order Asymptotic Properties of a Class of Test Statistics under the Existence of Nuisance Parameters. *Sci. Math. Jpn.* **61** 119–143.

[40] TAMAKI, K. and TANIGUCHI, M. (2005). Higher Order Asymptotic Option Valuation for Non-Gaussian Dependent Returns. *To appear in J. Statist. Plann. Inference*

[41] TAMAKI, K. (2005). Second order optimality for estimators in time series regression models. *Submitted for publication*.

[42] TAMAKI, K. (2005). Second order properties of locally stationary processes. *Submitted for publication*.

[43] TANIGUCHI, M. (1983) On the second order asymptotic efficiency of estimators of Gaussian ARMA processes. *Ann. Statist.* **11** 157–169.

[44] TANIGUCHI, M. (1988). Asymptotic expansion of the distribution of some test statistics for Gaussian ARMA processes. *J. Multivariate Anal.* **27** 494–511.

[45] TANIGUCHI, M. (1991). *Higher Order Asymptotic Theory for Time Series Analysis.* Lecture Notes in Statist., 68, Springer-Verlag, New York.

[46] TANIGUCHI, M. (1991). Third-order asymptotic properties of a class of test statistics under a local alternative. *J. Multivariate Anal.* **37** 223–238.

[47] TANIGUCHI, M., PURI, M. L. and KONDO, M. (1996). Nonparametric approach for non-Gaussian vector stationary processes. *J. Multivariate Anal.* **56** 259–283.

[48] TANIGUCHI, M. and KAKIZAWA, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Springer-Verlag, New York.

[49] TANIGUCHI, M., VAN GARDEREN, K. J. and PURI, M. L. (2003). Higher order asymptotic theory for minimum contrast estimators of spectral parameters of stationary processes. *Econometric Theory* **19** 984–1007.

[50] TOYOOKA, Y. (1985). Second-order risk comparison of SLSE with GLSE and MLE in a regression with serial correlation. *J. Multivariate Anal.* **17** 107–126.

[51] TOYOOKA, Y. (1986). Second-order risk structure of GLSE and MLE in a regression with a linear process. *Ann. Statist.* **14** 1214–1225.

[52] VELASCO, C. and ROBINSON, P. M. (2001). Edgeworth expansions for spectral density estimates and Studentized sample mean. *Econometric Theory* **17** 497–539.

[53] WIENER, N. (1933). *The Fourier Integral and Certain of Its Application.* Cambridge University Press, Cambridge.

[54] XIAO, Z. and PHILLIPS, P. C. B. (1998). Higher-order approximations for frequency domain time series regression. *J. Econometrics* **86** 297–336.

[55] XIAO, Z. and PHILLIPS, P. C. B. (2002). Higher order approximations for Wald statistics in time series regressions with integrated processes. *J. Econometrics* **108** 157–198.

List of Papers

1. TAMAKI, K. (2005). Second Order Asymptotic Properties of a Class of Test Statistics under the Existence of Nuisance Parameters. *Sci. Math. Jpn.* **61** 119–143.

2. TAMAKI, K. and TANIGUCHI, M. (2005). Higher Order Asymptotic Option Valuation for Non-Gaussian Dependent Returns. *To appear in J. Statist. Plann. Inference*.