SECOND ORDER ASYMPTOTIC PROPERTIES OF A CLASS OF TEST STATISTICS UNDER THE EXISTENCE OF NUISANCE PARAMETERS

Kenichiro TAMAKI*

Received May 27 2004; revised June 10, 2004

Abstract. Under the existence of nuisance parameters, we consider a class of tests $\mathscr{S}$ 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 $T \in \mathscr{S}$ under a sequence of local alternatives. This result and concrete examples illuminate some interesting features of effects due to nuisance parameters. Optimum properties for a modified likelihood ratio test proposed in Mukerjee [8] are shown under the criteria of second order local maximinity.

1. Introduction. In multivariate analysis, the second order asymptotic powers of various test statistics have been investigated by Hayakawa [5], and Harris and Peers [4]. 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 [9]). Under the existence of nuisance parameters, Eguchi [3] studied the effect of the composite null hypothesis from a geometric point of view. Mukerjee [8] suggested a test that is superior to the usual LR test with regard to second order local maximinity. The test proposed in Mukerjee [8] 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 [12] considered the problem of second order comparison of tests. He worked with a large class of tests that contains LR, R and Wald's (W) tests as special cases. Taniguchi [13] 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 [14] 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-type adjustment procedure has been elucidated in various directions. Cordeiro and Ferrari [2] 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 [6] considered the extension of Cordeiro and Ferrari's [2] adjustment to the case of order n^{-k}, where k is an integer $k \geq 1$. Rao and Mukerjee [10] compared various Bartlett-type adjustments for the R statistic. Rao and Mukerjee [11] 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 paper, under the existence of nuisance parameters, we consider the second order properties of a class of tests $\mathscr{S}$ 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 $T \in \mathscr{S}$ are equally

2000 *Mathematics Subject Classification.* 62F05; 62H15; (62N99).

Key words and phrases. local maximinity; local power; local unbiasedness; nuisance parameters; second order.

sensitive to perturbation of the nuisance parameter. In Section 3 we compare the second order local power. It is seen that the local average powers of all $T \in \mathscr{S}$ 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 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 5, we discuss the local unbiasedness of $T \in \mathscr{S}$. The results and their examples illuminate some interesting features of effects due to nuisance parameters. The proofs of theorems are relegated to Section 6.

2. Asymptotic expansion of a class of tests. Let $\mathbf{X}_n = (X_1, \ldots, X_n)$ be a collection of m-dimensional random vectors forming a stochastic process. Let $p_n(\boldsymbol{x}_n; \theta)$, $\boldsymbol{x}_n \in \mathbf{R}^{nm}$, be the probability density function of $\mathbf{X}_n$, 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 $\mathscr{S}$ which contains LR, W and R tests as special cases. In the presence of nuisance parameters, the powers and sizes of $T \in \mathscr{S}$ 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 [7], 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:

(A-1) $l_n(\theta) = \log p_n(\mathbf{X}_n; \theta)$ is continuously four times differentiable with respect to θ.

(A-2) The partial derivative $\partial_\alpha = \partial/\partial\theta^\alpha$ and the expectation E_θ with respect to $p_n(\boldsymbol{x}_n; \theta)$ are interchangeable.

(A-3) For an appropriate sequence $\{c_n\}$ satisfying $c_n \to +\infty$ as $n \to +\infty$, the asymptotic moments (cumulants) of

$$\begin{aligned} Z_\alpha(\theta) &= c_n^{-1}\partial_\alpha l_n(\theta), \\ Z_{\alpha\beta}(\theta) &= c_n^{-1}[\partial_\alpha\partial_\beta l_n(\theta) - E_\theta\{\partial_\alpha\partial_\beta l_n(\theta)\}], \end{aligned}$$

possess the following asymptotic expansions

$$\begin{aligned} E_\theta\{Z_\alpha(\theta)Z_\beta(\theta)\} &= I_{(\alpha\beta)}(\theta) + O(c_n^{-2}), \\ E_\theta\{Z_\alpha(\theta)Z_{\beta\gamma}(\theta)\} &= J_{\alpha,\beta\gamma}(\theta) + O(c_n^{-2}), \\ E_\theta\{Z_\alpha(\theta)Z_\beta(\theta)Z_\gamma(\theta)\} &= c_n^{-1}K_{\alpha,\beta,\gamma}(\theta) + O(c_n^{-3}), \end{aligned}$$

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

(A-4) (i) $I_{(\alpha\beta)}(\theta)$ is continuously two times differentiable with respect to θ.

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

(A-5) (i) $I(\theta) = \{I_{(\alpha\beta)}(\theta)\}$ is positive definite for all $\theta \in \Theta$.

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

Let $\hat{\theta} = (\hat{\theta}^1, \dots, \hat{\theta}^{p+q})'$ be the global maximum likelihood estimator of θ and $\tilde{\theta}_2 = (\tilde{\theta}^{p+1}, \dots, \tilde{\theta}^{p+q})'$ 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

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

Let

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

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:

$$\begin{aligned}
\mathscr{S} = \{T \mid T &= g^{ij}W_iW_j + c_n^{-1}a_1g^{i\alpha}g^{j\beta}W_{\alpha\beta}W_iW_j + 2c_n^{-1}g^{i\alpha}g^{rs}W_{\alpha r}W_iW_s \\
&+ c_n^{-1}a_2^{ijk}W_iW_jW_k - c_n^{-1}g^{i\alpha}g^{j\beta}g^{rs}K_{\alpha,\beta,r}W_iW_jW_s \\
&- c_n^{-1}g^{i\alpha}g^{rt}g^{su}(K_{\alpha,r,s} + J_{\alpha,rs})W_iW_tW_u + c_n^{-1}a_3^iW_i + o_p(c_n^{-1}), \\
&\text{under } H, \text{ where } a_1, a_2^{ijk} \text{ and } a_3^i \text{ are nonrandom constants}\}.
\end{aligned} \tag{1}$$

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

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

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

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), we obtain

$$2(\hat{l}_n - \tilde{l}_n) = \hat{g}_{ij}\eta^i\eta^j - c_n^{-1}\hat{Z}_{\alpha\beta}\eta^\alpha\eta^\beta - c_n^{-1}\left(\frac{1}{3}\hat{K}_{\alpha,\beta,\gamma} + \hat{J}_{\alpha,\beta\gamma}\right)\eta^\alpha\eta^\beta\eta^\gamma + o_p(c_n^{-1}). \tag{3}$$

By Taylor expansion around θ_0,

$$\hat{g}_{ij} = g_{ij} + g_{ik}g^{k\alpha}g_{jl}g^{l\beta}(K_{\alpha,\beta,\gamma} + J_{\alpha,\beta\gamma} + J_{\beta,\gamma\alpha})(\hat{\theta}^{\gamma} - \theta_0^{\gamma}) + o_p(c_n^{-1}). \tag{4}$$

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

$$\begin{aligned} c_n^{-1}(\hat{\theta}^{\alpha} - \theta_0^{\alpha}) &= g^{\beta\alpha}W_{\beta} + c_n^{-1}I^{\alpha\beta}g^{\delta\gamma}W_{\beta\gamma}W_{\delta} \\ &\quad - \frac{1}{2}c_n^{-1}I^{\alpha\alpha'}g^{\beta\beta'}g^{\gamma\gamma'}(K_{\alpha',\beta',\gamma'} + J_{\alpha',\beta'\gamma'})W_{\beta}W_{\gamma} + o_p(c_n^{-1}). \end{aligned} \tag{5}$$

Inserting (4) and (5) in (3) and noting $\hat{Z}_{\alpha\beta} = W_{\alpha\beta} + o_p(1)$, we have

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

Hence, LR belongs to $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathscr{S}$. 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 $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathscr{S}$ 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 $\mathrm{LR}^* = \mathrm{LR} + c_n^{-1}\tilde{g}^{i\alpha}\tilde{g}^{rs}(\tilde{K}_{\alpha,r,s} + \tilde{J}_{\alpha,rs})\tilde{Z}_i$ proposed in Mukerjee [8] belongs to $\mathscr{S}$ with the coefficients $a_1 = 1$, $a_2^{ijk} = -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha,\beta,\gamma}/3$ and $a_3^i = g^{i\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs})$.

Li [7] compared the sensitivities of LR, W_2 and R_2 statistics to nuisance parameters. In the one-parameter case, Taniguchi [14] discussed the third order asymptotic properties of a class of tests $\mathscr{S}_1$. Rao and Mukerjee [11] studied a wider class $\mathscr{S}_2(\supset \mathscr{S}_1)$ which enables us to compare the various Bartlett-type adjustments available for the members of $\mathscr{S}_1$. Our class $\mathscr{S}$ contains $\mathscr{S}_1$ and $\mathscr{S}_2$, hence the class $\mathscr{S}$ is sufficiently rich.

Remark 1. Test statistics in Example 1 are based on the maximum likelihood estimator. From (2) and (3), these statistics can be written as

$$\begin{aligned} T &= \hat{g}_{ij}\eta^i\eta^j + c_n^{-1}b_1\hat{Z}_{\alpha\beta}\eta^\alpha\eta^\beta + c_n^{-1}(b_2\hat{K}_{\alpha,\beta,\gamma} + b_3\hat{J}_{\alpha,\beta\gamma})\eta^\alpha\eta^\beta\eta^\gamma \\ &\quad + c_n^{-1}b_4\hat{g}^{rs}(\hat{K}_{\alpha,r,s} + \hat{J}_{\alpha,rs})\eta^\alpha + o_p(c_n^{-1}), \end{aligned} \tag{6}$$

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

$$\begin{array}{lllll} b_1 = -1, & b_2 = -1/3, & b_3 = -1, & b_4 = 0, & \text{for LR}, \\ b_1 = -1, & b_2 = -1/3, & b_3 = -1, & b_4 = 1, & \text{for LR}^*, \\ b_1 = 0, & b_2 = 0, & b_3 = 0, & b_4 = 0, & \text{for W}_1, \\ b_1 = 0, & b_2 = -1, & b_3 = -2, & b_4 = 0, & \text{for W}_2, \\ b_1 = -1, & b_2 = 0, & b_3 = 0, & b_4 = 0, & \text{for W}_3, \\ b_1 = -1, & b_2 = -1, & b_3 = -3, & b_4 = 0, & \text{for W}_4, \\ b_1 = -2, & b_2 = -1, & b_3 = -3, & b_4 = 0, & \text{for R}_1, \\ b_1 = -2, & b_2 = 0, & b_3 = -1, & b_4 = 0, & \text{for R}_2, \\ b_1 = -1, & b_2 = -1, & b_3 = -3, & b_4 = 0, & \text{for R}_3, \\ b_1 = -1, & b_2 = 0, & b_3 = 0, & b_4 = 0, & \text{for R}_4. \end{array} \tag{7}$$

Inserting (4) and (5) in (6), we obtain

$$\begin{aligned} T &= g^{ij}W_iW_j + c_n^{-1}(b_1+2)g^{i\alpha}g^{j\beta}W_{\alpha\beta}W_iW_j + 2c_n^{-1}g^{i\alpha}g^{rs}W_{\alpha r}W_iW_s \\ &\quad + c_n^{-1}g^{i\alpha}g^{j\beta}g^{k\gamma}\{b_2K_{\alpha,\beta,\gamma} + (b_3+1)J_{\alpha,\beta\gamma}\}W_iW_jW_k \\ &\quad - c_n^{-1}g^{i\alpha}g^{j\beta}g^{rs}K_{\alpha,\beta,r}W_iW_jW_s \\ &\quad - c_n^{-1}g^{i\alpha}g^{rt}g^{su}(K_{\alpha,r,s} + J_{\alpha,rs})W_iW_tW_u \\ &\quad + c_n^{-1}b_4g^{i\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs})W_i + o_p(c_n^{-1}). \end{aligned} \tag{8}$$

The class $\mathscr{S}$ in (1) is motivated from (8).

First, we give the second-order asymptotic expansion of the distribution function of $T \in \mathscr{S}$ 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 1. *The distribution function of $T \in \mathscr{S}$ under a sequence of local alternatives $\theta = \theta_0 + c_n^{-1}\varepsilon$ has the asymptotic expansion*

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

where

$$
\begin{aligned}
m_3 &= \frac{1}{6}K_{\alpha,\beta,\gamma}d^\alpha d^\beta d^\gamma + \frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'},\\
m_2 &= -\frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'} + \frac{1}{2}B^{\alpha\beta}K_{\alpha,\beta,\gamma}d^\gamma + \frac{1}{2}a_2^{ijk}[3]g_{il}g_{jk}d^l,\\
m_1 &= \frac{1}{2}J_{\alpha,\beta\gamma}d^\alpha d^\beta d^\gamma - \frac{1}{2}(K_{\alpha,\beta,r} + J_{\alpha,\beta r} + J_{\beta,\alpha r})d^\alpha d^\beta (d^r - \varepsilon^r)\\
&\quad - \frac{1}{2}B^{\alpha\beta}K_{\alpha,\beta,\gamma}d^\gamma - \frac{1}{2}a_2^{ijk}[3]g_{il}g_{jk}d^l\\
&\quad - \frac{1}{2}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs})d^\alpha + \frac{1}{2}a_3^i g_{ij}d^j,\\
m_0 &= -\frac{1}{6}(K_{\alpha,\beta,\gamma} + 3J_{\alpha,\beta\gamma})d^\alpha d^\beta d^\gamma\\
&\quad + \frac{1}{2}(K_{\alpha,\beta,r} + J_{\alpha,\beta r} + J_{\beta,\alpha r})d^\alpha d^\beta (d^r - \varepsilon^r)\\
&\quad + \frac{1}{2}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs})d^\alpha - \frac{1}{2}a_3^i g_{ij}d^j,
\end{aligned} \tag{9}
$$

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

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

Second, we consider the sensitivity of $T \in \mathscr{S}$ 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. (i) *For $T \in \mathscr{S}$, 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*

$$g_{ii'}g^{i'\alpha}g_{jj'}g^{j'\alpha}(K_{\alpha,\beta,r} + J_{\alpha,\beta r} + J_{\beta,\alpha r}) = 0, \tag{10}$$

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

Remark 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 (10) holds.

Remark 3. In the case of i.i.d. observations, Li [7] 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, we can see that the powers and sizes of all $T \in \mathscr{S}$ are equally sensitive to nuisance parameters. Hence, our results agree with that of Li [7].

Example 2. Suppose that X_i, $i = 1, \dots, n$ 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 (10) holds.

(ii) If $\theta_1 = \mu$ and $\theta_2 = \sigma^2$, then $g_{11}(\mu, \sigma^2) = (\sigma^2)^{-1}$. Hence, the condition (10) does not hold.

Example 3. Consider the nonlinear regression model

$$X_t = \alpha + \beta \cos(t-1)\lambda + u_t, \qquad t = 1, \dots, n, \tag{11}$$

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

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

For our model (11) we calculate $g_{11}(\theta)$. From (12)

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

which implies that the condition (10) does not hold.

3. Comparison of power. Taking $\varepsilon_1 = \mathbf{0}$ in (9), it can be seen that all $T \in \mathscr{S}$ have sizes $\alpha + o(c_n^{-1})$. Hence, it would be meaningful to compare $T \in \mathscr{S}$ in terms of power up to $o(c_n^{-1})$. From Theorem 1, we can see that there is no test which is second order uniformly most powerful in $\mathscr{S}$. Thus we attempt to compare the tests in $\mathscr{S}$ on the basis of their second order power. First, we derive the explicit formula to compare the local power of $T \in \mathscr{S}$. Note that the first order powers of all $T \in \mathscr{S}$ are identical and independent of ε_2. Write the power function of $T \in \mathscr{S}$ 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 1, we can state

Theorem 3. *For T_1 and $T_2 \in \mathscr{S}$ with the coefficient $(a_{11}, a_{21}^{ijk}, a_{31}^i)$ and $(a_{12}, a_{22}^{ijk}, a_{32}^i)$, 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

$$m'_2 = \frac{1}{2}(a_{21}^{ijk} - a_{22}^{ijk}) g_{ii'} g_{jj'} g_{kk'} d^{i'} d^{j'} d^{k'},$$
$$m'_1 = \frac{1}{2}(a_{21}^{ijk}[3] - a_{22}^{ijk}[3]) g_{il} g_{jk} d^l,$$
$$m'_0 = \frac{1}{2}(a_{31}^i - a_{32}^i) g_{ij} d^j.$$

Note that m'_2, m'_1 and m'_0 are independent of ε_2. From Theorem 3 we have

Corollary 1. *For T_1 and $T_2 \in \mathscr{S}$ with the coefficient $(a_{11}, a_{21}^{ijk}, a_{31}^{i})$ and $(a_{12}, a_{22}^{ijk}, a_{32}^{i})$, 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 3.

Example 4. Suppose that X_i, $i = 1, \dots, n$ 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

$$\begin{aligned} &g_{11}(\rho,\mu) = \frac{1+\rho^2}{(1-\rho^2)^2}, \quad K_{1,1,1}(\rho,\mu) = -\frac{6\rho + 2\rho^3}{(1-\rho^2)^3}, \quad J_{1,11}(\rho,\mu) = -K_{1,1,1}(\rho,\mu), \\ &g_{22}(\rho,\mu) = \frac{2}{1+\rho}, \quad K_{1,2,2}(\rho,\mu) = \frac{2}{(1+\rho)^2}, \quad J_{1,22}(\rho,\mu) = 0. \end{aligned} \tag{13}$$

For test statistics T_1 and T_2 in (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 $(i = 1, 2, 3, 4)$, R_i $(i = 1, 2, 3, 4)$, LR and LR*.

(i) (i)-1 If $\rho > 0$ and $\varepsilon_1 > 0$, then

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

(i)-2 If $\rho < 0$ and $\varepsilon_1 > 0$, then

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

(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 (13) in Example 4, it is seen that cumulants g_{11}, $K_{1,1,1}$ and $J_{1,11}$ tend to ∞ as $\rho \to \pm 1$, and g_{22} and $K_{1,2,2}$ tend to ∞ as $\rho \to -1$. Hence, we need to inspect second order power functions if ρ is close to ± 1. Note the relation

$$G_{p,\Delta}(z) - G_{p+2,\Delta}(z) = 2f_{p+2,\Delta}(z), \tag{14}$$

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 (13) and (14) it follows that second order powers of all test statistics in Example 1 converge to 0 as $\rho \to \pm 1$ at each fixed ε_1.

In Figure 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 4 with $\alpha = 0.05$, $\varepsilon_1 = 1$ and $-1 < \rho < 1$. Figure 1 illustrates that second order powers of these statistics converge to 0 as $\rho \to \pm 1$.

In Figure 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 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 1.

Figures 1 and 2 are about here.

Example 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,

$$\begin{aligned} & g_{11}(\psi, \sigma^2) = \frac{1}{1-\psi^2}, \quad K_{1,1,1}(\psi, \sigma^2) = -\frac{6\psi}{(1-\psi^2)^2}, \quad J_{1,11}(\psi, \sigma^2) = \frac{4\psi}{(1-\psi^2)^2}, \\ & g_{22}(\psi, \sigma^2) = \frac{1}{2\sigma^4}, \quad K_{1,2,2}(\psi, \sigma^2) = J_{1,22}(\psi, \sigma^2) = 0. \end{aligned} \tag{15}$$

Note that $g^{22}(K_{1,2,2} + J_{1,22}) = 0$. For test statistics T_1 and T_2 in (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 $(i = 1,2,3,4)$, 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 (15) in Example 5, it is seen that cumulants g_{11}, $K_{1,1,1}$ 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 (14) and (15) it follows that second order powers of all test statistics in Example 1 converge to 0 as $\psi \to \pm 1$ at each fixed ε_1.

In Figure 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 5 with $\alpha = 0.01$, $\varepsilon_1 = 6.5$ and $-1 < \psi < 1$. From Figure 3 we observe that second order powers of these statistics converge to 0 as $\psi \to \pm 1$.

In Figure 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 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 3.

Figures 3 and 4 are about here.

Example 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

$$\begin{aligned} &g_{11}(\rho, \sigma^2) = \frac{1}{1-\rho^2}, \quad K_{1,1,1}(\rho, \sigma^2) = \frac{6\rho}{(1-\rho^2)^2}, \quad J_{1,11}(\rho, \sigma^2) = -\frac{2\rho}{(1-\rho^2)^2}, \\ &g_{22}(\rho, \sigma^2) = \frac{1}{2\sigma^4}, \quad K_{1,2,2}(\rho, \sigma^2) = J_{1,22}(\rho, \sigma^2) = 0. \end{aligned} \tag{16}$$

Note that $g^{22}(K_{1,2,2} + J_{1,22}) = 0$. For test statistics T_1 and T_2 in (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 $(i = 1, 2, 3, 4)$, 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 (16) in Example 6, it is seen that cumulants g_{11}, $K_{1,1,1}$ 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 (14) and (16) it follows that second order powers of all test statistics in Example 1 converge to 0 as $\rho \to \pm 1$ at each fixed ε_1.

In Figure 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 6 with $\alpha = 0.01$, $\varepsilon_1 = 3$ and $-1 < \rho < 1$. From Figure 5 it is seen that second order powers of these statistics converge to 0 as $\rho \to \pm 1$.

In Figure 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 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 5.

Figures 5 and 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 1 it is easily seen that for each $T \in \mathscr{S}$,

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

It is, therefore, clear that the average powers of all $T \in \mathscr{S}$ are identical up to c_n^{-1}. However, even in this situation, with a more detailed analysis it is possible to compare tests in $\mathscr{S}$ in a meaningful way under suitable choice of criterion. Under the absence of nuisance parameters, Mukerjee [9] 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 4. *For $T \in \mathscr{S}$ whose coefficients do not satisfy $z(a_2^{ijk}[3]g_{jk} + g^{i\alpha}B^{\beta\gamma}K_{\alpha,\beta,\gamma}) + (p+2)\{a_3^i - g^{i\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs})\} = 0$ (the coefficients of LR* satisfy $a_2^{ijk}[3]g_{jk} = -g^{i\alpha}B^{\beta\gamma}K_{\alpha,\beta,\gamma}$ and $a_3^i = g^{i\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs})$), there exists a positive Δ_0 such that*

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

whenever $0 < \Delta < \Delta_0$.

Example 7. (i) In Example 4, W_1, W_3 and R_4 are most powerful in Example 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* tests in terms of second-order local maximinity. Note that the condition (10) holds. From Theorem 1 and Example 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 $0 < \Delta \leq 1$.

(ii) If $\rho < 0$ and $\varepsilon_1 > 0$, then R_1, R_3 and W_4 are most powerful in Example 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 $\rho = -1/2$, $\Delta \leq 1$ 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 $0 < \Delta \leq 1$.

Example 8. (i) In Example 5, W_1, W_3 and R_4 are most powerful in Example 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 5, the condition (10) holds. From Theorem 1, we obtain

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

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

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

Hence,

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

whenever $0 < \Delta \leq 1$.

(ii) If $\psi < 0$ and $\varepsilon_1 > 0$, then R_1, R_3 and W_4 are most powerful in Example 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 1, we get

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

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

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

Hence,

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

whenever $0 < \Delta \le 1$.

4. Effect of nuisance parameters. In this section, we consider the case where the nuisance parameter $\theta_2 = \theta_{20}$ is known. Let $\bar{\theta}_1 = (\bar{\theta}^1, \dots, \bar{\theta}^p)'$ 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 1 are given by

$$\begin{aligned} &\mathrm{LR}_0 = \mathrm{LR}_0^* = 2(\bar{l}_n - l_n), \\ &\mathrm{W}_{10} = \bar{I}_{(ij)}\tau^i\tau^j, \quad \mathrm{W}_{20} = I_{(ij)}\tau^i\tau^j, \quad \mathrm{W}_{30} = \bar{L}_{(ij)}\tau^i\tau^j, \quad \mathrm{W}_{40} = L_{(ij)}\tau^i\tau^j, \\ &\mathrm{R}_{10} = \bar{I}_0^{ij}Z_iZ_j, \quad \mathrm{R}_{20} = I_0^{ij}Z_iZ_j, \quad \mathrm{R}_{30} = \bar{L}_0^{ij}Z_iZ_j, \quad \mathrm{R}_{40} = L_0^{ij}Z_iZ_j, \end{aligned} \tag{17}$$

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 (17) are given by

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

where the coefficient (b_1, b_2, b_3) is the same as in (7) and $W'_{ij} = Z_{ij} - J_{k,ij}I_0^{kl}Z_l$. Hence, we consider the following class of tests:

$$\begin{aligned} \mathscr{S}_0 = \{T_0 \mid T_0 = {}& I_0^{ij}Z_iZ_j + c_n^{-1}a_1I_0^{ik}I_0^{jl}W'_{kl}Z_iZ_j \\ & + c_n^{-1}a_2^{ijk}Z_iZ_jZ_k + o_p(c_n^{-1}), \\ & \text{under } H, \text{ where } a_1 \text{ and } a_2^{ijk} \text{ are nonrandom constants}\}. \end{aligned}$$

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

(A-6) $I_{ir} = 0 \quad i = 1, \dots, p,\ r = p+1, \dots, p+q.$

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

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

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 5. (i) *Under* (A-6), *for* $T \in \mathscr{S}$ *and* $T_0 \in \mathscr{S}_0$ *with the same coefficient, the distribution functions of* T *are decomposed into*

(18)
$$\begin{aligned}P_{\theta_0 + c_n^{-1}\varepsilon}[T < z] = {} & P_{\theta_{10} + c_n^{-1}\varepsilon_1, \theta_{20}}[T_0 < z] \\ & + \frac{1}{2} c_n^{-1} (K_{i,j,r} + J_{i,jr} + J_{j,ir}) \varepsilon^i \varepsilon^j \varepsilon^r \{G_{p+2,\Delta}(z) - G_{p,\Delta}(z)\} \\ & + \frac{1}{2} c_n^{-1} \{I_{(ij)} a_3^j - g^{rs}(K_{i,r,s} + J_{i,rs})\} \varepsilon^i \{G_{p+2,\Delta}(z) - G_{p,\Delta}(z)\} + o(c_n^{-1}).\end{aligned}$$

(ii) *If*

$$K_{i,j,r} + J_{i,jr} + J_{j,ir} = 0, \tag{19}$$
$$g^{rs}(K_{i,r,s} + J_{i,rs}) = 0, \tag{20}$$

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

Remark 4. The condition (19) agree with (10) in Theorem 2 under (A-6). If the condition (20) holds, then LR test is second order asymptotically unbiased. Therefore, the third term of the right hand in (18) can be interpreted as second order local bias in the usual likelihood ratio test (see Mukerjee [8]). 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 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 $g^{22}(K_{1,2,2} + J_{1,22}) = 0$. Hence, the conditions (19) and (20) 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 $g^{22}(K_{1,2,2} + J_{1,22}) = (\theta_1)^{-1}$. Hence, the condition (19) holds and (20) does not hold.

Example 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,

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

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

$$K_{1,3,3}(\sigma^2, \rho, \psi) = \frac{2\sigma^{-2}}{1-\psi^2}, \quad J_{1,33}(\sigma^2, \rho, \psi) = -\frac{\sigma^{-2}}{1-\psi^2},$$

$$K_{1,2,3}(\sigma^2, \rho, \psi) = -\frac{2\sigma^{-2}}{1-\rho\psi}, \quad J_{1,23}(\sigma^2, \rho, \psi) = \frac{\sigma^{-2}}{1-\rho\psi}.$$

Hence, the condition (19) hold, and $g^{rs}(K_{1,r,s}+J_{1,rs}) = 2\sigma^{-2}$ shows that the condition (20) does not hold.

(ii) If $\theta_1 = (\rho, \psi)'$ 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 $g^{33}(K_{i,3,3} + J_{i,33}) = 0$. Hence, the conditions (19) and (20) hold.

5. Unbiased test. We discuss the local unbiasedness of $T \in \mathscr{S}$. 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 1, among the test statistics in Example 1, LR* test is the only one which is second order asymptotically unbiased unless $g_{ij}g^{j\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs}) = 0$. If $g_{ij}g^{j\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs}) = 0$, then LR $=$ LR$^* + o_p(c_n^{-1})$. Hence, LR test is locally unbiased. Since $T \in \mathscr{S}$ is not generally unbiased, we consider modification of $T \in \mathscr{S}$ 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 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 $T \in \mathscr{S}$, 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) $A^i = g^{i\alpha}g^{rs}(K_{\alpha,r,s} + J_{\alpha,rs}) - a_3^i$.

For $h(\theta_1)$ and A^i satisfying (i) and (ii), respectively, in Theorem 6, from Theorem 1, we can get the asymptotic expansion of the distribution function of T^*.

Theorem 7. Suppose that $h(\theta_1)$ and A^i satisfy (i) and (ii), respectively, in Theorem 6. Then, for $T \in \mathscr{S}$, 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

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

where

$$
\begin{aligned}
m_3^* &= \frac{1}{6}K_{\alpha,\beta,\gamma}d^\alpha d^\beta d^\gamma - \frac{1}{2(p+2)}g_{ij}B^{\alpha\beta}K_{\alpha,\beta,\gamma}d^\gamma d^i d^j \\
&\quad + \frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'} - \frac{1}{2(p+2)}a_2^{ijk}[3]g_{ii'}g_{jk}g_{j'k'}d^{i'}d^{j'}d^{k'}, \\
m_2^* &= \frac{1}{2(p+2)}g_{ij}B^{\alpha\beta}K_{\alpha,\beta,\gamma}d^\gamma d^i d^j \\
&\quad - \frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'} + \frac{1}{2(p+2)}a_2^{ijk}[3]g_{ii'}g_{jk}g_{j'k'}d^{i'}d^{j'}d^{k'}, \\
m_1^* &= \frac{1}{2}J_{\alpha,\beta\gamma}d^\alpha d^\beta d^\gamma - \frac{1}{2}(K_{\alpha,\beta,r} + J_{\alpha,\beta r} + J_{\beta,\alpha r})d^\alpha d^\beta (d^r - \varepsilon^r), \\
m_0^* &= -\frac{1}{6}(K_{\alpha,\beta,\gamma} + 3J_{\alpha,\beta\gamma})d^\alpha d^\beta d^\gamma \\
&\quad + \frac{1}{2}(K_{\alpha,\beta,r} + J_{\alpha,\beta r} + J_{\beta,\alpha r})d^\alpha d^\beta (d^r - \varepsilon^r).
\end{aligned} \tag{21}
$$

If $p = 1$, then

$$
a_2^{ijk}g_{ii'}g_{jj'}g_{kk'} - \frac{1}{(p+2)}a_2^{ijk}[3]g_{ii'}g_{jk}g_{j'k'} = 0. \tag{22}
$$

In this case we observe that the coefficients m_3^*, m_2^*, m_1^* and m_0^* in (21) are independent of $T \in \mathscr{S}$, 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 11. Consider the $ARMA(1,1)$ model in Example 10 (ii). For test statistics in (7),

$$
\begin{aligned}
a_2^{ijk}g_{i1}g_{j1}g_{k1} &= b_2K_{1,1,1} + (b_3+1)J_{1,11} \\
&= \frac{2\rho}{(1-\rho^2)^2}(3b_2 - b_3 - 1),
\end{aligned} \tag{23}
$$

and

$$
\begin{aligned}
\frac{1}{4}a_2^{ijk}[3]g_{i1}g_{jk}g_{11} &= \frac{3}{4}b_2K_{1,i,j}g^{ij}g_{11} + \frac{1}{4}(b_3+1)J_{1,ij}[3]g^{ij}g_{11} \\
&= \frac{3}{4}b_2\left\{\frac{6\rho}{(1-\rho^2)^2} + \frac{4\psi}{(1-\rho^2)(1-\rho\psi)}\right\} \\
&\quad + \frac{1}{4}(b_3+1)\frac{12\rho^3\psi - 10\rho^2\psi^2 - 8\rho^2 + 4\psi^2 + 2}{(1-\rho^2)^2(1-\rho\psi)(\rho-\psi)}.
\end{aligned} \tag{24}
$$

(23) and (24) show that (22) does not holds.

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

$$
T = g^{ij}T_iT_j + o_p(c_n^{-1}),
$$

where

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

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

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

Similarly, we consider factorizations of $T_0 \in \mathscr{S}_0$ 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 (A-6), A^i in Theorem 6 can be written as

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

Note that the third term of the right hand in (18) 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 $T \in \mathscr{S}$.

6. Proofs. In this section, we give the proofs of theorems.

Proof of Theorem 1. Since the actual calculation procedure is formidable, we give a sketch of the derivation. First, we evaluate the characteristic function of T,

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

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

$$\begin{aligned} \psi_n(\xi,\varepsilon) &= \int \exp\{tT(\boldsymbol{x}_n)\}L_n(\boldsymbol{x}_n)p_n(\boldsymbol{x}_n;\theta_0)d\boldsymbol{x}_n \\ &= E_{\theta_0}[\exp\{tT + \log L_n(\boldsymbol{x}_n)\}]. \end{aligned} \tag{25}$$

We expand $\log L_n(\boldsymbol{x}_n)$ in a Taylor series in $c_n^{-1}\varepsilon$, leading to

(26)
$$\begin{aligned}\log L_n(\boldsymbol{x}_n) &= W_i\varepsilon^i + g_{\alpha r}g^{rs}W_s\varepsilon^\alpha - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta + \frac{1}{2}c_n^{-1}W_{\alpha\beta}\varepsilon^\alpha\varepsilon^\beta \\ &\quad + \frac{1}{2}c_n^{-1}J_{\gamma,\alpha\beta}g^{\delta\gamma}W_\delta\varepsilon^\alpha\varepsilon^\beta - \frac{1}{6}c_n^{-1}(K_{\alpha,\beta,\gamma} + J_{\alpha,\beta\gamma}[3])\varepsilon^\alpha\varepsilon^\beta\varepsilon^\gamma + o_p(c_n^{-1}) \\ &= D_{ij}\varepsilon^iY^j + g_{\alpha r}g^{rs}D_{st}\varepsilon^\alpha Y^t - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta + \frac{1}{2}c_n^{-1}W_{\alpha\beta}\varepsilon^\alpha\varepsilon^\beta \\ &\quad + \frac{1}{2}c_n^{-1}J_{\gamma,\alpha\beta}g^{\delta\gamma}D_{\delta\zeta}\varepsilon^\alpha\varepsilon^\beta Y^\zeta - \frac{1}{6}c_n^{-1}(K_{\alpha,\beta,\gamma} + J_{\alpha,\beta\gamma}[3])\varepsilon^\alpha\varepsilon^\beta\varepsilon^\gamma + o_p(c_n^{-1}).\end{aligned}$$

Inserting (26) in $\exp\{tT + \log L_n(\boldsymbol{x}_n)\}$ we obtain, after further expansion and collection of terms,

(27)
$$\begin{aligned}\exp\{tT + \log L_n(\boldsymbol{x}_n)\} &= \exp\left\{t\sum_{i=1}^{p}(Y^i)^2 + D_{ij}\varepsilon^iY^j + g_{\alpha r}g^{rs}D_{st}\varepsilon^\alpha Y^t - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta\right\} \\ &\quad \times \left\{1 + c_n^{-1}q_1(Y^\alpha, W_{\beta\gamma})\right\} + o_p(c_n^{-1}),\end{aligned}$$

where $q_1(\cdot,\cdot)$ is a polynomial. In view of the assumption (A-3) 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

(28)
$$\begin{aligned}P_{\theta_0}(Y^\alpha < y^\alpha) &= \int_{-\infty}^{y^\alpha} f(y^\alpha)\left\{1 + \frac{1}{6}c_n^{-1}\sum_{\beta,\gamma,\delta=1}^{p+q}C_{\beta\gamma\delta}H_{\beta\gamma\delta}(y^\alpha)\right\}dy^\alpha + o(c_n^{-1}) \\ &= \int_{-\infty}^{y^\alpha} q(y^\alpha)dy^\alpha + o(c_n^{-1}),\end{aligned}$$

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

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

From (25), (27) and (28) it follows that

$$\begin{aligned}\psi_n(\xi,\varepsilon) &= \int \exp\left\{t\sum_{i=1}^{p}(y^i)^2 + D_{ij}\varepsilon^i y^j + g_{\alpha r}g^{rs}D_{st}\varepsilon^\alpha y^t - \frac{1}{2}I_{(\alpha\beta)}\varepsilon^\alpha\varepsilon^\beta\right\} \\ &\quad\times\left\{1 + c_n^{-1}q_1(y^\gamma,0)\right\}q(y^\delta)dy^\zeta + o(c_n^{-1}) \\ &= \exp\left(\frac{tg_{ij}\varepsilon^i\varepsilon^j}{1-2t}\right)(1-2t)^{-p/2}\left\{1 + c_n^{-1}\sum_{j=0}^{3}m_j(1-2t)^{-j}\right\} + o(c_n^{-1}).\end{aligned} \tag{29}$$

Inverting (29) by Fourier inverse transform we can prove Theorem 1. □

Proof of Theorem 2.

(i) Note that d^α is independent of ε_2. From Theorem 1, for $T \in \mathscr{S}$ we have

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

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. □

Proof of Theorem 3 and Corollary 1. From Theorem 1 we can see that

$$\begin{aligned}m_3 &= \frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'} + C_3, \\ m_2 &= -\frac{1}{2}a_2^{ijk}g_{ii'}g_{jj'}g_{kk'}d^{i'}d^{j'}d^{k'} + \frac{1}{2}a_2^{ijk}[3]g_{il}g_{jk}d^l + C_2, \\ m_1 &= -\frac{1}{2}a_2^{ijk}[3]g_{il}g_{jk}d^l + \frac{1}{2}a_3^i g_{ij}d^j + C_1, \\ m_0 &= -\frac{1}{2}a_3^i g_{ij}d^j + C_0,\end{aligned} \tag{30}$$

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 $\mathscr{S}$. Theorem 3 and Corollary 1 follow from (30). □

Proof of Theorem 4. Let a_2^{ijk} and a_3^i be the coefficients of $T \in \mathscr{S}$. Then, we can rewrite

$$\begin{aligned}P_2^T(\varepsilon) &= Q_{1,i'j'k'}(a_2^{ijk})\varepsilon^{i'}\varepsilon^{j'}\varepsilon^{k'} + Q_{2,ijr}\varepsilon^i\varepsilon^j\varepsilon^r \\ &\quad+ \frac{1}{2}g_{li}(g^{i\alpha}B^{\beta\gamma}K_{\alpha,\beta,\gamma} + a_2^{ijk}[3]g_{jk})\varepsilon^l\{G_{p+2,\Delta}(z)-G_{p+4,\Delta}(z)\} \\ &\quad+ \frac{1}{2}g_{ij}\{a_3^i - g^{i\alpha}g^{rs}(K_{\alpha,r,s}+J_{\alpha,rs})\}\varepsilon^j\{G_{p,\Delta}(z)-G_{p+2,\Delta}(z)\}.\end{aligned} \tag{31}$$

Note that $|\varepsilon^i| \le (\Delta/\lambda)^{1/2}$, where λ is the smallest eigenvalue of $I_{11\cdot 2}$. By (31)

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

where

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

Hence, we obtain

$$P^T_{\varepsilon_2}(\Delta) \le \Psi_1(\Delta, a_2^{ijk})\Delta^{3/2} + \Psi_{2r}(\Delta)\Delta \mid \varepsilon^r \mid + M(\Delta), \tag{32}$$

where

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

Similarly, we have

$$P^{\mathrm{LR}^*}_{\varepsilon_2}(\Delta) \ge -\Psi_1(\Delta, -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha,\beta,\gamma}/3)\Delta^{3/2} - \Psi_{2r}(\Delta)\Delta \mid \varepsilon^r \mid . \tag{33}$$

From (32) and (33),

$$\begin{aligned} P^{\mathrm{LR}^*}_{\varepsilon_2}(\Delta) - P^T_{\varepsilon_2}(\Delta) \ge & - \left\{ \Psi_1(\Delta, a_2^{ijk}) + \Psi_1(\Delta, -g^{i\alpha}g^{j\beta}g^{k\gamma}K_{\alpha,\beta,\gamma}/3) \right\} \Delta^{3/2} \\ & - 2\Psi_{2r}(\Delta)\Delta \mid \varepsilon^r \mid - M(\Delta). \end{aligned}$$

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

$$P^{\mathrm{LR}^*}_{\varepsilon_2}(\Delta) - P^T_{\varepsilon_2}(\Delta) > 0,$$

whenever $0 < \Delta < \Delta_0$. □

Proof of Theorem 5. The distribution function of $T_0 \in \mathscr{S}_0$ 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

$$\begin{aligned} m_{30} &= \left(\frac{1}{6}K_{i,j,k} + \frac{1}{2}a_2^{i'j'k'}I_{(i'i)}I_{(j'j)}I_{(k'k)} \right) \varepsilon^i\varepsilon^j\varepsilon^k, \\ m_{20} &= -\frac{1}{2}a_2^{i'j'k'}I_{(i'i)}I_{(j'j)}I_{(k'k)}\varepsilon^i\varepsilon^j\varepsilon^k + \frac{1}{2}I_0^{ij}K_{i,j,k}\varepsilon^k + \frac{1}{2}a_2^{ijk}[3]I_{(il)}I_{(jk)}\varepsilon^l, \\ m_{10} &= \frac{1}{2}J_{i,j,k}\varepsilon^i\varepsilon^j\varepsilon^k - \frac{1}{2}I_0^{ij}K_{i,j,k}\varepsilon^k - \frac{1}{2}a_2^{ijk}[3]I_{(il)}I_{(jk)}\varepsilon^l, \\ m_{00} &= -\frac{1}{6}(K_{i,j,k} + 3J_{i,jk})\varepsilon^i\varepsilon^j\varepsilon^k. \end{aligned} \tag{34}$$

Note that, under (A-6), $d^r = 0$ 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 1 can be written as

$$\begin{aligned}
m_3 &= m_{30}, \\
m_2 &= m_{20}, \\
m_1 &= m_{10} + \frac{1}{2}(K_{i,j,r} + J_{i,jr} + J_{j,ir})\varepsilon^i\varepsilon^j\varepsilon^r \\
&\quad + \frac{1}{2}\{a_3^i I_{(ij)} - g^{rs}(K_{j,r,s} + J_{j,rs})\}\varepsilon^j, \\
m_0 &= m_{00} - \frac{1}{2}(K_{i,j,r} + J_{i,jr} + J_{j,ir})\varepsilon^i\varepsilon^j\varepsilon^r \\
&\quad - \frac{1}{2}\{a_3^i I_{(ij)} - g^{rs}(K_{j,r,s} + J_{j,rs})\}\varepsilon^j.
\end{aligned} \tag{35}$$

The comparison of (34) and (35) leads to Theorem 5. □

Proof of Theorem 6 and 7. Note that $\tilde{Z}_i = W_i + o_p(1)$. Expand T^* as

$$\begin{aligned}
T^* &= h(\hat{\theta}_1)T + c_n^{-1}A^i\tilde{Z}_i \\
&= (1 + c_n^{-1}h_i\eta^i)T + c_n^{-1}A^iW_i + o_p(c_n^{-1}).
\end{aligned} \tag{36}$$

Inserting (5) in (36) we obtain

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

where

$$\begin{aligned}
a_2^{*ijk} &= a_2^{ijk} + h_lg^{li}g^{jk}. \\
a_3^{*i} &= a_3^i + A^i.
\end{aligned} \tag{37}$$

This implies $T^* \in \mathscr{S}$, and hence a necessary and sufficient condition for its locally unbiasedness is that the coefficients in (37) satisfy

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

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

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 (i) and (ii) with respect to h_i and A^i, we obtain the relations in Theorem 6. Theorem 7 follows from the above argument and Theorem 1. □

Acknowledgments The author would like to express his sincere thanks to Professor Masanobu Taniguchi for his encouragement and guidance. Also thanks are extended to Professor Takeru Suzuki for many helpful suggestions and advice throughout the period in which he supervised his research.

References

[1] 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.

[2] 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.

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

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

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

[6] Kakizawa, Y. (1997). Higher-order Bartlett-type adjustment. *J. Statist. Plan. Inf.* **65.** 269–280.

[7] Li, B. (2001). Sensitivity of Rao's score test, the Wald test and the likelihood ratio test to nuisance parameters. *J. Statist. Plan. Inf.* **97.** 57–66.

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

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

[10] Rao, C. R. and Mukerjee, R. (1995). Comparison of Bartlett-type adjustments for the efficient score statistic. *J. Statist. Plan. Inf.* **46.** 137–146.

[11] 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.

[12] Taniguchi, M. (1988). Asymptotic expansion of the distribution of some test statistics for Gaussian ARMA processes. *J. Multivariate. Anal.* **27.** 494–511.

[13] Taniguchi, M. (1991). *Higher Order Asymptotic Theory for Time Series Analysis.* Lecture Notes in Statistics. Vol. 68. Heidelberg: Springer-Verlag.

[14] Taniguchi, M. (1991). Third-order asymptotic properties of a class of test statistics under a local alternative. *J. Multivariate. Anal.* **37.** 223–238.

*Department of Mathematical Sciences School of Science & Engineering
Waseda University, 3-4-1, Okubo, Shinjuku-ku, Tokyo, 169-8555, Japan
E-mail: k-tamaki@toki.waseda.jp

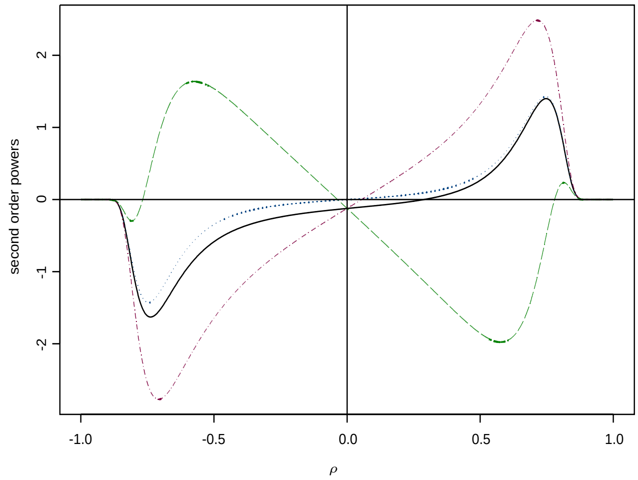

Figure 1: For the bivariate normal model with correlation coefficient $\theta_1 = \rho$, both means $\theta_2 = \mu$ and both variances 1 in Example 4, second order powers of LR, LR*, R_1 and W_1 statistics are plotted. $P_2^{\mathrm{LR}}(\varepsilon)$ (solid line), $P_2^{\mathrm{LR}^*}(\varepsilon)$ (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 $\varepsilon_1 = 1$.

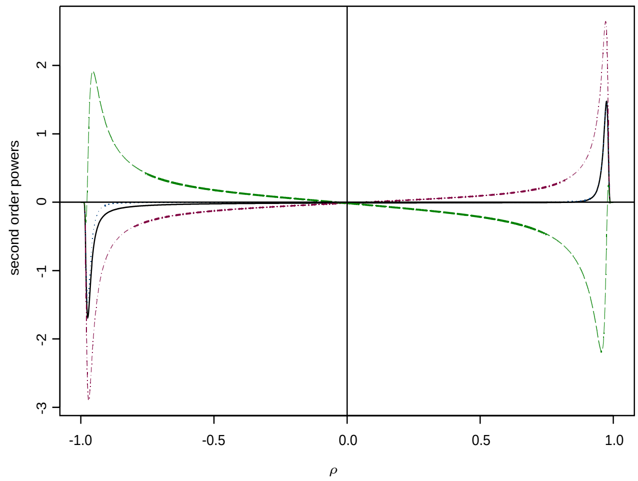

Figure 2: For the bivariate normal model with correlation coefficient $\theta_1 = \rho$, both means $\theta_2 = \mu$ and both variances 1 in Example 4, second order powers of LR, LR*, R_1 and W_1 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{LR}^*}$ (dotted line), $P_2^{\mathrm{R}_1}$ (dashed line) and $P_2^{\mathrm{W}_1}$ (dash-dotted line) with $\alpha = 0.05$ and $\varepsilon_1 = 0.1$.

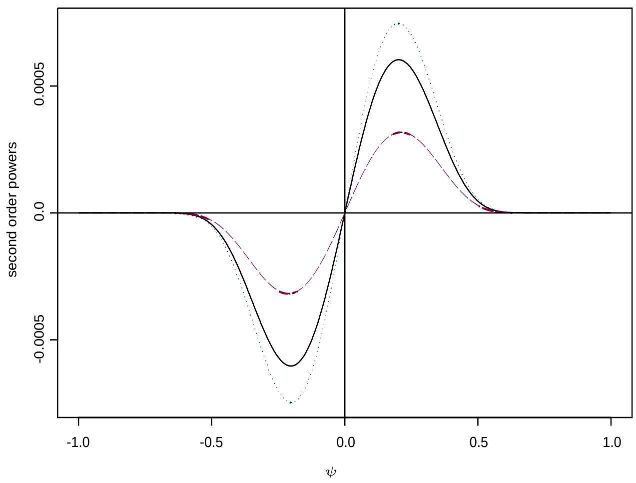

Figure 3: For $MA(1)$ model in Example 5, second order powers of LR, W_1 and R_1 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_1}$ (dotted line) and $P_2^{\mathrm{R}_1}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 6.5$.

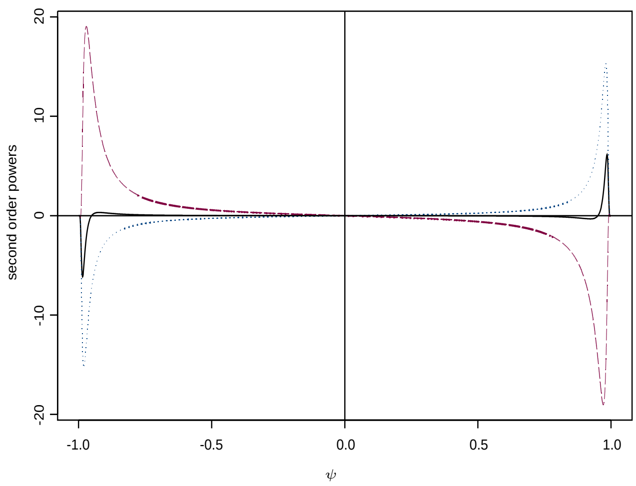

Figure 4: For $MA(1)$ model in Example 5, second order powers of LR, W_1 and R_1 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_1}$ (dotted line) and $P_2^{\mathrm{R}_1}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 0.65$.

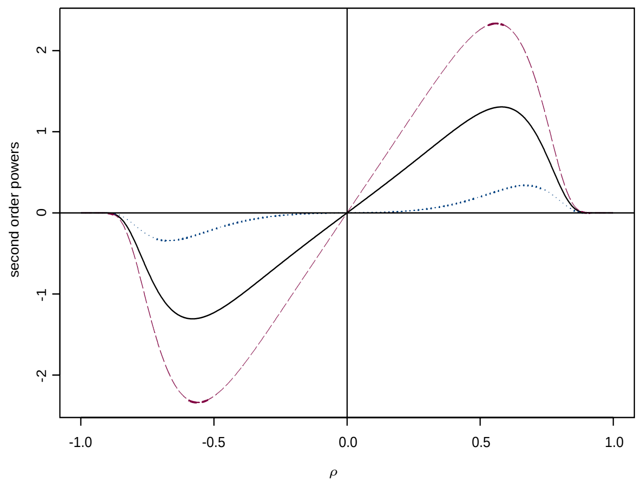

Figure 5: For $AR(1)$ model in Example 6, second order powers of LR, W_2 and R_2 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_2}$ (dotted line) and $P_2^{\mathrm{R}_2}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 3$.

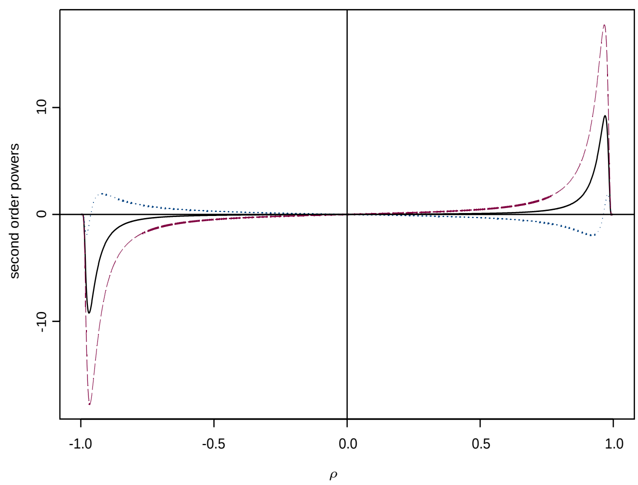

Figure 6: For $AR(1)$ model in Example 6, second order powers of LR, W_2 and R_2 statistics are plotted. P_2^{LR} (solid line), $P_2^{\mathrm{W}_2}$ (dotted line) and $P_2^{\mathrm{R}_2}$ (dashed line) with $\alpha = 0.01$ and $\varepsilon_1 = 0.8$.