EXACT INFORMATION LOSS IN THE MULTIVARIATE GAMMA DISTRIBUTION

Etsuo Kumagai and Nobuo Inagaki

Received February 18, 2005

Abstract. Exact information losses with respect to the maximum likelihood estimator(MLE) in various models are hardly explored. The aim of this paper is to reconsider a multivariate Efron's information curvature from the viewpoint of the regression and to investigate an exact Fisher information loss with respect to MLE in a parameterization in the multivariate gamma distribution. The exact information loss was explicitly calculated in detail and the results were also a natural extension for the Nile problem.

1 Introduction To investigate an exact information loss of a various model is important to study the efficiency of estimators in detail, especially MLE. But, for the exact loss, there were few studies, for example, Fisher [4] studied the Nile problem which showed the exact information loss explicitly all, and Inagaki and Kumagai [6] and furthermore Inoue, Inagaki and Kumagai [7] studied the spherical model in the normal distribution which showed the exact information loss explicitly except a part of it. The Nile problem by Fisher showed the exact information loss with a hyperbolic curve as a parameter in the two dimensional exponential distribution.

In this paper, we propose a generalized hyperbolic model as a natural extension of the Nile problem, reconsider a multivariate Efron's information curvature from the viewpoint of the regression, which Efron's information curvature means a natural extension of the information loss with Efron's statistical curvature [3]. The projection matrix by regression has an important role in this calculation, which is partly based on the projection with respect to the conditional expectation given MLE by Inagaki [5]. Also we explicitly investigate the exact Fisher information loss with respect to MLE in the generalized hyperbolic model in detail, so that we find that the limitation of the exact Fisher information loss is equivalent to Efron's information curvature, which is the same as the result of Amari's information geometric approach [1].

2 Generalized hyperbolic model Let k be a positive integer and let random variables $\{X_i\}_{i=1}^k$ be independent mutually and be each distributed with the gamma distribution $\{G_A(q_i, \alpha_i^{-1})\}_{i=1}^k$, where $\{q_i\}_{i=1}^k$ and $\{\alpha_i\}_{i=1}^k$ are constants and parameters respectively which are all positive and finite. We consider the k-dimensional gamma distribution $(X_1, \dots, X_k)$ whose probability density function is

$$\begin{aligned} p(x_1, \dots, x_k) &= \exp\{\langle \boldsymbol{\alpha}, \boldsymbol{x} \rangle - \psi(\boldsymbol{\alpha})\}\, h(\boldsymbol{x}) \\ &= \exp\left\{ -\sum_{i=1}^k \alpha_i \; x_i + \sum_{i=1}^k \log \alpha_i^{q_i} \right\} \prod_{i=1}^k \frac{x_i^{q_i - 1}}{\Gamma(q_i)} \ , \end{aligned}$$

2000 *Mathematics Subject Classification.* 62F12; 62F10.

Key words and phrases. Exact information loss; Maximum likelihood estimator; Gamma distribution; Efron's information curvature; Nile problem.

where

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}, \quad \boldsymbol{x} = -\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}, \quad \psi(\boldsymbol{\alpha}) = -\sum_{i=1}^{k} q_i \log \alpha_i \ , \quad h(\boldsymbol{x}) = \prod_{i=1}^{k} \frac{x_i^{q_i - 1}}{\Gamma(q_i)} \ .$$

Thus some properties of the exponential family bring the mean vector $\boldsymbol{\mu}(\boldsymbol{\alpha}) = \nabla\psi(\boldsymbol{\alpha})$ and the covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\alpha}) = {}^t\nabla\nabla\psi(\boldsymbol{\alpha})$ as follows;

$$\boldsymbol{\mu}(\boldsymbol{\alpha}) \ = \ -\begin{pmatrix} q_1\,\alpha_1^{-1} \\ \vdots \\ q_k\,\alpha_k^{-1} \end{pmatrix}, \quad \boldsymbol{\Sigma}(\boldsymbol{\alpha}) \ = \ \begin{pmatrix} q_1\,\alpha_1^{-2} & 0 & \cdots & 0 \\ 0 & q_2\,\alpha_2^{-2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & q_k\,\alpha_k^{-2} \end{pmatrix},$$

where the notations ∇ and ^{t} mean the differentiation by $\boldsymbol{\alpha}$ and the transposition, respectively.

We consider the k-dimensional gamma distribution whose parameters satisfy the condition $\psi(\boldsymbol{\alpha}(\boldsymbol{\theta})) = 0$, that is, parameters have the relationships as follows:

$$\text{(2.1)} \qquad \boldsymbol{\alpha}(\boldsymbol{\theta}) = \begin{pmatrix} \alpha_1(\boldsymbol{\theta}) \\ \vdots \\ \alpha_{k-1}(\boldsymbol{\theta}) \\ \alpha_k(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{k-1} \\ \vartheta^{-1} \end{pmatrix}, \quad \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{k-1} \end{pmatrix},$$

where

$$\text{(2.2)} \qquad \vartheta = \prod_{i=1}^{k-1} \theta_i^{r_i}, \quad 0 < \theta_1, \ldots, \theta_{k-1} < \infty, \quad r_1 = \frac{q_1}{q_k}, \ \ldots\ , \ r_{k-1} = \frac{q_{k-1}}{q_k}.$$

Note that $r_1, \ldots, r_{k-1}$ are positive constants. We call this *a generalized hyperbolic model*. The above condition with respect to the cumulant generating function also makes this model an "equipotential" curved exponential model. Remark that this model is based on the parameterization in the k-dimensional gamma distribution and this is different from a generalized hyperbolic distribution, for instance, in Blæsild [2], and that this restriction is independent of the former results, for example, Inagaki and Kumagai [6] and Inoue, Inagaki and Kumagai [7], because they just considered the dimensional relationships between the parameter space $\{\boldsymbol{\alpha}\}$ with the dimension k and the parameter subspace $\{\boldsymbol{\theta}\}$ with the dimension $k-1$ with respect to the subspace $\{\boldsymbol{\alpha}(\boldsymbol{\theta})\}$.

Thus the mean vector and the covariance matrix parameterized with $\boldsymbol{\theta}$ are represented as follows:

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = -\begin{pmatrix} q_1\,\theta_1^{-1} \\ \vdots \\ q_{k-1}\,\theta_{k-1}^{-1} \\ q_k\,\vartheta \end{pmatrix}, \quad \boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{pmatrix} q_1\,\theta_1^{-2} & 0 & \cdots & 0 \\ 0 & q_2\,\theta_2^{-2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & q_k\,\vartheta^2 \end{pmatrix}.$$

The probability density function of the generalized hyperbolic model is

$$\text{(2.3)} \qquad f(x_1, \ldots, x_k \mid \boldsymbol{\theta}) = \exp\left\{ -\sum_{i=1}^{k-1} \theta_i\,x_i - \vartheta^{-1}\ x_k \right\} h(\boldsymbol{x}),$$

and the log-likelihood function and its derivatives are

$$\begin{aligned}
\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{k-1} \theta_i x_i - \vartheta^{-1} x_k + \log h(\boldsymbol{x}), \\
\dot{\ell}(\boldsymbol{\theta})_i &= \frac{\partial}{\partial \theta_i} \ell(\boldsymbol{\theta}) = -x_i + r_i \vartheta^{-1} \theta_i^{-1} x_k, \\
\ddot{\ell}(\boldsymbol{\theta})_{ij} &= \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell(\boldsymbol{\theta}) = \begin{cases} -r_i(r_i+1)\vartheta^{-1}\theta_i^{-2} x_k & (i = j), \\ -r_i r_j \vartheta^{-1} \theta_i^{-1} \theta_j^{-1} x_k & (i \neq j), \end{cases}
\end{aligned}$$

so that, since $E(X_k) = q_k \vartheta$, each component of Fisher information matrix $\boldsymbol{I}(\boldsymbol{\theta})$ is represented by

$$\boldsymbol{I}(\boldsymbol{\theta})_{ij} = -E\{\ddot{\ell}(\boldsymbol{\theta})_{ij}\} = \begin{cases} q_k r_i (r_i+1) \theta_i^{-2} & (i = j), \\ q_k r_i r_j \theta_i^{-1} \theta_j^{-1} & (i \neq j). \end{cases} \tag{2.4}$$

In the differentiations of $\boldsymbol{\alpha}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, the first derivative is

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\alpha}(\boldsymbol{\theta})}{\partial\, {}^t\boldsymbol{\theta}} = (\dot{\boldsymbol{\alpha}}_1(\boldsymbol{\theta}) \ \cdots \ \dot{\boldsymbol{\alpha}}_{k-1}(\boldsymbol{\theta})), \quad \dot{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0} \\ 1 \\ \mathbf{0} \\ -r_i \vartheta^{-1} \theta_i^{-1} \end{pmatrix} \begin{matrix} \\ < i \\ \\ \\ \end{matrix}$$

and the second is

$$\ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{\partial \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\partial^2 \boldsymbol{\alpha}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\, \partial\, {}^t\boldsymbol{\theta}} = \begin{pmatrix} \ddot{\boldsymbol{\alpha}}_{11}(\boldsymbol{\theta}) & \cdots & \ddot{\boldsymbol{\alpha}}_{1\,k-1}(\boldsymbol{\theta}) \\ \vdots & \cdots & \vdots \\ \ddot{\boldsymbol{\alpha}}_{k-1\,1}(\boldsymbol{\theta}) & \cdots & \ddot{\boldsymbol{\alpha}}_{k-1\,k-1}(\boldsymbol{\theta}) \end{pmatrix},$$

where

$$\ddot{\boldsymbol{\alpha}}_{ii}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0} \\ r_i(r_i+1)\,\vartheta^{-1}\,\theta_i^{-2} \end{pmatrix}, \quad \ddot{\boldsymbol{\alpha}}_{ij}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0} \\ r_i r_j \vartheta^{-1}\,\theta_i^{-1}\theta_j^{-1} \end{pmatrix} \quad (i \neq j).$$

The following theorem holds with respect to this generalized hyperbolic model:

Theorem 1 *In the k-dimensional generalized hyperbolic model, let $\boldsymbol{P}^*(\boldsymbol{\theta})$ be the projection matrix into the subspace spanned by $\boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})$, that is,*

$$\boldsymbol{P}^*(\boldsymbol{\theta}) = \boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \left({}^t\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta})\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\right)^{-1} {}^t\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta}),$$

so that the generalized statistical curvature $\boldsymbol{\Gamma}(\boldsymbol{\theta})^2$ is represented as follows:

$$\begin{aligned}
&\boldsymbol{\Gamma}(\boldsymbol{\theta})^2 \\
&\quad = \left({}^t\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta})\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\right)^{-2} \left\{ {}^t\ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \left[\mathbf{I}_{k-1} \otimes \boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})(\mathbf{I}_k - \boldsymbol{P}^*(\boldsymbol{\theta}))\boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})\right] \ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \right\} \\
&\quad = \frac{1}{q_1 + \cdots + q_k}\, \mathbf{I}_{k-1}, \qquad (2.5)
\end{aligned}$$

where $\mathbf{I}_{k-1}$ and $\mathbf{I}_k$ are the $k-1$ and k dimensional identity matrices, respectively, and the notation $\otimes$ means the Kronecker product.

Before the proof, note that the derivation of $\mathbf{\Gamma}(\boldsymbol{\theta})^2$ is on the analogy of the relationship between the mathematical curvature and the statistical curvature of a surface in the three dimensional space, and that the formulation of $\mathbf{\Gamma}(\boldsymbol{\theta})^2$ will be also justified by the limitation of the exact information loss in the next section.

Proof. We define the term

$$\begin{aligned}\mathbf{\Lambda}(\boldsymbol{\theta})^2 &= {}^t\ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \left[\mathbf{I}_{k-1} \otimes \mathbf{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})(\mathbf{I}_k - \boldsymbol{P}^*(\boldsymbol{\theta}))\mathbf{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})\right] \ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \\ &= \boldsymbol{\nu}_{22}(\boldsymbol{\theta}) - \boldsymbol{\nu}_{12}(\boldsymbol{\theta}) \left[\mathbf{I}_{k-1} \otimes \boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1}\right] \ {}^t\boldsymbol{\nu}_{12}(\boldsymbol{\theta}),\end{aligned}$$

where $\boldsymbol{\nu}_{11}(\boldsymbol{\theta}) = {}^t\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\, \mathbf{\Sigma}(\boldsymbol{\theta})\, \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})$, $\boldsymbol{\nu}_{12}(\boldsymbol{\theta}) = {}^t\ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\, [\mathbf{I}_{k-1} \otimes \mathbf{\Sigma}(\boldsymbol{\theta})\, \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})]$ and $\boldsymbol{\nu}_{22}(\boldsymbol{\theta}) = {}^t\ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\, [\mathbf{I}_{k-1} \otimes \mathbf{\Sigma}(\boldsymbol{\theta})]\, \ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta})$. We shall investigate the above three terms and $\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-2}$ in turn.

First, we calculate the term $\boldsymbol{\nu}_{11}(\boldsymbol{\theta})$. Because of the term

$$\mathbf{\Sigma}(\boldsymbol{\theta})\, \dot{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0} \\ q_i \theta_i^{-2} \\ \mathbf{0} \\ -q_i \vartheta\, \theta_i^{-1} \end{pmatrix} \begin{matrix} \\ < i \\ \\ \\ \end{matrix} \quad ,$$

each component of $\boldsymbol{\nu}_{11}(\boldsymbol{\theta})$ is represented by

$$\boldsymbol{\nu}_{11}(\boldsymbol{\theta})_{ij} = {}^t\dot{\boldsymbol{\alpha}}_i(\boldsymbol{\theta})\, \mathbf{\Sigma}(\boldsymbol{\theta})\, \dot{\boldsymbol{\alpha}}_j(\boldsymbol{\theta}) = \begin{cases} q_i(r_i+1)\theta_i^{-2}, & (i=j), \\ q_i r_j \theta_i^{-1}\theta_j^{-1}, & (i \neq j), \end{cases}$$

so that the determinant is

$$\det(\boldsymbol{\nu}_{11}(\boldsymbol{\theta})) = \prod_{i=1}^{k-1} q_i\, \theta_i^{-2} \left(1 + \sum_{j=1}^{k-1} r_j\right) = r_* \prod_{i=1}^{k-1} q_i\, \theta_i^{-2} \ ,$$

where $r_* = 1 + \sum_{i=j}^{k-1} r_j$. Thus components of the inverse matrix are

$$\left(\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1}\right)_{ij} = \begin{cases} \dfrac{r_* - r_i}{q_i\, r_*}\, \theta_i^2, & (i=j), \\[2ex] -\dfrac{1}{q_k\, r_*}\, \theta_i\, \theta_j, & (i \neq j). \end{cases}$$

Secondly we calculate the term $\boldsymbol{\nu}_{12}(\boldsymbol{\theta})$. Because of

$$\mathbf{\Sigma}(\boldsymbol{\theta})\, \ddot{\boldsymbol{\alpha}}_{ii}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0} \\ q_k r_i (r_i+1) \vartheta\, \theta_i^{-2} \end{pmatrix} \text{ and } \mathbf{\Sigma}(\boldsymbol{\theta})\, \ddot{\boldsymbol{\alpha}}_{ij}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0} \\ q_k r_i r_j \vartheta\, \theta_i^{-1} \theta_j^{-1} \end{pmatrix},$$

it holds that

$${}^t\ddot{\boldsymbol{\alpha}}_{ji}(\boldsymbol{\theta})\, \mathbf{\Sigma}(\boldsymbol{\theta})\, \dot{\boldsymbol{\alpha}}_a(\boldsymbol{\theta}) = \begin{cases} -q_k r_i (r_i+1) r_a\, \theta_i^{-2} \theta_a^{-1}, & (i=j), \\ -q_k r_i r_j r_a\, \theta_i^{-1} \theta_j^{-1} \theta_a^{-1}, & (i \neq j), \end{cases}$$

so that each component of $\boldsymbol{\nu}_{12}(\boldsymbol{\theta})$ is the following $1 \times (k-1)$ vector:

$$\begin{aligned}\boldsymbol{\nu}_{12}(\boldsymbol{\theta})_{ij} &= {}^t\ddot{\boldsymbol{\alpha}}_{ij}(\boldsymbol{\theta})\, \mathbf{\Sigma}(\boldsymbol{\theta})\, \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \\ &= \begin{cases} \left(-q_k r_i (r_i+1) \theta_i^{-2}\, r_a \theta_a^{-1}\right)_{a=1,\dots,k-1}, & (i=j), \\ \left(-q_k r_i r_j \theta_i^{-1} \theta_j^{-1}\, r_a \theta_a^{-1}\right)_{a=1,\dots,k-1}, & (i \neq j). \end{cases}\end{aligned}$$

Thus, since

$$\boldsymbol{\nu}_{12}(\boldsymbol{\theta})_{ij}\ \boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1} = \begin{cases} \left(-\frac{r_i(r_i+1)}{r_*}\theta_i^{-2}\theta_b\right)_{b=1,\dots,k-1}, & (i=j), \\ \left(-\frac{r_i r_j}{r_*}\,\theta_i^{-1}\theta_j^{-1}\theta_b\right)_{b=1,\dots,k-1}, & (i\neq j) \end{cases}$$

and

$$\boldsymbol{\nu}_{12}(\boldsymbol{\theta})_{ji}\ \boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1}\ {}^t\boldsymbol{\nu}_{12}(\boldsymbol{\theta})_{ja} = \begin{cases} q_k\,\dfrac{r_*-1}{r_*}\,\dfrac{r_i^2(r_i+1)^2}{\theta_i^4}, & (i=j=a), \\ q_k\,\dfrac{r_*-1}{r_*}\,\dfrac{r_i^2 r_j^2}{\theta_i^2\theta_j^2}, & (i=a\neq j), \\ q_k\,\dfrac{r_*-1}{r_*}\,\dfrac{r_i^2(r_i+1)r_a}{\theta_i^3\theta_a}, & (i=j\neq a), \\ q_k\,\dfrac{r_*-1}{r_*}\,\dfrac{r_i r_j^2 r_a}{\theta_i\theta_j^2\theta_a}, & (i\neq j\neq a), \end{cases}$$

it holds that

$$\begin{aligned} \left(\boldsymbol{\nu}_{12}(\boldsymbol{\theta})\left[\mathbf{I}_{k-1}\otimes\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1}\right]\ {}^t\boldsymbol{\nu}_{12}(\boldsymbol{\theta})\right)_{ia} &= \sum_{j=1}^{k-1}\boldsymbol{\nu}_{12}(\boldsymbol{\theta})_{ji}\ \boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1}\ {}^t\boldsymbol{\nu}_{12}(\boldsymbol{\theta})_{ja} \\ &= \begin{cases} q_k\,\frac{r_*-1}{r_*}\left[\frac{r_i^2(r_i+1)^2}{\theta_i^4}+\frac{r_i^2}{\theta_i^2}\sum_{j\neq i}\frac{r_j^2}{\theta_j^2}\right], & (i=a), \\ q_k\frac{r_*-1}{r_*}\left[\frac{r_i^2(r_i+1)r_a}{\theta_i^3\theta_a}+\frac{r_a^2(r_a+1)r_i}{\theta_a^3\theta_i}+\frac{r_i r_a}{\theta_i\theta_a}\sum_{j\neq i,a}\frac{r_j^2}{\theta_j^2}\right], & (i\neq a). \end{cases} \end{aligned}$$

Thirdly we calculate the term $\boldsymbol{\nu}_{22}(\boldsymbol{\theta})$. Since

$${}^t\ddot{\boldsymbol{\alpha}}_{ji}(\boldsymbol{\theta})\,\boldsymbol{\Sigma}(\boldsymbol{\theta})\,\ddot{\boldsymbol{\alpha}}_{ja}(\boldsymbol{\theta}) = \begin{cases} q_k\,\dfrac{r_i^2(r_i+1)^2}{\theta_i^4}, & (i=j=a), \\ q_k\,\dfrac{r_i^2 r_j^2}{\theta_i^2\theta_j^2}, & (i=a\neq j), \\ q_k\,\dfrac{r_i^2(r_i+1)r_a}{\theta_i^3\theta_a}, & (i=j\neq a), \\ q_k\,\dfrac{r_i r_j^2 r_a}{\theta_i\theta_j^2\theta_a}, & (i\neq j\neq a) \end{cases}$$

it holds that

$$\begin{aligned} \boldsymbol{\nu}_{22}(\boldsymbol{\theta})_{ia} &= \sum_{j=1}^{k-1}{}^t\ddot{\boldsymbol{\alpha}}_{ji}(\boldsymbol{\theta})\,\boldsymbol{\Sigma}(\boldsymbol{\theta})\,\ddot{\boldsymbol{\alpha}}_{ja}(\boldsymbol{\theta}) \\ &= \begin{cases} q_k\left[\frac{r_i^2(r_i+1)^2}{\theta_i^4}+\frac{r_i^2}{\theta_i^2}\sum_{j\neq i}\frac{r_j^2}{\theta_j^2}\right], & (i=a), \\ q_k\left[\frac{r_i^2(r_i+1)r_a}{\theta_i^3\theta_a}+\frac{r_a^2(r_a+1)r_i}{\theta_a^3\theta_i}+\frac{r_i r_a}{\theta_i\theta_a}\sum_{j\neq i,a}\frac{r_j^2}{\theta_j^2}\right], & (i\neq a). \end{cases} \end{aligned}$$

Thus we obtain that

$$\begin{aligned} \boldsymbol{\Lambda}(\boldsymbol{\theta})_{ia}^2 &= \boldsymbol{\nu}_{22}(\boldsymbol{\theta})_{ia}-\left(\boldsymbol{\nu}_{12}(\boldsymbol{\theta})\left[\mathbf{I}_{k-1}\otimes\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-1}\right]\ {}^t\boldsymbol{\nu}_{12}(\boldsymbol{\theta})\right)_{ia} \\ &= \begin{cases} \frac{q_k}{r_*}\,\frac{r_i^2}{\theta_i^2}\left(\frac{1+2r_i}{\theta_i^2}+\xi\right), & (i=a), \\ \frac{q_k}{r_*}\,\frac{r_i r_a}{\theta_i\theta_a}\left(\frac{r_i}{\theta_i^2}+\frac{r_a}{\theta_a^2}+\xi\right), & (i\neq a), \end{cases} \end{aligned}$$

where $\xi = \sum_{j=1}^{k-1} r_j^2 \theta_j^{-2}$.

Last it holds that

$$\left(\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-2}\right)_{ia} = \begin{cases} \frac{1}{q_k^2 r_*^2} \frac{\theta_i^2}{r_i^2} \{(r_*^2 - 2r_* r_i)\theta_i^2 + r_i^2 \zeta\}, & (i = a), \\ -\frac{1}{q_k^2 r_*^2} \frac{\theta_i \theta_a}{r_i r_a} \{r_* r_a \theta_i^2 + r_* r_i \theta_a^2 - r_i r_a \zeta\}, & (i \neq a), \end{cases}$$

where $\zeta = \sum_{j=1}^{k-1} \theta_j^2$. Therefore, the generalized statistical curvature

$$\boldsymbol{\Gamma}(\boldsymbol{\theta})^2 = \boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-2}\, \boldsymbol{\Lambda}(\boldsymbol{\theta})^2$$

has the following components:

$$\begin{aligned} q_k r_*^3 \left(\boldsymbol{\Gamma}(\boldsymbol{\theta})^2\right)_{ii} &= q_k r_*^3 \left(\left(\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-2}\right)_{ii} \boldsymbol{\Lambda}(\boldsymbol{\theta})^2_{ii} + \sum_{a \neq i} \left(\boldsymbol{\nu}_{11}(\boldsymbol{\theta})^{-2}\right)_{ia} \boldsymbol{\Lambda}(\boldsymbol{\theta})^2_{ai} \right) \\ &= r_*^2 \end{aligned}$$

and, for $i \neq j$,

$$\begin{aligned} & q_k r_*^3 \frac{r_i \theta_j}{\theta_i r_j} \left(\boldsymbol{\Gamma}(\boldsymbol{\theta})^2\right)_{ij} \\ = \; & q_k r_*^3 \frac{r_i \theta_j}{\theta_i r_j} \left(\left(\boldsymbol{\nu}(\boldsymbol{\theta})_{11}^{-2}\right)_{ii} \boldsymbol{\Lambda}(\boldsymbol{\theta})^2_{ij} + \left(\boldsymbol{\nu}(\boldsymbol{\theta})_{11}^{-2}\right)_{ij} \boldsymbol{\Lambda}(\boldsymbol{\theta})^2_{jj} + \sum_{a \neq i,j} \left(\boldsymbol{\nu}(\boldsymbol{\theta})_{11}^{-2}\right)_{ia} \boldsymbol{\Lambda}(\boldsymbol{\theta})^2_{aj} \right) \\ = \; & 0, \end{aligned}$$

so that we obtain the generalized statistical curvature

$$\boldsymbol{\Gamma}(\boldsymbol{\theta})^2 = \frac{1}{q_k r_*} \mathbf{I}_{k-1} = \frac{1}{q_1 + \cdots + q_k} \mathbf{I}_{k-1}.$$

This is what we required. □

Definition 1 *We define Efron's information curvature,* $\mathbf{EIC}(\boldsymbol{\theta})$*, by the multiplication of the Fisher information matrix (2.4) and the generalized statistical curvature matrix (2.5), that is,*

$$\begin{aligned} (2.6)\quad \mathbf{EIC}(\boldsymbol{\theta}) &\equiv \boldsymbol{I}(\boldsymbol{\theta})\, \boldsymbol{\Gamma}(\boldsymbol{\theta})^2 \\ &= \left({}^t\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \boldsymbol{\Sigma}(\boldsymbol{\theta}) \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\right)^{-1} \left\{ {}^t\ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \left[\mathbf{I}_{k-1} \otimes \boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta})(\mathbf{I}_k - \boldsymbol{P}^*(\boldsymbol{\theta}))\boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{\theta}) \right] \ddot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \right\}. \end{aligned}$$

This definition has a similarity to the formulation which is partly based on the projection with respect to the conditional expectation given MLE by Inagaki [5]. And this has a meaning with respect to the use of projection matrix like a regression form which is different from the approach by the information geometric framework (for example, Amari [1]) even if the result is the same.

¿From Theorem 1, we easily obtain (2.6) for the generalized hyperbolic model as follows:

Theorem 2 *For the generalized hyperbolic model, the* (i, j)*-th element of* $\mathbf{EIC}(\boldsymbol{\theta})$ *is represented by*

$$\mathrm{EIC}(\boldsymbol{\theta})_{ij} = \frac{1}{q_1 + \cdots + q_k} \boldsymbol{I}(\boldsymbol{\theta})_{ij} = \begin{cases} \dfrac{q_k r_i (r_i + 1)}{(q_1 + \cdots + q_k)\, \theta_i^{-2}} & (i = j), \\ \dfrac{q_k r_i r_j}{(q_1 + \cdots + q_k)\, \theta_i^{-1} \theta_j^{-1}} & (i \neq j). \end{cases}$$

Proof. By calculating the definition of (2.6) with the Fisher information (2.4) and the statistical curvature (2.5), we easily obtain the result. □

3 Exact information loss Apart from the generalized statistical curvature, we directly investigate the exact information loss in the k-dimensional hyperbolic model. For n random vectors

$$\begin{pmatrix} X_{11} \\ \vdots \\ X_{1k} \end{pmatrix}, \quad \cdots \quad, \begin{pmatrix} X_{n1} \\ \vdots \\ X_{nk} \end{pmatrix}$$

which are independently distributed with k-dimensional hyperbolic model (2.3), we redefine random vectors as follows:

$$\boldsymbol{X}_1 = (X_{11}, \ldots, X_{n1}), \quad \cdots, \quad \boldsymbol{X}_k = (X_{1k}, \ldots, X_{nk}).$$

Note that the length of each $\boldsymbol{X}_i$ is n. Then the joint probability density function is

$$f_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \,|\, \boldsymbol{\theta}) \;=\; \exp\left\{-\sum_{j=1}^{k-1}\left(\theta_j \sum_{i=1}^{n} x_{ij}\right) - \vartheta^{-1}\sum_{i=1}^{n} x_{ik}\right\} \prod_{i=1}^{k} h(\boldsymbol{x}_i),$$

so that the log-likelihood and its derivative for $a, b = 1, \ldots, k-1$ are

$$\begin{aligned} \ell_n(\boldsymbol{\theta}) &= -\sum_{j=1}^{k-1}\left(\theta_j \sum_{i=1}^{n} x_{ij}\right) - \vartheta^{-1}\sum_{i=1}^{n} x_{ik} + \sum_{i=1}^{k} \log h(\boldsymbol{x}_i), \\ \dot{\ell}_n(\boldsymbol{\theta})_a &= \frac{\partial}{\partial\theta_a}\ell_n(\boldsymbol{\theta}) \;=\; -\sum_{i=1}^{n} x_{ia} + r_a\vartheta^{-1}\theta_a^{-1}\sum_{i=1}^{n} x_{ik}, \\ \ddot{\ell}_n(\boldsymbol{\theta})_{ab} &= \frac{\partial^2}{\partial\theta_a\partial\theta_b}\ell_n(\boldsymbol{\theta}) \;=\; \begin{cases} -r_a(r_a+1)\vartheta^{-1}\theta_a^{-2}\sum_{i=1}^{n} x_{ik}, & (a=b), \\ -r_a r_b \vartheta^{-1}\theta_a^{-1}\theta_b^{-1}\sum_{i=1}^{n} x_{ik}, & (a \neq b). \end{cases} \end{aligned}$$

Here we newly redefine $\{X_i\}_{i=1}^{k}$ as follows:

$$X_1 = \sum_{i=1}^{n} X_{i1}, \quad \ldots, \quad X_{k-1} = \sum_{i=1}^{n} X_{i\,k-1}, \quad X_k = \sum_{i=1}^{n} X_{ik}. \tag{3.1}$$

Then, by the reproducibility of the gamma distribution, it holds that each X_i is distributed with the gamma distribution $G_A(nq_1, \theta_1^{-1})$, …, $G_A(nq_{k-1}, \theta_{k-1}^{-1})$, $G_A(nq_k, \vartheta)$, respectively. By the relation $\dot{\ell}_n(\boldsymbol{\theta}) = \boldsymbol{0}$, the maximum likelihood estimator(MLE) $\widehat{\boldsymbol{\theta}} = {}^t(\widehat{\theta}_1, \ldots, \widehat{\theta}_{k-1})$ is, for $i = 1, \ldots, k-1$,

$$\widehat{\theta}_i \;=\; \left(\frac{r_i X_k}{X_i}\right)^{\frac{r_*-r_i}{r_*}} \prod_{j\neq i}\left(\frac{r_j X_k}{X_j}\right)^{-\frac{r_j}{r_*}} \;=\; \left(\frac{X_i}{r_i}\right)^{-1} X_k^{\frac{1}{r_*}} \prod_{j=1}^{k-1}\left(\frac{X_j}{r_j}\right)^{\frac{r_j}{r_*}}. \tag{3.2}$$

For the joint density function of $(X_1, \ldots, X_k)$

$$f(x_1, \ldots, x_k) = \left(\prod_{j=1}^{k} \frac{x_j^{nq_j-1}}{\Gamma(nq_j)}\right) \exp\left\{-\sum_{i=1}^{k-1}\theta_i x_i - \vartheta^{-1}x_k\right\},$$

we suppose the variable transformation $\boldsymbol{t} = {}^t(t_1, \dots, t_{k-1}, t_k)$ as follows:

$$t_i = \left(\frac{x_i}{r_i}\right)^{-1} x_k^{\frac{1}{r_*}} \prod_{j=1}^{k-1} \left(\frac{x_j}{r_j}\right)^{\frac{r_j}{r_*}} \quad (i = 1, \dots, k-1), \quad t_k = x_k^{\frac{1}{r_*}} \prod_{j=1}^{k-1} \left(\frac{x_j}{r_j}\right)^{\frac{r_j}{r_*}} .$$

Remark that MLE (3.2) is equivalent to $(t_1, \dots, t_{k-1})$, that is,

$$t_i \;=\; \widehat{\theta}_i \qquad (i = 1, \dots, k-1).$$

Thus the relationship between $\boldsymbol{x}$ and $\boldsymbol{t}$ is

$$\boldsymbol{x} \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ x_k \end{pmatrix} \;=\; \begin{pmatrix} r_1\ t_1^{-1}\ t_k \\ \vdots \\ r_{k-1}\ t_{k-1}^{-1}\ t_k \\ t_k\ \prod_{i=1}^{k-1} t_i^{r_i} \end{pmatrix},$$

and the Jacobian is

$$\left|\frac{\partial(x_1, \dots, x_k)}{\partial(t_1, \dots, t_k)}\right| \;=\; \left(\prod_{j=1}^{k-1} r_j\right) r_* \left(\prod_{j=1}^{k-1} t_j^{r_j - 2}\right) t_k^{k-1},$$

so that the joint density function of $(T_1, \dots, T_k)$ is represented by

$$\begin{aligned} & g(t_1, \dots, t_k) \\ &\quad = \frac{r_* \left(\prod_{j=1}^{k-1} \frac{r_j^{nq_j}}{t_j}\right)}{\prod_{j=1}^{k} \Gamma(nq_j)}\, t_k^{n(q_1 + \cdots + q_k) - 1} \exp\left\{-t_k \left(\sum_{j=1}^{k-1} r_j \frac{\theta_j}{t_j} + \prod_{j=1}^{k-1} \left(\frac{t_j}{\theta_j}\right)^{r_j}\right)\right\}. \end{aligned}$$

Hence, by the probability density function of the gamma distribution, the marginal density function of $(T_1, \dots, T_{k-1})$, that is, the density function of MLE $\widehat{\boldsymbol{\theta}}$ (3.2) is

$$\begin{aligned} (3.3)\, g_{1\cdots k-1}(t_1, \dots, t_{k-1}) &= \int_0^\infty g(t_1, \dots, t_k) dt_k \\ = \frac{\Gamma(n(q_1 + \cdots + q_k))}{\prod_{j=1}^{k} \Gamma(nq_j)}\, r_* & \left(\prod_{j=1}^{k-1} \frac{r_j^{nq_j}}{t_j}\right) \left(\sum_{j=1}^{k-1} r_j \frac{\theta_j}{t_j} + \prod_{j=1}^{k-1} \left(\frac{t_j}{\theta_j}\right)^{r_j}\right)^{-n(q_1 + \cdots + q_k)} \end{aligned}$$

and the conditional density function of T_k given $\widehat{\boldsymbol{\theta}} = (T_1, \dots, T_{k-1})$ is

$$\begin{aligned} g_k(t_k | t_1, \dots, t_{k-1}) &= \frac{g(t_1, \dots, t_k)}{g_{1\cdots k-1}(t_1, \dots, t_{k-1})} \\ &= \frac{\left(\sum_{j=1}^{k-1} r_j \frac{\theta_j}{t_j} + \prod_{j=1}^{k-1} \left(\frac{t_j}{\theta_j}\right)^{r_j}\right)^{-n(q_1 + \cdots + q_k)}}{\Gamma(n(q_1 + \cdots + q_k))} \\ &\qquad \times t_k^{n(q_1 + \cdots + q_k) - 1} \exp\left\{-t_k \left(\sum_{j=1}^{k-1} r_j \frac{\theta_j}{t_j} + \prod_{j=1}^{k-1} \left(\frac{t_j}{\theta_j}\right)^{r_j}\right)\right\}. \end{aligned}$$

This is just the probability density function of the following gamma distribution :

$$G_A\left(n(q_1+\cdots+q_k),\ \left(\sum_{j=1}^{k-1} r_j\frac{\theta_j}{t_j}+\prod_{j=1}^{k-1}\left(\frac{t_j}{\theta_j}\right)^{r_j}\right)^{-1}\right).$$

Here we investigate the density (3.3) in detail, because it is the density of MLE and we use it to calculate the Fisher information by MLE.

Lemma 1 *The marginal density (3.3) is equivalent to the density of k-dimensional Dirichlet distribution $Diri(nq_1,\ldots,nq_{k-1};nq_k)$.*

Proof. In (3.3), we consider the transformation as follows:

$$t_i \;=\; \theta_i\left(\frac{w_i}{r_i}\right)^{-1} w_k^{\frac{1}{r_*}}\prod_{j=1}^{k-1}\left(\frac{w_j}{r_j}\right)^{\frac{r_j}{r_*}} \qquad (i=1,\ldots,k-1),$$

where we assume that $w_k = 1-\sum_{i=1}^{k-1} w_i$, so that the Jacobian is

$$\frac{\partial(t_1,\ldots,t_{k-1})}{\partial(w_1,\ldots,w_{k-1})} \;=\; \frac{\prod_{i=1}^{k-1} t_i}{r_*\,\prod_{i=1}^{k} w_i}.$$

Since

$$\sum_{j=1}^{k-1} r_j\frac{\theta_j}{t_j}+\prod_{j=1}^{k-1}\left(\frac{t_j}{\theta_j}\right)^{r_j} \;=\; \left(w_k^{\frac{1}{r_*}}\prod_{j=1}^{k-1}\left(\frac{w_j}{r_j}\right)^{\frac{r_j}{r_*}}\right)^{-1}$$

and $n(q_1+\cdots+q_k) = nq_kr_*$, it holds that

$$h_{1\cdots k-1}(w_1,\ldots,w_{k-1}) \;=\; \frac{\Gamma(n(q_1+\cdots+q_k))}{\prod_{j=1}^{k}\Gamma(nq_j)}\prod_{j=1}^{k} w_j^{nq_j-1}.$$

This is just the probability density function of k-dimensional Dirichlet distribution $Diri(nq_1,\ldots,nq_{k-1};nq_k)$. □

Note that the assumption $w_k = 1-\sum_{i=1}^{k-1} w_i$ in the proof of the above lemma is based on the property of the variables in k-dimensional Dirichlet distribution.

Lemma 2 *In the k-dimensional Dirichlet distribution $Diri(nq_1,\ldots,nq_{k-1};nq_k)$, we have the following expectations:*

$$\begin{aligned}
E(W_i) &= \frac{q_i}{q_1+\cdots+q_k} \quad (i=1,\ldots,k),\\
E(W_i^2) &= \frac{q_i(nq_i+1)}{(q_1+\cdots+q_k)(n(q_1+\cdots+q_k)+1)} \quad (i=1,\ldots,k),\\
E(W_iW_j) &= \frac{nq_iq_j}{(q_1+\cdots+q_k)(n(q_1+\cdots+q_k)+1)} \quad (i\neq j).
\end{aligned}$$

Proof. For arbitrary non-negative integers a,b,c and $i,j=1,\ldots,k-1$, it holds that

$$E(W_i^a\,W_j^b\,W_k^c) \;=\; \frac{\Gamma(n(q_1+\cdots+q_k))}{\prod_{j=1}^{k}\Gamma(nq_j)}\,\frac{\Gamma(nq_i+a)\,\Gamma(nq_j+b)\,\Gamma(nq_k+c)}{\Gamma(n(q_1+\cdots+q_k)+a+b+c)},$$

so that we obtain the results easily. □

Theorem 3 *The exact Fisher information of MLE $\widehat{\boldsymbol{\theta}}$ is represented by*

$$\boldsymbol{I}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \frac{n(q_1+\cdots+q_k)}{n(q_1+\cdots+q_k)+1}\ n\boldsymbol{I}(\boldsymbol{\theta}).$$

Proof. The derivative of log-likelihood function of MLE $\widehat{\boldsymbol{\theta}} = {}^t(T_1,\dots,T_{k-1})$ is

$$\dot{\ell}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) \ = \ -n(q_1+\cdots+q_k)\left(\frac{\frac{r_i}{t_i}-\frac{r_i}{\theta_i}\prod_{j=1}^{k-1}\left(\frac{t_j}{\theta_j}\right)^{r_j}}{\sum_{j=1}^{k-1} r_j\frac{\theta_j}{t_j}+\prod_{j=1}^{k-1}\left(\frac{t_j}{\theta_j}\right)^{r_j}}\right)_{i=1,\dots,k-1}$$

and is transformed by $\boldsymbol{W}$ as follows:

$$\dot{\ell}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = -n(q_1+\cdots+q_k)\left(\frac{w_i-r_iw_k}{\theta_i}\right)_{i=1,\dots,k-1}.$$

Hence the exact Fisher information is

$$\begin{aligned}\boldsymbol{I}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) &= E\left\{\dot{\ell}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\ {}^t\dot{\ell}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\right\}\\ &= \left(E\{(W_i-r_iW_k)(W_j-r_jW_k)\}\frac{(n(q_1+\cdots+q_k))^2}{\theta_i\theta_j}\right)_{i,j=1,\dots,k-1}\end{aligned}$$

and each component of the information is

$$\begin{aligned}E\{(W_i-r_iW_k)^2\} &= \frac{q_kr_i(r_i+1)}{(q_1+\cdots+q_k)(n(q_1+\cdots+q_k)+1)},\\ E\{(W_i-r_iW_k)(W_j-r_jW_k)\} &= \frac{q_kr_ir_j}{(q_1+\cdots+q_k)(n(q_1+\cdots+q_k)+1)}.\end{aligned}$$

Since the whole Fisher information is (2.4), we have the required result. □

Theorem 4 *The exact Fisher information loss of k-dimensional hyperbolic model is represented by*

$$\boldsymbol{I}_n(\boldsymbol{\theta})-\boldsymbol{I}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) \ = \ \frac{n}{n(q_1+\cdots+q_k)+1}\boldsymbol{I}(\boldsymbol{\theta}) \tag{3.4}$$

and this converges to $\mathbf{EIC}(\boldsymbol{\theta})$ *as follows:*

$$\lim_{n\to\infty}\left\{\boldsymbol{I}_n(\boldsymbol{\theta})-\boldsymbol{I}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\right\} \ = \ \frac{1}{q_1+\cdots+q_k}\boldsymbol{I}(\boldsymbol{\theta}) \ = \ \mathbf{EIC}(\boldsymbol{\theta}).$$

Proof. Because of Theorem 3, the exact Fisher information loss is

$$\boldsymbol{I}_n(\boldsymbol{\theta})-\boldsymbol{I}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \left(1-\frac{n(q_1+\cdots+q_k)}{n(q_1+\cdots+q_k)+1}\right)n\boldsymbol{I}(\boldsymbol{\theta}) = \frac{n}{n(q_1+\cdots+q_k)+1}\boldsymbol{I}(\boldsymbol{\theta}),$$

so that the convergence holds easily. □

The main contribution in this theorem is the exact formulation (3.4) of the exact information loss of k-dimensional hyperbolic model, because the exact representation does not seem to be appeared in the multi-dimensional model like this. Remark that the asymptotic representation of information loss in the above theorem is equivalent to Theorem 7 in Amari [1] which calculated the asymptotic information loss for Fisher-efficient estimators in multi-parameter curved exponential families. Note that if the statistical curvature is defined as the determinant of the residual by the projection, the convergence in Theorem 4 does not hold, that is, the limitation of Fisher's exact information loss is not equal to Efron's information curvature.

4 Discussion We explicitly investigated the exact information loss in the generalized hyperbolic model as an extension of the Nile problem. This would be useful for studying an exact information loss with respect to MLE for various models. But our parameterization in this model has a small problem which is the assumption of vanishing the cumulant generating function in it. Thus, in a generalized parameterization for the k-dimensional gamma distribution, the exact information loss could be a problem in the future. Further it is an interesting issue whether a similar result holds or not if other "equipotential" curved exponential model is considered, so that we need to investigate the properties and effectiveness of equipotentiality in the "equipotential" curved exponential model.

Acknowledgments

This article is partly supported by both grants-in-aid from Japan 15740059 for the first author and grants-in-aid from Japan 14540115 for the second author.

References

[1] S.Amari, Differential geometry of curved exponential families Efron — curvatures and information loss, Ann. Statist. 10 (1982) 357–385.

[2] P. Blæsild, The two-dimensional hyperbolic distribution and related distributions, with an application to Johannsen's bean data, Biometrika 68 (1981) 251–263.

[3] B.Efron, Defining the curvature of a statistical problem (with applications to second order efficiency), Ann. Statist. 6 (1975) 1189–1242.

[4] R.A.Fisher, *Statistical Methods and Scientific Inference*, Hafner Press, New York, 1973.

[5] N.Inagaki, The asymptotic distribution of a conditional likelihood ratio function given a maximum likelihood estimator in *Statistical Theory and Data Analysis (K.Matusita(Editor))*, North-Holland, Amsterdam, 1985.

[6] N.Inagaki and E.Kumagai, Exact information loss in Fisher's circle model, Mathematica Japonica, 44 (1996) 455–467.

[7] K.Inoue, N.Inagaki and E.Kumagai, The exact information loss of the maximum likelihood estimator in the k-dimensional sphere model, Mathematica Japonica, 52 (2000) 289–310.

Etsuo KUMAGAI,

Department of Financial Statistics, Graduate School of Engineering Science, Osaka University, Osaka, 560-8531, JAPAN

Nobuo INAGAKI,

Department of Financial Statistics, Graduate School of Engineering Science, Osaka University, Osaka, 560-8531, JAPAN