Ann. Inst. Statist. Math.
Vol. 52, No. 2, 197–214 (2000)

QUANTITATIVE APPROXIMATION TO THE ORDERED DIRICHLET DISTRIBUTION UNDER VARYING BASIC PROBABILITY SPACES

TOMOYA YAMADA[1] AND TADASHI MATSUNAWA[2]

[1] *The Graduate Universities for Advanced Studies, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan*
[2] *The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan*

(Received May 1, 1998; revised April 8, 1999)

Abstract. An approximate expansion of a sequence of ordered Dirichlet densities is given under the set-up with varying dimensions of the relating basic probability spaces. The problem is handled as the approximation to the joint distribution of an increasing number of selected order statistics based on the random sample drawn from the uniform distribution $U(0,1)$. Some inverse factorial series to the expansion of logarithmic function enable us to give quantitative error evaluations to our problem. With the help of them the relating modified K-L information number, which is defined on an approximate main domain and different from the usual ones, is accurately evaluated. Further, the proof of the approximate joint normality of the selected order statistics is more systematically presented than those given in existing works. Concerning the approximate normality the modified affinity and the half variation distance are also evaluated.

Key words and phrases: Ordered Dirichlet distribution, approximate distribution, sample quantiles, modified K-L information number, modified affinity, half variation distance, approximate main domain, approximate joint normality.

1. Introduction

The purpose of this paper is to give an approximate expansion of ordered Dirichlet distributions defined on the varying basic probability spaces when the underlying sample sizes increase. Such subject is generally difficult comparing with the case where the dimension of the basic spaces are fixed independent of the sample sizes. To overcome the difficulty we chiefly make use of the analytic tools developed by the second author of the present paper.

This research is motivated by the understanding of the concept of Kullback-Leibler mean information number remarked by Matsunawa (1995). In this paper, however, the quantity discussed mainly is an extended one which will be called a modified K-L information number: $I^*(P_n, Q_n; A_n) = \int_{A_n} \ln(dP_n/dQ_n)dP_n$, where P_n and Q_n are probability distributions which are absolutely continuous with respect to some σ-finite measure over some measurable set A_n, for each n. When $A_n = R_n$ (=the whole space, for each n), Matsunawa (1995) noticed from statistical and physical point of views that it is reasonable to consider Q_n as a basic (or an *original*) distribution and P_n as the approximate (or *developed*) distribution of Q_n. Thus, he interpreted the K-L information $I(P_n, Q_n)$ as the mean number with respect to P_n which should be understood as a relative number of the density dP_n of the approximate distribution to the density of dQ_n of the basic distribution. He stressed that it is required for us to recognize strict distinction between approximate and basic distributions. Namely, $I^*(P_n, Q_n; A_n)$and

$I^*(Q_n, P_n; A_n)$ have complete different meaning, although both of them are useful to estimate some uniform distance between P_n and Q_n. If we also adopt his understanding in this paper, it is quite natural to consider the approximation of joint sample quantiles with a Dirichlet distribution Q_n by a certain approximation distribution P_n based on the information $I^*(P_n, Q_n; A_n)$. On the contrary, if we make use of $I^*(Q_n, P_n; A_n)$, it means that we reversely consider the Dirichlet approximation of Q_n to the distribution P_n. The latter case was investigated by Ikeda and Matsunawa (1972) provided that P_n is a certain multi-dimensional normal distribution, where the Kullback-Leibler information was handled as a only tool to evaluate the uniform approximation based on a generalized Kolmogorov-Smirnov distance. Such being the case, the present authors have been aware that the use of the information in the previous joint work was not suitable and the direction characteristics of approximation associated with the above information should be paid more attentions from statistical and physical point of views (cf. Matsunawa (1995), Yamada and Matsunawa (1998)). So, we investigate the quantitative approximation to the joint distribution of sample quantiles by making allowance for the understanding of the related information numbers. In that case we will see that the evaluation of the probability, $\min\{P_n(A_n), Q_n(A_n)\}$, also becomes essentially important problems, and it will be treated as evaluation of probabilities on approximate main domains.

Let $\{\boldsymbol{X}_{n(k)}\}_{(n=1,2,\dots)}$ and $\{\boldsymbol{Y}_{n(k)}\}_{(n=1,2,\dots)}$ be two sequences of k-dimensional random vectors defined on the sequence of measurable space $(R_{n(k)}, \boldsymbol{B}_{n(k)})_{(n=1,2,\dots)}$, where $R_{(k)}$ is k-dimensional real space and $\boldsymbol{B}_{(k)}$ is Borel field of subsets of $R_{(k)}$, for each n. The dimension of the space k may vary as n increases . Denote their corresponding sequences of probability distributions by $\{P_n\}_{(n=1,2,\dots)}$ and $\{Q_n\}_{(n=1,2,\dots)}$, respectively. For each n, assume that these distributions are absolutely continuous with respect to Lebesgue measure μ on the measurable space and designate their respective densities by p_n and q_n. Under these set-up the following quantities are considered as measures of discrepancies between the two sequences of probability distributions:

(i) strong Kolmogorov-Smirnov uniform distance (=half variation distance)

$$D(\boldsymbol{X}_{n(k)}, \boldsymbol{Y}_{n(k)}; \boldsymbol{B}_{n(k)}) = \sup_{E \in \boldsymbol{B}_{(k)}} |P_n(E) - Q_n(E)|, \tag{1.1}$$

(ii) a modified Kullback-Leibler information number (=K-L information in short)

$$I^*(\boldsymbol{X}_{n(k)}, \boldsymbol{Y}_{n(k)}; C_{n(k)}) = \int_{C_{n(k)}} p_n(\boldsymbol{x}_{(k)}) \ln \frac{p_n(\boldsymbol{x}_{(k)})}{q_n(\boldsymbol{x}_{(k)})} d\boldsymbol{x}_{(k)}, \qquad C_{n(k)} \in \boldsymbol{B}_{n(k)}, \tag{1.2}$$

(iii) a modified Matusita affinity

$$\rho^*(\boldsymbol{X}_{n(k)}, \boldsymbol{Y}_{n(k)}; C_{n(k)}) = \int_{C_{n(k)}} \sqrt{p_n(\boldsymbol{x}_{(k)}) q_n(\boldsymbol{x}_{(k)})} d\boldsymbol{x}_{(k)}, \qquad C_{n(k)} \in \boldsymbol{B}_{n(k)}. \tag{1.3}$$

Concerning above quantities it is known that the following three statements are equivalent (cf. Matsunawa (1982)):

(1) $D(\boldsymbol{X}_{n(k)}, \boldsymbol{Y}_{n(k)}; \boldsymbol{B}_{n(k)}) \to 0$ as $n \to \infty$.

(2) There exists a sequence of measurable sets $\{A_{n(k)} \in \boldsymbol{B}_{n(k)}\}_{(n=1,2,\dots)}$ such that $P(A_{n(k)}) \to 1$ as $n \to \infty$ and that simultaneously $I^*(\boldsymbol{X}_{n(k)}, \boldsymbol{Y}_{n(k)}; A_{n(k)}) \to 0$ as $n \to \infty$.

(3) There exists a sequence of measurable sets $\{A_{n(k)} \in \boldsymbol{B}_{n(k)}\}_{(n=1,2,\dots)}$ such that $\rho^*(\boldsymbol{X}_{n(k)}, \boldsymbol{Y}_{n(k)}; A_{n(k)}) \to 1$ as $n \to \infty$.

As an interesting example of these equivalence we will ascertain that our uniform asymptotic result in this paper satisfies the fact. The above statements are also useful to analyze the uniform approximate equivalence of two sequences of probability distributions, even if the parameter n is not so large and if the dimension k varies according as n increases.

Now, we begin considering our approximation problem on the ordered Dirichlet distribution based on the order statistics from a uniform distribution. Let $U_{n1} < U_{n2} < \cdots < U_{nn}$ be the order statistics based on the random sample of size n drawn from the uniform distribution $U(0,1)$. Consider to select $k = k(n)$ $(< n)$ order statistics from the above whole order statistics and denote them as $U_{nn_1} < U_{nn_2} < \cdots < U_{nn_k}$ $(n_1 < n_2 < \cdots < n_k)$. In what follows we put conventions as $n_0 = 0$, $n_{k+1} = n+1$, $U_0 = 0$, $U_{k+1} = 1$. Then the joint pdf. of the random vector $\boldsymbol{U}_{n(k)} = (U_{nn_1}, \ldots, U_{nn_k})$ is given by the ordered Dirichlet density

$$h_n(\boldsymbol{z}_{n(k)}) = \left\{ n! \Big/ \prod_{i=1}^{k+1} d_i! \right\} \cdot \prod_{i=1}^{k+1} (z_i - z_{i-1})^{d_i}, \qquad (0 = z_0 < z_1 < \cdots < z_k < z_{k+1} = 1), \tag{1.4}$$

where $\boldsymbol{z}_{n(k)} = (z_1, \ldots, z_k) \in R_{(k)}$ and $d_i = n_i - n_{i-1} - 1$, $(i = 1, \ldots, k+1)$. Corresponding to $\boldsymbol{U}_{n(k)}$ let us consider the normal random vector $\boldsymbol{Z}_{n(k)} = (Z_{n1}, \ldots, Z_{nk})$ having the joint pdf.

$$g_n(\boldsymbol{z}_{n(k)}) = \left(\frac{1}{2\pi}\right)^{k/2} |\boldsymbol{L}_{n(k)}|^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{z}_{n(k)} - \boldsymbol{\ell}_{n(k)})^t \boldsymbol{L}_{n(k)}^{-1} (\boldsymbol{z}_{n(k)} - \boldsymbol{\ell}_{n(k)}) \right] \qquad (-\infty < z_i < \infty, i = 1, \ldots, k), \tag{1.5}$$

where $\boldsymbol{\ell}_{n(k)} = (\ell_{n1}, \ldots, \ell_{nk})^t$, with $\ell_{ni} = n_i/(n+1)$, $i = 1, \ldots, k$, $\ell_{n0} = 0$, $\ell_{nk+1} = 1$;

$$\boldsymbol{L}_{n(k)} := \frac{1}{n+2} (\ell_{ni}(1 - \ell_{nj}))_{k \times k} \ (1 \le i \le j \le k) \quad \text{and} \quad |\boldsymbol{L}_{n(k)}| = (n+2)^{-k} \prod_{i=1}^{k+1} (\ell_{ni} - \ell_{ni-1}). \tag{1.6}$$

In case of fixed k Mosteller (1946) and Walker (1968) gave the limiting joint normality of $\boldsymbol{U}_{n(k)}$ to $\boldsymbol{Z}_{n(k)}$ in the sense of weak convergence. In case of varying k Weiss (1969), Ikeda and Matsunawa (1972) and Reiss (1975) gave the asymptotic joint normality of the corresponding problem in the sense of variation distance between the probability distributions of the two random vectors. In this paper we also consider the approximation in the same sense. However, our approach mainly based on the modified K-L information of $g_n(\boldsymbol{z}_{n(k)})$ with respect to $h_n(\boldsymbol{z}_{n(k)})$ is very different from the above works. By making use of the information we can give an error evaluation on the approximate joint normality of k sample quantiles. In the process of the approximation a certain approximate main domain is introduced and it plays an important role in our approach.

In next section necessary lemmas are prepared. The first lemma is an expression on a logarithmic function. That is very useful and is expected to have wide applicabilities. With the help of the lemma an approximate expansion of ordered Dirichlet density is given in Section 3. In Section 4 an approximate joint normality of the selected sample quantiles is investigated under the situation of varying basic probability spaces. To get the result we make use of a modified information criterion with a certain approximate main domain. In Section 5 the same approximation problem as in Section 4 is discussed based on the modified affinity and the half variation distance.

2. Necessary lemmas

In this section two lemmas are presented. The first one is an expression for logarithmic functions which are needed to get an approximate expansion of the ordered Dirichlet density in Section 3. Another one is utilized to evaluate a certain measurable set whether it becomes an approximate main domain of our problem discussed in Section 4.

LEMMA 2.1. *Let L and M be certain given finite positive constants. Then, for any u such that $u > L$ or $u < -1 - M$ we can represent*

$$\ln\left(1+\frac{1}{u}\right) = \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \alpha\frac{1}{u^4} \tag{2.1}$$

where

$$\frac{1}{6}\left\{1 - \frac{1}{(L+1)^2}\right\} < \alpha = \alpha(L,M) < \max\left\{\frac{1}{3}, \frac{1}{4}\left(1+\frac{1}{M}\right)\right\}. \tag{2.2}$$

Remark 2.1. The usefulness of the above expression for us is to use it without knowing the sign of the variable u.

PROOF. (a) In case of $u > 0$ we make use of the following inequalities

$$\frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u^2(u+1)} \leq \ln\left(1+\frac{1}{u}\right) \leq \frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u(u+1)^2} \quad (u > 0), \tag{2.3}$$

which can be derived as follows: Let us put

$$\kappa(u) := \ln\left(1+\frac{1}{u}\right) - \left\{\frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u^2(u+1)}\right\} \quad \text{and}$$
$$\lambda(u) := \ln\left(1+\frac{1}{u}\right) - \left\{\frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u(u+1)^2}\right\},$$

then

$$\kappa'(u) = -\frac{1}{3u^3(u+1)^2} < 0, \quad \lim_{u\to 0}\kappa(u) = \infty \quad \text{and} \quad \lim_{u\to\infty}\kappa(u) = 0 \ \therefore \kappa(u) > 0$$
$$\lambda'(u) = \frac{3u+2}{3u^2(u+1)^3} > 0, \quad \lim_{u\to 0}\lambda(u) = -\infty \quad \text{and} \quad \lim_{u\to\infty}\lambda(u) = 0 \ \therefore \lambda(u) < 0,$$

from which we get the inequality (2.3). Modifying the upper bound of (2,3)

$$\begin{aligned}
&\frac{1}{u} - \frac{1}{2u^2} + \left\{\frac{1}{2u^2} - \frac{1}{2u(u+1)}\right\} + \left\{\frac{1}{3u^3} - \frac{1}{3u^3}\right\} - \frac{1}{6u(u+1)^2} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{6u^4} + \frac{1}{6u^4(u+1)^2} \\
&< \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{6}\left\{1 - \frac{1}{(L+1)^2}\right\}\frac{1}{u^4} \quad (\because u > 0).
\end{aligned} \tag{2.4}$$

Modification of the lower bound of (2.3) yields

$$\begin{aligned}
&\frac{1}{u}-\frac{1}{2u^2}+\left\{\frac{1}{2u^2}-\frac{1}{2u(u+1)}\right\}+\left\{\frac{1}{3u^3}-\frac{1}{3u^3}\right\}-\frac{1}{6u^2(u+1)}\\
&\quad=\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{3u^3(u+1)}\\
&\quad>\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{3u^4}\qquad(\because u>0). \qquad (2.5)
\end{aligned}$$

Thus, from (2.4) and (2.5)

$$(2.6)\quad \ln\left(1+\frac{1}{u}\right)=\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{\theta}{u^4}\left(u>L>0;\frac{1}{6}\left\{1-\frac{1}{(L+1)^2}\right\}<\theta<\frac{1}{3}\right).$$

(b) If $u<0$ in expression of $\ln(1+1/u)$, it is only meaningful when $u<-1$. Since $0<-1/u<1$, by applying Maclaurin's formula

$$\begin{aligned}
\ln\left(1+\frac{1}{u}\right)&=\ln\left(1-\left(-\frac{1}{u}\right)\right)\\
&=-\left(-\frac{1}{u}\right)-\frac{1}{2}\left(-\frac{1}{u}\right)^2-\frac{1}{3}\left(-\frac{1}{u}\right)^3-\frac{1}{4}\left(-\frac{1}{u}\right)^4-\frac{1}{5}\left(-\frac{1}{u}\right)^5-\cdots\\
&<\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{4u^4},\qquad(\because u<-1). \qquad (2.7)
\end{aligned}$$

On the other hand, for $u<-1-M$, we have

$$(2.8)\quad \ln\left(1+\frac{1}{u}\right)>\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{4u^4}\left(\frac{u}{u+1}\right)>\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{4u^4}\left(1+\frac{1}{M}\right).$$

Thus, form (2.7) and (2.8)

$$(2.9)\quad \ln\left(1+\frac{1}{u}\right)=\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{\vartheta}{4u^4},\qquad\left(u<-1-M;\frac{1}{4}<\vartheta<\frac{1}{4}\left(1+\frac{1}{M}\right)\right).$$

Consequently, combining (2.6) and (2.9), we have desired result (2.1) with (2.2).

Remark 2.2. The above lemma can be improved by sharpening the bounds of (2.3) with the help of more accurate inequalities developed in Matsunawa (1976).

The following lemma is utilized to evaluate the magnitude of the approximate main domain in our approximation problem discussed in Section 4.

LEMMA 2.2. (The probability of the half normal integral) *For $x>0$ it holds that*

$$\begin{aligned}
(2.10)\qquad &\frac{1}{2}\left(1-\frac{1}{2}e^{-x^2/2}-\frac{1}{2}e^{-(2-\sqrt{2})x^2}\right)^{1/2}\\
&\qquad<N(x)<\frac{1}{2}\left(1-\frac{1}{2}e^{-x^2}-\frac{1}{2}e^{-(2-\sqrt{2})x^2}\right)^{1/2},
\end{aligned}$$

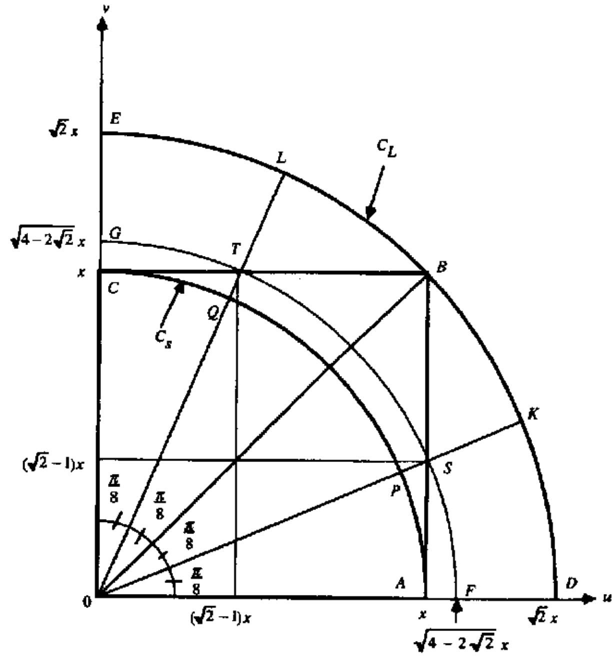

Fig. 1. Graphical evaluations of the half normal integral.

where

$$N(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\nu^2/2} d\nu. \tag{2.11}$$

PROOF. The inequalities (2.10) can be obtained by a simple geometrical argument based on the joint distribution of two independent normal random variables according to D'Ortenzio's idea introduced in the book by Johnson *et al.* (1994).

Consider to evaluate the following integral graphically

$$\vartheta(x) := \int_0^x \int_0^x e^{-(u^2+\nu^2)/2} du d\nu = \int_0^x e^{-u^2/2} du \int_0^x e^{-\nu^2/2} d\nu.$$

With the help of Fig. 1 we can observe that $\vartheta(x)$ is larger than the integral over the small circle C_s with radius x and smaller than the integral over the large circle C_L with radius $\sqrt{2}x$, in the first quadrant of the u-ν rectangular coordinate system. Putting $u = r\cos\theta$, $\nu = r\sin\theta$ with $r > 0$ and $0 \le \theta \le \pi/2$. Then, $u^2 + \nu^2 = r^2$ and the Jacobian of this transformation J is

$$J = \frac{D(u,\nu)}{D(r,\theta)} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} \quad \text{and} \quad |J| = r,$$

hence we have

$$\int_0^x e^{-r^2/2} r dr \int_0^{\pi/2} d\theta < \vartheta(x) < \int_0^{\sqrt{2}x} e^{-r^2/2} r dr \int_0^{\pi/2} d\theta.$$

Since

$$\int_0^x e^{-r^2/2} r dr = \int_0^x d(-e^{-r^2/2}) = 1 - e^{-x^2/2},$$

$$\int_0^{\sqrt{2}x} e^{-r^2/2} r dr = 1 - e^{-x^2} \quad \text{and} \quad \int_0^{\pi/2} d\theta = \frac{\pi}{2},$$

then it holds that

$$\frac{\pi}{2}(1-e^{-x^2/2}) < \vartheta(x) < \frac{\pi}{2}(1-e^{-x^2}). \tag{2.12}$$

Next, we try to improve the inequalities (2.12). The double integral of $e^{-(u^2+\nu^2)/2}$ over the area $PQTS$ in Fig. 1 is given by

$$\begin{aligned}\Delta^+ &:= \int_x^{\sqrt{4-2\sqrt{2}x}} e^{-r^2/2} r dr \int_{\tan^{-1}(\sqrt{2}-1)}^{\tan^{-1}(\sqrt{2}+1)} d\theta = [-e^{-r^2/2}]_x^{\sqrt{4-2\sqrt{2}x}} \cdot [\theta]_{\tan^{-1}(\sqrt{2}-1)}^{\tan^{-1}(\sqrt{2}+1)} \\ &= (e^{-x^2/2} - e^{-(2-\sqrt{2})x^2}) \cdot \{\tan^{-1}(\sqrt{2}+1) - \tan^{-1}(\sqrt{2}+1)\} \\ &= \frac{\pi}{4}(e^{-x^2/2} - e^{-(2-\sqrt{2})x^2}),\end{aligned}$$

because $\tan^{-1}(\sqrt{2}+1) = 3\pi/8$ and $\tan^{-1}(\sqrt{2}-1) = \pi/8$. Therefore, the lower bound of (2.12) can be graphically improved as

$$\vartheta(x) > \frac{\pi}{2}(1-e^{-x^2/2}) + \Delta^+ = \frac{\pi}{2}\left(1 - \frac{1}{2}e^{-x^2/2} - \frac{1}{2}e^{-(2-\sqrt{2})x^2}\right). \tag{2.13}$$

Similarly, the double integral of $e^{-(u^2+\nu^2)/2}$ over the area $DFSK$ in Fig. 1 is given by

$$\begin{aligned}\Delta_1^- &:= \int_{\sqrt{4-2\sqrt{2}x}}^{\sqrt{2}x} e^{-r^2/2} r dr \int_0^{\pi/8} d\theta = [-e^{-r^2/2}]_{\sqrt{4-2\sqrt{2}x}}^{\sqrt{2}x} \cdot \left\{\frac{\pi}{8} - 0\right\} \\ &= \frac{\pi}{8}(e^{-(2-\sqrt{2})x^2} - e^{-x^2}),\end{aligned}$$

and the double integral of $e^{-(u^2+\nu^2)/2}$ over the area $EGTL$ in Fig. 1 is given by

$$\begin{aligned}\Delta_2^- &:= \int_{\sqrt{4-2\sqrt{2}x}}^{\sqrt{2}x} e^{-r^2/2} r dr \int_{\tan^{-1}(\sqrt{2}+1)}^{\pi/2} d\theta = (e^{-(2-\sqrt{2})x^2} - e^{-x^2}) \cdot \left\{\frac{\pi}{2} - \frac{3\pi}{8}\right\} \\ &= \frac{\pi}{8}(e^{-(2-\sqrt{2})x^2} - e^{-x^2}).\end{aligned}$$

Therefore, the upper bound of (2.12) can be graphically improved as

$$\vartheta(x) < \frac{\pi}{2}(1-e^{-x^2}) - \Delta_1^- - \Delta_2^- = \frac{\pi}{2}\left(1 - \frac{1}{2}e^{-x^2} - \frac{1}{2}e^{-(2-\sqrt{2})x^2}\right). \tag{2.14}$$

Noticing the fact $N(x) = \sqrt{\vartheta(x)/(2\pi)}$, we get the desired inequalities (2.10).

Remark 2.3. The above lemma implies the following inequalities. For $x > 0$, the cdf. of the standard normal distribution $\Phi(x)$ is evaluated as

$$\begin{aligned}&\frac{1}{2}\left\{1 + \left(1 - \frac{1}{2}e^{-x^2/2} - \frac{1}{2}e^{-(2-\sqrt{2})x^2}\right)^{1/2}\right\} \\ &\qquad < \Phi(x) < \frac{1}{2}\left\{1 + \left(1 - \frac{1}{2}e^{-x^2} - \frac{1}{2}e^{-(2-\sqrt{2})x^2}\right)^{1/2}\right\}.\end{aligned} \tag{2.15}$$

These bounds are improved ones to those given by D'Ortenzio of the form

$$\frac{1}{2}[1 + (1-e^{-x^2/2})^{1/2}] < \Phi(x) < \frac{1}{2}[1 + (1-e^{-x^2})^{1/2}], \qquad (x > 0),$$

which can be derived from (2.12). It is possible to get more accurate bounds than those of (2.10) and (2.15) along the similar lines adopted above, although such approach becomes complicated.

3. An approximate expansion of ordered Dirichlet density

In this section we try to expand the ordered Dirichlet density $h_n(\boldsymbol{z}_{(k)})$ defined by (1.4). This was roughly carried out by Mosteller (1946). Different from his derivation we apply more accurate expansion to a logarithmic function given in Lemma 2.1. We also make use of a useful expression of Stirling's formula with accurate error bounds. Let

$$Q^0_{n(k)} = \{\boldsymbol{z}_{(k)} = (z_1, \dots, z_k); 0 \equiv z_0 < z_1 < \cdots < z_k < z_{k+1} \equiv 1\}$$

and for certain given finite positive constants L and M let us define

$$Q^{L,M}_{n(k)} = \left\{ \boldsymbol{z}_{(k)} = (z_1, \dots, z_k); 1 < \frac{z_1 - z_{i-1}}{\ell_{ni} - \ell_{ni-1}} < 1 + \frac{1}{L} \text{ or } \right.$$
$$\left. \frac{M}{M+1} < \frac{z_i - z_{i-1}}{\ell_{ni} - \ell_{ni-1}} < 1, i = 0, 1, \dots, k+1 \right\},$$

where as in Section 1 $\ell_{ni} = n_i/(n+1)$, $i = 1, \dots, k$, $\ell_{n0} = 0$, $\ell_{nk+1} = 1$ and $0 \equiv n_0 < n_1 < n_2 < \cdots < n_k < n_{k+1} \equiv n+1$. Under the same notations as those in Section 1 we have the following result:

THEOREM 3.1. *For $\boldsymbol{z}_{(k)} \in Q^0_{n(k)} \cap Q^{L,M}_{n(k)} =: A^{L,M}_{n(k)}$, the ordered Dirichlet distribution density can be expressed by*

$$(3.1) \quad h_n(\boldsymbol{z}_{(k)}) = g^*_n(\boldsymbol{z}_{(k)}) \cdot \left(1 + \frac{1}{n+1}\right)^{-k/2}$$
$$\times \exp\left[+ \frac{1}{12(n+1)} \left(1 - \sum_{i=1}^{k+1} \frac{1}{\ell_{ni+1} - \ell_{ni}}\right) - R(n+1) \right.$$
$$+ \sum_{i=1}^{k+1} R(d_i + 1) + \frac{1}{2(n+2)} (\boldsymbol{z}_{(k)} - \boldsymbol{\ell}_{n(k)})^t \boldsymbol{L}^{-1}_{n(k)} (\boldsymbol{z}_{(k)} - \boldsymbol{\ell}_{n(k)})$$
$$- \sum_{i=1}^{k+1} \frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}}$$
$$+ (n+1) \sum_{i=1}^{k+1} \left\{ \frac{\{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})\}^3}{3(\ell_{ni} - \ell_{ni-1})^2} \right\}$$
$$+ \frac{1}{2} \sum_{i=1}^{k+1} \left(\frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}} \right)^2$$
$$- \frac{1}{3} \sum_{i=1}^{k+1} \left(\frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}} \right)^3$$
$$- \theta(n+1) \sum_{i=1}^{k+1} \frac{\{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})\}^4}{(\ell_{ni} - \ell_{ni-1})^3}$$
$$\left. + \theta \sum_{i=1}^{k+1} \left(\frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}} \right)^4 \right],$$

where for $\boldsymbol{z}_{(k)} \in Q^0_{n(k)} \cap Q^{L,M}_{n(k)}$

$$g^*_n(\boldsymbol{z}_{(k)}) = (2\pi)^{-k/2} \cdot |\boldsymbol{L}_{n(k)}|^{-k/2} \cdot \exp\left[-\frac{1}{2}(\boldsymbol{z}_{(k)} - \boldsymbol{\ell}_{n(k)})^t \boldsymbol{L}^{-1}_{n(k)}(\boldsymbol{z}_{(k)} - \boldsymbol{\ell}_{n(k)})\right], \tag{3.2}$$

$$R(x) = \sum_{i=2}^{\infty} \frac{a_{i+1}}{x(x+1)\cdots(x+i)} \quad \text{with} \tag{3.3}$$

$$a_r = \frac{1}{r}\int_0^1 t(1-t)(2-t)\cdots(r-1-t)\left(\frac{1}{2}-t\right)dt,$$

and where for given finite positive constants L *and* M *and for any positive integer* n

$$\frac{1}{6}\left\{1 - \frac{1}{(L+1)^2}\right\} < \theta = \theta(n; L, M) < \max\left\{\frac{1}{3}, \frac{1}{4}\left(1 + \frac{1}{M}\right)\right\}. \tag{3.4}$$

Remark 3.1. It should be noted that $g^*_n(\boldsymbol{z}_{(k)})$ in the theorem has the same functional form as that of $g_n(\boldsymbol{z}_{(k)})$ in (1.5), but their domain is different. The function $R(x)$ can be evaluated by the inequalities in (3.7) given later.

PROOF. Let us put

$$t_i := \sqrt{n+1}(z_i - \ell_{ni}), \qquad (i = 0, 1, \ldots, k+1), \tag{3.5}$$

then $dz_i = (n+1)^{-1/2} dt_i$ and

$$\left\{n! \Big/ \prod_{i=1}^{k+1} d_i!\right\} dz_1 \cdots dz_k$$
$$= \frac{(n+1)!}{(n+1)n_1!(n_2-n_1)!\cdots(n_k-n_{k-1})!(n_{k+1}-n_k)!}\prod_{i=1}^{k+1}(n_i - n_{i-1})^{1/2}$$
$$\cdot (n+1)^{1/2}\prod_{i=1}^{k+1}(\ell_{ni} - \ell_{ni-1})^{1/2} \cdot dt_1 \cdots dt_k,$$

where we have used the relations $d_i = n_i - n_{i-1} - 1$, $(i = 1, \ldots, k+1)$ with $n_0 = 0$, $n_{k+1} = n+1$. Since for any positive x the following representation holds

$$\Gamma(x+1) = \sqrt{2\pi}x^{x+1/2}\exp\left\{-x + \frac{1}{12x} - R(x)\right\} \tag{3.6}$$

where

$$\frac{1}{360x(x+1)(x+2)} < R(x) < \frac{1}{360x(x+1)(x+2)} + \frac{1}{32x^2(x+1)(x+2)} \tag{3.7}$$

(cf. Matsunawa (1976)). Thus, after some manipulations we have

$$\left\{n! \Big/ \prod_{i=1}^{k+1} d_i!\right\} dz_1 \cdots dz_k \tag{3.8}$$
$$= \frac{\prod_{i=1}^{k+1}(\ell_{ni} - \ell_{ni-1})^{-(n_i - n_{i-1} - 1/2)}}{(2\pi)^{k/2}}$$
$$\cdot \exp\left(\frac{1}{12}\left(\frac{1}{n+1} - \sum_{i=1}^{k+1}\frac{1}{d_i+1}\right) - R(n+1) + \sum_{i=1}^{k+1} R(d_i+1)\right) \cdot dt_1 \cdots dt_k.$$

Next, we try to evaluate the following term

$$\prod_{i=1}^{k+1}(z_i - z_{i-1})^{d_i}$$
$$= \prod_{i=1}^{k+1}(\ell_{ni} - \ell_{ni-1})^{n_i - n_{i-1} - 1} \cdot \prod_{i=1}^{k+1}\left\{1 + \frac{t_i - t_{i-1}}{\sqrt{n+1}(\ell_{ni} - \ell_{ni-1})}\right\}^{(n+1)(\ell_{ni} - \ell_{ni-1}) - 1}.$$

Put

$$w_i = w_i(n) = \sqrt{n+1}(\ell_{ni} - \ell_{ni-1})/(t_i - t_{i-1}), \qquad (i = 1, \ldots, k+1) \tag{3.9}$$

then by Lemma 2.1 for given finite positive constants L and M and for any positive w_i such that $w_i > L$ or $w_i < -1 - M$ $(i = 1, \ldots, k+1)$ we can represent

$$\ln\left\{1 + \frac{t_i - t_{i-1}}{\sqrt{n+1}(\ell_{ni} - \ell_{ni-1})}\right\}^{(n+1)(\ell_{ni} - \ell_{ni-1})}$$
$$= \sqrt{n+1}(t_i - t_{i-1}) - \frac{1}{2} \cdot \frac{(t_i - t_{i-1})^2}{\ell_{ni} - \ell_{ni-1}}$$
$$+ \frac{(t_i - t_{i-1})^3}{3\sqrt{n+1}(\ell_{ni} - \ell_{ni-1})^2} - \theta\frac{(t_i - t_{i-1})^4}{(n+1)(\ell_{ni} - \ell_{ni-1})^3}.$$

Thus,

$$\prod_{i=1}^{k+1}\left\{1 + \frac{t_i - t_{i-1}}{\sqrt{n+1}(\ell_{ni} - \ell_{ni-1})}\right\}^{(n+1)(\ell_{ni} - \ell_{ni-1}) - 1}$$
$$= \prod_{i=1}^{k+1}\exp\cdot\left[\ln\left\{1 + \frac{t_i - t_{i-1}}{\sqrt{n+1}(\ell_{ni} - \ell_{ni-1})}\right\}^{(n+1)(\ell_{ni} - \ell_{ni-1})}\right]$$
$$\cdot\prod_{i=1}^{k+1}\exp\cdot\left[-\ln\left\{1 + \frac{t_i - t_{i-1}}{\sqrt{n+1}(\ell_{ni} - \ell_{ni-1})}\right\}\right]$$
$$= \exp\left[-\frac{1}{2}\sum_{i=1}^{k+1}\frac{(t_i - t_{i-1})^2}{\ell_{ni} - \ell_{ni-1}} - \frac{1}{\sqrt{n+1}}\sum_{i=1}^{k+1}\left\{\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}} - \frac{(t_i - t_{i-1})^3}{3(\ell_{ni} - \ell_{ni-1})^2}\right\}\right.$$
$$+ \frac{1}{n+1}\sum_{i=1}^{k+1}\left\{\frac{1}{2}\left(\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}}\right)^2 - \theta\frac{(t_i - t_{i-1})^4}{(\ell_{ni} - \ell_{ni-1})^3}\right\}$$
$$\left. - \frac{1}{3(n+1)^{3/2}}\sum_{i=1}^{k+1}\left(\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}}\right)^3 + \frac{\theta}{(n+1)^2}\sum_{i=1}^{k+1}\left(\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}}\right)^4\right]$$

$$= \exp\left[-\frac{1}{2}\left\{\sum_{i=1}^{k+1}\frac{\ell_{ni+1} - \ell_{ni-1}}{(\ell_{ni+1} - \ell_{ni})(\ell_{ni} - \ell_{ni-1})}t_i^2 - 2\sum_{i=1}^{k+1}\frac{1}{\ell_{ni} - \ell_{ni-1}}t_i t_{i-1}\right\}\right. \tag{3.10}$$
$$+ \frac{1}{\sqrt{n+1}}\left\{-\sum_{i=1}^{k+1}\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}}\right.$$
$$+ \frac{1}{3}\sum_{i=1}^{k+1}\frac{(\ell_{ni+1} - \ell_{ni-1})(\ell_{ni+1} - 2\ell_{ni} + \ell_{ni-1})}{(\ell_{ni+1} - \ell_{ni})^2(\ell_{ni} - \ell_{ni-1})^2}t_i^3$$

$$- \sum_{i=1}^{k+1} \frac{t_i^2 t_{i-1}}{(\ell_{ni} - \ell_{ni-1})^2} + \sum_{i=1}^{k+1} \frac{t_i t_{i-1}^2}{(\ell_{ni} - \ell_{ni-1})^2} \Bigg\}$$

$$+ \frac{1}{n+1} \sum_{i=1}^{k+1} \left\{ \frac{1}{2} \left(\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}} \right)^2 - \theta \frac{(t_i - t_{i-1})^4}{(\ell_{ni} - \ell_{ni-1})^3} \right\}$$

$$- \frac{1}{3(n+1)^{3/2}} \sum_{i=1}^{k+1} \left(\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}} \right)^3 + \frac{\theta}{(n+1)^2} \sum_{i=1}^{k+1} \left(\frac{t_i - t_{i-1}}{\ell_{ni} - \ell_{ni-1}} \right)^4 \Bigg].$$

Therefore, from (3.9) and (3.10), and reverting t_i to z_i $(i = 0, 1, \dots, k+1)$, we have an expression

$$h_n(\boldsymbol{z}_{(k)}) = (2\pi)^{-k/2} \cdot \left(\frac{n+1}{n+2} \right)^{k/2} (n+2)^{k/2} \prod_{i=1}^{k+1} (\ell_{ni} - \ell_{ni-1})^{-1/2}$$

$$\times \exp \Bigg[- \left(\frac{n+1}{n+2} \right) \frac{n+1}{2} \sum_{i=1}^{k+1} \Bigg\{ \frac{\ell_{ni+1} - \ell_{ni-1}}{(\ell_{ni+1} - \ell_{ni})(\ell_{ni} - \ell_{ni-1})} (z_i - \ell_{ni})^2$$

$$- \frac{2}{\ell_{ni} - \ell_{ni-1}} (z_i - \ell_{ni})(z_{i-1} - \ell_{ni-1}) \Bigg\}$$

$$+ \frac{1}{12(n+1)} \left(1 - \sum_{i=1}^{k+1} \frac{1}{\ell_{ni+i} - \ell_{ni}} \right) - R(n+1) + \sum_{i=1}^{k+1} R(d_i + 1)$$

$$- \sum_{i=1}^{k+1} \frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}}$$

$$+ \frac{1}{3}(n+1) \sum_{i=1}^{k+1} \frac{\{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})\}^3}{(\ell_{ni} - \ell_{ni-1})^2}$$

$$+ \frac{1}{2} \sum_{i=1}^{k+1} \left\{ \left(\frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}} \right)^2 \right\}$$

$$- \frac{1}{3} \sum_{i=1}^{k+1} \left(\frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}} \right)^3$$

$$- \theta(n+1) \sum_{i=1}^{k+1} \frac{\{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})\}^4}{(\ell_{ni} - \ell_{ni-1})^3}$$

$$+ \theta \sum_{i=1}^{k+1} \left(\frac{(z_i - \ell_{ni}) - (z_{i-1} - \ell_{ni-1})}{\ell_{ni} - \ell_{ni-1}} \right)^4 \Bigg],$$

from which and noticing the facts (1.6) we have the desired expression (3.1). □

4. An approximate joint normality of the selected $k(n)$ sample quantiles

In this section we proceed with considering the joint approximate normality of $k = k(n)$ selected sample quantiles of n order statistics based on random sample drawn from the uniform distribution $U(0, 1)$.

First, we consider evaluation of the modified Kullback-Leibler information number when we take the set $A_{n(k)}^{L,M} := Q_{n(k)}^0 \cap Q_{n(k)}^{L,M}$ as the approximate main domain of our problem. Namely, we need to evaluate

$$I^*(g_n, h_n; A^{L,M}_{n(k)}) = \int_{A^{L,M}_{n(k)}} g_n(\mathbf{z}_{(k)}) \ln \frac{g_n(\mathbf{z}_{(k)})}{h_n(\mathbf{z}_{(k)})} d\mathbf{z}_{(k)}, \tag{4.1}$$

where, as in Section 2, $g_n(\mathbf{z}_{(k)})$ is the pdf. of the joint normal random vector $\mathbf{Z}_{n(k)}$ given by (1.5), and where $h_n(\mathbf{z}_{(k)})$ is the joint pdf. of the selected $k(n)$ sample quantiles $\mathbf{U}_{n(k)}$ given by (1.4). In what follows we denote the probability distributions of $\mathbf{U}_{n(k)}$ and $\mathbf{Z}_{n(k)}$ by $P^{\mathbf{U}_{n(k)}}$ and $P^{\mathbf{Z}_{n(k)}}$, respectively. It should be noted that Ikeda and Matsunawa (1972) evaluated the related K-L information $I(h_n, g_n; Q^0_{n(k)})$. This information is more feasible than $I^*(g_n, h_n; A^{L,M}_{n(k)})$ in calculation. However, it seems not to be adequate to handle our approximation problem from the aspects of directionality of approximation and related physical point of view (cf. Matsunawa (1995)). On the contrary, adopting $I^*(g_n, h_n; A^{L,M}_{n(k)})$ needs our careful treatment of $A^{L,M}_{n(k)}$, which is treated as the approximation main domain.

We have a theorem on the approximate normality of ordered Dirichlet random vector $\mathbf{U}_{n(k)}$:

THEOREM 4.1. *If the following condition is satisfied*

$$\varepsilon_n := \frac{k(n)}{\min_{1\le i\le k+1}(n_i - n_{i-1})} \to 0 \quad \text{as} \quad n \to \infty, \tag{4.2}$$

then $P^{\mathbf{U}_{n(k)}}$ is asymptotically normally distributed to $P^{\mathbf{Z}_{n(k)}}$ in the sense of

$$I^*(g_n, h_n; A^{L,M}_{n(k)}) = \int_{A^{L,M}_{n(k)}} g_n(\mathbf{z}_{(k)}) \ln \frac{g_n(\mathbf{z}_{(k)})}{h_n(\mathbf{z}_{(k)})} d\mathbf{z}_{(k)} \to 0 \quad \text{as} \quad n \to \infty, \tag{4.3}$$

and simultaneously

$$P^{\mathbf{Z}_{n(k)}}(A^{L,M}_{n(k)}) \to 1 \quad \text{as} \quad n \to \infty, \tag{4.4}$$

where for given finite positive constants L and M,

$$A^{L,M}_{n(k)} = \left\{ \mathbf{z}_{(k)} = (z_1, \dots, z_k) \;\middle|\; \begin{array}{l} 0 \equiv z_0 < z_1 < \cdots < z_k < z_{k+1} \equiv 1, \\ 1 < \dfrac{z_i - z_{i-1}}{\ell_{ni} - \ell_{ni-1}} < 1 + \dfrac{1}{L} \text{ or} \\ \dfrac{M}{M+1} < \dfrac{z_i - z_{i-1}}{\ell_{ni} - \ell_{ni-1}} < 1, i = 0, 1, \dots, k+1 \end{array} \right\}. \tag{4.5}$$

PROOF. Making use of Theorem 3.1, for $\mathbf{z}_{n(k)} \in A^{L,M}_{n(k)}$ we have

$$\begin{aligned} |I^*(g_n, h_n; A^{L,M}_{n(k)})| &< \int_{A_{n(k)}} g_n(\mathbf{z}_{n(k)}) |\ln(g_n(\mathbf{z}_{n(k)})/h_n(\mathbf{z}_{n(k)}))| d\mathbf{z}_{n(k)} \\ &=: E^*_g[|\ln(g_n(\mathbf{Z}_{n(k)})/h_n(\mathbf{Z}_{n(k)}))|] \\ &< \frac{k}{2}\left\{\frac{1}{n+1} - \frac{1}{6k(n+1)} - \frac{1}{2(n+1)(n+2)} - \frac{1}{6(n+1)(n+2)^2}\right\} \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{12(n+1)}\sum_{i=1}^{k+1}\frac{1}{\ell_{ni}-\ell_{ni-1}}+R(n+1)\\
&+\frac{1}{2(n+2)}E_g^*[(\boldsymbol{Z}_{n(k)}-\boldsymbol{\ell}_{n(k)})^t\boldsymbol{L}_{n(k)}^{-1}(\boldsymbol{Z}_{n(k)}-\boldsymbol{\ell}_{n(k)})]\\
&+\sum_{i=1}^{k+1}\frac{1}{\ell_{ni}-\ell_{ni-1}}E_g^*[|(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})|]\\
&+\frac{1}{2}\sum_{i=1}^{k+1}\frac{1}{(\ell_{ni}-\ell_{ni-1})^2}E_g^*[|(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})|^2]\\
&+\frac{1}{3}\sum_{i=1}^{k+1}\left(\frac{n+1}{(\ell_{ni}-\ell_{ni-1})^2}+\frac{1}{(\ell_{ni}-\ell_{ni-1})^3}\right)\\
&\cdot E_g^*[|(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})|^3]\\
&+|\theta|\sum_{i=1}^{k+1}\left\{\frac{n+1}{(\ell_{ni}-\ell_{ni-1})^3}+\frac{1}{(\ell_{ni}-\ell_{ni-1})^4}\right\}\\
&\cdot E_g^*[\{(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})\}^4].
\end{aligned}
$$

Making use of the following inequalities

$$
\begin{aligned}
&E_g^*[(\boldsymbol{Z}_{n(k)}-\boldsymbol{\ell}_{n(k)})^t\boldsymbol{L}_{n(k)}^{-1}(\boldsymbol{Z}_{n(k)}-\boldsymbol{\ell}_{n(k)})]\\
&\qquad<E_g[(\boldsymbol{Z}_{n(k)}-\boldsymbol{\ell}_{n(k)})^t\boldsymbol{L}_{n(k)}^{-1}(\boldsymbol{Z}_{n(k)}-\boldsymbol{\ell}_{n(k)})]=k,\\
&E_g^*[|(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})|]\\
&\qquad<\frac{\sqrt{2}}{\sqrt{\pi}(n+2)^{1/2}}[(\ell_{ni}-\ell_{ni-1})\{1-(\ell_{ni}-\ell_{ni-1})\}]^{1/2},\\
&E_g^*[|(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})|^2]<\frac{1}{n+2}[(\ell_{ni}-\ell_{ni-1})\{1-(\ell_{ni}-\ell_{ni-1})\}],\\
&E_g^*[|(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})|^3]\\
&\qquad<\frac{2\sqrt{2}}{\sqrt{\pi}(n+2)^{3/2}}[(\ell_{ni}-\ell_{ni-1})\{1-(\ell_{ni}-\ell_{ni-1})\}]^{3/2},\\
&E_g^*[\{(Z_{ni}-\ell_{ni})-(Z_{ni-1}-\ell_{ni-1})\}^4]<\frac{3}{(n+2)^2}[(\ell_{ni}-\ell_{ni-1})\{1-(\ell_{ni}-\ell_{ni-1})\}]^2,
\end{aligned}
$$

we have

$$
\begin{aligned}
&|I^*(g_n,h_n;A_{n(k)}^{L,M})|\\
&\quad<\frac{k}{n+1}\left\{1-\frac{1}{12k}-\frac{3}{4(n+2)}-\frac{1}{12(n+2)^2}\right\}+\frac{1}{360(n+1)(n+2)(n+3)}\\
&\quad+\frac{1}{32(n+1)^2(n+2)(n+2)}+\left\{\frac{1}{12}+3\theta\left(\frac{n+1}{n+2}\right)^2+\frac{n+1}{2(n+2)}\right\}\left(1+\frac{1}{k}\right)\varepsilon_n\\
&\quad+\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{1/2}\left\{1+\frac{2}{3}\left(\frac{n+1}{n+2}\right)\right\}\varepsilon_n^{1/2}\\
&\quad+\frac{2}{3}\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{3/2}\left(1+\frac{1}{k}\right)\frac{\varepsilon_n^{3/2}}{\sqrt{k}}+3\theta\left(\frac{n+1}{n+2}\right)^2\left(1+\frac{1}{k}\right)\frac{\varepsilon_n^2}{k}
\end{aligned}
$$

$$(4.6)\quad \to 0 (n\to\infty),\quad \text{if}\quad \varepsilon_n \to 0 (n\to\infty)$$
$$\left(\text{since}\frac{1}{6}\left\{1-\frac{1}{(L+1)^2}\right\} < \theta < \max\left\{\frac{1}{3},\frac{1}{4}\left(1+\frac{1}{M}\right)\right\}\right).$$

Next, we prove (4.4). If the following statement holds, we call the set $A^{L,M}_{n(k)}$ an approximate main domain of our problem. "For certain given finite positive constants L and M, and for any sequence of positive numbers $\{\varepsilon_n\}_{(n=1,2,\dots)}$ such that $\varepsilon_n \to 0$ as $n\to\infty$, there exist (1) a sequence of positive valued functions $\{\eta_n(\varepsilon_n; L, M)\}_{(n=1,2,\dots)}$ such that $0<\eta_n<1$ and $\eta_n \to 0$ as $n\to\infty$, (2) a sequence of measurable subsets $\{A^{L,M}_{n(k)} \in \boldsymbol{B}_{n(k)}\}_{(n=1,2,\dots)}$, and (3) a positive integer $n_0 = n_0(\varepsilon_n; L, M)$, such that for any $n \geq n_0$

$$P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)}) \geq 1-\eta(\varepsilon_n).\text{"} \tag{4.7}$$

Thus, if (4.7) is obtained, we can show the condition (4.4) as $n\to\infty$.

Let us decompose the domain $A^{L,M}_{n(k)}$ as $A^{L,M}_{n(k)} = A^{+L}_{n(k)} \cup A^{-M}_{n(k)}$, where for any given positive integers L and M the sub-domains are given by

$$A^{+L}_{n(k)} = Q^0_{n(k)} \cap \left\{\boldsymbol{z}_{(k)} = (z_1,\dots,z_k); 1 < \frac{z_i - z_{i-1}}{\ell_{ni}-\ell_{ni-1}} < 1+\frac{1}{L}, i=0,1,\dots,k+1\right\}, \tag{4.8}$$

$$A^{-M}_{n(k)} = Q^0_{n(k)} \cap \left\{\boldsymbol{z}_{(k)} = (z_1,\dots,z_k); \frac{M}{M+1} < \frac{z_i - z_{i-1}}{\ell_{ni}-\ell_{ni-1}} < 1, i=0,1,\dots,k+1\right\}, \tag{4.9}$$

here $Q^0_{n(k)}$ is the set defined in Section 3. In order to estimate the probability we make use of the following transformation. For each n, let $\boldsymbol{Z}_{n(k)} = (Z_{n1},\dots,Z_{nk})$ be a normal random vector which is distributed according to $N(\boldsymbol{\ell}_{n(k)}, \boldsymbol{L}_{n(k)})$. Consider the following transformed random vector $\boldsymbol{V}_{n(k)} = (V_{n1},\dots,V_{nk})$ of $\boldsymbol{Z}_{n(k)}$ through

$$V_{ni} := \frac{\sqrt{n+2}\{(Z_{ni}-\ell_{ni}) - (Z_{ni-1}-\ell_{ni-1})\}}{\sqrt{\ell_{ni}-\ell_{ni-1}}} \quad (i=1,\dots,k). \tag{4.10}$$

Then,

$$Z_{ni}-\ell_{ni} = \sum_{j=1}^{i}(\ell_{nj}-\ell_{nj-1})^{1/2}V_{ni}\Big/\sqrt{n+2} \quad (i=1,\dots,k), \tag{4.11}$$

and thus the pdf. of $\boldsymbol{V}_{n(k)}$ is given by for $\boldsymbol{\nu}_{(k)} = (\nu_1,\dots,\nu_k)$

$$\varphi(\boldsymbol{\nu}_{(k)}) = (2\pi)^{-k/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{k}\nu_i^2\right), \quad (-\infty<\nu_i<\infty, i=1,\dots,k), \tag{4.12}$$

since the Jacobian of the transformation is $J(\boldsymbol{z}_{(k)} \to \boldsymbol{\nu}_{(k)}) = \prod_{i=1}^{k}(\ell_{ni}-\ell_{ni-1})^{1/2}/\sqrt{n+2}$. It is noted that each components of $\boldsymbol{V}_{n(k)}$ is independently and identically distributed according to the standard normal distribution. Under the transformation

(4.11) let us consider the sets corresponding to $A^{+L}_{n(k)}$ and $A^{-M}_{n(k)}$ given by (4.8) and (4.9), respectively:

$$(4.13)\quad B^{+L}_{n(k)} = \left\{ \nu_{(k)} = (\nu_1, \ldots, \nu_k)^t; 0 \le \nu_i < \sqrt{\frac{n+2}{n+1}} \frac{\sqrt{n_i - n_{i-1}}}{L}, (i = 1, \ldots, k) \right\},$$

$$(4.14)\quad B^{-M}_{n(k)} = \left\{ \nu_{(k)} = (\nu_1, \ldots, \nu_k)^t; -\sqrt{\frac{n+2}{n+1}} \frac{\sqrt{n_i - n_{i-1}}}{1+M} < \nu_i \le 0, (i = 1, \ldots, k) \right\}.$$

Therefore,

$$\begin{aligned} P^{\mathbf{Z}_{n(k)}}(A^{L,M}_{n(k)}) &= P^{\mathbf{V}_{n(k)}}(B^{+L}_{n(k)}) + P^{\mathbf{V}_{n(k)}}(B^{-M}_{n(k)}) \\ &= (2\pi)^{-k/2} \int_{B^{+L}_{n(k)} \cup B^{-M}_{n(k)}} \exp\left[-\frac{1}{2} \sum_{i=1}^{k} \nu_i^2 \right] d\nu_{(k)} \\ &= \prod_{j=1}^{k} \left\{ \frac{1}{\sqrt{2\pi}} \int_0^{\sqrt{(n+2)/(n+1)}(\sqrt{n_i - n_{i-1}}/L)} \exp\left(-\frac{\nu_i^2}{2} \right) \cdot d\nu_j \right. \\ &\qquad \left. + \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{(n+2)/(n+1)}(\sqrt{n_i - n_{i-1}}/(1+M))}^{0} \exp\left(-\frac{\nu_i^2}{2} \right) \cdot d\nu_j \right\} \\ &=: \prod_{j=1}^{k} \left\{ N\left(\sqrt{\frac{n+2}{n+1}} \frac{\sqrt{n_i - n_{i-1}}}{L} \right) + N\left(\sqrt{\frac{n+2}{n+1}} \frac{\sqrt{n_i - n_{i-1}}}{1+M} \right) \right\}, \end{aligned}$$

where

$$(4.15)\quad N(x) := \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\nu^2/2} d\nu, \qquad (x > 0),$$

and

$$(4.16)\quad \frac{1}{2}\left(1 - \frac{1}{2}e^{-x^2/2} - \frac{1}{2}e^{-(2-\sqrt{2})x^2}\right)^{1/2} < N(x) < \frac{1}{2}\left(1 - \frac{1}{2}e^{-x^2} - \frac{1}{2}e^{-(2-\sqrt{2})x^2}\right)^{1/2}.$$

Consequently, making use of Lemma 2.2 and the condition (4.2), we have

$$P^{\mathbf{Z}_{n(k)}}(A^{L,M}_{n(k)}) > \left(\frac{1}{2}\right)^k \prod_{j=1}^{k} \left[\begin{aligned} &\left[1 - \frac{1}{2}\exp\left\{ -\frac{1}{2}\left(\frac{n+2}{n+1} \cdot \frac{n_j - n_{j-1}}{L^2} \right) \right\} \right. \\ &\quad \left. -\frac{1}{2}\exp\left\{ -(2-\sqrt{2})\left(\frac{n+2}{n+1} \cdot \frac{n_j - n_{j-1}}{L^2} \right) \right\} \right]^{1/2} \\ &+ \left[1 - \frac{1}{2}\exp\left\{ -\frac{1}{2}\left(\frac{n+2}{n+1} \cdot \frac{n_j - n_{j-1}}{(1+M)^2} \right) \right\} \right. \\ &\quad \left. -\frac{1}{2}\exp\left\{ -(2-\sqrt{2})\left(\frac{n+2}{n+1} \cdot \frac{n_j - n_{j-1}}{(1+M)^2} \right) \right\} \right]^{1/2} \end{aligned} \right]$$

$$(4.17)\qquad \to 1 \quad \text{as} \quad n \to \infty$$

which completes the proof of the theorem. □

5. Approximations based on the modified affinity and the half variation distance

In this section we also assume the condition (4.2) in Theorem 4.1. As is pointed out in Section 1 the asymptotic character described in the theorem can be seen by the modified affinity on $A^{L,M}_{n(k)}$ and by the half variation distance as follows.

$$\begin{aligned}(1 \geq)\rho^*(g_n, h_n; A^{L,M}_{n(k)}) &= \int_{A^{L,M}_{n(k)}} \sqrt{g_n(\boldsymbol{z}_{(k)})h_n(\boldsymbol{z}_{(k)})}d\boldsymbol{z}_{(k)} \\ &\geq P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)}) \exp\left\{ \frac{1}{P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)})} \int_{A^{L,M}_{n(k)}} g_n(\boldsymbol{z}_{(k)}) \cdot \ln \sqrt{\frac{h_n(\boldsymbol{z}_{(k)})}{g_n(\boldsymbol{z}_{(k)})}} d\boldsymbol{z}_{(k)} \right\} \\ &\qquad (\because \text{Jensen's Inequality}) \\ &= P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)}) \exp\left\{ -\frac{1}{2P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)})} I^*(g_n, h_n; A^{L,M}_{n(k)}) \right\} \\ &\geq P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)}) \cdot \exp\left\{ -\frac{1}{2P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)})} |I^*(g_n, h_n; A^{L,M}_{n(k)})| \right\} \end{aligned} \tag{5.1}$$

In view of (4.6) and (4.17) ,

$$\rho^*(g_n, h_n; A^{L,M}_{n(k)}) \to 1 \quad \text{as} \quad n \to \infty, \tag{5.2}$$

which shows the statement (3) in Section 1.

Next, we evaluate the half-variation distance between the two distributions:

$$\begin{aligned} D(\boldsymbol{U}_{n(k)}, \boldsymbol{Z}_{n(k)}; \boldsymbol{B}_{n(k)}) &= \sup_{E \in \boldsymbol{B}_{(k)}} |P^{\boldsymbol{U}_{n(k)}}(E) - P^{\boldsymbol{Z}_{n(k)}}(E)| \\ &= \frac{1}{2} \int |h_n(\boldsymbol{z}_{(k)}) - g_n(\boldsymbol{z}_{(k)})| d\boldsymbol{z}_{(k)} \\ &\leq \sqrt{1 - \left(\int_{R_{n(k)}} \sqrt{g_n(\boldsymbol{z}_{(k)})h_n(\boldsymbol{z}_{(k)})} d\boldsymbol{z}_{(k)} \right)^2} \\ &\leq \sqrt{1 - (\rho^*(g_n, h_n; A^{L,M}_{n(k)}))^2} \to 0, \quad \text{as} \quad n \to \infty, \end{aligned} \tag{5.3}$$

which coincides with the statement (1) in Section 1.

Next, it can be shown (cf. Matsunawa (1986)) that

$$\begin{aligned} D(\boldsymbol{U}_{n(k)}, \boldsymbol{Z}_{n(k)}; \boldsymbol{B}_{n(k)}) &= \sup_{E \in \boldsymbol{B}_{(k)}} |P^{\boldsymbol{U}_{n(k)}}(E) - P^{\boldsymbol{Z}_{n(k)}}(E)| \\ &\geq \ell(a_n^+) I^*(g_n, h_n; A^{L,M}_{n(k)} \cap F^+_{n(k)}) \\ &\quad + (P^{\boldsymbol{U}_{n(k)}}(A^{L,M}_{n(k)}) - P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)}))^- \end{aligned} \tag{5.4}$$

$$\begin{aligned} &= \ell(a_n^+)(|I^*(g_n, h_n; A^{L,M}_{n(k)})|)^+ \\ &\quad + (P^{\boldsymbol{U}_{n(k)}}(A^{L,M}_{n(k)}) - P^{\boldsymbol{Z}_{n(k)}}(A^{L,M}_{n(k)}))^- \geq 0, \end{aligned} \tag{5.5}$$

where $(x)^- = \max(-x, 0)$ and $(x)^+ = \max(x, 0)$, and where

$$F^+_{n(k)} = \{z_{(k)}; g_n(z_{(k)}) \geq h_n(z_{(k)})\} \quad \text{and}$$
$$a^+_n := \inf\{h_n(z_{(k)})/g_n(z_{(k)}); z_{(k)} \in A^{L,M}_{n(k)} \cap F^+_{n(k)}\}$$

and where for $t > 0$

$$(5.6) \qquad \ell(t) = 210t^{1/3}(1+t^{1/3})^7(1+t^{1/3}+t^{2/3}) \Big/ \begin{pmatrix} 35 + 1832t^{1/3} + 7796t^{2/3} + 18968t + 23378t^{4/3} \\ +18968t^{5/3} + 7796t^2 + 1832t^{7/3} + 35t^{8/3} \end{pmatrix}.$$

(cf. Matsunawa (1986)) which is easily seen a monotone increasing function in t. For each n, since $0 < a^+_n \leq 1$, then $0 < \ell(a^+_n) \leq \ell(1) = 1$, from (5.5) we have the result: If the following statement $(\alpha) D(U_{n(k)}, Z_{n(k)}; B_{n(k)}) \to 0$ as $n \to \infty$ holds, then from the inequality (5.5) we have $(\beta) |I^*(g_n, h_n; A^{L,M}_{n(k)})| \to 0$ as $n \to \infty$ and simultaneously $(\gamma) P^{U_{n(k)}}(A^{L,M}_{n(k)}) - P^{Z_{n(k)}}(A^{L,M}_{n(k)}) \to 0$ as $n \to \infty$. Finally, under (α) we show that the condition (4.2) holds: $P^{Z_{n(k)}}(A^{L,M}_{n(k)}) \to 1$ as $n \to \infty$. From (α) the uniform asymptotic equivalence $Z_{n(k)} \sim U_{n(k)}$, $(n \to \infty)$ holds for all Borel sets in $B_{n(k)}$. So, corresponding to $V_{n(k)} = (V_{n1}, \ldots, V_{nk})$ with (4.10), the random vector $S_{n(k)} := (S_{n1}, \ldots, S_{nk})$ with the components

$$(5.7) \qquad S_{ni} := \frac{\sqrt{n+2}\{(U_{ni} - \ell_{ni}) - (U_{ni-1} - \ell_{ni-1})\}}{\sqrt{\ell_{ni} - \ell_{ni-1}}} \qquad (i = 1, \ldots, k),$$

is asymptotically normally distributed according to (4.12) as $n \to \infty$. Thus,

$$P^{U_{n(k)}}(A^{L,M}_{n(k)}) = P^{S_{n(k)}}(B^{+M}_{n(k)}) + P^{S_{n(k)}}(B^{-M}_{n(k)})$$
$$\approx P^{V_{n(k)}}(B^{+M}_{n(k)}) + P^{V_{n(k)}}(B^{-M}_{n(k)}) \to 1 \quad \text{as} \quad n \to \infty.$$

Therefore, from (γ) we have the desired result. Consequently, in case of our approximation problem, we have proved the equivalent relations among three statements in Section 1.

References

Ikeda, S. and Matsunawa, T. (1972). On the uniform asymptotic joint normality of sample quantiles, *Ann. Inst. Statist. Math.*, **24**, 33–52.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., Wiley, New York.

Matsunawa, T. (1976). Some inequalities based on inverse factorial series, *Ann. Inst. Statist. Math.*, **28**, 291–305.

Matsunawa, T. (1982). Uniform ϕ - equivalence of probability distributions based on information and related measures of discrepancy, *Ann. Inst. Statist. Math.*, **34**, 1–17.

Matsunawa, T. (1986). Modified information criteria for a uniform approximate equivalence of probability distributions, *Ann. Inst. Statist. Math.*, **38**, 205–222.

Matsunawa, T. (1995). Development of distributions — The Legendre transformation and canonical information criteria —, *Pro. Inst. Statist. Math.*, **43**, 293–311(in Japanese).

Mosteller, F. (1946). On some useful "inefficient" statistics, *Ann. Math. Statist.*, **17**, 377–408.

Reiss, R. D. (1974). The asymptotic normality and asymptotic expansions for the joint distribution of several order statistics, *Colloq. Math. Soc. János Bolyai*, 297–340.

Walker, A. M. (1968). A note on the asymptotic distribution of sample quantiles, *J. Roy. Statist. Soc. Ser. B*, **30**, 570–575.

Weiss, L. (1969). The asymptotic joint distribution of an increasing number of sample quantiles, *Ann. Inst. Statist. Math.*, **21**, 257–263.

Yamada, T. and Matsunawa, T. (1998). Uniform approximations to probability distributions based on K-L information defined on approximate main domains — With applications to quantitative evaluations of fluctuations in multivariate general exponential families —, *Proc. Inst. Statist. Math.*, **46**, 461–476 (in Japanese).