95-F-24

Asymptotic Expansion of Null Distribution of Likelihood Ratio Statistic in Multiparameter Exponential Family to an Arbitrary Order

by

Satoshi Kuriki
The Institute of Statistical Mathematics

Akimichi Takemura
Faculty of Economics, University of Tokyo

October 1995

Asymptotic Expansion of Null Distribution of Likelihood Ratio Statistic in Multiparameter Exponential Family to an Arbitrary Order

SATOSHI KURIKI
The Institute of Statistical Mathematics
4-6-7 Minami-azabu, Minato-ku, Tokyo 106, Japan

and

AKIMICHI TAKEMURA
Faculty of Economics, University of Tokyo
7-3-1 Bunkyo-ku, Tokyo 113, Japan

ABSTRACT

Consider likelihood ratio test of a simple null hypothesis in a multiparameter exponential family. We study the asymptotic expansion of the null distribution of log likelihood ratio statistic to an arbitrary order. Bartlett correctability of the $O(n^{-1})$ term is well known. We show that higher order terms exhibit a similar simplicity. Moreover we give a combinatorially explicit expression for all terms of the asymptotic expansion of the characteristic function of log likelihood ratio statistic.
Key words: Bartlett correction, Generalized Hermite polynomial, Higher order asymptotic expansion, Likelihood ratio test, Multivariate Lagrange inversion.

1. Introduction

It is by now very well known that likelihood ratio statistic is Bartlett correctable. We briefly summarize basic references on Bartlett correction. The first general treatment was given by Lawley (1956). Later Hayakawa's extensive calculation (Hayakawa (1977), Hayakawa (1987)) gave a proof of Bartlett correctability. Cordeiro (1987) showed that Hayakawa's 1977 calculation is consistent with Lawley's result. Barndorff-Nielsen and Cox (1984) derived Bartlett correction from Barndorff-Nielsen's formula (Barndorff-Nielsen (1983)) on conditional distribution of the maximum likelihood estimator. Bickel and Ghosh (1990) gave a Bayesian proof of the Bartlett correction (see also Chapter 8 of Ghosh (1994)). See Takemura and Kuriki (1995) for more on Bartlett correction.

Bartlett correction is concerned with the term of order $O(n^{-1})$. Here we consider asymptotic expansion of the null distribution of log likelihood ratio statistic to an arbitrary order.

Among various asymptotically chi-square statistics, the null distribution of the likelihood ratio statistic exhibits remarkably simple asymptotic expansion. For the term of order $O(n^{-1})$ this simplicity results in Bartlett correctability. However higher order terms show similar simplicity and there seems to be a beautiful combinatorial simplification hidden in the asymptotic expansion of likelihood ratio statistic. In Takeuchi and Takemura (1989) we have showed this for the case of one-parameter exponential family. In this paper we generalize our previous result to the multiparameter exponential family. Furthermore we give an explicit expression of general order terms of the asymptotic expansion of the characteristic function of the likelihood ratio statistic.

A fundamental work of Bickel and Ghosh (1990) has showed that in general the asymptotic expansion of the characteristic function of the likelihood ratio statistic possesses the remarkable simplicity stated in our Theorem 2.1, whereas our treatment is restricted to exponential family. The proof of Bickel and Ghosh (1990) is based on Bayes argument and does not use any structure of higher order moments and cumulants of the derivatives of log likelihood function. One drawback of their approach is that one can not obtain explicit forms of asymptotic expansion from their argument.

Our framework in this paper is restricted to the case of simple null hypothesis of a multiparameter exponential family. On the other hand we can derive explicit forms of asymptotic expansion. Actually our Theorem 2.2 below gives combinatorially explicit forms of all the terms of the asymptotic expansion.

Even in the simple framework of exponential family we need fairly complicated combinatorics and tensor notations (McCullagh (1987)). One of the main purposes of this paper is to show what kind of combinatorics are involved in higher order asymptotic expansion of the null distribution of likelihood ratio statistic.

2. Main results

In this section we will show the main results of this paper without proofs. The proofs as well as some other results are given in Section 3.

Suppose that n i.i.d. samples are observed from the p-dimensional continuous exponential family with the density function

$$f(x,\theta) = \exp\{\langle x,\theta\rangle - \psi(\theta)\}, \tag{1}$$

where $x = (x_1,\dots,x_p)$, $\theta = (\theta^1,\dots,\theta^p)$, and $\langle x,\theta\rangle = \sum_{i=1}^p x_i\theta^i$.

Consider the likelihood ratio test for testing

$$H:\theta=\theta^0 \quad \text{against} \quad K:\theta\in\Theta,$$

where θ^0 is an inner point of the p-dimensional natural parameter space Θ. We discuss here the asymptotic expansion of the null distribution of twice the log likelihood ratio test statistic $2\log\lambda$ to an arbitrary order of n. The results are described in terms of the characteristic function.

Theorem 2.1 *The characteristic function*

$$\varphi(t) = E(\mathrm{e}^{2it\log\lambda}\,|\,\theta^0) \tag{2}$$

is expanded as

$$\varphi(t) = \xi^{\frac{p}{2}}\left\{1 + \frac{f_1(\xi)}{n} + \frac{f_2(\xi)}{n^2} + \cdots\right\} \qquad with\ \ \xi = (1-2it)^{-\frac{1}{2}}\ ,$$

where $f_j(\xi)$ *is a polynomial of degree* j *in* ξ *.*

According to the same argument of Lemma 6 of Takeuchi and Takemura (1989), next corollary is proved immediately.

Corollary 2.1 $2\log\lambda$ *has the Cornish-Fisher type expansion*

$$2\log\lambda \quad \sim \quad Y\left\{1 + \frac{1}{n}B_1(Y) + \frac{1}{n^2}B_2(Y) + \cdots\right\},$$

where Y *is a random variable distributed according to the chi-square distribution with* p *degrees of freedom, and* $B_j(Y)$ *is a polynomial of degree* $j-1$ *in* Y *.*

Remark 2.1 $\deg B_1(Y) = 0$ *(i.e.* $B_1(Y) = B_1 = \text{const}$ *) implies the Bartlett correctability of order* $O(n^{-1})$ *, since*

$$\frac{2\log\lambda}{1 + B_1/n} \quad \sim \quad Y + O(n^{-2}).$$

Remark 2.2 *Theorem 2.1 and Corollary 2.1 are multivariate versions of Lemma 5 and Lemma 6 of Takeuchi and Takemura (1989), respectively.*

Remark 2.3 *The asymptotic expansion in Theorem 2.1 is valid up to an arbitrary order because for our case the log likelihood ratio statistic is a smooth function of the sample mean* $\bar{X}$ *. See Chandra and Ghosh (1979).*

Furthermore we derive an explicit expression of the asymptotic expansion of the characteristic function to an arbitrary order. We treat $\log\varphi(t)$ instead of $\varphi(t)$, because $\log\varphi(t)$ is somewhat of simpler than $\varphi(t)$. For the notation of set partition see Appendix A.

Theorem 2.2 *Let*

$$\psi_{\underbrace{ijk\cdots}_{m}} = \psi_{ijk\cdots}(\theta^0) = \frac{\partial^m}{\partial\theta^i\partial\theta^j\partial\theta^k\cdots}\,\psi(\theta)\Big|_{\theta=\theta^0}$$

be the cumulant of the distribution of (1), and let

$$\psi^{ijk\cdots} = \psi^{i\alpha}\psi^{j\beta}\psi^{k\gamma}\cdots\psi_{\alpha\beta\gamma\cdots},$$

where $\psi^{ij} = (\psi_{ij})^{-1}$ *is the inverse matrix. Put* $\psi_J = \psi_{ijk\ldots}$ *and* $\psi^J = \psi^{ijk\ldots}$ *for a set of indices* $J = \{i, j, k, \ldots\}$ *. Then, the logarithm of the characteristic function in (2) is expressed as*

$$\log\varphi(t) = \log\xi^{\frac{p}{2}} + \sum_{v\geq 3}\frac{1}{v!}\sum_{\substack{I_1|\cdots|I_u\in\mathcal{P}(I)\\|I_i|\geq 3}}\prod_{i=1}^{u}\sum_{\mathcal{I}_i}\psi^{I_{i1}}\cdots\psi^{I_{il_i}}\sum_{\mathcal{B}_i}\psi_{B_{i1}}\cdots\psi_{B_{is_i}}$$
$$\times\sum_{\mathcal{C}}\psi_{C_1}\cdots\psi_{C_h}\frac{1}{n^{v-s-l-h}}(-1)^s\prod_{i=1}^{u}\left(1-\xi^{s_i-l_i}\right)\xi^{v-2s-h} \qquad (3)$$

with $\xi = (1-2it)^{-1}$ *; where* $I = \{i, j, \ldots\}$ *is a set of running variables such that* $|I| = v$ *; the summation* $\sum_{\mathcal{I}_i}$ *is over the partitions*

$$\mathcal{I}_i = I_{i1}|\cdots|I_{il_i} \in \mathcal{P}(I_i)$$

such that $|I_{ij}| \geq 3$ *; the summation* $\sum_{\mathcal{B}_i}$ *is over the partitions*

$$\mathcal{B}_i = B_{i1}|\cdots|B_{is_i}|a_{i1}|\cdots|a_{iv_i-2s_i} \in \mathcal{P}(I_i) \quad (v_i = |I_i|)$$

such that $|B_{ij}| = 2$, $[v_i/2] \geq s_i \geq l_i + 1$,

$$\mathcal{B}_i \vee \mathcal{I}_i = \mathbf{1}(I_i),$$

and the summation $\sum_{\mathcal{C}}$ *is over the partitions*

$$\mathcal{C} = C_1|\cdots|C_h \in \mathcal{P}(A)$$

such that $|C_i| \geq 2$,

$$\mathcal{C} \vee A_1|\cdots|A_u = \mathbf{1}(A) \qquad \text{with } A_i = \{a_{i1}, \ldots, a_{iv_i-2s_i}\}\,,\ A = \textstyle\bigcup_{i=1}^{u} A_i\,;$$

and $s = \sum_{i=1}^{u} s_i$, $l = \sum_{i=1}^{u} l_i$. *(For* $A = \emptyset$ *, we assume* $h = 0$ *and ignore* $\sum_{\mathcal{C}}$ *.)*

Remark 2.4 *In the expression (3), the index set* I *is partitioned doubly, i.e.* $I_1|\cdots|I_u \in \mathcal{P}(I)$ *and* $I_{i1}|\cdots|I_{il_i} \in \mathcal{P}(I_i)$, $i = 1, \ldots, u$.

Remark 2.5 *The highest degree of the polynomial* $\prod_{i=1}^{u}(1-\xi^{s_i-l_i})\xi^{v-2s-h}$ *in* ξ *is* $v - s - l - h$.

Listing out all terms up to $O(n^{-2})$ we obtain the following corollary.

Corollary 2.2 *The asymptotic expansion of* $\log\varphi(t)$ *up to the order* $O(n^{-2})$ *is as follows:*

$$\begin{aligned}
\log\varphi(t) &= \log\xi^{\frac{p}{2}} + \frac{1}{n}g_1(\xi) + \frac{1}{n^2}g_2(\xi) + O(n^{-3}),\\
g_1(\xi) &= (1-\xi)\{\frac{1}{8}\psi^{iijj} - \frac{1}{24}(3\psi^{iij}\psi^{jkk} + 2\psi^{ijk}\psi^{ijk})\},\\
g_2(\xi) &= (1-\xi)\xi\{\frac{1}{16}\psi^{iijjkk} - \frac{1}{48}(9\psi^{iijjk}\psi^{kll} + 16\psi^{iijkl}\psi^{jkl})\\
&\quad -\frac{1}{12}(3\psi^{iijk}\psi^{jkll} + \psi^{ijkl}\psi^{ijkl}) + \frac{1}{48}(24\psi^{iijk}\psi^{jkl}\psi^{lmm} + 9\psi^{iijk}\psi^{jll}\psi^{kmm}\\
&\quad +16\psi^{ijkl}\psi^{ijk}\psi^{lmm} + 30\psi^{iijk}\psi^{jlm}\psi^{klm} + 30\psi^{ijkl}\psi^{ijm}\psi^{klm})\\
&\quad -\frac{1}{16}(4\psi^{iij}\psi^{jkl}\psi^{klm}\psi^{mnn} + \psi^{iij}\psi^{jkl}\psi^{kmm}\psi^{lnn} + 10\psi^{iij}\psi^{jkl}\psi^{kmn}\psi^{lmn}\\
&\quad +6\psi^{ijk}\psi^{ijl}\psi^{kmn}\psi^{lmn} + 4\psi^{ijk}\psi^{ilm}\psi^{kln}\psi^{jmn})\}\\
&\quad +(1-\xi^2)\{-\frac{1}{48}\psi^{iijjkk} + \frac{1}{48}(3\psi^{iijjk}\psi^{kll} + 4\psi^{iijkl}\psi^{jkl})\\
&\quad +\frac{1}{48}(3\psi^{iijk}\psi^{jkll} + \psi^{ijkl}\psi^{ijkl}) - \frac{1}{48}(6\psi^{iijk}\psi^{jkl}\psi^{lmm} + 3\psi^{iijk}\psi^{jll}\psi^{kmm}\\
&\quad +4\psi^{ijkl}\psi^{ijk}\psi^{lmm} + 6\psi^{iijk}\psi^{jlm}\psi^{klm} + 6\psi^{ijkl}\psi^{ijm}\psi^{klm})\\
&\quad +\frac{1}{48}(3\psi^{iij}\psi^{jkl}\psi^{klm}\psi^{mnn} + \psi^{iij}\psi^{jkl}\psi^{kmm}\psi^{lnn} + 6\psi^{iij}\psi^{jkl}\psi^{kmn}\psi^{lmn}\\
&\quad +3\psi^{ijk}\psi^{ijl}\psi^{kmn}\psi^{lmn} + 2\psi^{ijk}\psi^{ilm}\psi^{kln}\psi^{jmn})\}.
\end{aligned}$$

Here we use the notational convention

$$\psi^{\cdots i \cdots i \cdots} = \psi^{\cdots i \cdots j \cdots}\psi_{ij}, \quad \psi^{\cdots i \cdots}\psi^{\cdots i \cdots} = \psi^{\cdots i \cdots}\psi^{\cdots j \cdots}\psi_{ij}, \tag{4}$$

which is also used in Takemura and Kuriki (1995).

Corollary 2.3 *In the univariate case, i.e. the case of* $p=1$ *, the expression of the asymptotic expansion reduces to the following formula:*

$$\begin{aligned}
&\log\varphi(t) = \log\xi^{\frac{1}{2}}\\
&\quad +\frac{1}{n}(1-\xi)(\frac{1}{8}\kappa_4 - \frac{5}{24}\kappa_3{}^2)\\
&\quad +\frac{1}{n^2}\Big\{(1-\xi)\xi(\frac{1}{16}\kappa_6 - \frac{25}{48}\kappa_5\kappa_3 - \frac{1}{3}\kappa_4{}^2 + \frac{109}{48}\kappa_4\kappa_3{}^2 - \frac{25}{16}\kappa_3{}^4)\\
&\qquad +(1-\xi^2)(-\frac{1}{48}\kappa_6 + \frac{7}{48}\kappa_5\kappa_3 + \frac{1}{12}\kappa_4{}^2 - \frac{25}{48}\kappa_4\kappa_3{}^2 + \frac{5}{16}\kappa_3{}^4)\Big\}\\
&\quad +O(n^{-3})
\end{aligned}$$

with

$$\kappa_j = \psi^{\overbrace{1\ldots1}^{j}}\psi_{11}{}^{\frac{j}{2}}.$$

The result in Corollary 2.3 is consistent with the result in Lemma 5 of Takeuchi and Takemura (1989).

3. Proofs

Our proof is organized as follows. First we give stochastic expansion of log likelihood ratio statistic in terms of the sufficient statistic to an arbitrary order. Next we characterize asymptotic expansion of the log density of the sufficient statistic to an arbitrary order. Combining these two expansions we prove our Theorem 2.1. Theorem 2.1 can be proved by looking at only main order terms of the asymptotic expansions. Then we proceed to prove our Theorem 2.2 by considering explicit forms of all the terms of these expansions.

3.1. Asymptotic expansion of log likelihood ratio statistic

Let $\hat{\theta}$ be the maximum likelihood estimate and $\phi(x)$ be the dual function of $\psi(\theta)$,

$$\phi(x) = \max_{\theta} \ell(x,\theta) = \max_{\theta}\{\langle x,\theta\rangle - \psi(\theta)\},$$

where

$$\ell(x,\theta) = \log f(x,\theta).$$

Let $\bar{X}$ be the sample mean of n i.i.d. observations from the distribution (1). Then twice the log likelihood ratio statistic is written as

$$\begin{aligned} 2\log\lambda &= 2n\,\{\ell(\bar{X},\hat{\theta}) - \ell(\bar{X},\theta^0)\} \\ &= 2n\,\{\phi(\bar{X}) - \phi(x^0) - \langle \bar{X} - x^0, \theta^0\rangle\}, \end{aligned} \tag{5}$$

where $x^0 = E_{\theta^0}(\bar{X})$ since

$$\phi(x^0) = \langle x^0, \theta^0\rangle - \psi(\theta^0).$$

Denote normalized sufficient statistic as

$$Z = \sqrt{n}(\bar{X} - x^0).$$

Expanding (5) in Taylor series in terms of Z we obtain

$$\begin{aligned} 2\log\lambda &= 2\left\{ \frac{1}{2}\phi^{ij}Z_iZ_j + \frac{1}{3!\sqrt{n}}\phi^{ijk}Z_iZ_jZ_k \right. \\ &\quad \left. + \frac{1}{4!\,n}\phi^{ijkl}Z_iZ_jZ_kZ_l + \frac{1}{5!\,n\sqrt{n}}\phi^{ijklm}Z_iZ_jZ_kZ_lZ_m + \cdots \right\} \end{aligned}$$

where

$$\overbrace{\phi^{ijk\cdots}}^{m} = \phi^{ijk\cdots}(x^0) = \frac{\partial^m}{\partial x_i \partial x_j \partial x_k \cdots} \phi(x)\Big|_{x=x^0}.$$

We want to express $\phi^{ij}, \phi^{ijk}, \ldots,$ in terms of the higher order cumulants $\psi_{ijk\cdots}$ of the sufficient statistic defined in Theorem 2.2.

Note that

$$x_i(\theta) = \frac{\partial}{\partial \theta^i}\psi(\theta) \quad \text{and} \quad \theta^i(x) = \frac{\partial}{\partial x_i}\phi(x)$$

are inverse transformations of each other. Therefore

$$\phi^{ij}(x) = \frac{\partial \theta^i}{\partial x_j} = \Big(\frac{\partial x_j}{\partial \theta^i}\Big)^{-1} = \Big(\psi_{ij}(\theta)\Big)^{-1}.$$

Put

$$\psi^{ij} = \Big(\psi_{ij}\Big)^{-1}, \quad \psi^{ijk\cdots} = \psi^{i\alpha}\psi^{j\beta}\psi^{k\gamma}\cdots\psi_{\alpha\beta\gamma\cdots}.$$

Using the well known result on differentiating the inverse matrix

$$\frac{\partial}{\partial \theta^k}\psi^{ij}(\theta) = -\psi^{i\alpha}\psi^{j\beta}\psi_{\alpha\beta k}, \tag{6}$$

we obtain

$$\phi^{ijk} = \frac{\partial \theta^\gamma}{\partial x_k}\frac{\partial}{\partial \theta^\gamma}\psi^{ij}(\theta) = -\psi^{k\gamma}\psi^{i\alpha}\psi^{j\beta}\psi_{\alpha\beta\gamma} = -\psi^{ijk}. \tag{7}$$

In order to express further derivatives of ϕ it is convenient to use the notational convention (4). Now differentiating (7) successively using (6) it is easily shown that

$$\phi^{ijkl} = -\psi^{ijkl} + \psi^{ij\alpha}\psi^{kl\alpha}[3], \tag{8}$$

$$\phi^{ijklm} = -\psi^{ijklm} + \psi^{ijk\alpha}\psi^{lm\alpha}[10] - \psi^{ij\alpha}\psi^{k\alpha\beta}\psi^{lm\beta}[15]. \tag{9}$$

The general order term $\phi^{i_1\cdots i_m}$ can be characterized using terminology of graph theory. For illustration consider ϕ^{ijklm} of (9). The three terms of the right hand side correspond to the trees in Figure 1, Figure 2, and Figure 3 respectively. The fixed indices $i, j, k, \ldots$ correspond to the leaf edges (edges connecting the terminal vertices) of the tree and the running indices $\alpha, \beta, \ldots$ correspond to the intermediate edges connecting non-leaf vertices. Note that each running index appears twice using our notational convention (4). The numbers of symmetric terms, i.e. 10 for the second term and 15 for the third term, correspond to the number of different trees. It should be noted that in counting different trees, the leaf edges are labeled with fixed and distinct labels $i, j, k, \ldots$. On the other hand the non-leaf edges are not labeled.

Consider undirected unrooted tree T with m leaf edges $a_1, \ldots, a_m$. Here we use indices $a_1, a_2, \ldots,$ to make our description of the tree consistent with the set theoretic notation of Theorem 2.2. We denote the corresponding terminal vertices by $z_{a_1}, \ldots, z_{a_m}$ to make the notation consistent with the notation of the generalized

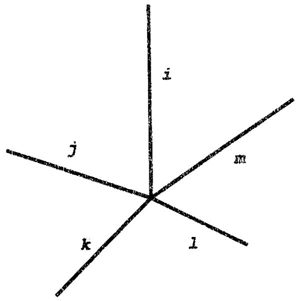


Figure 1

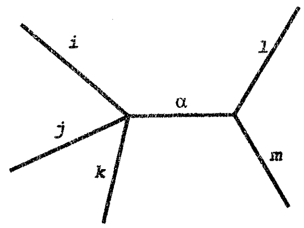


Figure 2

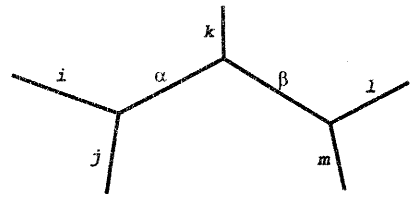


Figure 3

Hermite polynomials in Appendix B. Furthermore let n be the number of non-terminal nodes denoted by $x_1, \ldots, x_n$. Let $d(x)$ denote the degree of vertex x, i.e., the number of edges connected to the vertex. Then $d(z_{a_1}) = \cdots = d(z_{a_m}) = 1$. Further more we only consider a tree such that $d(x_1) \geq 3, \ldots, d(x_n) \geq 3$.

As mentioned above, in counting non-isomorphic trees, we consider that the leaf edges $a_1, \ldots, a_m$ are labeled and these labels are fixed and distinct. However the non-leaf edges are considered non-labeled.

In summary we consider the following set of trees.

$$\mathcal{T}(a_1, \ldots, a_m; n) = \{ T \text{ with } n \text{ non-terminal nodes of degree } \geq 3 \text{ and } m \text{ labeled leaf vertices } a_1, \ldots, a_m \}$$

Note that since there are exactly $n-1$ edges connecting non-terminal nodes $x_1, \ldots, x_n$, we have

$$n - 1 = \frac{d(x_1) + \cdots + d(x_n) - m}{2} \geq \frac{3n - m}{2}.$$

Therefore $1 \leq n \leq m-2$, and

$$\mathcal{T}(a_1, \ldots, a_m) = \cup_{n=1}^{m-2} \mathcal{T}(a_1, \ldots, a_m; n)$$

is a finite set.

Let a tree $T \in \mathcal{T}(a_1, \ldots, a_m; n)$ be given. We arbitrarily label the $n-1$ edges connecting non-terminal nodes by $b_1, \ldots, b_{n-1}$. Let A_j be the set of leaf edges connected to x_j and let B_j be the set of non-leaf edges connected to x_j. Then $(A_1, \ldots, A_n)$ is a partition of $\{a_1, \ldots, a_m\}$ and each b_i belongs to exactly 2 of $B_1, \ldots, B_n$. $b_1, \ldots, b_{n-1}$ correspond to the running variables $\alpha, \beta, \ldots$ in (8) and (9).

To this T we can associate a term of the form

$$\psi^{A_1 B_1} \cdots \psi^{A_n B_n}$$

using the summation convention (4). Now we can describe general order term as follows.

Lemma 3.1

$$\phi^{a_1 \dots a_m} = \sum_{n=1}^{m-2} (-1)^n \sum_{\mathcal{T}(a_1,\dots,a_m;n)} \psi^{A_1 B_1} \dots \psi^{A_n B_n}$$

Proof. The proof is by induction on m. Assume that the lemma holds up to a particular value of m. Then

$$\phi^{a_1 \dots a_m a_{m+1}} = \frac{\partial}{\partial x^{a_{m+1}}} \phi^{a_1 \dots a_m} = \psi^{a_{m+1} a'_{m+1}} \frac{\partial}{\partial \theta_{a'_{m+1}}} \phi^{a_1 \dots a_m}.$$

Consider differentiating a term of $\phi^{a_1 \dots a_m}$ corresponding to a particular tree. For the purpose of proof we need to write out the terms explicitly without using the summation convention (4). Note that in original quantities our summation convention is written out as

$$\psi^{\dots\alpha\dots} \psi^{\dots\alpha\dots} = \psi^{b_1 b'_1} \psi^{\dots}{}_{b_1}{}^{\dots} \psi^{\dots}{}_{b'_1}{}^{\dots}.$$

Now without loss of generality consider

$$\psi^{a_1 a'_1} \psi^{a_2 a'_2} \dots \psi^{a_{m_1} a'_{m_1}} \psi_{a'_1 \dots a'_{m_1} b_1 \dots b_{n_1}} \psi^{b_1 b'_1} \dots \psi^{b_{n_1} b'_{n_1}} \dots \tag{10}$$

There are three types in differentiating (10) with respect to $\theta_{a'_{m+1}}$ (Figure 4). The first type is differentiating $\psi_{a'_1 \dots a'_{m_1} b_1 \dots b_{n_1}}$ corresponding to a particular non-terminal vertex x. Differentiating this we obtain

$$\psi^{a_{m+1} a'_{m+1}} \psi_{a'_1 \dots a'_{m_1} a'_{m+1} b_1 \dots b_n}.$$

In terms of the tree this amounts to adding a leaf edge a_{m+1} and connecting it to x. The second type is differentiating $\psi^{b_1 b'_1}$ corresponding to a particular non-leaf edge α. By (6) this results in

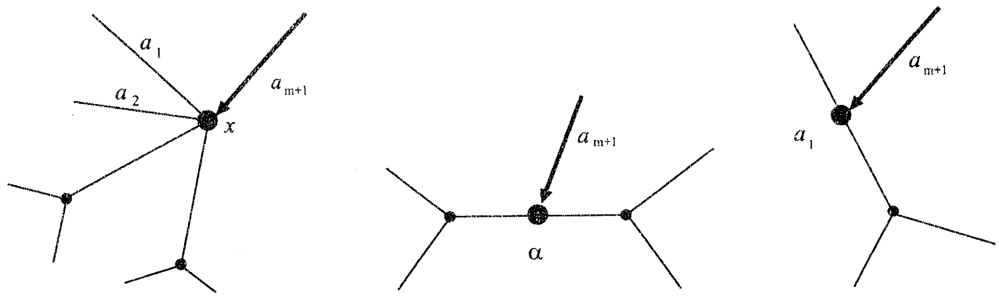
$$-\psi^{a_{m+1} a'_{m+1}} \psi^{b_1 c_1} \psi^{b'_1 c_2} \psi_{c_1 c_2 a'_{m+1}}.$$

This amounts to adding a new vertex on the edge α and and connecting a new leaf edge a_{m+1} to the new vertex. The third type is differentiating $\psi^{a_1 a'_1}$. As in the case of the second type this results in

$$-\psi^{a_{m+1} a'_{m+1}} \psi^{a_1 c_1} \psi^{a'_1 c_2} \psi_{c_1 c_2 a'_{m+1}}.$$

This amounts to adding a new non-terminal vertex on the leaf edge a_1 and connecting a new leaf edge a_{m+1} to the new vertex.

We see that differentiation of (10) generates terms corresponding to different trees in $\mathcal{T}(a_1, \dots, a_{m+1}; n)$ or $\mathcal{T}(a_1, \dots, a_{m+1}; n+1)$. To see that all trees of $\mathcal{T}(a_1, \dots, a_{m+1})$ are generated, consider removing the leaf edge a_{m+1} from a tree in $\mathcal{T}(a_1, \dots, a_{m+1}; n)$. Then we obtain a tree in $\mathcal{T}(a_1, \dots, a_m; n)$ or $\mathcal{T}(a_1, \dots, a_m; n-1)$. Therefore by the induction assumption on m, we see that the above process of adding the new leaf edge a_{m+1} generates each tree of $\mathcal{T}(a_1, \dots, a_{m+1})$ exactly once. This proves the lemma. □



Type 1 Type 2 Type 3

Figure 4

Remark 3.6 *Lemma 3.1 is a particular form of multivariate Lagrange inversion formula. Corollary 2 of Haiman and Schmitt (1989) already gave the same type inversion formula in the case that (ψ_{ij}) is identity matrix. However, our proof is more direct and brief than Haiman and Schmitt (1989) which needs preparations in incidence algebra.*

3.2. Asymptotic expansion of log density function

The second step of our proof is to characterize a general order term of the Edgeworth expansion of the log density $\log p_n(z)$ of the sufficient statistic Z. The principle of the Edgeworth expansion is given in Section 5.2.2 of McCullagh (1987). However here we need to characterize a general order term more explicitly and determine its combinatorial coefficient.

The log density function of Z given by McCullagh (1987) is

$$
\begin{aligned}
\log p_n(z) \quad = \quad & \text{const} - \frac{1}{2}\psi^{ij} z_i z_j \\
& + \frac{1}{\sqrt{n}}\frac{1}{3!}\psi^{ijk} h_{ijk}(z) \\
& + \frac{1}{n}\left(\frac{1}{4!}\psi^{ijkl} h_{ijkl}(z) + \frac{1}{6!}\psi^{ijk}\psi^{lmn} h_{ijk,lmn}(z)[10]\right) \\
& + \frac{1}{n\sqrt{n}}\Bigg(\frac{1}{5!}\psi^{ijklm} h_{ijklm}(z) + \frac{1}{7!}\psi^{ijkl}\psi^{mno} h_{ijkl,mno}(z)[35] \\
& \qquad + \frac{1}{9!}\psi^{ijk}\psi^{lmn}\psi^{opq} h_{ijk,lmn,opq}(z)[280]\Bigg) \\
& + \cdots . \qquad (11)
\end{aligned}
$$

Here h 's are the generalized Hermite polynomials given in Appendix B. The explicit expression of $\log p_n(z)$ of any order of n is given as follows.

Lemma 3.2 *The logarithm of the joint density function of* Z *is*

$$\log p_n(z) = \text{const} - \frac{1}{2}\psi^{ij} z_i z_j + \sum_{v\geq 3}\frac{1}{v!}\sum_{\substack{I_1|\cdots|I_u\in\mathcal{P}(I)\\ |I_i|\geq 3}} n^{-\frac{1}{2}(v-2u)}\psi^{I_1}\cdots\psi^{I_u}h_{I_1,\ldots,I_u}(z), \tag{12}$$

where $I=\{i,j,\ldots\}$ *is a set of running variables such that* $|I|=v$ *.*

Remark 3.7 *Takemura and Takeuchi (1988) gives the essentially same expression for the log density* $\log p_n(z)$ *in the univariate case.*

From Lemma 3.2, the term of order $n^{-\frac{1}{2}(m-2)}$ is shown to be a polynomial of degree m in z . Next lemma is concerning the coefficients of the leading terms (i.e. the terms of degree m) of this polynomial.

Lemma 3.3 *The terms of order* $n^{-\frac{1}{2}(m-2)}$ *is a polynomial in* z *of the form:*

$$-\frac{1}{m!}\phi^{a_1\ldots a_m}z_{a_1}\cdots z_{a_m} + \textit{terms of degree } m-2, m-4, \ldots .$$

Proof. As explained in Remark B.2 the highest degree term corresponds to $s=u-1$ in (24) and the summation over partition can be reduced the summation over trees. Fix $m=v-2s$ and the indices $a_1,\ldots,a_{v-2s}$ in (24). For a particular partition $I_1,\ldots,I_u$ and particular values of $b_1,\ldots,b_{2s}$, the summand in the RHS of (12) is

$$\frac{1}{v!}n^{-\frac{1}{2}(m-2)}(-1)^{u-1}\psi^{I_1}\cdots\psi^{I_u}\psi_{b_1b_2}\cdots\psi_{b_{2s-1}b_{2s}}z_{a_1}\cdots z_{a_m}. \tag{13}$$

Writing $I_j=A_jB_j$ we see that (13) corresponds to a particular tree in Lemma 3.1. It is clear that all the trees in Lemma 3.1 appear in this form. The remaining question is how many times a particular tree T is counted in

$$\sum_{\substack{I_1|\cdots|I_u\in\mathcal{P}(I)\\ |I_i|\geq 3}} \psi^{I_1}\cdots\psi^{I_u}h_{I_1,\ldots,I_u}(z).$$

Remove the labels of leaf edges $a_1,\ldots,a_m$ from T and let $\tilde{T}$ be the resulting non-labeled tree. There are $v!$ ways of placing v indices in $\tilde{T}$. On the other hand there are $m!$ ways of permuting among $a_1,\ldots,a_m$. Therefore T is counted $v!/m!$ times and the coefficient of $z_{a_1}\ldots z_{a_m}$ is

$$-\frac{v!}{m!}\frac{1}{v!}\phi^{a_1\ldots a_m} = -\frac{1}{m!}\phi^{a_1\ldots a_m}.$$

This proves the lemma. □

3.3. Proof of Theorem 2.1

From Lemma 3.2 and Lemma 3.3, the log density of Z can be written as

$$\begin{aligned}\log p_n(z) &= \text{const} - \frac{1}{2}\phi^{ij} z_i z_j \\ &+\frac{1}{\sqrt{n}}\Big(-\frac{1}{3!}\phi^{ijk} z_i z_j z_k + q_1(z)\Big) \\ &+\frac{1}{n}\Big(-\frac{1}{4!}\phi^{ijkl} z_i z_j z_k z_l + q_2(z)\Big) + \cdots, \end{aligned} \tag{14}$$

where $q_j(z)$ is an even (odd) polynomial of degree j in z for j even (odd, respectively). Using (14), the characteristic function of $2\log\lambda$ is

$$\begin{aligned}\varphi(t) &= \int \exp\Big\{2it\log\lambda + \log p_n(z)\Big\} dz \\ &\propto \int \exp\Big\{-(1-2it)\,\frac{1}{2}\phi^{ij} z_i z_j \\ &\quad -(1-2it)\,\frac{1}{3!\sqrt{n}}\phi^{ijk} z_i z_j z_k + \frac{1}{\sqrt{n}} q_1(z) \\ &\quad -(1-2it)\,\frac{1}{4!n}\phi^{ijkl} z_i z_j z_k z_l + \frac{1}{n} q_2(z) + \cdots\Big\} dz. \end{aligned}$$

Letting $(1-2it)^{-1} = \xi$, and making a change of variable $z := \xi^{\frac{1}{2}} z$, we have

$$\begin{aligned}\varphi(t) \propto \xi^{\frac{p}{2}} \cdot \int e^{-\frac{1}{2}\phi^{ij} z_i z_j} \exp\Big\{&-\frac{\xi^{\frac{1}{2}}}{3!\sqrt{n}}\phi^{ijk} z_i z_j z_k + \frac{1}{\sqrt{n}} q_1(\xi^{\frac{1}{2}} z) \\ &-\frac{\xi}{4!n}\phi^{ijkl} z_i z_j z_k z_l + \frac{1}{n} q_2(\xi^{\frac{1}{2}} z) + \cdots\Big\} dz. \end{aligned} \tag{15}$$

Theorem 2.1 now follows by expanding the exponential function in (15) and integrating term by term using the fact

$$\int e^{-\frac{1}{2}\phi^{ij} z_i z_j} \underbrace{z_k z_l \ldots z_m}_{\text{odd times}} dz = 0.$$

3.4. Proof of Theorem 2.2

The key idea is that the integrand in (15) is nearly equal to $\exp\{\log p_n(z)\}$ by formally setting $n := n\xi^{-1}$.

Let $m = n\xi^{-1}$. Then $\varphi(t)$ in (15) can be written as

$$\begin{aligned}\varphi(t) &= \xi^{\frac{p}{2}} \cdot \int \exp\Big\{\log p_m(z) \\ &\quad + \frac{1}{\sqrt{n}}\Big(q_1(\xi^{\frac{1}{2}}z) - \xi^{\frac{1}{2}}q_1(z)\Big) + \frac{1}{n}\Big(q_2(\xi^{\frac{1}{2}}z) - \xi q_2(z)\Big) + \cdots\Big\}dz \\ &= \xi^{\frac{p}{2}} \cdot E_m\Big[\exp\Big\{\frac{1}{\sqrt{n}}\Big(q_1(\xi^{\frac{1}{2}}z) - \xi^{\frac{1}{2}}q_1(z)\Big) \\ &\quad + \frac{1}{n}\Big(q_2(\xi^{\frac{1}{2}}z) - \xi q_2(z)\Big) + \cdots\Big\}\Big], \end{aligned} \tag{16}$$

where $E_m[\,\cdot\,]$ denotes the expectation operator with the sample size m . Since $q_1(z)$ is a linear polynomial without constant terms, and $q_2(z)$ is a linear polynomial without linear terms, we see that

$$\begin{aligned} q_1(\xi^{\frac{1}{2}}z) - \xi^{\frac{1}{2}}q_1(z) &= 0, \\ q_2(\xi^{\frac{1}{2}}z) - \xi q_2(z) &= (1-\xi)q_2(0). \end{aligned}$$

Hence (16) reduces to

$$\begin{aligned}\varphi(t) &= \xi^{\frac{p}{2}} \cdot E_m\Big[\exp\Big\{\frac{1}{n}(1-\xi)q_2(0) + O(n^{-2})\Big\}\Big] \\ &= \xi^{\frac{p}{2}} \cdot \Big\{1 + \frac{1}{n}(1-\xi)q_2(0) + O(n^{-2})\Big\}. \end{aligned}$$

Remark 3.8 *The Bartlett correction coefficient is* $B_1 = -(2/p)q_2(0)$.

Using this technique, we can give an explicit expression of the characteristic function of any order of n . Let $\bar{h}_{I_1,\ldots,I_l}(z)$ be the generalized Hermite polynomial from which the terms of the highest degree (i.e. degree $\sum_{j=1}^{l}|I_j| - 2(l-1)$) are removed. Let

$$\tilde{h}_{I_1,\ldots,I_l}(z;\xi) = \bar{h}_{I_1,\ldots,I_l}(\xi^{\frac{1}{2}}z) - \xi^{\frac{1}{2}(\sum_{j=1}^{l}|I_j|-2l)}\bar{h}_{I_1,\ldots,I_l}(z). \tag{17}$$

In terms of $\tilde{h}$'s in (17), (16) is written as

$$\begin{aligned}\varphi(t) = \xi^{\frac{p}{2}} \cdot E_m\Big[\exp\Big\{&\eta^i\tilde{h}_i \\ &+ \frac{1}{2!}\eta^{i,j}\tilde{h}_{ij} + \frac{1}{2!}\eta^i\eta^j\tilde{h}_{i,j} \\ &+ \frac{1}{3!}\eta^{i,j,k}\tilde{h}_{ijk} + \frac{1}{3!}\eta^i\eta^{j,k}\tilde{h}_{i,jk}[3] + \frac{1}{3!}\eta^i\eta^j\eta^k\tilde{h}_{i,j,k} + \cdots\Big\}\Big], \end{aligned} \tag{18}$$

where

$$\begin{aligned} \eta^{\overbrace{i,j,k,\ldots}^{v}} &= n^{-\frac{1}{2}(v-2)}\psi^{\overbrace{ijk\ldots}^{v}} && \text{if } v \geq 3\,, \\ &= 0 && \text{if } v = 1,2\,. \end{aligned}$$

Expanding the exponential function in (18), taking the expectation $E_m[\cdot]$, and then taking the logarithm, we get

$$\begin{aligned}\log\varphi(t) &= \log\xi^{\frac{p}{2}} + \eta^i E_m(\tilde{h}_i)\\ &+\frac{1}{2!}\eta^{i,j}E_m(\tilde{h}_{ij}) + \frac{1}{2!}\eta^i\eta^j E_m(\tilde{h}_{i,j}) + \frac{1}{2!}\eta^i\eta^j \mathrm{cum}_m(\tilde{h}_i,\tilde{h}_j)\\ &+\frac{1}{3!}\eta^{i,j,k}E_m(\tilde{h}_{ijk}) + \frac{1}{3!}\eta^i\eta^{j,k}E_m(\tilde{h}_{i,jk})[3] + \frac{1}{3!}\eta^i\eta^j\eta^k E_m(\tilde{h}_{i,j,k})\\ &+\frac{1}{3!}\eta^i\eta^{j,k}\mathrm{cum}_m(\tilde{h}_i,\tilde{h}_{jk})[3] + \frac{1}{3!}\eta^i\eta^j\eta^k\mathrm{cum}_m(\tilde{h}_i,\tilde{h}_{j,k})[3]\\ &+\frac{1}{3!}\eta^i\eta^j\eta^k\mathrm{cum}_m(\tilde{h}_i,\tilde{h}_j,\tilde{h}_k) + \cdots, \end{aligned} \tag{19}$$

where cum_m denotes the cumulant when the sample size is m. The general terms of (19) is expressed as follows. Let $I = \{i, j, \ldots\}$ be a set of running variables such that $|I| = v$. Then

$$\log\varphi(t) = \log\xi^{\frac{p}{2}} + \sum_{v\geq 3}\frac{1}{v!}\sum_{\substack{I_1|\cdots|I_u\in\mathcal{P}(I)\\ |I_i|\geq 3}}$$
$$\prod_{i=1}^{u}\sum_{\substack{I_{i1}|\cdots|I_{il_i}\in\mathcal{P}(I_i)\\ |I_{ij}|\geq 3}}\eta^{I_{i1}}\cdots\eta^{I_{il_i}}\mathrm{cum}_m\left\{\tilde{h}_{I_{11},\ldots,I_{1l_1}},\ldots,\tilde{h}_{I_{u1},\ldots,I_{ul_u}}\right\}. \tag{20}$$

$\tilde{h}$'s in (20) are also expressed explicitly in terms of the set partition. According to (24) in Appendix B, it holds that

$$\bar{h}_{I_{i1},\ldots,I_{il_i}} = \sum_{\mathcal{B}_i}(-1)^{s_i}\psi_{B_{i1}}\cdots\psi_{B_{is_i}}z_{a_{i1}}\cdots z_{a_{iv_i-2s_i}}\quad (v_i = |I_i|), \tag{21}$$

where the summation $\sum_{\mathcal{B}_i}$ is over the set partition

$$\mathcal{B}_i = B_{i1}|\cdots|B_{is_i}|a_{i1}|\cdots|a_{iv_i-2s_i}\in\mathcal{P}(I_i)$$

such that $|B_{ij}| = 2$, $[v_i/2]\geq s_i\geq(l_i-1)+1$, and

$$\mathcal{B}_i\vee I_{i1}|\cdots|I_{il_i} = \mathbf{1}(I_i).$$

(Note that the summation $\sum_{\mathcal{B}_i}$ not over $s_i\geq l_i-1$ but over $s_i\geq(l_i-1)+1$ because the leading terms were removed.)

Putting

$$\tilde{h}_{I_{i1},\ldots,I_{il_i}} = \bar{h}_{I_{i1},\ldots,I_{il_i}}(\xi^{\frac{1}{2}}z) - \xi^{\frac{1}{2}(v_i-2l_i)}\bar{h}_{I_{i1},\ldots,I_{il_i}}(z)$$

as well as (21) into the cumulant in (20), we have

$$
\begin{aligned}
&\mathrm{cum}_m\left\{\tilde{h}_{I_{11},\ldots,I_{1l_1}},\ldots,\tilde{h}_{I_{u1},\ldots,I_{ul_u}}\right\} \\
&=\prod_{i=1}^{u}\sum_{\mathcal{B}_i}(-1)^{s_i}\psi_{B_{i1}}\cdots\psi_{B_{is_i}}\left(\xi^{\frac{1}{2}(v_i-2s_i)}-\xi^{\frac{1}{2}(v_i-2l_i)}\right) \\
&\times\mathrm{cum}_m\left\{z_{a_{11}}\cdots z_{a_{1v_1-2s_1}},\ldots,z_{a_{u1}}\cdots z_{a_{uv_u-2s_u}}\right\}. \qquad (22)
\end{aligned}
$$

Remark 3.9 *In the summation $\sum_{\mathcal{B}_i}$ in (22), the contribution of $s_i=l_i$ is zero. So we can restrict the summation $\sum_{\mathcal{B}_i}$ to be over $s_i\geq l_i+1$.*

Finally, we have to evaluate the part of the generalized cumulant in (22).

Let $A_i=\{a_{i1},\ldots,a_{iv_i-2s_i}\}$ and $A=\bigcup_{i=1}^{u}A_i$. From Section 3 of McCullagh (1987), the generalized cumulant in (22) is written as

$$
\begin{aligned}
&\mathrm{cum}_m\left\{z_{a_{11}}\cdots z_{a_{1v_1-2s_1}},\ldots,z_{a_{u1}}\cdots z_{a_{uv_u-2s_u}}\right\} \\
&=\sum_{\mathcal{C}}\prod_{j=1}^{h}m^{-\frac{1}{2}(|C_j|-2)}\psi_{C_j} \\
&=\sum_{\mathcal{C}}(n\xi^{-1})^{-\frac{1}{2}(\sum_{j=1}^{h}|C_j|-2h)}\psi_{C_1}\cdots\psi_{C_h} \\
&=\sum_{\mathcal{C}}(n\xi^{-1})^{-\frac{1}{2}(v-2s-2h)}\psi_{C_1}\cdots\psi_{C_h}, \qquad (23)
\end{aligned}
$$

where the summation is over the partition

$$\mathcal{C}=C_1|\cdots|C_h\in\mathcal{P}(A)$$

such that $|C_i|\geq 2$ and

$$\mathcal{C}\vee A_1|\cdots|A_u=\mathbf{1}(A);$$

and $s=\sum_{i=1}^{u}s_i$, $l=\sum_{i=1}^{u}l_i$.

Combining the equations (20), (22), and (23), we complete the proof of Theorem 2.2.

Appendix A. Set partition and lattice

Let I be a finite set of indices, and let $\mathcal{P}(I)$ denote the set of all partition of I . Partial order $\preceq$ is defined by $\mathcal{I}_1\preceq\mathcal{I}_2$ if $\mathcal{I}_1\in\mathcal{P}(I)$ is a sub-partition of $\mathcal{I}_2\in\mathcal{P}(I)$. It is well known that the poset $\mathcal{P}(I)$ forms a lattice, in which the least upper bound and the greatest lower bound $(\vee,\wedge)$ are well defined. Let $\mathbf{1}(I)$ and $\mathbf{0}(I)$ be the greatest and the least elements in $\mathcal{P}(I)$, respectively. For example, $\mathcal{P}(I) = \{\, i|j|k,\ i|jk,\ j|ik,\ k|ij,\ ijk \,\}$ for $I = \{i, j, k\}$, and $i|j|k = \mathbf{0}(I)$, $ijk = \mathbf{1}(I)$. See Section 3.6 of McCullagh (1987) for more.

Appendix B. Generalized Hermite polynomial

Following arguments in Section 5.4.3 of McCullagh (1987), the contravariant version of the generalized Hermite polynomials indexed by the set partition

$$\begin{aligned} \mathcal{I} &= I_1|\cdots|I_u \\ &= i_{11}\ldots i_{1l_1}|\cdots|i_{u1}\ldots i_{ul_u} \in \mathcal{P}(I) \end{aligned}$$

is defined by

$$\begin{aligned} h_{I_1,\ldots,I_u}(z) &= h_{i_{11}\ldots i_{1l_1},\cdots,i_{u1}\ldots i_{ul_u}}(z) \\ &= \sum_{\mathcal{B}}(-1)^s \psi_{b_1b_2}\cdots\psi_{b_{2s-1}b_{2s}} z_{a_1}\cdots z_{a_{|I|-2s}} \end{aligned} \tag{24}$$

where the summation $\sum_{\mathcal{B}}$ is over the set partition

$$\mathcal{B} = b_1b_2|\cdots|b_{2s-1}b_{2s}|a_1|\cdots|a_{|I|-2s} \in \mathcal{P}(I)$$

such that $[|I|/2] \geq s \geq u-1$ and

$$\mathcal{B} \vee \mathcal{I} = \mathbf{1}(I).$$

This is an even (odd) polynomial of degree $d = |I| - 2(u-1)$ in z for d even (odd, respectively).

Here we give some examples of the generalized Hermite polynomials which appear in (11):

$$\begin{aligned}
&h_{ijk} = z_iz_jz_k - z_i\psi_{jk}[3], \\
&h_{ijkl} = z_iz_jz_kz_l - z_iz_j\psi_{kl}[6] + \psi_{ij}\psi_{kl}[3], \\
&h_{ijklm} = z_iz_jz_kz_lz_m - z_iz_jz_k\psi_{lm}[10] + z_i\psi_{jk}\psi_{lm}[15], \\
&h_{ijk,lmn} = -z_iz_jz_lz_m\psi_{kn}[9] + z_iz_l\psi_{jm}\psi_{kn}[18] + z_iz_j\psi_{kl}\psi_{mn}[18] \\
&\qquad -\psi_{ij}\psi_{lm}\psi_{kn}[9] - \psi_{il}\psi_{jm}\psi_{kn}[6], \\
&h_{ijkl,mno} = -z_iz_jz_kz_mz_n\psi_{lo}[12] \\
&\qquad +z_iz_jz_m\psi_{kn}\psi_{lo}[36] + z_iz_jz_k\psi_{mn}\psi_{lo}[12] + z_iz_mz_n\psi_{jk}\psi_{lo}[36] \\
&\qquad -z_i\psi_{jm}\psi_{kn}\psi_{lo}[24] - z_m\psi_{ij}\psi_{kn}\psi_{lo}[36] - z_i\psi_{jk}\psi_{lo}\psi_{mn}[36], \\
&h_{ijk,lmn,opq} = z_iz_jz_lz_oz_p\psi_{km}\psi_{nq}[162] \\
&\qquad -z_iz_lz_o\psi_{jm}\psi_{kp}\psi_{nq}[216] - z_iz_jz_l\psi_{ko}\psi_{mp}\psi_{nq}[324] - z_iz_jz_l\psi_{km}\psi_{no}\psi_{pq}[324] \\
&\qquad +z_i\psi_{jl}\psi_{km}\psi_{no}\psi_{pq}[324] + z_i\psi_{jl}\psi_{ko}\psi_{mn}\psi_{pq}[162] + z_i\psi_{jl}\psi_{ko}\psi_{mp}\psi_{nq}[324].
\end{aligned}$$

Remark B.1 *The summation in (24) can be interpreted as follows: There are u "islands" on the sea which are named by I_j , $j = 1, \ldots, u$. In the j -th island, there are v_j "villages." Suppose that s "bridges" are constructed between two villages. (These two villages may or may not be on the same island.) The summation $\sum_{\mathcal{B}}$ means to sum over all distinct ways of constructing the bridges such that all islands are connected by the bridges. This interpretation can be illustrated by considering $h_{ijk,lmn}$. In Figure 5 there are 2 islands $I_1 = \{i, j, k\}$ and $I_2 = \{k, l, m\}$. These islands have to be connected at least by one bridge. Then there are 5 types of connecting 2 islands. The first type is connecting the islands with just 1 bridge and there are 9 ways of doing this resulting in the term $-z_i z_j z_l z_m \psi_{kn}[9]$. The second type is connecting 2 islands doubly by 2 bridges resulting in the term $z_i z_l \psi_{jm} \psi_{kn}[18]$ with 18 possible ways of doing this bridging. For other 3 types see Figure 5.*

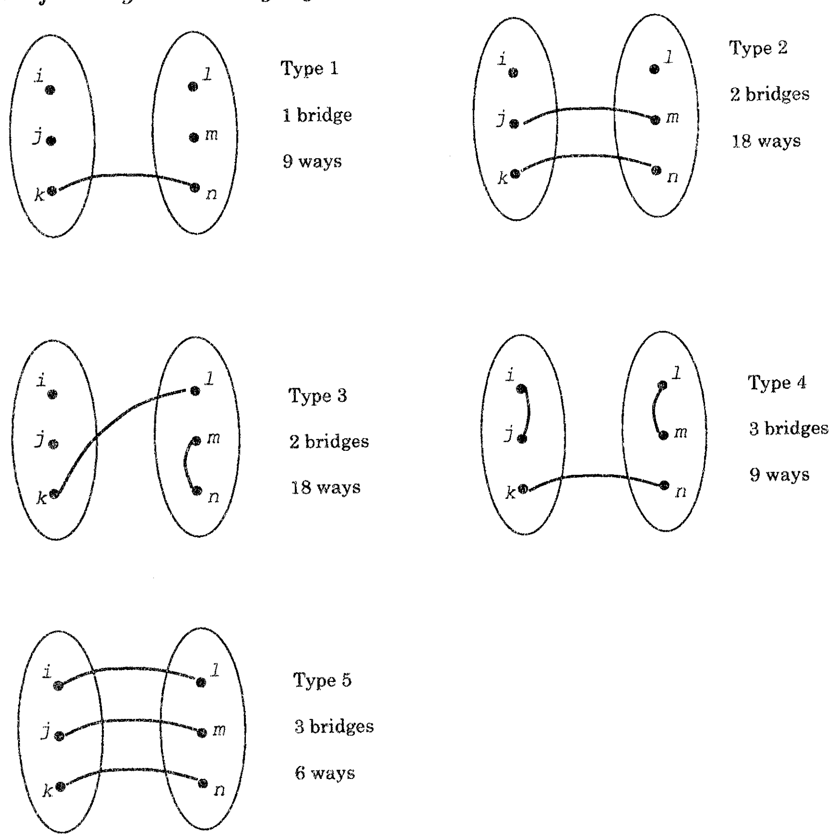


Figure 5

Remark B.2 *The highest degree terms in (24) are terms with $s = u - 1$. For this case the islands are minimally connected, i.e., if islands are considered as vertices and bridges are considered as edges of a graph, then the graph is a tree with u vertices and s edges $\alpha_1 = (b_1, b_2), \ldots, \alpha_s = (b_{2s-1}, b_{2s})$. Furthermore consider moving non-bridged villages $a_1, \ldots, a_{v-2s}$ ($v = \sum_{j=1}^{u} v_j$) outside the islands and making these villages non-leaf edges connected to the corresponding islands. Then we obtain an* *equivalent tree of the type considered in Lemma 3.1. For our problem it seems to be easier to think using the metaphor of "islands", "villages" and "bridges." However this terminology is not standard and we have used more standard terminology of tree in stating Lemma 3.1. This process is illustrated in Figure 6.*

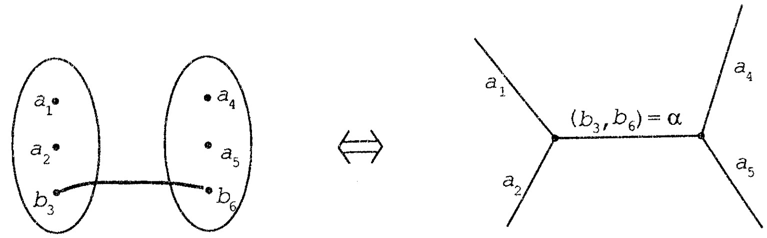


Figure 6

References

[1] Barndorff-Nielsen, O.E. (1983). On a formula for the conditional distribution of the maximum likelihood estimator. *Biometrika*, **70**, 343–365.

[2] Barndorff-Nielsen, O.E. and Cox, D.R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. Roy. Statist. Soc. Ser. B*, **46**, 483–495.

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

[4] Chandra, T.K. and Ghosh, J.K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables. *Sankhya Ser. A*, **41**, 22–47.

[5] Cordeiro, G.M. (1987). On the correction to the likelihood ratio statistics. *Biometrika*, **74**, 265–274.

[6] Ghosh, J.K. (1994). *Higher Order Asymptotics*. NSF-CBMS Regional Conference Series in Probability and Statistics, Volume 4, Institute of Mathematical Statistics and American Statistical Association.

[7] Haiman, M. and Schmitt, W. (1989). Incidence algebra antipodes and Lagrange inversion in one and several variables. *J. Combin. Theory Ser. A*, **50**, 172-185.

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

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

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

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

[12] Takemura, A. and Kuriki, S. (1995). A proof of independent Bartlett cor-

rectability of nested likelihood ratio tests, Discussion Paper Series 95-F-16, Faculty of Economics, University of Tokyo.

[13] 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.

[14] Takeuchi, K. and Takemura, A. (1989). On Cornish-Fisher type expansion of likelihood ratio statistic in one parameter exponential family. In *Probability theory and mathematical statistics* (Kyoto, 1986), Lecture Notes in Math., 1299, 492–501, Springer.