Ann. Inst. Statist. Math.
Vol. 48, No. 4, 603–620 (1996)

A PROOF OF INDEPENDENT BARTLETT CORRECTABILITY OF NESTED LIKELIHOOD RATIO TESTS

AKIMICHI TAKEMURA[1] AND SATOSHI KURIKI[2]

[1] *Faculty of Economics, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan*
[2] *The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan*

(Received July 21, 1995; revised March 8, 1996)

Abstract. It is well known that likelihood ratio statistic is Bartlett correctable. We consider decomposition of a likelihood ratio statistic into 1 degree of freedom components based on sequence of nested hypotheses. We give a proof of the fact that the component likelihood ratio statistics are distributed mutually independently up to the order $O(1/n)$ and each component is independently Bartlett correctable. This was implicit in Lawley (1956, *Biometrika*, **43**, 295–303) and proved in Bickel and Ghosh (1990, *Ann. Statist.*, **18**, 1070–1090) using a Bayes method. We present a more direct frequentist proof.

Key words and phrases: Likelihood ratio test, Bartlett correction, nested hypotheses, component likelihood ratio statistic.

1. Introduction

It has been now well established that likelihood ratio test under the null hypothesis is Bartlett correctable. The first general treatment of the distribution of the likelihood ratio test was Lawley (1956). Despite Lawley's result, Bartlett correctability of likelihood ratio tests did not seem to be a generally accepted fact for a long time. Later Hayakawa's extensive calculation (Hayakawa (1977, 1987)) gave a proof of Bartlett correctability. Harris (1986) pointed out an incompleteness of Hayakawa's 1977 proof. Chesher and Smith (1995) fixed Hayakawa's formula, and showed that the corrected formula is consistent with Lawley's result. Furthermore, Bickel and Ghosh (1990) gave a proof based on Bayes approach. Ghosh and Mukerjee (1991, 1992) discuss conditions where frequentist and Bayesian Bartlett correction factors coincide. Cordeiro (1993) gives formulae for computing Bartlett correction factor. For a brief survey on log likelihood ratio and Bartlett correction see Jensen (1993).

As in Lawley (1956) and Bickel and Ghosh (1990), in this article we consider decomposition of overall likelihood ratio statistic into 1 degree of freedom components based on nested hypotheses. Let $\Theta \subset R^p$ denote the parameter space and consider sequence of nested subspaces of Θ:

$$\Theta_0 \subset \Theta_1 \subset \cdots \subset \Theta_p = \Theta, \qquad \text{with} \quad \dim \Theta_j = j.$$

Let

$$H_j : \theta \in \Theta_j$$

and let λ_j be the likelihood ratio statistic for testing H_{j-1} vs. H_j, $j = 1, \ldots, p$. Note that $2 \log \lambda_j$ is asymptotically distributed according to chi-square distribution with 1 degree of freedom under H_{j-1}. Let the overall likelihood ratio statistic for testing H_0 vs. H_p be denoted by λ. Then λ is decomposed as

$$\lambda = \lambda_1 \cdots \lambda_p.$$

We give a proof of the fact that under H_0, λ_j, $j = 1, \ldots, p$, are mutually independently distributed up to the order $O(1/n)$ and they are independently Bartlett correctable. This was implicit in Lawley (1956) and proved in Bickel and Ghosh (1990) using Bayes method. Our proof is more direct frequentist proof. As we have decomposed the overall likelihood ratio statistic into 1 degree of freedom components, it follows immediately from our result that likelihood ratio statistic for intermediate composite hypothesis H_k vs. H_m, $k < m$, is Bartlett correctable as well.

Our result is based on formal asymptotic expansion of the joint characteristic function of the component likelihood ratio statistics under the null hypothesis. We do not treat the validity aspect of the asymptotic expansion.

In Section 2, we state our result in terms of characteristic function and in Section 3 we give our proof. By considering joint characteristic function of the 1 degree of freedom components λ_j, our proof of the independent Bartlett correctability of λ_j's became much harder than the proof of Bartlett correctability of the overall statistic λ. During the course of our proof in Section 3, we point out added complexities in the form of remarks.

2. Main result

Before stating our result, we set up our framework somewhat more precise. Let θ be the p-dimensional parameter vector. We assume that independent and identically distributed observations $x_1, \ldots, x_n$ are obtained from a density $f(x, \theta)$. The likelihood ratio statistic λ_j for testing H_{j-1} vs. H_j is defined as

$$\lambda_j = \frac{\max_{\theta \in \Theta_j} \prod_{i=1}^n f(x_i, \theta)}{\max_{\theta \in \Theta_{j-1}} \prod_{i=1}^n f(x_i, \theta)}, \qquad j = 1, \ldots, p.$$

We state our main theorem in terms of joint characteristic function of λ_j, $j = 1, \ldots, p$.

THEOREM 2.1. *Under H_0, $\lambda_1, \ldots, \lambda_p$ are mutually independently distributed up to the order $O(1/n)$ and independently Bartlett correctable. Namely, there exist constants $c_1, \ldots, c_p$ (depending only on H_0) such that*

$$\begin{aligned} (2.1) \qquad & E_{H_0}[\exp(it_1 2 \log \lambda_1 + \cdots + it_p 2 \log \lambda_p)] \\ & = \prod_{j=1}^p \left[(1 - 2it_j)^{-1/2} \left(1 + \frac{c_j}{n} \left(\frac{1}{1 - 2it_j} - 1 \right) \right) \right] + o(1/n). \end{aligned}$$

Note that up to the order $O(1/n)$ (2.1) is equivalent to

$$
\begin{aligned}
(2.2)\quad & E_{H_0}[\exp(it_1 2\log\lambda_1 + \cdots + it_p 2\log\lambda_p)] \\
& = \left(\prod_{j=1}^{p}(1-2it_j)^{-1/2}\right)\left(1+\frac{1}{n}\sum_{j=1}^{p} c_j\left(\frac{1}{1-2it_j}-1\right)\right) + o(1/n).
\end{aligned}
$$

Setting $t_1 = \cdots = t_k = t_{m+1} = \cdots = t_p = 0$ and $t_{k+1} = \cdots = t_m = t$, $k < m$, the following corollary follows immediately from (2.2).

COROLLARY 2.1. *Consider the likelihood ratio statistic $\lambda_{k,m}$ for testing H_k* vs. *H_m, $k < m$. $\lambda_{k,m}$ is Bartlett correctable under H_k.*

3. Proof

Here we give our proof of Theorem 2.1. We divide our proof into 4 parts. First, we setup necessary notations. Second, we discuss choosing appropriate parameterization to make our calculation simpler. Third, we give stochastic expansion of $2\log\lambda_j$. Finally, we evaluate joint characteristic function of $2\log\lambda_1, \ldots, 2\log\lambda_p$.

3.1 *Notation*

Let $\theta = (\theta^1, \ldots, \theta^p) \in \Theta_p$ be the parameter vector. We use tensor notation and we index parameter components by superscripts. Although we mostly follow standard tensor notation as in McCullagh (1987), we shall later introduce some simplifying notational convention for convenience. Let $\theta^0 = (\theta^{10}, \ldots, \theta^{p0})$ be the true parameter vector, i.e. $\Theta_0 = \{\theta^0\}$.

We denote higher order derivatives of the log likelihood function and related quantities as follows. Let

$$
\ell_{j_1\cdots j_k} = \ell_{j_1\cdots j_k}(x;\theta) = \frac{\partial^k}{\partial\theta^{j_1}\cdots\partial\theta^{j_k}}\log f(x,\theta),
$$

and

$$
\begin{aligned}
\mathcal{L}_{j_1\cdots j_k} &= \frac{1}{n}\sum_{i=1}^{n}\ell_{j_1\cdots j_k}(x_i;\theta^0), \\
L_{j_1\cdots j_k} &= E_{\theta^0}[\ell_{j_1\cdots j_k}(x;\theta^0)], \\
Z_{j_1\cdots j_k} &= \sqrt{n}(\mathcal{L}_{j_1\cdots j_k} - L_{j_1\ldots j_k}). \qquad (3.1)
\end{aligned}
$$

Since the dimensionality of x_i's is irrelevant, subscript i for x is used to index the observation.

Denote the higher order mixed cumulants and moments by

$$
\begin{aligned}
\kappa_{i_1\cdots i_{m_1}, j_1\cdots j_{m_2}, \ldots, k_1\cdots k_{m_h}} &= \mathrm{cum}_\theta(\ell_{i_1\cdots i_{m_1}}, \ell_{j_1\cdots j_{m_2}}, \ldots, \ell_{k_1\cdots k_{m_h}}), \\
L_{i_1\cdots i_{m_1}, j_1\cdots j_{m_2}, \ldots, k_1\cdots k_{m_h}} &= E_\theta(\ell_{i_1\cdots i_{m_1}}\ell_{j_1\cdots j_{m_2}}\cdots\ell_{k_1\cdots k_{m_h}}).
\end{aligned}
$$

Note that κ's and L's are functions of θ. However, we usually use these quantities evaluated at θ^0 and in that case omit θ.

Differentiating the identity

$$E_\theta \left(\frac{\partial}{\partial \theta^i} \log f(x, \theta) \right) = 0$$

the following well known relations on the third order and fourth order mixed derivatives can be easily established:

$$\begin{aligned}
&L_{ijk} + L_{ij,k}[3] + L_{i,j,k} = 0, \\
&\kappa_{ijk} + \kappa_{ij,k}[3] + \kappa_{i,j,k} = 0, \\
&L_{ijkl} + L_{ijk,l}[4] + L_{ij,kl}[3] + L_{ij,k,l}[6] + L_{i,j,k,l} = 0, \\
&\kappa_{ijkl} + \kappa_{ijk,l}[4] + \kappa_{ij,kl}[3] + \kappa_{ij,k,l}[6] + \kappa_{i,j,k,l} = 0.
\end{aligned}$$

General result of this type is given in Skovgaard (1986).

In addition to the standard tensor notation and summation convention, we introduce further notational convention for convenience. We shall later assume that the Fisher information is the identity at θ^0. Because of this assumption, we often encounter terms of the following general form

$$\delta^{ij} \mu_{\cdots i \cdots} \mu_{\cdots j \cdots},$$

where δ^{ij} is the Kronecker's delta. In this case we simply write

$$\mu_{\cdots i \cdots} \mu_{\cdots i \cdots}.$$

Furthermore, in order to discuss joint characteristic function, we need to consider terms of the form $\sum_{i=1}^p t_i z_i z_i$. Omitting the summation sign we simply write this as $t_i z_i z_i$.

More formally, we introduce the following notational convention for our proof.

Notational convention on summation. Indices appearing more than once as subscripts are interpreted as running variables and summed over.

3.2 *Parameterization*

Here we try to choose some canonical parameterization, which makes our derivation simpler. First by considering $\theta - \theta^0$, θ^0 can be taken to be the origin, i.e., $\theta^0 = (0, \ldots, 0)$. Then in some neighborhood of the origin we can choose parameterization such that

$$\begin{aligned}
&\Theta_0 = \{(0, 0, \ldots, 0)\}, \\
&\Theta_1 = \{(\theta^1, 0, \ldots, 0) \mid \theta^1 : \text{free}\}, \\
&\qquad \cdots \\
&\Theta_{p-1} = \{(\theta^1, \ldots, \theta^{p-1}, 0) \mid \theta^1, \ldots, \theta^{p-1} : \text{free}\}, \\
&\Theta_p = \{(\theta^1, \ldots, \theta^p) \mid \theta^1, \ldots, \theta^p : \text{free}\}.
\end{aligned} \tag{3.2}$$

Now, considering appropriate triangular linear transformation $\theta^i \mapsto t^i_j\theta^j$ where $t^i_j = 0$ for $i > j$, we can assume without loss of generality that the Fisher information at the origin is the identity (matrix), i.e.,

$$\kappa_{i,j} = -\kappa_{ij} = -L_{ij} = \delta_{ij}, \tag{3.3}$$

where δ_{ij} is the Kronecker's delta.

Further simplification is possible by considering nonlinear reparameterization in a neighborhood of the origin. Define new parameter vector $\tau = (\tau^1, \ldots, \tau^p)$ by the following relation

$$\theta^i = \tau^i + \frac{1}{2}a^i_{jk}\tau^j\tau^k + \frac{1}{6}a^i_{jkl}\tau^j\tau^k\tau^l + \cdots. \tag{3.4}$$

Here the coefficients $a^i_{jk}, a^i_{jkl}, \ldots$ are invariant under the permutation of subscripts. Note that the Jacobian of (3.4) is the identity at the origin and (3.4) is 1-to-1 in some neighborhood of the origin. Furthermore for our purpose (3.4) can be taken as a polynomial with finite but sufficiently high degree and there is no problem of convergence.

The Fisher information in terms of τ at the origin remains to be the identity and (3.3) is satisfied. Next lemma specifies the form of nonlinear reparameterization in (3.4) such that (3.2) remains to be satisfied.

LEMMA 3.1. (i) *and* (ii) *are equivalent.*

(i) *For all* $m \le p$,

$$(\theta^m, \theta^{m+1}, \ldots, \theta^p) = (0, 0, \ldots, 0) \Leftrightarrow (\tau^m, \tau^{m+1}, \ldots, \tau^p) = (0, 0, \ldots, 0).$$

(ii)

$$a^{i_1}_{i_2\cdots i_k} = 0 \quad \textit{if} \quad \max(i_2, \ldots, i_k) < i_1.$$

PROOF. Consider $m = p$. We want

$$\theta^p = 0 \Leftrightarrow \tau^p = 0 \tag{3.5}$$

for arbitrary values of $\theta^1, \ldots, \theta^{p-1}$. We claim that a necessary and sufficient condition for (3.5) is

$$a^p_{i_1\cdots i_k} = 0 \quad \text{if} \quad \max(i_1, \ldots, i_k) < p. \tag{3.6}$$

Note that (3.6) holds if and only if θ^p can be written as

$$\theta^p = \tau^p(1 + b_j\tau^j + b_{jk}\tau^j\tau^k + \cdots).$$

This holds if and only if $\tau^p = 0 \Rightarrow \theta^p = 0$. Conversely, writing

$$\tau^p = \theta^p(1 + b_j\tau^j + b_{jk}\tau^j\tau^k + \cdots)^{-1}$$

and expanding and expressing the right hand side in terms of θ, τ^p can be written as

$$\tau^p = \theta^p(1 + c_j\theta^j + c_{jk}\theta^j\theta^k + \cdots).$$

Therefore $\theta^p = 0 \Rightarrow \tau^p = 0$.

Next consider $m = p - 1$. Assume (3.5) or equivalently (3.6). We want to ensure that

$$(\theta^{p-1}, \theta^p) = (0,0) \Leftrightarrow (\tau^{p-1}, \tau^p) = (0,0). \tag{3.7}$$

We claim that a necessary and sufficient condition for (3.7) is

$$a^{p-1}_{i_1\cdots i_k} = 0 \quad \text{if} \quad \max(i_1, \ldots, i_k) < p - 1, \tag{3.8}$$

which is equivalent to

$$\theta^{p-1} = \tau^{p-1}(1 + A) + \tau^p B$$

for some polynomials A, B. Hence (3.8) holds if and only if $(\tau^{p-1}, \tau^p) = (0,0) \Rightarrow (\theta^{p-1}, \theta^p) = (0,0)$. Conversely, expressing the right hand side of

$$\tau^{p-1} = \theta^{p-1}(1 + A)^{-1} - \tau^p B(1 + A)^{-1}$$

in terms of θ, we see that $(\theta^{p-1}, \theta^p) = (0,0) \Rightarrow (\tau^{p-1}, \tau^p) = (0,0)$.

Arguing recursively, we prove the lemma. □

From Lemma 3.1, if $\max(i_2, \ldots, i_k) \geq i_1$ then we can choose the value of $a^{i_1}_{i_2\cdots i_k}$ for our convenience.

Now by choosing appropriate nonlinear reparameterization, we can make some of the higher order cumulants vanish. Consider the following relation.

$$\ell_{jk}(x;\tau) = \frac{\partial^2}{\partial\tau^j\partial\tau^k}\log f(x, \theta(\tau)) = \frac{\partial\theta^\alpha}{\partial\tau^j}\frac{\partial\theta^\beta}{\partial\tau^k}\ell_{\alpha\beta}(x;\theta) + \frac{\partial^2\theta^\alpha}{\partial\tau^j\partial\tau^k}\ell_\alpha(x;\theta).$$

Evaluating this at the origin we obtain

$$\ell_{jk}(x;\tau) = \ell_{jk}(x;\theta) + a^\alpha_{jk}\ell_\alpha(x;\theta) \qquad (\text{at } \tau = 0, \ \theta = 0).$$

Therefore at the origin

$$\mathrm{Cov}_{\tau=0}(\ell_{jk}(x;\tau), \ell_i(x;\tau)) = \kappa_{i,jk} + a^\alpha_{jk}\delta_{i\alpha}.$$

Letting

$$a^\alpha_{jk}\delta_{i\alpha} = a^i_{jk} = -\kappa_{i,jk}$$

we can make $\kappa_{i,jk}$ vanish for (i,j,k) such that $\max(j,k) \geq i$. Similarly from

$$\ell_{jkl}(x;\tau) = \frac{\partial\theta^\alpha}{\partial\tau^j}\frac{\partial\theta^\beta}{\partial\tau^k}\frac{\partial\theta^\gamma}{\partial\tau^l}\ell_{\alpha\beta\gamma}(x;\theta) + \frac{\partial^2\theta^\alpha}{\partial\tau^j\partial\tau^k}\frac{\partial\theta^\beta}{\partial\tau^l}\ell_{\alpha\beta}(x;\theta)[3]$$
$$+ \frac{\partial^3\theta^\alpha}{\partial\tau^j\partial\tau^k\partial\tau^l}\ell_\alpha(x;\theta)$$

we obtain at the origin

$$\mathrm{Cov}_{\tau=0}(\ell_{jkl}(x;\tau), \ell_i(x;\tau)) = \kappa_{i,jkl} + a^{\alpha}_{jk} k_{i,\alpha l}[3] + a^{\alpha}_{jkl}\delta_{i\alpha}.$$

Hence for $\max(j,k,l) \geq i$ letting

$$a^i_{jkl} = -\kappa_{i,jkl} - a^{\alpha}_{jk} k_{i,\alpha l}[3]$$

we can make $\kappa_{i,jkl}$ vanish. Arguing recursively we have the following lemma.

LEMMA 3.2. *Consider the sequence of nested hypotheses in* (3.2). *Without loss of generality we can choose parameterization such that at* $\theta^0 = (0, 0, \ldots, 0)$

$$\kappa_{i_1, i_2 \cdots i_k} = 0 \quad \text{if} \quad \max(i_2, \ldots, i_k) \geq i_1. \tag{3.9}$$

The simplification in (3.9) is very useful for calculation of $O(1/n)$ terms needed to prove our result.

REMARK 3.1. Consider the following quantities obtained by the procedures described above:

$$\begin{aligned}
&\ell_{jk}(x;\theta) + a^{\alpha}_{jk}\ell_{\alpha}(x;\theta) \mid_{\theta=0}, \\
&\ell_{jkl}(x;\theta) + a^{\alpha}_{jk}\ell_{\alpha l}(x;\theta)[3] + a^{\alpha}_{jkl}\ell_{\alpha}(x;\theta) \mid_{\theta=0}, \\
&\ldots
\end{aligned}$$

where

$$a^i_{jkl\cdots} = 0 \quad \text{if} \quad \max(j,k,l,\ldots) < i$$

and otherwise defined recursively by the relations

$$\begin{aligned}
&\kappa_{i,jk} + a^{\alpha}_{jk}\delta_{i\alpha} = 0, \\
&\kappa_{i,jkl} + a^{\alpha}_{jk}\kappa_{i,\alpha l}[3] + a^{\alpha}_{jkl}\delta_{i\alpha} = 0, \\
&\ldots .
\end{aligned}$$

It can be easily shown that these quantities are invariant under any nonlinear transformation of the coordinate system from $\theta = (\theta^1, \ldots, \theta^p)$ to $\tilde{\tau} = (\tilde{\tau}^1, \ldots, \tilde{\tau}^p)$ of the form

$$\theta^i = \tilde{\tau}^i + \frac{1}{2}\tilde{a}^i_{jk}\tilde{\tau}^j\tilde{\tau}^k + \frac{1}{6}\tilde{a}^i_{jkl}\tilde{\tau}^j\tilde{\tau}^k\tilde{\tau}^l + \cdots \tag{3.10}$$
$$(\tilde{a}^i_{jkl\cdots} = 0 \quad \text{for } \max(j,k,l,\ldots) < i).$$

This can be proved in an analogous way as Subsection 7.2.3 of McCullagh (1987). However, in our case the nonlinear transformation of the coordinate system, (3.10), has to be adapted to the nested subspaces in (3.2), whereas McCullagh (1987) considers nonlinear transformations with no restrictions.

REMARK 3.2. In this paper we are considering the finest sequence of nested subspaces in (3.2). More often, one considers just one intermediate subspace Θ_q, $0 < q < p$, corresponding to a composite null hypothesis, or even no intermediate subspace. Similar argument to the proof of Lemma 3.2 shows that we can choose, without loss of generality, parameterizations such that

$$\kappa_{i_1,i_2\cdots i_k} = 0 \quad \text{if} \quad i_1 \le q \text{ or } \max(i_2,\ldots,i_k) > q$$

at θ^0 corresponding to the nesting $\Theta_0 \subset \Theta_q \subset \Theta_p$, and

$$\kappa_{i_1,i_2\cdots i_k} = 0 \quad \text{for all indices} \tag{3.11}$$

at θ^0 corresponding to the nesting $\Theta_0 \subset \Theta_p$. The possibility to choose a parameterization satisfying (3.11) for testing a simple null hypothesis is well known (e.g., McCullagh and Cox (1986), Subsection 7.2.3 of McCullagh (1987)).

3.3 *Stochastic expansion of log likelihood ratio*

Here we give a stochastic expansion of $2\log\lambda$ in terms of the random variables $Z_{i_1\cdots i_k}$ defined in (3.1). From Section 2 of Hayakawa (1977) and Section 7.4 of McCullagh (1987) we have

$$\begin{aligned} 2\log\lambda = {} & Z_iZ_i + \frac{1}{\sqrt{n}}Z_{ij}Z_iZ_j + \frac{1}{3\sqrt{n}}L_{ijk}Z_iZ_jZ_k + \frac{1}{n}Z_{ia}Z_{ja}Z_iZ_j \\ & + \frac{1}{3n}Z_{ijk}Z_iZ_jZ_k + \frac{1}{n}L_{jka}Z_{ia}Z_iZ_jZ_k \\ & + \frac{1}{12n}(L_{ijkl} + L_{ija}L_{kla}[3])Z_iZ_jZ_kZ_l + o_p(1/n). \end{aligned} \tag{3.12}$$

Note that in (3.12) indices $i, j, k, a, \ldots$ run from 1 through p. Now the stochastic expansion of the 1 degree of freedom component λ_q for H_{q-1} vs. H_q can be obtained from (3.12) by the following simple argument. Consider the likelihood ratio statistic λ_{0r} for H_0 vs. H_r. Because of (3.2), the stochastic expansion for $2\log\lambda_{0r}$ is the same as in (3.12) except for the range of indices, which is now 1 up to r. Therefore, the stochastic expansion for

$$2\log\lambda_q = 2\log\lambda_{0q} - 2\log\lambda_{0,q-1}$$

is as in (3.12), where at least one of the running variables equals q. From this argument it follows that the stochastic expansion of $\sum_{i=1}^{p} t_i 2\log\lambda_i$, needed to evaluate the characteristic function, can be written as

$$\begin{aligned} 2t_i\log\lambda_i = {} & t_iZ_iZ_i + \frac{1}{\sqrt{n}}Z_{ij}Z_iZ_jt_{\max(i,j)} + \frac{1}{3\sqrt{n}}L_{ijk}Z_iZ_jZ_kt_{\max(i,j,k)} \\ & + \frac{1}{n}Z_{ia}Z_{ja}Z_iZ_jt_{\max(i,j,a)} + \frac{1}{3n}Z_{ijk}Z_iZ_jZ_kt_{\max(i,j,k)} \\ & + \frac{1}{n}L_{jka}Z_{ia}Z_iZ_jZ_kt_{\max(i,j,k,a)} \\ & + \frac{1}{12n}(L_{ija}L_{kla}[3])Z_iZ_jZ_kZ_lt_{\max(i,j,k,l,a)} \\ & + \frac{1}{12n}L_{ijkl}Z_iZ_jZ_kZ_lt_{\max(i,j,k,l)} + o_p(1/n). \end{aligned} \tag{3.13}$$

3.4 *Evaluation of the joint characteristic function*

We now evaluate the joint characteristic function of $\lambda_1, \dots, \lambda_p$ using the stochastic expansion (3.13). From now on, we omit the pure imaginary number i for convenience and write $2t_j \log \lambda_j$ instead of $2it_j \log \lambda_j$. We take the expectation in two steps. First, we consider conditional expectation given the first order derivatives $Z_1, \dots, Z_p$ and then we evaluate the expectation with respect to $Z_1, \dots, Z_p$:

$$(3.14) \qquad E[\exp(2t_j \log \lambda_j)] = E[E(\exp(2t_j \log \lambda_j) \mid Z_1, \dots, Z_p)].$$

For the first step we need the conditional expectation. Relevant conditional expectations (see Section 5.6 of McCullagh (1987)) are

$$\begin{aligned}
(3.15) \quad E(Z_{ij} \mid Z_1, \dots, Z_p) &= \kappa_{k,ij} Z_k + \frac{1}{2\sqrt{n}} (\kappa_{ij,a,b} - \kappa_{k,ij}\kappa_{k,a,b})(Z_a Z_b - \delta_{ab}) \\
&\quad + o(n^{-1/2}), \\
(3.16) \quad E(Z_{ij} Z_{kl} \mid Z_1, \dots, Z_p) &= E(Z_{ij} \mid Z_1, \dots, Z_p) E(Z_{kl} \mid Z_1, \dots, Z_p) \\
&\quad + \mathrm{Cov}(Z_{ij}, Z_{kl} \mid Z_1, \dots, Z_p) \\
&= \kappa_{a,ij} Z_a \kappa_{b,kl} Z_b + \kappa_{ij,kl} - \kappa_{a,ij}\kappa_{a,kl} + o(1).
\end{aligned}$$

(3.15) and (3.16) can be easily derived from asymptotic expansion of the joint density of Z_i and Z_{ij}.

Now we can carry out the calculation of the conditional expectation of the joint characteristic function. By expanding the exponential function in (3.14), and taking the conditional expectation of (3.15) and (3.16), we can see that

$$\begin{aligned}
(3.17) \quad & E[\exp(2t_i \log \lambda_i \mid z_1, \dots, z_p)] \\
&= \exp(t_i z_i z_i + A_1 + A_2) \times (1 + B_1 + B_2 + \dots + B_7) + o(1/n),
\end{aligned}$$

where

$$\begin{aligned}
A_1 &= \frac{1}{3\sqrt{n}} L_{ijk} z_i z_j z_k t_{\max(i,j,k)}, \\
A_2 &= \frac{1}{\sqrt{n}} \kappa_{k,ij} z_i z_j z_k t_{\max(i,j)}, \\
B_1 &= \frac{1}{2n} (\kappa_{ij,k,l} - \kappa_{m,ij}\kappa_{m,k,l})(z_k z_l - \delta_{kl}) z_i z_j t_{\max(i,j)}, \\
B_2 &= \frac{1}{2n} z_i z_j z_k z_l (\kappa_{ij,kl} - \kappa_{a,ij}\kappa_{a,kl}) t_{\max(i,j)} t_{\max(k,l)}, \\
B_3 &= \frac{1}{n} z_i z_j t_{\max(i,j,a)} (\kappa_{c,ia}\kappa_{d,ja} z_c z_d + \kappa_{ai,aj} - \kappa_{c,ia}\kappa_{c,ja}), \\
B_4 &= \frac{1}{3n} \kappa_{a,ijk} z_a z_i z_j z_k t_{\max(i,j,k)}, \\
B_5 &= \frac{1}{n} L_{jka} \kappa_{b,ia} z_i z_j z_k z_b t_{\max(i,j,k,a)}, \\
B_6 &= \frac{1}{12n} (L_{aij} L_{akl}[3]) z_i z_j z_k z_l t_{\max(i,j,k,l,a)}, \\
B_7 &= \frac{1}{12n} L_{ijkl} z_i z_j z_k z_l t_{\max(i,j,k,l)}.
\end{aligned}$$

We combine (3.17) with the Edgeworth expansion of the density of $z_1, \ldots, z_p$ and take the expectation. The Edgeworth expansion of the the density of $z_1, \ldots, z_p$ can be written as follows (see Takemura and Takeuchi (1988)).

$$
\begin{aligned}
(3.18)\quad f(z_1, \ldots, z_p) = \frac{1}{(2\pi)^{p/2}} \exp\bigg[& -\frac{1}{2} z_i z_i + \frac{1}{6\sqrt{n}} \kappa_{i,j,k} z_i z_j z_k + \frac{1}{\sqrt{n}} q_1(z) \\
& + \frac{1}{24n} (\kappa_{i,j,k,l} - \kappa_{i,j,a}\kappa_{k,l,a}[3]) z_i z_j z_k z_l \\
& + \frac{1}{n} q_2(z) \bigg] + o(1/n),
\end{aligned}
$$

where q_1 is linear in $z_1, \ldots, z_p$ without the constant term and q_2 is a second degree polynomial in $z_1, \ldots, z_p$ without the linear terms. Concrete forms of q_1 and q_2 are irrelevant for establishing our result. Denote

$$
\begin{aligned}
C_1 &= \frac{1}{6\sqrt{n}} \kappa_{i,j,k} z_i z_j z_k + \frac{1}{\sqrt{n}} q_1(z), \\
C_2 &= \frac{1}{24n} (\kappa_{i,j,k,l} - \kappa_{i,j,a}\kappa_{k,l,a}[3]) z_i z_j z_k z_l + \frac{1}{n} q_2(z).
\end{aligned}
$$

We combine (3.17) and (3.18) and our problem is reduced to evaluating the following integration term by term:

$$
\begin{aligned}
(3.19)\quad & E(\exp(2t_i \log \lambda_i)) \\
& = \int \cdots \int \exp\left(-\frac{1}{2}(1 - 2t_i) z_i z_i + A_1 + A_2 + C_1 \right) \\
& \quad \times \{1 + B_1 + B_2 + \cdots + B_7 + C_2\} dz_1 \cdots dz_p + o(1/n).
\end{aligned}
$$

At this point the following simple recursive argument is useful.

LEMMA 3.3. *In order to prove Theorem* 2.1 *it is sufficient to prove that all the* $O(1/n)$ *terms containing* t_p *in* (3.19) *do not contain* t_i, $i < p$, *and are linear (i.e. first degree polynomial) in* $1/(1 - 2t_p)$.

PROOF. If the assertion is true, then for some c_p we can write

$$
\begin{aligned}
E(\exp(2t_i \log \lambda_i)) = \prod (1 - 2t_i)^{-1/2} & \left\{ h(t_1, \ldots, t_{p-1}) + \frac{c_p}{n} \left(\frac{1}{1 - 2t_p} - 1 \right) \right\} \\
& + o(1/n).
\end{aligned}
$$

Now put $t_p = 0$. Then because of the recursive nature of the subspaces in (3.2), we have exactly the same problem with dimensionality reduced by 1. Therefore

$$
h(t_1, \ldots, t_{p-1}) = \tilde{h}(t_1, \ldots, t_{p-2}) + \frac{c_{p-1}}{n} \left(\frac{1}{1 - 2t_{p-1}} - 1 \right).
$$

This recursive argument implies (2.2). □

REMARK 3.3. For our proof we have to eliminate not only the terms of the form $1/(1-2t_p)^k$, $k \geq 2$, but also terms of the form $(1-2t_i)/(1-2t_p)$. This is an added complexity in considering joint characteristic function of the component likelihood ratio statistics.

Our proof now consists of exhaustive verification of each term of (3.19) that each term of the order $O(1/n)$ containing t_p is linear in $1/(1-2t_p)$.

3.4.1 *Terms containing third order cumulants*

We begin by considering the term $\exp(A_1 + A_2 + C_1)$. Using $L_{ijk} = \kappa_{ijk} = -\kappa_{ij,k}[3] - \kappa_{i,j,k}$, we have

$$\begin{aligned}\exp(A_1 + A_2 + C_1) &= \exp\left(D_1 + D_2 + D_3 + \frac{1}{\sqrt{n}}q_1(z)\right)\\ &= 1 + \sum D_i + \frac{1}{\sqrt{n}}q_1(z) + \frac{1}{2}\sum (D_i)^2 + \sum_{i<j} D_i D_j\\ &\quad + \frac{1}{\sqrt{n}}q_1(z)\sum D_i + \frac{1}{2n}q_1(z)^2 + o(1/n),\end{aligned}$$

where

$$\begin{aligned}D_1 &= \frac{1}{6\sqrt{n}}\kappa_{i,j,k} z_i z_j z_k (1 - 2t_{\max(i,j,k)}),\\ D_2 &= \frac{1}{2\sqrt{n}}\kappa_{k,ij} z_i z_j z_k (1 - 2t_{\max(i,j,k)}),\\ D_3 &= -\frac{1}{2\sqrt{n}}\kappa_{k,ij} z_i z_j z_k (1 - 2t_{\max(i,j)}).\end{aligned}$$

Note that $\sum D_i + q_1(z)/\sqrt{n}$ is an odd polynomial in z and this vanishes by integration. Furthermore, the index for t agrees with one of i, j, k and hence integration of terms $q_1(z)\sum D_i$ yields only linear terms in $1/(1-2t_a)$, $a \leq p$. Also $q_1(z)^2$ is quadratic in z and yields only linear terms in $1/(1-2t_a)$, $a \leq p$. We see that $q_1(z)$ is irrelevant for our argument. We note here that $q_2(z)$ is quadratic as well and irrelevant for our proof.

Therefore, integrations of only 6 terms $(D_1)^2$, $(D_2)^2$, $(D_3)^2$, D_1D_2, D_1D_3, D_2D_3 require close inspection.

These terms consist of basic terms of the form

$$z_i z_j z_k (1 - 2t_\alpha) z_l z_m z_n (1 - 2t_\beta),$$

where $\alpha \in \{i, j, k\}$ and $\beta \in \{l, m, n\}$. We only need to consider the case where each distinct index appears even times.

Suppose that $\alpha \neq \beta$. Then both α and β have to appear at least twice. These lead to terms linear in $1/(1-2t_a)$, $a \leq p$. If $\alpha = \beta$ and if there are at least 4 α's in $\{i, j, k, l, m, n\}$, then again only terms linear in $1/(1-2t_a)$, $a \leq p$, appear.

We see that the only essentially difficult terms to check are of the form $(1-2t_\alpha)^2 (z_\alpha)^2 z_a z_b z_c z_d$, where $a, b, c, d \neq \alpha$. Integrating $(z_\alpha)^2$ out we have

$(1-2t_\alpha)z_a z_b z_c z_d$, $a,b,c,d \neq \alpha$. We have to verify that terms of this type in $D_i D_j$ cancel somewhere in the entire expression of the joint characteristic function.

Consider $(D_1)^2$, $(D_2)^2$, $(D_3)^2$, D_1D_2, D_1D_3, D_2D_3 in turn.

1. $(D_1)^2$: In

$$\frac{1}{2}(D_1)^2 = \frac{1}{72n}\kappa_{i,j,k}\kappa_{l,m,h}z_i z_j z_k z_l z_m z_h(1-2t_{\max(i,j,k)})(1-2t_{\max(l,m,h)}),$$

we need to consider the case $\max(i,j,k) = \max(l,m,h)$. If these are less than p, then t_p does not appear. Therefore, we can restrict our attention to the following term, which remains to be canceled.

$$\frac{1}{8n}\kappa_{p,i,j}\kappa_{p,k,l}(1-2t_p)z_i z_j z_k z_l \qquad (i,j,k,l<p). \tag{3.20}$$

2. $(D_2)^2$: Noting $\kappa_{i,pj} = 0$ by (3.9), similar reasoning applied to $(D_2)^2/2$ yields the following term yet to be canceled.

$$\frac{1}{8n}\kappa_{p,ij}\kappa_{p,kl}(1-2t_p)z_i z_j z_k z_l \qquad (i,j,k,l<p). \tag{3.21}$$

3. D_1D_2: Similarly D_1D_2 yields

$$\frac{1}{4n}\kappa_{p,i,j}\kappa_{p,kl}(1-2t_p)z_i z_j z_k z_l \qquad (i,j,k,l<p). \tag{3.22}$$

4. $(D_3)^2$: In

$$\frac{1}{2}(D_3)^2 = \frac{1}{8n}\kappa_{k,ij}\kappa_{l,mh}z_i z_j z_k z_l z_m z_h(1-2t_{\max(i,j)})(1-2t_{\max(m,h)}), \tag{3.23}$$

we need to consider the case $\max(i,j) = \max(m,h) = a$ (say). If $a = p$ then $\kappa_{k,ij} = 0$ and the term vanishes. Therefore $a < p$ and relevant terms in (3.23) are

$$\frac{1}{8n}\kappa_{k,aa}\kappa_{l,aa}z_k z_l(z_a)^4(1-2t_a)^2 \qquad (k,l \le p \text{ and } a<p)$$
$$+\frac{1}{2n}\kappa_{k,ai}\kappa_{l,am}z_k z_l z_i z_m(z_a)^2(1-2t_a)^2 \qquad (k,l\le p \text{ and } i,m<a<p).$$

t_p appears only from the case $k=l=p$ and $i=m$, and we have

$$\frac{1}{8n}\kappa_{p,aa}^2(z_p)^2(z_a)^4(1-2t_a)^2 \qquad (a<p)$$
$$+\frac{1}{2n}\kappa_{p,ai}^2(z_i)^2(z_p)^2(z_a)^2(1-2t_a)^2 \qquad (i<a<p).$$

Integrating z_p and z_a out yields

$$\frac{3}{8n}\kappa_{p,aa}^2\frac{1}{1-2t_p} \qquad (a<p)$$
$$+\frac{1}{2n}\kappa_{p,ai}^2(z_i)^2\frac{1-2t_a}{1-2t_p} \qquad (i<a<p).$$

The first term is linear in $1/(1-2t_p)$. Therefore remaining term to be canceled is

$$\frac{1}{2n}\kappa_{p,ai}^2(z_i)^2\frac{1-2t_a}{1-2t_p} \qquad (i<a<p). \tag{3.24}$$

5. D_1D_3: D_1D_3 is irrelevant. Actually in

$$\kappa_{i,j,k}z_iz_jz_k(1-2t_{\max(i,j,k)})\kappa_{l,mh}z_lz_mz_h(1-2t_{\max(m,h)})$$

we set $\max(m,h)=\max(i,j,k)=a$. If $a=p$ then $\kappa_{l,mh}=0$. On the other hand if $a<p$ then t_p does not appear.

6. D_2D_3: Similarly D_2D_3 is irrelevant for our proof.

The terms coming from D_iD_j's which have to be canceled are (3.20), (3.21), (3.22), and (3.24). These terms have to be canceled by terms in $B_1,\dots,B_7$ and C_2. We now pick up terms from $B_1,\dots,B_7$, C_2, which contain third order cumulants.

1. B_1: Consider

$$\begin{aligned}&-\frac{1}{2n}\kappa_{m,ij}\kappa_{m,k,l}(z_kz_l-\delta_{kl})z_iz_jt_{\max(i,j)}\\ &\qquad=-\frac{1}{2n}\kappa_{m,ij}\kappa_{m,k,l}z_kz_lz_iz_jt_{\max(i,j)}+\frac{1}{2n}\kappa_{m,ij}\kappa_{m,k,k}z_iz_jt_{\max(i,j)}.\end{aligned}$$

The second term on the right hand side obviously yields only linear term and can be ignored. In the first term, if $\max(i,j)=p$ then $\kappa_{m,ij}=0$. Therefore, we only need to consider $\max(i,j)<p$. The only case where t_p appears is $k=l=p$ and $i=j<p$, and integrating z_p out we are left with

$$-\frac{1}{2n}\kappa_{m,ii}\kappa_{m,p,p}\frac{t_i}{1-2t_p}(z_i)^2 \qquad (m\le p \text{ and } i<p). \tag{3.25}$$

2. B_2: B_2 is irrelevant. Actually if t_p appears from

$$z_iz_jz_kz_l\kappa_{a,ij}\kappa_{a,kl}t_{\max(i,j)}t_{\max(k,l)},$$

then either $\max(i,j)=p$ or $\max(k,l)=p$ and hence $\kappa_{a,ij}\kappa_{a,kl}=0$.

3. B_3:

$$\frac{1}{n}z_iz_jz_cz_dt_{\max(i,j,a)}\kappa_{c,ia}\kappa_{d,ja}-\frac{1}{n}z_iz_jt_{\max(a,i,j)}\kappa_{c,ia}\kappa_{c,ja}.$$

In the second term if $\max(a,i,j)=p$ then $\kappa_{c,ia}\kappa_{c,ja}=0$, otherwise if $\max(a,i,j)<p$ then t_p does not appear. Therefore, the second term is irrelevant. Consider the first term. Again if $\max(a,i,j)=p$ then $\kappa_{c,ia}\kappa_{d,ja}=0$. Therefore, let $\max(a,i,j)<p$. The only possibility where t_p appears is when $c=d=p$ and $i=j<p$. Integrating z_p out, the remaining term to be canceled is

$$\frac{1}{n}\kappa_{p,ai}^2\frac{t_{\max(a,i)}}{1-2t_p}(z_i)^2 \qquad (a,i<p). \tag{3.26}$$

4. B_4, B_7: B_4 and B_7 do not contain third order cumulants.
5. B_5:

$$\frac{1}{n} L_{jka} \kappa_{b,ia} z_i z_j z_k z_b t_{\max(i,j,k,a)}.$$

If $a = p$ then $\kappa_{b,ia} = 0$. Therefore, let $a < p$. If $\max(i,j,k) < p$ then because b has to be equal to one of i, j, k, t_p can not appear. Therefore, the only remaining possibilities are either $b = j = p > i = k$ or $b = k = p > i = j$. Therefore, we have

$$\frac{2}{n} L_{api} \kappa_{p,ai} (z_p)^2 (z_i)^2 t_p.$$

Integrating z_p out and using $L_{api} = -\kappa_{p,ai} - \kappa_{p,a,i}$, we obtain the remaining term as

$$-\frac{1}{n}(\kappa_{p,ai}^2 + \kappa_{p,ai}\kappa_{p,a,i}) \frac{2t_p}{1-2t_p} (z_i)^2 \qquad (a, i < p). \tag{3.27}$$

6. B_6:

$$\frac{1}{12n}(L_{aij} L_{akl}[3]) z_i z_j z_k z_l t_{\max(i,j,k,l,a)} \tag{3.28}$$

is the hardest term to look at. If $\max(i,j,k,l,a) < p$ in (3.28) t_p does not appear. Therefore let $\max(i,j,k,l,a) = p$.

First consider a particular case, where $i = j = k = l = p$ and $a \leq p$. Then by integrating z_p out and letting $L_{app} = -\kappa_{a,p,p}$, (3.28) becomes

$$\frac{3}{4n} \frac{t_p}{(1-2t_p)^2} \kappa_{a,p,p}^2 \qquad (a \leq p). \tag{3.29}$$

Next consider the case where not all of i, j, k, l are equal to p. The relevant subcases are of the following 2 types: (i) $i, j, k, l < p$, or (ii) two of i, j, k, l equal to p. For the subcase (i), we need $a = p$ for t_p to appear. Then $L_{pij} = -\kappa_{p,ij} - \kappa_{p,i,j}$, $L_{pkl} = -\kappa_{p,kl} - \kappa_{p,k,l}$ and (3.28) is

$$\frac{1}{4n}(\kappa_{p,ij} + \kappa_{p,i,j})(\kappa_{p,kl} + \kappa_{p,k,l}) z_i z_j z_k z_l t_p \qquad (i, j, k, l < p). \tag{3.30}$$

For the subcase (ii), let $i = j < k = l = p$ and $a \leq p$. Then $L_{app} = -\kappa_{a,p,p}$, $L_{aij} = -\kappa_{a,ii} - \kappa_{a,i,i}$, $L_{api} = L_{apj} = -\kappa_{p,ai} - \kappa_{a,p,i}$. Therefore, for this case (3.28) becomes

$$\frac{1}{12n}\{\kappa_{a,p,p}(\kappa_{a,ii} + \kappa_{a,i,i}) + 2(\kappa_{p,ai} + \kappa_{a,p,i})^2\} t_p (z_p)^2 (z_i)^2.$$

Considering the symmetry, there are 6 possibilities of this type. Therefore, the term to be canceled is

$$\frac{1}{2n}\frac{t_p}{1-2t_p}\{\kappa_{a,p,p}(\kappa_{a,ii} + \kappa_{a,i,i}) + 2(\kappa_{p,ai} + \kappa_{a,p,i})^2\}(z_i)^2 \qquad (i < p \text{ and } a \leq p). \tag{3.31}$$

We saw 3 types of terms (3.29), (3.30) and (3.31) to be canceled from B_6.

7. C_2: Consider

$$-\frac{1}{24n}(\kappa_{i,j,a}\kappa_{k,l,a}[3])z_i z_j z_k z_l.$$

If $i = j = k = l = p$ and $a \le p$, then integrating z_p out we have

$$-\frac{3}{8n}\kappa_{p,p,a}^2\frac{1}{(1-2t_p)^2} \qquad (a \le p). \tag{3.32}$$

Otherwise for t_p to appear two of i, j, k, l have to be equal to p. Let $i = j < k = l = p$ and $a \le p$. Then we obtain

$$-\frac{1}{24n}(\kappa_{p,p,a}\kappa_{i,i,a} + 2\kappa_{p,i,a}^2)(z_i)^2\frac{1}{1-2t_p} \qquad (i < p \text{ and } a \le p).$$

Considering the symmetry, there are 6 possibilities of this type. Therefore, the remaining term to be canceled is

$$-\frac{1}{4n}(\kappa_{p,p,a}\kappa_{i,i,a} + 2\kappa_{p,i,a}^2)(z_i)^2\frac{1}{1-2t_p} \qquad (i < p \text{ and } a \le p). \tag{3.33}$$

We have enumerated all the remaining terms containing third order cumulants. Our list of these terms which have to be canceled (including terms coming from D_iD_j's) is as follows: (3.20), (3.21), (3.22), (3.24), (3.25), (3.26), (3.27), (3.29), (3.30), (3.31), (3.32), and (3.33).

Adding together (3.20), (3.21), (3.22), and (3.30), we see that t_p vanishes. Sum of (3.29) and (3.32) reduces to a linear term in $1/(1-2t_p)$. Adding (3.33) to (3.31) cancels some terms in (3.31) and (3.31) is reduced to

$$\frac{1}{2n}\frac{t_p}{1-2t_p}(\kappa_{a,p,p}\kappa_{a,ii} + 2\kappa_{p,ai}^2 + 4\kappa_{p,ai}\kappa_{a,p,i})(z_i)^2 \qquad (i < p \text{ and } a \le p). \tag{3.34}$$

Our reduced list of remaining terms is now: (3.24), (3.25), (3.26), (3.27), and (3.34). Rewrite (3.25) as

$$-\frac{1}{4n}\kappa_{a,ii}\kappa_{a,p,p}\frac{1}{1-2t_p}(z_i)^2 + \frac{1}{4n}\kappa_{a,ii}\kappa_{a,p,p}\frac{1-2t_i}{1-2t_p}(z_i)^2 \qquad (a \le p \text{ and } i < p).$$

The second term becomes linear in $1/(1-2t_p)$ when z_i is integrated out and can be ignored. Add the first term to (3.34). Then the first term within the parentheses of (3.34) no longer contains t_p. Now add (3.27) to (3.34). Then (3.34) is reduced to

$$-\frac{1}{n}\frac{t_p}{1-2t_p}\kappa_{p,ai}^2(z_i)^2 \qquad (i, a < p). \tag{3.35}$$

Here we ignored the case $a = p$ since then $\kappa_{p,pi} = 0$.

Now remaining terms to be canceled are: (3.24), (3.26), and (3.35). Write

$$t_{\max(a,i)} = t_p + \frac{1}{2}(1-2t_p) - \frac{1}{2}(1-2t_{\max(a,i)})$$

in (3.26). Then the first term cancels (3.35). The second term is irrelevant. Now consider the third term:

$$-\frac{1}{2n}\kappa_{p,ai}^2(z_i)^2\frac{1-2t_{\max(a,i)}}{1-2t_p}. \tag{3.36}$$

If $i < a$ then $\max(a,i) = a$ and this cancels (3.24). We are left with the case $i \geq a$. Then (3.36) is

$$-\frac{1}{2n}\kappa_{p,ai}^2(z_i)^2\frac{1-2t_i}{1-2t_p},$$

which is linear in $1/(1-2t_p)$ after integrating z_i out.

We have now checked all terms containing the third order cumulants and verified that these terms yield only terms linear in $1/(1-2t_p)$.

3.4.2 *Terms containing fourth order cumulants*

Verifying terms containing fourth order cumulants is much simpler than the last subsection. Picking up relevant terms we have the following list of terms.

$$B_1: \quad \frac{1}{2n}\kappa_{ij,k,l}(z_k z_l - \delta_{kl})z_i z_j t_{\max(i,j)}, \tag{3.37}$$

$$B_2: \quad \frac{1}{2n}z_i z_j z_k z_l \kappa_{ij,kl} t_{\max(i,j)} t_{\max(k,l)}, \tag{3.38}$$

$$B_3: \quad \frac{1}{n}z_i z_j t_{\max(i,j,a)}\kappa_{ai,aj}, \tag{3.39}$$

$$B_7: \quad \frac{1}{12n}L_{ijkl}z_i z_j z_k z_l t_{\max(i,j,k,l)}, \tag{3.40}$$

$$C_2: \quad \frac{1}{24n}\kappa_{i,j,k,l}z_i z_j z_k z_l. \tag{3.41}$$

Using $L_{ijkl} = -\kappa_{ijk,l}[4] - \kappa_{ij,kl}[3] - \kappa_{ij,k,l}[6] - \kappa_{i,j,k,l}$ expand (3.40). If $\max(i,j,k,l) < p$ in (3.40), then t_p can not appear. Therefore, we only need to consider the case $\max(i,j,k,l) = p$ in (3.40). Then there are at least two p's among i, j, k, l and $\kappa_{ijk,l}[4] = 0$. Therefore in (3.40) we can let $L_{ijkl} = -\kappa_{ij,kl}[3] - \kappa_{ij,k,l}[6] - \kappa_{i,j,k,l}$, $\max(i,j,k,l) = p$, and reduce (3.40) to the following form

$$-\frac{1}{4n}\kappa_{ij,kl}z_i z_j z_k z_l t_p - \frac{1}{2n}\kappa_{ij,k,l}z_i z_j z_k z_l t_p - \frac{1}{12n}\kappa_{i,j,k,l}z_i z_j z_k z_l t_p \qquad (\max(i,j,k,l) = p). \tag{3.42}$$

Now we examine cumulants $\kappa_{i,j,k,l}$, $\kappa_{ij,k,l}$ and $\kappa_{ij,kl}$ in turn.

1. $\kappa_{i,j,k,l}$: Adding the last term of (3.42) to (3.41) we have

$$\frac{1}{24n}\kappa_{i,j,k,l}z_i z_j z_k z_l(1-2t_p) \qquad (\max(i,j,k,l) = p).$$

Integrating z_p out, this term yields only linear terms in $1/(1-2t_a)$, $a \le p$.

2. $\kappa_{ij,k,l}$: First we take care of δ_{kl} in (3.37). Consider $\kappa_{ij,k,k} z_i z_j t_{\max(i,j)}$. This obviously yields terms linear in $1/(1-2t_i)$. In (3.37) and (3.42) we are left with

$$\frac{1}{2n}\kappa_{ij,k,l} z_i z_j z_k z_l (t_{\max(i,j)} - t_p) \qquad (\max(i,j,k,l) = p).$$

This is non zero only if $p > \max(i,j)$. Because the indices have to appear in pairs, $p > i = j$. Therefore we have

$$\frac{1}{2n}\kappa_{ii,p,p}(t_i - t_p)(z_i)^2(z_p)^2 = -\frac{1}{4n}\kappa_{ii,p,p}\{(1-2t_i) - (1-2t_p)\}(z_i)^2(z_p)^2.$$

Integrating z_i and z_p out we get a term linear in $1/(1-2t_i)$ and a term linear in $1/(1-2t_p)$.

3. $\kappa_{ij,kl}$: Consider

$$\frac{1}{2n}\kappa_{ij,kl} z_i z_j z_k z_l t_{\max(i,j)} t_{\max(k,l)} + \frac{1}{n}\kappa_{ai,aj} z_i z_j t_{\max(a,i,j)} - \frac{1}{4n}\kappa_{ij,kl} z_i z_j z_k z_l t_p.$$

(i) $\kappa_{pp,pp}$ appears only when $i = j = k = l = p$. In this case by integrating z_p out the coefficient for $\kappa_{pp,pp}$ becomes $t_p/\{4n(1-2t_p)\}$, which is linear in $1/(1-2t_p)$; (ii) $\kappa_{pp,ii}$ with $i < p$ appears in the form

$$\frac{1}{n}\kappa_{pp,ii}(z_p)^2(z_i)^2 t_p t_i - \frac{1}{2n}\kappa_{pp,ii}(z_p)^2(z_i)^2 t_p.$$

Integrating this out, the coefficient for $\kappa_{pp,ii}$ becomes $-t_p/\{2n(1-2t_p)\}$, which is linear in $1/(1-2t_p)$; (iii) $\kappa_{pi,pi}$ with $i < p$ appears in the form

$$\frac{2}{n}\kappa_{pi,pi}(t_p)^2(z_i)^2(z_p)^2 + \frac{1}{n}\kappa_{pi,pi}(z_i)^2 t_p - \frac{1}{n}\kappa_{pi,pi}(z_i)^2(z_p)^2 t_p.$$

Integrating z_i and z_p out, the coefficient for $\kappa_{pi,pi}$ is shown to be 0.

We have now exhausted all relevant terms and completed our proof of Theorem 2.1.

References

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.

Chesher, A. and Smith, R. J. (1995). Bartlett corrections to likelihood ratio tests, *Biometrika*, **82**, 433–436.

Cordeiro, G. M. (1993). General matrix formulae for computing Bartlett corrections, *Statist. Probab. Lett.*, **16**, 11–18.

Ghosh, J. K. and Mukerjee, R. (1991). Characterization of priors under which Bayesian and frequentist Bartlett corrections are equivalent in the multiparameter case, *J. Multivariate Anal.*, **38**, 385–393.

Ghosh, J. K. and Mukerjee, R. (1992). Bayesian and frequentist Bartlett corrections for likelihood ratio and conditional likelihood ratio tests, *J. Roy. Statist. Soc. Ser. B*, **54**, 867–875.

Harris, P. (1986). A note on Bartlett adjustments to likelihood ratio tests, *Biometrika*, **73**, 735–737.

Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution, *Ann. Inst. Statist. Math.*, **29**, 359–378 (Correction: *ibid.* (1987). **39**, p. 681).

Jensen, J. L. (1993). A historical sketch and some new results on the improved log likelihood ratio statistic, *Scand. J. Statist.*, **20**, 1–15.

Lawley, D. N. (1956). A general method of approximating to the distribution of likelihood ratio criteria, *Biometrika*, **43**, 295–303.

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

McCullagh, P. and Cox, D. R. (1986). Invariants and likelihood ratio statistics, *Ann. Statist.*, **14**, 1419–1430.

Skovgaard, I. (1986). A note on the differentiation of cumulants of log likelihood derivatives, *Internat. Statist. Rev.*, **54**, 29–32.

Takemura, A. and Takeuchi, K. (1988). Some results on univariate and multivariate Cornish-Fisher expansion, algebraic properties and validity, *Sankhyā Ser. A*, **50**, 111–136.