95-F-16

A proof of independent Bartlett correctability of nested likelihood ratio tests

by

Akimichi Takemura
University of Tokyo

and

Satoshi Kuriki
The Institute of Statistical Mathematics

May 1995

A proof of independent Bartlett correctability of nested likelihood ratio tests

Akimichi Takemura
Faculty of Economics, University of Tokyo
and
Satoshi Kuriki
The Institute of Statistical Mathematics

May, 1995

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 first shown in Bickel and Ghosh (1990) using a Bayes method. We present a more direct frequentist proof.

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

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), Hayakawa (1987)) gave a proof of Bartlett correctability. Harris (1986) pointed out an incompleteness of Hayakawa's 1977 proof. Cordeiro (1987) showed that Hayakawa's 1977 calculation is consistent with Lawley's result. Furthermore Bickel and Ghosh (1990) gave a proof based on Bayes approach. See Cordeiro (1993) for more recent work.

As in 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 :

$$\Theta_0 \subset \Theta_1 \subset \cdots \subset \Theta_p = \Theta, \quad \text{with} \quad \dim \Theta_j = j. \tag{1}$$

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 already stated and proved in Bickel and Ghosh (1990) using Bayes method. Our proof is more direct frequentist proof. Because 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 precisely. 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

$$\lambda_j = \frac{\max_{\theta \in \Theta_j} \prod_{i=1}^n f(x_i, \theta)}{\max_{\theta \in \Theta_{j-1}} \prod_{i=1}^n f(x_i, \theta)}, \quad j = 1, \ldots, p. \tag{2}$$

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*

$$\begin{aligned} &E_{H_0}[\exp(it_1 2 \log \lambda_1 + \cdots + it_p 2 \log \lambda_p)] \\ &\quad = \prod_{j=1}^{p} \Big[(1 - 2it_j)^{-1/2} \Big(1 + \frac{c_j}{n} \big(\frac{1}{1 - 2it_j} - 1 \big) \Big) \Big] + o(1/n). \end{aligned} \tag{3}$$

Note that up to the order $O(1/n)$ (3) is equivalent to

$$E_{H_0}[\exp(it_1 2\log\lambda_1 + \cdots + it_p 2\log\lambda_p)]$$
$$= \Big(\prod_{j=1}^{p}(1-2it_j)^{-1/2}\Big)\Big(1+\frac{1}{n}\sum_{j=1}^{p}c_j(\frac{1}{1-2it_j}-1)\Big)+o(1/n). \qquad (4)$$

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 (4).

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

$$\ell_{j_1\ldots j_k} = \ell_{j_1\ldots j_k}(x;\theta) = \frac{\partial^k}{\partial\theta^{j_1}\cdots\partial\theta^{j_k}}\log f(x,\theta), \qquad (5)$$

and

$$\mathcal{L}_{j_1\ldots j_k} = \frac{1}{n}\sum_{i=1}^{n}\ell_{j_1\ldots j_k}(x_i;\theta^0), \qquad (6)$$
$$L_{j_1\ldots j_k} = E_{\theta^0}[\ell_{j_1\ldots j_k}(x;\theta^0)], \qquad (7)$$
$$Z_{j_1\ldots j_k} = \sqrt{n}(\mathcal{L}_{j_1\ldots j_k} - L_{j_1\ldots j_k}). \qquad (8)$$

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

$$\kappa_{i_1\ldots i_{m_1}, j_1\ldots j_{m_2}, \ldots, k_1\ldots k_{m_h}} = \mathrm{cum}_\theta(\ell_{i_1\ldots i_{m_1}}, \ell_{j_1\ldots j_{m_2}}, \ldots, \ell_{k_1\ldots k_{m_h}}) \qquad (9)$$
$$L_{i_1\ldots i_{m_1}, j_1\ldots j_{m_2}, \ldots, k_1\ldots k_{m_h}} = E_\theta(\ell_{i_1\ldots i_{m_1}}\ell_{j_1\ldots j_{m_2}}\cdots\ell_{k_1\ldots k_{m_h}}) \qquad (10)$$

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

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

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

$$L_{ijk}+L_{ij,k}[3]+L_{i,j,k}=0, \tag{11}$$
$$\kappa_{ijk}+\kappa_{ij,k}[3]+\kappa_{i,j,k}=0, \tag{12}$$
$$L_{ijkl}+L_{ijk,l}[4]+L_{ij,kl}[3]+L_{ij,k,l}[6]+L_{i,j,k,l}=0, \tag{13}$$
$$\kappa_{ijkl}+\kappa_{ijk,l}[4]+\kappa_{ij,kl}[3]+\kappa_{ij,k,l}[6]+\kappa_{i,j,k,l}=0. \tag{14}$$

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

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

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

$$\mu_{\ldots i\ldots}\mu_{\ldots i\ldots}\ . \tag{15}$$

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

$$\begin{aligned}
\Theta_0 &= \{\ (0,0,\ldots,0)\ \},\\
\Theta_1 &= \{\ (\theta^1,0,\ldots,0)\mid \theta^1:\text{free}\ \},\\
&\ \ \cdots\\
\Theta_{p-1} &= \{\ (\theta^1,\ldots,\theta^{p-1},0)\mid \theta^1,\ldots,\theta^{p-1}:\text{free}\ \},\\
\Theta_p &= \{\ (\theta^1,\ldots,\theta^p)\mid \theta^1,\ldots,\theta^p:\text{free}\ \}.
\end{aligned} \tag{16}$$

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

$$\kappa_{i,j}=-\kappa_{ij}=-L_{ij}=\delta_{ij}, \tag{17}$$

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

$$\theta^i = \tau^i + \frac{1}{2} a^i_{jk} \tau^j \tau^k + \frac{1}{6} a^i_{jkl} \tau^j \tau^k \tau^l + \cdots . \tag{18}$$

Here the coefficients $a^i_{jk}, a^i_{jkl}, \ldots$ are invariant under the permutation of subscripts. Note that the Jacobian of (18) is the identity at the origin and (18) is 1-to-1 in some neighborhood of the origin. Furthermore for our purpose (18) 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 (17) is satisfied. Now we want to choose τ such that (16) remains to be satisfied. Consider Θ_{p-1}. We want

$$\theta^p = 0 \Leftrightarrow \tau^p = 0 \tag{19}$$

for arbitrary values of $\theta^1, \ldots, \theta^{p-1}$. We claim that a necessary and sufficient condition for (19) is

$$a^p_{i_1 \ldots i_k} = 0 \quad \text{if} \quad \max(i_1, \ldots, i_k) < p. \tag{20}$$

Note that (20) holds if and only if θ^p can be written as

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

Then obviously $\tau^p = 0 \Rightarrow \theta^p = 0$. 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 Θ_{p-2}. We want to ensure that

$$(\theta^{p-1}, \theta^p) = (0, 0) \Leftrightarrow (\tau^{p-1}, \tau^p) = (0, 0). \tag{21}$$

A necessary and sufficient condition for (21) is

$$a^{p-1}_{i_1 \ldots i_k} = 0 \quad \text{if} \quad \max(i_1, \ldots, i_k) < p - 1 \tag{22}$$

If (22) holds,

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

for some polynomials A, B. Hence $(\tau^{p-1}, \tau^p) = (0, 0) \Rightarrow (\theta^{p-1}, \theta^p) = (0, 0)$. 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 see that any nonlinear reparameterization of the form (18) satisfying

$$a^{i_1}_{i_2\ldots i_k} = 0 \quad \text{for} \quad \max(i_2, \ldots, i_k) < i_1 \tag{23}$$

satisfies (16). In other words, if $\max(i_2, \ldots, i_k) \geq i_1$ then we can choose the value of $a^{i_1}_{i_2\ldots i_k}$ 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

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

Letting

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

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

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

we obtain at the origin

$$\mathrm{Cov}_{\tau=0}(\ell_{jkl}(x;\tau), \ell_i(x;\tau)) = \kappa_{i,jkl} + a^\alpha_{jk}k_{i,\alpha l}[3] + \delta_{i\alpha}a^\alpha_{jkl}. \tag{24}$$

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, it follows that we can choose parameterization such that

$$\kappa_{i_1,i_2\ldots i_k} = 0 \quad \text{if} \quad \max(i_2, \ldots, i_k) \geq i_1. \tag{25}$$

The simplification in (25) is very useful for calculation of $O(1/n)$ terms needed to prove our result.

3.3 Stochastic expansion of log likelihood ratio

Here give a stochastic expansion of $2\log\lambda$ in terms of the random variables $Z_{i_1\ldots i_k}$ defined in (8). Equivalent expansions are already given for example in Hayakawa (1977) or Section 7.4 of McCullagh (1987), but we give the expansion here for reference. Let $\hat\theta$ be the maximum likelihood estimate. Expanding the likelihood equation

$$\frac{\partial}{\partial\theta^i}\sum_{t=1}^{n}\log f(x_t,\hat\theta)=0$$

around the origin, we get

$$0=\mathcal{L}_i+\hat\theta^j\mathcal{L}_{ij}+\frac{1}{2}\hat\theta^j\hat\theta^k\mathcal{L}_{ijk}+\frac{1}{6}\hat\theta^j\hat\theta^k\hat\theta^l\mathcal{L}_{ijkl}+\cdots. \tag{26}$$

Let $V^i=\sqrt{n}\hat\theta^i$. Multiplying (26) by $\sqrt{n}$, (26) is rewritten as

$$0=Z_i+V^j\mathcal{L}_{ij}+\frac{1}{2\sqrt{n}}V^jV^k\mathcal{L}_{ijk}+\frac{1}{6n}V^jV^kV^l\mathcal{L}_{ijkl}+o_p(1/n). \tag{27}$$

Let $\mathcal{L}^{ij}$ be the inverse (matrix) of $\mathcal{L}_{ij}$. Because

$$\mathcal{L}_{ij}=L_{ij}+\frac{1}{\sqrt{n}}Z_{ij}=-\delta_{ij}+\frac{1}{\sqrt{n}}Z_{ij},$$

$\mathcal{L}^{ij}$ can be expanded as

$$\mathcal{L}^{ij}=-\delta^{ij}-\frac{1}{\sqrt{n}}\delta^{ia}\delta^{jb}Z_{ab}-\frac{1}{n}\delta^{ia}\delta^{jb}\delta^{cd}Z_{ac}Z_{bd}+o_p(1/n). \tag{28}$$

Solving (27) for V^i and substituting (28), V^i can be expressed (using our notational convention) as

$$\begin{aligned}V^i &= Z_i+\frac{1}{\sqrt{n}}Z_{ij}Z_j+\frac{1}{2\sqrt{n}}L_{ijk}Z_jZ_k\\ &\quad+\frac{1}{n}Z_{ia}Z_{ja}Z_j+\frac{1}{2n}Z_{ijk}Z_jZ_k+\frac{3}{2n}L_{jka}Z_{ia}Z_jZ_k\\ &\quad+\frac{1}{6n}(L_{ijkl}+L_{ija}L_{kla}[3])Z_jZ_kZ_l+o_p(1/n).\end{aligned} \tag{29}$$

Now consider log likelihood ratio statistic. For the moment we look at the overall likelihood ratio statistic λ for H_0 vs. H_p. Expanding $2\sum_{t=1}^{n}(\log f(x_t,\hat\theta)-\log f(x_t,0))$ around 0 we obtain

$$\begin{aligned}2\log\lambda &= 2V^iZ_i+V^iV^j\mathcal{L}_{ij}+\frac{1}{3\sqrt{n}}V^iV^jV^k\mathcal{L}_{ijk}+\frac{1}{12n}V^iV^jV^kV^l\mathcal{L}_{ijkl}+o_p(1/n)\\ &= 2V^iZ_i-V^iV^j\delta_{ij}+\frac{1}{\sqrt{n}}V^iV^jZ_{ij}+\frac{1}{3\sqrt{n}}V^iV^jV^kL_{ijk}\\ &\quad+\frac{1}{3n}V^iV^jV^kZ_{ijk}+\frac{1}{12n}V^iV^jV^kV^lL_{ijkl}+o_p(1/n).\end{aligned} \tag{30}$$

Substituting (29) into (30), the stochastic expansion of $2\log\lambda$ in terms of Z 's is obtained as follows:

$$
\begin{aligned}
2\log\lambda &= Z_iZ_i + \frac{1}{\sqrt{n}}Z_{ij}Z_iZ_j + \frac{1}{3\sqrt{n}}L_{ijk}Z_iZ_jZ_k + \frac{1}{n}Z_{ia}Z_{ja}Z_iZ_j \\
&\quad + \frac{1}{3n}Z_{ijk}Z_iZ_jZ_k + \frac{1}{n}L_{jka}Z_{ia}Z_iZ_jZ_k \\
&\quad + \frac{1}{12n}(L_{ijkl} + L_{ija}L_{kla}[3])Z_iZ_jZ_kZ_l + o_p(1/n).
\end{aligned}
\tag{31}
$$

Note that (31) is the expansion for the overall log likelihood ratio λ , hence the range of the running variables $i, j, k, a, \ldots$, are 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 (31) by the following simple argument. Consider the likelihood ratio statistic λ_{0r} for H_0 vs. H_r . Because of (16), the stochastic expansion for $2\log\lambda_{0r}$ is the same as in (31) except for the range of running variables, 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 (31), 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

$$
\begin{aligned}
2t_i\log\lambda_i &= t_iZ_iZ_i + \frac{1}{\sqrt{n}}Z_{ij}Z_iZ_jt_{\max(i,j)} + \frac{1}{3\sqrt{n}}L_{ijk}Z_iZ_jZ_kt_{\max(i,j,k)} \\
&\quad + \frac{1}{n}Z_{ia}Z_{ja}Z_iZ_jt_{\max(i,j,a)} + \frac{1}{3n}Z_{ijk}Z_iZ_jZ_kt_{\max(i,j,k)} \\
&\quad + \frac{1}{n}L_{jka}Z_{ia}Z_iZ_jZ_kt_{\max(i,j,k,a)} + \frac{1}{12n}(L_{ija}L_{kla}[3])Z_iZ_jZ_kZ_lt_{\max(i,j,k,l,a)} \\
&\quad + \frac{1}{12n}L_{ijkl}Z_iZ_jZ_kZ_lt_{\max(i,j,k,l)} + o_p(1/n).
\end{aligned}
\tag{32}
$$

3.4 Evaluation of the joint characteristic function

We now evaluate the joint characteristic function of $\lambda_1, \ldots, \lambda_p$ using the stochastic expansion (32). From now on we omit the pure imaginary number i for convenience and write $2t_j\log\lambda_j$ instead of $2it_j\log\lambda_j$. By doing this we can use i again for index. We take the expectation in two steps. First we consider conditional expectation given the first order derivatives $Z_1, \ldots, Z_p$ and then we evaluate the expectation with respect to $Z_1, \ldots, Z_p$:

$$E[\exp(2t_j\log\lambda_j)] = E[E(\exp(2t_j\log\lambda_j)|Z_1, \ldots, Z_p)]. \tag{33}$$

For the first step we need the conditional expectation $E[Z_{ij}|Z_1, \ldots, Z_p]$ to the order $O(n^{-1/2})$. For this we use the following lemma.

Lemma 3.1 *Let* $Y = Y(n) = (Y_1, \ldots, Y_m)$ *be an asymptotically normal random vector with* $E(Y_i) = 0, E(Y_iY_j) = \delta_{ij}, \mathrm{cum}(Y_i, Y_j, Y_k) = n^{-1/2}\kappa_{i,j,k}$. *Then*

$$E(Y_m|Y_1, \ldots, Y_{m-1}) = \frac{1}{2\sqrt{n}} \sum_{a,b=1}^{m-1} \kappa_{m,a,b}(Y_aY_b - \delta_{ab}) + o(n^{-1/2}). \tag{34}$$

Proof. From multivariate Edgeworth expansion of the density function of Y we obtain

$$\begin{aligned} f(y_m|y_1, \ldots, y_{m-1}) &= \frac{f(y_1, \ldots, y_m)}{f(y_1, \ldots, y_{m-1})} \\ &= \phi(y_m)\{1 + \frac{1}{6\sqrt{n}} \sum_{\max(a,b,c)=m} \kappa_{a,b,c}(y_ay_by_c - \delta_{ab}y_c[3])\} \\ &\quad +o(n^{-1/2}). \end{aligned}$$

Therefore

$$\begin{aligned} E(y_m|y_1, \ldots, y_{m-1}) &= \int_{-\infty}^{\infty} y_m f(y_m|y_1, \ldots, y_{m-1})dy_m + o(n^{-1/2}) \\ &= \frac{1}{2\sqrt{n}} \sum_{a,b=1}^{m-1} \kappa_{m,a,b}(y_ay_b - \delta_{ab}) + o(n^{-1/2}). \end{aligned}$$

Now consider $Z_1, \ldots, Z_p$ and $s^{-1}(Z_{ij} - \kappa_{k,ij}Z_k)$ with $s^2 = \mathrm{Var}(Z_{ij} - \kappa_{k,ij}Z_k)$. These $p+1$ variables are uncorrelated and satisfy condition of Lemma 3.1. Therefore

$$\begin{aligned} &E(s^{-1}(Z_{ij} - \kappa_{k,ij}Z_k)|Z_1, \ldots, Z_p) \\ &\quad = \frac{1}{2}\mathrm{cum}(s^{-1}(Z_{ij} - \kappa_{k,ij}Z_k), Z_a, Z_b)(Z_aZ_b - \delta_{ab}) + o(n^{-1/2}) \\ &\quad = \frac{1}{2\sqrt{n}}s^{-1}(\kappa_{ij,a,b} - \kappa_{k,ij}\kappa_{k,a,b})(Z_aZ_b - \delta_{ab}) + o(n^{-1/2}), \end{aligned}$$

or

$$E(Z_{ij}|Z_1, \ldots, Z_p) = \kappa_{k,ij}Z_k + \frac{1}{2\sqrt{n}}(\kappa_{ij,a,b} - \kappa_{k,ij}\kappa_{k,a,b})(Z_aZ_b - \delta_{ab}) + o(n^{-1/2}). \tag{35}$$

For our purpose the conditional expectation appearing in terms of the order $O(1/n)$ can be evaluated as if $Z_{ij...k}$ are normal random variables. Therefore for example it suffices to write

$$\begin{aligned} &E(Z_{ij}Z_{kl}|Z_1, \ldots, Z_p) \\ &\quad = E(Z_{ij}|Z_1, \ldots, Z_p)E(Z_{kl}|Z_1, \ldots, Z_p) + \mathrm{Cov}(Z_{ij}, Z_{kl}|Z_1, \ldots, Z_p) \\ &\quad = \kappa_{a,ij}Z_a\kappa_{b,kl}Z_b + \kappa_{ij,kl} - \kappa_{a,ij}\kappa_{a,kl} + o(1). \end{aligned} \tag{36}$$

For the last equality we used the usual formula on residual variance.

Now we can carry out the calculation of the conditional expectation of the joint characteristic function.

$$
\begin{aligned}
&E(\exp(2t_i \log \lambda_i)|Z_1, \dots, Z_p) \qquad (37)\\
&= E[\exp(t_i Z_i Z_i)\{1 + \frac{1}{\sqrt{n}} Z_{ij} Z_i Z_j t_{\max(i,j)} + \frac{1}{3\sqrt{n}} L_{ijk} Z_i Z_j Z_k t_{\max(i,j,k)}\\
&+ \frac{1}{2n} Z_{ij} Z_i Z_j Z_{kl} Z_k Z_l t_{\max(i,j)} t_{\max(k,l)}\\
&+ \frac{1}{18n} L_{ijk} Z_i Z_j Z_k L_{lmh} Z_l Z_m Z_h t_{\max(i,j,k)} t_{\max(l,m,h)}\\
&+ \frac{1}{3n} Z_{ij} Z_i Z_j L_{klm} Z_k Z_l Z_m t_{\max(i,j)} t_{\max(k,l,m)}\\
&+ \frac{1}{n} Z_{ia} Z_{ja} Z_i Z_j t_{\max(i,j,a)} + \frac{1}{3n} Z_{ijk} Z_i Z_j Z_k t_{\max(i,j,k)}\\
&+ \frac{1}{n} L_{jka} Z_{ia} Z_i Z_j Z_k t_{\max(i,j,k,a)} + \frac{1}{12n}(L_{ija} L_{kla}[3]) Z_i Z_j Z_k Z_l t_{\max(i,j,k,l,a)}\\
&+ \frac{1}{12n} L_{ijkl} Z_i Z_j Z_k Z_l t_{\max(i,j,k,l)}\} \mid Z_1, \dots, Z_p] + o(1/n) \qquad (38)\\
&= \exp(t_i Z_i Z_i)\{1 + \frac{1}{3\sqrt{n}} L_{ijk} Z_i Z_j Z_k t_{\max(i,j,k)} + \frac{1}{\sqrt{n}} \kappa_{k,ij} Z_i Z_j Z_k t_{\max(i,j)}\\
&+ \frac{1}{2n}(\kappa_{ij,k,l} - \kappa_{m,ij}\kappa_{m,k,l})(Z_k Z_l - \delta_{kl}) Z_i Z_j t_{\max(i,j)}\\
&+ \frac{1}{2n} Z_i Z_j Z_k Z_l (\kappa_{a,ij}\kappa_{b,kl} Z_a Z_b + \kappa_{ij,kl} - \kappa_{a,ij}\kappa_{a,kl}) t_{\max(i,j)} t_{\max(k,l)}\\
&+ \frac{1}{18n} L_{ijk} L_{lmh} Z_i Z_j Z_k Z_l Z_m Z_h t_{\max(i,j,k)} t_{\max(l,m,h)}\\
&+ \frac{1}{3n} Z_i Z_j Z_k Z_l Z_m \kappa_{a,ij} Z_a L_{klm} t_{\max(i,j)} t_{\max(k,l,m)}\\
&+ \frac{1}{n} Z_i Z_j t_{\max(i,j,a)} (\kappa_{c,ia}\kappa_{d,ja} Z_c Z_d + \kappa_{ai,aj} - \kappa_{c,ia}\kappa_{c,ja})\\
&+ \frac{1}{3n} \kappa_{a,ijk} Z_a Z_i Z_j Z_k t_{\max(i,j,k)}\\
&+ \frac{1}{n} L_{jka} \kappa_{b,ia} Z_i Z_j Z_k Z_b t_{\max(i,j,k,a)}\\
&+ \frac{1}{12n}(L_{aij} L_{akl}[3]) Z_i Z_j Z_k Z_l t_{\max(i,j,k,l,a)}\\
&+ \frac{1}{12n} L_{ijkl} Z_i Z_j Z_k Z_l t_{\max(i,j,k,l)}\} + o(1/n). \qquad (39)
\end{aligned}
$$

It remains to take expectation of (39) with respect to $Z_1, \dots, Z_p$.

Remark 3.1 *Compared to the proof of Bartlett correctability of just the overall likelihood ratio statistic* λ *, we have added complexity in (39) due to the fact that* t *'s are indexed by the maximum of the individual indices. Therefore we need to keep track of the maximum value of the indices in the calculation below.*

Now we rearrange (39) as follows. Let

$$\begin{aligned}
A_1 &= \frac{1}{3\sqrt{n}} L_{ijk} z_i z_j z_k t_{\max(i,j,k)}, \\
A_2 &= \frac{1}{\sqrt{n}} \kappa_{k,ij} z_i z_j z_k t_{\max(i,j)}, \\
B_1 &= \frac{1}{2n} (\kappa_{ij,k,l} - \kappa_{m,ij}\kappa_{m,k,l})(z_k z_l - \delta_{kl}) z_i z_j t_{\max(i,j)}, \\
B_2 &= \frac{1}{2n} z_i z_j z_k z_l (\kappa_{ij,kl} - \kappa_{a,ij}\kappa_{a,kl}) t_{\max(i,j)} t_{\max(k,l)}, \\
B_3 &= \frac{1}{n} z_i z_j t_{\max(i,j,a)} (\kappa_{c,ia}\kappa_{d,ja} z_c z_d + \kappa_{ai,aj} - \kappa_{c,ia}\kappa_{c,ja}), \\
B_4 &= \frac{1}{3n} \kappa_{a,ijk} z_a z_i z_j z_k t_{\max(i,j,k)}, \\
B_5 &= \frac{1}{n} L_{jka} \kappa_{b,ia} z_i z_j z_k z_b t_{\max(i,j,k,a)}, \\
B_6 &= \frac{1}{12n} (L_{aij} L_{akl}[3]) z_i z_j z_k z_l t_{\max(i,j,k,l,a)}, \\
B_7 &= \frac{1}{12n} L_{ijkl} z_i z_j z_k z_l t_{\max(i,j,k,l)}. \qquad (40)
\end{aligned}$$

Then we can write

$$\begin{aligned}
&E[\exp(2t_i \log \lambda_i \mid z_1, \ldots, z_p)] \\
&\quad = \exp(t_i z_i z_i + A_1 + A_2) \times (1 + B_1 + B_2 + \cdots + B_7) + o(1/n). \qquad (41)
\end{aligned}$$

We combine (41) 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)).

$$\begin{aligned}
f(z_1, \ldots, z_p) &= \frac{1}{(2\pi)^{p/2}} \exp[-\frac{1}{2} z_i z_i + \frac{1}{6\sqrt{n}} \kappa_{i,j,k} z_i z_j z_k + \frac{1}{\sqrt{n}} q_1(z) \\
&\quad + \frac{1}{24n} (\kappa_{i,j,k,l} - \kappa_{i,j,a}\kappa_{k,l,a}[3]) z_i z_j z_k z_l + \frac{1}{n} q_2(z)] + o(1/n), \qquad (42)
\end{aligned}$$

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

$$\begin{aligned}
C_1 &= \frac{1}{6\sqrt{n}} \kappa_{i,j,k} z_i z_j z_k + \frac{1}{\sqrt{n}} q_1(z), \\
C_2 &= \frac{1}{24n} (\kappa_{i,j,k,l} - \kappa_{i,j,a}\kappa_{k,l,a}[3]) z_i z_j z_k z_l + \frac{1}{n} q_2(z). \qquad (43)
\end{aligned}$$

We combine (41) and (42) and our problem is reduced to evaluating the following integration term by term:

$$\begin{aligned}
E(\exp(2t_i \log \lambda_i)) &= \int \cdots \int \exp(-\frac{1}{2}(1 - 2t_i) z_i z_i + A_1 + A_2 + C_1) \\
&\quad \times \{1 + B_1 + B_2 + \cdots + B_7 + C_2\} dz_1 \ldots dz_p + o(1/n). \qquad (44)
\end{aligned}$$

At this point the following simple recursive argument is useful.

Lemma 3.2 *In order to prove Theorem 2.1 it is sufficient to prove that all the* $O(1/n)$ *terms containing* t_p *in (44) 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

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

Now put $t_p = 0$. Then because of the recursive nature of the subspaces in (16), we have exactly the same problem with dimensionality reduced by 1. Therefore

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

This recursive argument implies (4).

Remark 3.2 *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.*

For the term by term integration we just use

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} y^k \exp(-\frac{1}{2}y^2(1-2t))dy = \begin{cases} \frac{1\cdot 3\cdots(k-1)}{(1-2t)^{(k+1)/2}} & k: \text{even}, \\ 0 & k: \text{odd}. \end{cases} \tag{45}$$

Our proof now consists of exhaustive verification of each term of (44) 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 obtain

$$\begin{aligned} A_1 + A_2 + C_1 &= \frac{1}{6\sqrt{n}} \kappa_{i,j,k} z_i z_j z_k (1 - 2t_{\max(i,j,k)}) \\ &\quad + \frac{1}{6\sqrt{n}} (\kappa_{i,jk}[3]) z_i z_j z_k (1 - 2t_{\max(i,j,k)}) \\ &\quad - \frac{1}{2\sqrt{n}} \kappa_{k,ij} z_i z_j z_k (1 - 2t_{\max(i,j)}) + \frac{1}{\sqrt{n}} q_1(z) \\ &= D_1 + D_2 + D_3 + \frac{1}{\sqrt{n}} q_1(z), \end{aligned}$$

where

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

Then we again expand $\exp(D_1 + D_2 + D_3 + q_1(z)/\sqrt{n})$ as

$$
1 + \sum D_i + q_1(z)/\sqrt{n} + \frac{1}{2} \sum (D_i)^2 + \sum_{i<j} D_i D_j + \frac{1}{\sqrt{n}} q_1(z) \sum D_i + \frac{1}{2n} q_1(z)^2 + o(1/n).
$$

Note that $\sum D_i + q_1(z)/\sqrt{n}$ is odd polynomial in z and this vanishes by integration. Furthermore the index for t agrees with one of i, j, k and hence in view of (45) integration of terms $q_1(z) \sum D_i$ yields only linear terms in $1/(1 - 2t_a)$, $a = 1, \ldots, p$. Also $q_1(z)^2$ is quadratic in z and yields only linear terms in $1/(1 - 2t_a)$, $a = 1, \ldots, p$. 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 integration 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 cases where each distinct indices appears even times.

Suppose that $\alpha \neq \beta$. Then both α and β have to appear at least twice. In view of (45) these lead to terms linear in $1/(1 - 2t_a)$, $a = 1, \ldots, p$. 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 $1/(1 - 2t_a)$, $a = 1, \ldots, p,$ 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

$$
(1 - 2t_\alpha) z_a z_b z_c z_d \qquad (a, b, c, d \neq \alpha). \tag{46}
$$

We have to verity that terms of this type in D_iD_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. $(D_1)^2$:

$$\frac{1}{2}(D_1)^2 = \frac{1}{72n}\kappa_{i,j,k}\kappa_{l,m,h}z_iz_jz_kz_lz_mz_h(1-2t_{\max(i,j,k)})(1-2t_{\max(l,m,h)}).$$

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.

$$\frac{1}{8n}\kappa_{p,i,j}\kappa_{p,k,l}(1-2t_p)z_iz_jz_kz_l \qquad (i,j,k,l<p). \tag{47}$$

2. $(D_2)^2$:

Noting $\kappa_{i,pj}=0$ by (25), similar reasoning applied to $(D_2)^2/2$ yields the following term yet to be canceled.

$$\frac{1}{8n}\kappa_{p,ij}\kappa_{p,kl}(1-2t_p)z_iz_jz_kz_l \qquad (i,j,k,l<p) \tag{48}$$

3. D_1D_2 :

Similarly D_1D_2 yields

$$\frac{1}{4n}\kappa_{p,i,j}\kappa_{p,kl}(1-2t_p)z_iz_jz_kz_l \qquad (i,j,k,l<p). \tag{49}$$

4. $(D_3)^2$:

Now let us take a look at $(D_3)^2/2$. In

$$\frac{1}{2}(D_3)^2 = \frac{1}{8n}\kappa_{k,ij}\kappa_{l,mh}z_iz_jz_kz_lz_mz_h(1-2t_{\max(i,j)})(1-2t_{\max(m,h)}) \tag{50}$$

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 (50) are

$$\begin{aligned}&\frac{1}{8n}\kappa_{k,aa}\kappa_{l,aa}z_kz_l(z_a)^4(1-2t_a)^2\\ &\qquad\qquad\qquad (k,l\le p \text{ and } a<p)\\ &+\frac{1}{2n}\kappa_{k,ai}\kappa_{l,am}z_kz_lz_iz_m(z_a)^2(1-2t_a)^2\\ &\qquad\qquad\qquad (k,l\le p \text{ and } i,m<a<p).\end{aligned}$$

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

$$\begin{aligned}&\frac{1}{8n}\kappa_{p,aa}\kappa_{p,aa}(z_p)^2(z_a)^4(1-2t_a)^2 \qquad (a<p)\\ &+\frac{1}{2n}\kappa_{p,ai}\kappa_{p,am}z_iz_m(z_p)^2(z_a)^2(1-2t_a)^2 \qquad (i,m<a<p).\end{aligned}$$

Integrating z_p and z_a out yields

$$\begin{aligned}&\frac{3}{8n}\kappa_{p,aa}\kappa_{p,aa}\frac{1}{1-2t_p} \qquad (a<p)\\ &+\frac{1}{2n}\kappa_{p,ai}\kappa_{p,am}z_iz_m\frac{1-2t_a}{1-2t_p} \qquad (i,m<a<p).\end{aligned}$$

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

$$\frac{1}{2n}\kappa_{p,ai}\kappa_{p,am}z_i z_m \frac{1-2t_a}{1-2t_p} \qquad (i,m<a<p). \tag{51}$$

5. D_1D_3:

   D_1D_3 is irrelevant. Actually in

   $$\kappa_{i,j,k}z_i z_j z_k(1-2t_{\max(i,j,k)})\kappa_{l,mh}z_l z_m z_h(1-2t_{\max(m,h)})$$

   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.

Summarizing above examination gives the following list of terms yet to be canceled.

$$[(D_1)^2] \qquad \frac{1}{8n}\kappa_{p,i,j}\kappa_{p,k,l}(1-2t_p)z_i z_j z_k z_l \qquad (i,j,k,l<p), \tag{52}$$

$$[(D_2)^2] \qquad \frac{1}{8n}\kappa_{p,ij}\kappa_{p,kl}(1-2t_p)z_i z_j z_k z_l \qquad (i,j,k,l<p), \tag{53}$$

$$[D_1D_2] \qquad \frac{1}{4n}\kappa_{p,i,j}\kappa_{p,kl}(1-2t_p)z_i z_j z_k z_l \qquad (i,j,k,l<p), \tag{54}$$

$$[(D_3)^2] \qquad \frac{1}{2n}\kappa_{p,ai}\kappa_{p,am}z_i z_m \frac{1-2t_a}{1-2t_p} \qquad (i,m<a<p). \tag{55}$$

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

   $$-\frac{1}{2n}\kappa_{m,ij}\kappa_{m,k,l}(z_k z_l-\delta_{kl})z_i z_j t_{\max(i,j)}$$
   $$=-\frac{1}{2n}\kappa_{m,ij}\kappa_{m,k,l}z_k z_l z_i z_j t_{\max(i,j)}+\frac{1}{2n}\kappa_{m,ij}\kappa_{m,k,k}z_i z_j t_{\max(i,j)}$$

   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 integrating z_p out we are left with

   $$-\frac{1}{2n}\kappa_{m,ij}\kappa_{m,p,p}\frac{t_{\max(i,j)}}{1-2t_p}z_i z_j \qquad (m\le p \text{ and } i,j<p). \tag{56}$$

2. B_2:

   B_2 is irrelevant. Actually if t_p appears from

   $$z_i z_j z_k z_l \kappa_{a,ij}\kappa_{a,kl}t_{\max(i,j)}t_{\max(k,l)}$$

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

3. B_3 :

$$\frac{1}{n} z_i z_j z_c z_d t_{\max(i,j,a)} \kappa_{c,ia} \kappa_{d,ja} - \frac{1}{n} z_i z_j t_{\max(a,i,j)} \kappa_{c,ia} \kappa_{c,ja}.$$

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 $c = d = p$. Integrating z_p out, the remaining term to be canceled is

$$\frac{1}{n} \kappa_{p,ai} \kappa_{p,aj} \frac{t_{\max(a,i,j)}}{1-2t_p} z_i z_j \qquad (a,i,j < p) \tag{57}$$

4. B_4 :

   B_4 does 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

$$\frac{2}{n} L_{apj} \kappa_{p,aj} (z_p)^2 (z_j)^2 t_p.$$

Integrating z_p out we obtain

$$\frac{1}{n} \frac{2t_p}{1-2t_p} L_{apj} \kappa_{p,aj} (z_j)^2.$$

Since $L_{apj} = -\kappa_{p,aj} - \kappa_{p,a,j}$ the remaining term to be canceled is

$$-\frac{1}{n} (\kappa_{p,aj}\kappa_{p,aj} + \kappa_{p,aj}\kappa_{p,a,j}) \frac{2t_p}{1-2t_p} (z_j)^2 \qquad (a,j < p). \tag{58}$$

6. B_6 :

   B_6 is the hardest term to look at. If $\max(i,j,k,l,a) < p$ in

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

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 $L_{app} = -\kappa_{a,p,p}$ and we have

$$\frac{3}{12n} L_{app} L_{app} (z_p)^4 t_p = \frac{1}{4n} \kappa_{a,p,p} \kappa_{a,p,p} (z_p)^4 t_p.$$

Integrating z_p out, the remaining term to be canceled is

$$\frac{3}{4n}\frac{t_p}{(1-2t_p)^2}\kappa_{a,p,p}\kappa_{a,p,p} \qquad (a \le p). \tag{59}$$

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: 1) $i,j,k,l < p$, or 2) two of i,j,k,l equal to p.

For the subcase 1), 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 the remaining term to be canceled is

$$\frac{1}{12n}(L_{aij}L_{akl}[3])z_i z_j z_k z_l t_p = \frac{1}{4n}(\kappa_{p,ij}+\kappa_{p,i,j})(\kappa_{p,kl}+\kappa_{p,k,l})z_i z_j z_k z_l t_p \quad (i,j,k,l < p). \tag{60}$$

For the subcase 2), let $k,l < i = j = p$ and $a \le p$. Then

$$L_{app} = -\kappa_{a,p,p}, \quad L_{akl} = -\kappa_{a,kl} - \kappa_{k,al} - \kappa_{l,ak} - \kappa_{a,k,l}$$
$$L_{apk} = -\kappa_{p,ak} - \kappa_{a,p,k}, \quad L_{apl} = -\kappa_{p,al} - \kappa_{a,p,l}$$

Therefore for this case

$$\begin{aligned}\frac{1}{12n}(L_{aij}L_{akl}[3])z_i z_j z_k z_l t_p \quad &= \quad \frac{1}{12n}\Big\{\kappa_{a,p,p}(\kappa_{a,kl}+\kappa_{k,al}+\kappa_{l,ak}+\kappa_{a,k,l}) \\ &\quad + 2(\kappa_{p,ak}+\kappa_{a,p,k})(\kappa_{p,al}+\kappa_{a,p,l})\Big\}t_p(z_p)^2 z_k z_l\end{aligned}$$

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

$$\begin{aligned}&\frac{1}{2n}\frac{t_p}{1-2t_p}\Big\{\kappa_{a,p,p}(\kappa_{a,ij}+\kappa_{i,aj}+\kappa_{j,ai}+\kappa_{a,i,j}) \\ &+ 2(\kappa_{p,ai}+\kappa_{a,p,i})(\kappa_{p,aj}+\kappa_{a,p,j})\Big\}z_i z_j \qquad (i,j < p \text{ and } a \le p).\end{aligned} \tag{61}$$

We saw 3 types of terms (59), (60) and (61) to be canceled from B_6.

7. B_7:

   B_7 does not contain third order cumulants.

8. 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

$$-\frac{3}{8n}\kappa_{p,p,a}\kappa_{p,p,a}\frac{1}{(1-2t_p)^2} \qquad (a \le p). \tag{62}$$

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

$$-\frac{1}{24n}(\kappa_{p,p,a}\kappa_{k,l,a}+2\kappa_{p,l,a}\kappa_{p,k,a})z_l z_k\frac{1}{1-2t_p}. \qquad (k,l<p \text{ and } a\le p).$$

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

$$-\frac{1}{4n}(\kappa_{p,p,a}\kappa_{i,j,a}+2\kappa_{p,i,a}\kappa_{p,j,a})z_i z_j\frac{1}{1-2t_p}. \qquad (i,j<p \text{ and } a\le p) \tag{63}$$

We have enumerated all the remaining terms containing third order cumulants. Our list of these terms (including terms coming from D_iD_j 's) is as follows.

$[(D_1)^2]$ $$\frac{1}{8n}\kappa_{p,i,j}\kappa_{p,k,l}(1-2t_p)z_iz_jz_kz_l \qquad (i,j,k,l<p) \tag{64}$$

$[(D_2)^2]$ $$\frac{1}{8n}\kappa_{p,ij}\kappa_{p,kl}(1-2t_p)z_iz_jz_kz_l \qquad (i,j,k,l<p) \tag{65}$$

$[D_1D_2]$ $$\frac{1}{4n}\kappa_{p,i,j}\kappa_{p,kl}(1-2t_p)z_iz_jz_kz_l \qquad (i,j,k,l<p) \tag{66}$$

$[(D_3)^2]$ $$\frac{1}{2n}\kappa_{p,ai}\kappa_{p,am}z_iz_m\frac{1-2t_a}{1-2t_p} \qquad (i,m<a<p). \tag{67}$$

$[B_1]$ $$-\frac{1}{2n}\kappa_{m,ij}\kappa_{m,p,p}\frac{t_{\max(i,j)}}{1-2t_p}z_iz_j \qquad (m\le p \text{ and } i,j<p) \tag{68}$$

$[B_3]$ $$\frac{1}{n}\kappa_{p,ai}\kappa_{p,aj}\frac{t_{\max(a,i,j)}}{1-2t_p}z_iz_j \qquad (a,i,j<p) \tag{69}$$

$[B_5]$ $$-\frac{1}{n}(\kappa_{p,aj}\kappa_{p,aj}+\kappa_{p,aj}\kappa_{p,a,j})\frac{2t_p}{1-2t_p}(z_j)^2 \qquad (a,j<p) \tag{70}$$

$[B_6]$ $$\frac{3}{4n}\frac{t_p}{(1-2t_p)^2}\kappa_{a,p,p,}\kappa_{a,p,p} \qquad (a\le p) \tag{71}$$

$[B_6]$ $$\frac{1}{4n}(\kappa_{p,ij}+\kappa_{p,i,j})(\kappa_{p,kl}+\kappa_{p,k,l})z_iz_jz_kz_lt_p \qquad (i,j,k,l<p) \tag{72}$$

$[B_6]$ $$\frac{1}{2n}\frac{t_p}{1-2t_p}\Big\{\kappa_{a,p,p}(\kappa_{a,ij}+\kappa_{i,aj}+\kappa_{j,ai}+\kappa_{a,i,j}) +2(\kappa_{p,ai}+\kappa_{a,p,i})(\kappa_{p,aj}+\kappa_{a,p,j})\Big\}z_iz_j \qquad (i,j<p \text{ and } a\le p) \tag{73}$$

$[C_2]$ $$-\frac{3}{8n}\kappa_{p,p,a}\kappa_{p,p,a}\frac{1}{(1-2t_p)^2} \qquad (a\le p) \tag{74}$$

$[C_2]$ $$-\frac{1}{4n}(\kappa_{p,p,a}\kappa_{i,j,a}+2\kappa_{p,i,a}\kappa_{p,j,a})z_iz_j\frac{1}{1-2t_p} \qquad (i,j<p \text{ and } a\le p). \tag{75}$$

Adding together (64), (65), (66), and (72), we see that t_p vanishes.

Sum of (71) and (74) reduces to a linear term in $1/(1-2t_p)$.

Adding (75) to (73) cancels some terms in (73) and (73) is reduced to

$$\frac{1}{2n}\frac{t_p}{1-2t_p}\{\kappa_{a,p,p}(\kappa_{a,ij}+\kappa_{i,aj}+\kappa_{j,ai})$$
$$+2(\kappa_{p,ai}\kappa_{p,aj}+\kappa_{p,ai}\kappa_{a,p,j}+\kappa_{a,p,i}\kappa_{p,aj})\}z_iz_j \qquad (i,j<p \text{ and } a\le p)$$

Because only the the case $i=j$ matters and $\kappa_{i,ai}=0$, (73) can be further reduced to

$$\frac{1}{2n}\frac{t_p}{1-2t_p}(\kappa_{a,p,p}\kappa_{a,ii}+2\kappa_{p,ai}\kappa_{p,ai}+4\kappa_{p,ai}\kappa_{a,p,i})(z_i)^2 \qquad (i<p \text{ and } a\le p). \qquad (76)$$

Changing some indices and letting $i=j$ in (67), (68), (69), (70), our reduced list of the remaining terms is now

$[(D_3)^2]$ $$\frac{1}{2n}\kappa_{p,ai}\kappa_{p,ai}(z_i)^2\frac{1-2t_a}{1-2t_p} \qquad (i<a<p). \qquad (77)$$

$[B_1]$ $$-\frac{1}{2n}\kappa_{a,ii}\kappa_{a,p,p}\frac{t_i}{1-2t_p}(z_i)^2 \qquad (a\le p \text{ and } i<p) \qquad (78)$$

$[B_3]$ $$\frac{1}{n}\kappa_{p,ai}\kappa_{p,ai}\frac{t_{\max(a,i)}}{1-2t_p}(z_i)^2 \qquad (a,i<p) \qquad (79)$$

$[B_5]$ $$-\frac{1}{n}(\kappa_{p,ai}\kappa_{p,ai}+\kappa_{p,ai}\kappa_{p,a,i})\frac{2t_p}{1-2t_p}(z_i)^2 \qquad (i,a<p) \qquad (80)$$

$[B_6]$ $$\frac{1}{2n}\frac{t_p}{1-2t_p}(\kappa_{a,p,p}\kappa_{a,ii}+2\kappa_{p,ai}\kappa_{p,ai}+4\kappa_{p,ai}\kappa_{a,p,i})(z_i)^2$$
$$(i<p \text{ and } a\le p). \qquad (81)$$

Rewrite (78) as

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

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 (81). Then the first term within the parentheses of (81) no longer contains t_p. Now add (80) to (81). Then (81) is reduced to

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

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

Now our further reduced list is

$[(D_3)^2]$ $$\frac{1}{2n}\kappa_{p,ai}\kappa_{p,ai}(z_i)^2\frac{1-2t_a}{1-2t_p} \qquad (i<a<p). \qquad (82)$$

$[B_3]$ $$\frac{1}{n}\kappa_{p,ai}\kappa_{p,ai}\frac{t_{\max(a,i)}}{1-2t_p}(z_i)^2 \qquad (a,i<p) \qquad (83)$$

$[B_6]$ $$-\frac{1}{n}\frac{t_p}{1-2t_p}\kappa_{p,ai}\kappa_{p,ai}(z_i)^2 \qquad (i,a<p) \qquad (84)$$

Write

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

in (83). Then the first term cancels (84). The second term is irrelevant. Now consider the third term:

$$-\frac{1}{2n}\kappa_{p,ai}\kappa_{p,ai}(z_i)^2\frac{1-2t_{\max(a,i)}}{1-2t_p}. \tag{85}$$

If $i < a$ then $\max(a,i) = a$ and this cancels (82). We are left with the case $i \geq a$. Then (85) is

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

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 yields 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.

$$[B_1] \qquad \frac{1}{2n}\kappa_{ij,k,l}(z_k z_l - \delta_{kl})z_i z_j t_{\max(i,j)} \tag{86}$$

$$[B_2] \qquad \frac{1}{2n}z_i z_j z_k z_l \kappa_{ij,kl} t_{\max(i,j)} t_{\max(k,l)} \tag{87}$$

$$[B_3] \qquad \frac{1}{n}z_i z_j t_{\max(i,j,a)}\kappa_{ai,aj} \tag{88}$$

$$[B_7] \qquad \frac{1}{12n}L_{ijkl} z_i z_j z_k z_l t_{\max(i,j,k,l)} \tag{89}$$

$$[C_2] \qquad \frac{1}{24n}\kappa_{i,j,k,l} z_i z_j z_k z_l \tag{90}$$

Using

$$L_{ijkl} = -\kappa_{ijk,l}[4] - \kappa_{ij,kl}[3] - \kappa_{ij,k,l}[6] - \kappa_{i,j,k,l}$$

expand (89). If $\max(i,j,k,l) < p$ in (89), then t_p can not appear. Therefore we only need to consider the case $\max(i,j,k,l) = p$ in (89). Then there are at least two p's among i,j,k,l and $\kappa_{ijk,l}[4] = 0$. Therefore in (89) we can let

$$L_{ijkl} = -\kappa_{ij,kl}[3] - \kappa_{ij,k,l}[6] - \kappa_{i,j,k,l} \qquad (\max(i,j,k,l) = p)$$

and reduce (89) to the following form

$$-\frac{1}{4n}\kappa_{ij,kl} z_i z_j z_k z_l t_p - \frac{1}{2n}\kappa_{ij,k,l} z_i z_j z_k z_l t_p - \frac{1}{12n}\kappa_{i,j,k,l} z_i z_j z_k z_l t_p$$
$$(\max(i,j,k,l) = p). \tag{91}$$

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 (91) to (90) we have

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

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

2. $\kappa_{ij,k,l}$:

First we take care of δ_{kl} in (86). Consider

$$\kappa_{ij,k,k}z_iz_jt_{\max(i,j)}$$

This obviously yields terms linear in $1/(1-2t_i)$.

In (86) and (91) we are left with

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

The right hand side is non zero only if $p>\max(i,j)$. Because the indices have to appear in pairs, then $p>i=j$. 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 terms linear in $1/(1-2t_i)$ and terms linear in $1/(1-2t_p)$.

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

Consider

$$\frac{1}{2n}\kappa_{ij,kl}z_iz_jz_kz_lt_{\max(i,j)}t_{\max(k,l)}+\frac{1}{n}\kappa_{ai,aj}z_iz_jt_{\max(a,i,j)}-\frac{1}{4n}\kappa_{ij,kl}z_iz_jz_kz_lt_p.$$

$\kappa_{pp,pp}$ appears only when $i=j=k=l=p$. In this case by integrating z_p out the coefficient for $\kappa_{pp,pp}$ is

$$\frac{3}{4n(1-2t_p)^2}(2(t_p)^2-t_p)+\frac{t_p}{n(1-2t_p)}=\frac{t_p}{4n(1-2t_p)}$$

which is linear in $1/(1-2t_p)$.

$\kappa_{pp,ii}$ with $i<p$ appears in the form

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

Integrating this out, the coefficient for $\kappa_{pp,ii}$ is

$$\frac{1}{n}\frac{t_pt_i}{(1-2t_i)(1-2t_p)}-\frac{1}{2n}\frac{t_p}{(1-2t_i)(1-2t_p)}=-\frac{t_p}{2n(1-2t_p)}$$

which is linear in $1/(1-2t_p)$.

$\kappa_{pi,pi}$ with $i<p$ appears in the form

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

Integrating z_i and z_p out, the coefficient for $\kappa_{pi,pi}$ is

$$\frac{2(t_p)^2}{n(1-2t_i)(1-2t_p)}+\frac{t_p}{n(1-2t_i)}-\frac{t_p}{n(1-2t_i)(1-2t_p)}=0$$

and this term vanishes.

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

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. (1987). On the correction to the likelihood ratio statistics. *Biometrika*, **74**, 265–274.

[3] Cordeiro, G.M. (1993). General matrix formulae for computing Bartlett corrections. *Statistics & Probability Letters*, **16**, 11–18.

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

[5] Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Ann. Inst. Statist. Math.*, **29**, 359–378.

[6] Hayakawa, T. (1987). Correction to "The likelihood ratio criterion and the asymptotic its distribution". *Ann. Inst. Statist. Math.*, **39**, 681.

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

[8] McCullagh, P. (1987). *Tensor Methods in Statistics*, Chapman and Hall.

[9] Skovgaard, I. (1986). A note on the differentiation of cumulants of log likelihood derivatives. *International Statistical Review*, **54**, 29–32.

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