Ann. Inst. Statist. Math.
34 (1982), Part A, 1-17

UNIFORM ϕ-EQUIVALENCE OF PROBABILITY DISTRIBUTIONS BASED ON INFORMATION AND RELATED MEASURES OF DISCREPANCY

T. MATSUNAWA

(Received Jan. 30, 1976; revised Oct. 26, 1981)

Summary

Some criteria based on K-L information number and W-divergence are presented for a certain type of uniform approximate equivalence of two probability distributions. As applications, some necessary and sufficient conditions are also given for the corresponding uniform asymptotic equivalence of two random sequences.

1. Introduction

Measuring the discrepancy between two probability distributions is of fundamental importance in many statistical problems. Lots of measures of discrepancy have been presented and their properties have been investigated by many authors under various situations. But, so far as the present author knows, almost such works are mainly concerned with limiting or asymptotic cases. From practical point of views, however, it seems to be very important to evaluate the discrepancy not in limiting or asymptotic manners but in approximate ones by using inequalities available in small sample cases. Such approach also has theoretical merits of improving the asymptotic results so far obtained. Such being the case the present article is designed to give error evaluations relating to some important measures of discrepancy.

Let X and Y be two random variables defined on a measurable space $(R, \boldsymbol{B})$, where R is any abstract space and $\boldsymbol{B}$ is a σ-field of subsets of R. Let us designate by P^X and P^Y the corresponding probability distributions of X and Y, respectively. Moreover, let A be a measurable set belonging to $\boldsymbol{B}$ and let $\delta^*=\delta^*(X, Y; A)$ be a measure of discrepancy on the set A between the two distributions.

DEFINITION 1.1. Two random variables X and Y are said to be *uniformly ϕ-equivalent with respect to δ^* in the sense of type* $(\boldsymbol{B})_d$ and are denoted as

$$X \overset{\phi}{\sim} Y , \qquad (\boldsymbol{B})_d \tag{1.1}$$

if for any $\varepsilon \geqq 0$ there exists a non-negative monotone function of ε, $\phi(\varepsilon ; \delta^*)$, such that $\phi(\varepsilon ; \delta^*) \to 0$ as $\varepsilon \to 0$ and

$$D(X, Y; \boldsymbol{B}) \equiv \sup_{E \in \boldsymbol{B}} |P^X(E) - P^Y(E)| \leqq \phi(\varepsilon ; \delta^*) , \tag{1.2}$$

whenever $\delta^* \leqq \varepsilon$.

It should be remarked that the above notion is an extension of S. Ikeda's definition of type $(\boldsymbol{B})_d$ asymptotic equivalence of probability distributions (cf. [1], [2]). We have, of course, several types of ϕ-equivalence by considering subclasses of $\boldsymbol{B}$ in parallel with his notion. For example, let R be Euclidean space and taking the class $\boldsymbol{M}$ consisting of all infinite intervals which are right-opened, then we have other type of approximation between two random variables, which may be called ϕ-equivalence in the sense of type $(\boldsymbol{M})_d$.

Now, suppose that both X and Y are absolutely continuous with respect to a σ-finite measure μ over $(R, \boldsymbol{B})$. Then, X and Y have the gpdf's (μ) f and g, respectively. Under these set-up let us take some familiar measures of discrepancy between the two probability distributions;

$$K(X, Y; R) \equiv 1 - \rho(X, Y; R) \tag{1.3}$$

with

$$\rho(X, Y; R) \equiv \int_R \sqrt{fg}\, d\mu , \tag{1.4}$$

$$V(X, Y; R) \equiv \int_R |f-g|\, d\mu , \tag{1.5}$$

$$I(X, Y; R) \equiv \int_R f \ln (f/g) d\mu , \tag{1.6}$$

and

$$W(X, Y; R) \equiv \int_R \left(\frac{f}{g} - 1\right)^2 g d\mu , \tag{1.7}$$

which is sometimes called W-divergence or χ^2-divergence (cf. [4], [8]). It is known that the following inequalities hold among these quantities:

$$0 \leqq \rho(X, Y; R) \leqq 1 , \tag{1.8}$$

$$1 - \rho(X, Y; R) \leqq D(X, Y; \boldsymbol{B}) = \frac{V(X, Y; R)}{2} \leqq \sqrt{1 - \rho^2(X, Y; R)} \tag{1.9}$$

$$\leqq\sqrt{\frac{I(X, Y; R)}{2}}\leqq\sqrt{\frac{1}{2}\ln\{1+W(X, Y; R)\}}\ ,$$

which will be considered under more general situation in next section. Then, as is well-known, we can see that

$$D(X, Y; \boldsymbol{B})=0 \qquad \text{iff } \rho(X, Y; R)=1\ , \tag{1.10}$$

and

$$D(X, Y; \boldsymbol{B})=0 \qquad \text{if } I(X, Y; R)=0\ , \tag{1.11}$$

or

$$D(X, Y; \boldsymbol{B})=0 \qquad \text{if } W(X, Y; R)=0\ . \tag{1.12}$$

Judging from the above, it seems that the affinity $\rho(X, Y; R)$ is better than the Kullback-Leibler information number $I(X, Y; R)$ as a criterion for type $(\boldsymbol{B})_d$ equivalence problem of X and Y. This is certainly true in general discussion. However, $I(X, Y; R)$ is sometimes manageable in practical calculations and has the physical meaning of the negative entropy of X with respect to Y. As for $W(X, Y; R)$, it is apparently related to K. Pearson's χ^2-statistic. Therefore, seeking other necessary conditions for $D(X, Y; \boldsymbol{B})=0$ based on the two quantities, if exists, is strongly desired from both theoretical and practical reasons.

Concerning this problem, Pinsker [7] gave the following inequality:

$$P^X\left(\left\{x;\ \left|\ln\frac{f(x)}{g(x)}\right|\geqq\varepsilon,\ x\in R\right\}\right)\leqq\frac{1+\varepsilon}{\varepsilon}V(X, Y; R)\ , \tag{1.13}$$

for any given $\varepsilon>0$. This result suggests us that vanishing of K-L information number or its modified quantity is also useful to the necessary condition, if there exists a certain measurable set on which almost all mass of X and Y distributes. Similar guess arises to W-divergence.

In the following section the above problem is investigated in terms of uniform ϕ-equivalence in the sense of type $(\boldsymbol{B})_d$. In Section 3, as applications, we give some necessary and sufficient conditions for type $(\boldsymbol{B})_d$ asymptotic equivalence of two random sequences.

2. Criteria for type $(\boldsymbol{B})_d$ ϕ-equivalence

Let, as before, X and Y be two random variables defined on an abstract measurable space $(R, \boldsymbol{B})$, where R is any abstract space and $\boldsymbol{B}$ is a σ-field of subsets of R. Let μ be a σ-finite measure defined over the measurable space and A a measurable set in $\boldsymbol{B}$. Denote the corresponding probability distributions of X and Y by P^X and P^Y, respectively. Let, further, $f^*>0$ and $g^*>0$ be the respective Radon-Nikodym derivatives of P^X and P^Y with respect to μ over the set A. Then, for every measurable set E belonging to the sub-σ-field of $\boldsymbol{B}$ generated by the set A, it can be represented as

$$P^X(E)=\int_E f^* d\mu \qquad \text{and} \qquad P^Y(E)=\int_E g^* d\mu .$$

The set A is not necessarily identical with the whole space R, in which case X and Y may or may not be dominated by μ outside the set A. Throughout this section, the sets A_{ε_i} should always be considered to be the subsets of A. As the set A, one may take, for example, the common domain of the supports of X and Y, if R is the k-dimensional Euclidean space and each supports of X and Y is a k-dimensional right-opened interval. Such cases often appear in measuring discrepancies among probability distributions. Another example is associated with the approximate joint normality of k sample quantiles from the uniform distribution $U(0, 1)$, where we can take as A a k-dimensional set $\{x_{(k)};\ x_{(k)}=(x_{n_1},\cdots, x_{n_k}),\ 0<x_{n_1}<\cdots<x_{n_k}<1 \text{ and } 0<n_1<\cdots<n_k\}$.

Now as in [3], let us consider the quantities corresponding to (1.3)–(1.7) as $\delta^*(X, Y; A)$;

$$K^*(X, Y; A)=1-\rho^*(X, Y; A) \tag{2.1}$$

with

$$\rho^*(X, Y; A)=\int_A (f^*g^*)^{1/2} d\mu , \tag{2.2}$$

$$V^*(X, Y; A)=\int_A |f^*-g^*| d\mu , \tag{2.3}$$

$$I^*(X, Y; A)=\int_A f^* \ln (f^*/g^*) d\mu , \tag{2.4}$$

and

$$W^*(X, Y; A)=\int_A (f^*/g^*-1)^2 g^* d\mu . \tag{2.5}$$

In the case where the set A is taken to be the whole space R, we shall not asterisk the above quantities as in (1.3)–(1.7).

Among these measures of discrepancy the following result is useful:

LEMMA 2.1. (i) *If for any given* ε_1 $[0\leqq\varepsilon_1\leqq 3-2\sqrt{2}\,]$ *there exists a measurable set* $A_{\varepsilon_1}\in \boldsymbol{B}$ *satisfying*

$$[\rho^*(X, Y; A_{\varepsilon_1})]^2\geqq 1-\varepsilon_1 \tag{2.6}$$

or equivalently

$$K^*(X, Y; A_{\varepsilon_1}) \leqq 1-\sqrt{1-\varepsilon_1}\,, \tag{2.6$'$}$$

then it holds that

$$X \overset{\phi_1}{\sim} Y\,, \qquad (\boldsymbol{B})_d\,, \tag{2.7}$$

where

$$\phi_1 = \phi_1(\varepsilon_1; K^*) \leqq 2\sqrt{\varepsilon_1} + \varepsilon_1\,. \tag{2.8}$$

(ii) *If for any given* ε_2 $[0 \leqq \varepsilon_2 \leqq 1/3]$

$$D(X, Y; \boldsymbol{B}) \leqq \varepsilon_2\,, \tag{2.9}$$

then there exists a measurable set $A_{\varepsilon_2} \in \boldsymbol{B}$ *such that*

$$\rho^*(X, Y; A_{\varepsilon_2}) \geqq 1-3\varepsilon_2\,. \tag{2.10}$$

PROOF. (i) Assume that (2.6) is satisfied, then

$$\min(P^X(A_{\varepsilon_1}), P^Y(A_{\varepsilon_1})) \geqq P^X(A_{\varepsilon_1})\cdot P^Y(A_{\varepsilon_1}) \geqq [\rho^*(X, Y; A_{\varepsilon_1})]^2 \geqq 1-\varepsilon_1\,,$$

which implies

$$P^X(A_{\varepsilon_1}) \geqq 1-\varepsilon_1 \quad \text{and} \quad P^Y(A_{\varepsilon_1}) \geqq 1-\varepsilon_1\,.$$

Then, we can evaluate as

$$\begin{aligned} D(X, Y; \boldsymbol{B}) &\leqq \sup_{F\subset A_{\varepsilon_1}} |P^X(F)-P^Y(F)| + \sup_{G\subset R-A_{\varepsilon_1}} |P^X(G)-P^Y(G)| \\ &\leqq V^*(X, Y; A_{\varepsilon_1}) + \varepsilon_1 \\ &\leqq [\{P^X(A_{\varepsilon_1})+P^Y(A_{\varepsilon_1})\}^2 - 4(\rho^*(X, Y; A_{\varepsilon_1}))^2]^{1/2} + \varepsilon_1 \\ &\leqq \sqrt{4[1-(\rho^*(X, Y; A_{\varepsilon_1}))^2]} + \varepsilon_1 \leqq 2\sqrt{\varepsilon_1} + \varepsilon_1\,. \end{aligned}$$

(ii) Under the assumptions of the lemma it is obvious that for an arbitrary ε_2' $(0 \leqq \varepsilon_2 \leqq \varepsilon_2' \leqq 1/2)$ we can take a measurable set $A_{\varepsilon_2} \in \boldsymbol{B}$ such that $P^X(A_{\varepsilon_2}) \geqq 1-\varepsilon_2'$. Moreover, from (2.9) it follows that $P^Y(A_{\varepsilon_2}) \geqq P^X(A_{\varepsilon_2}) - \varepsilon_2 \geqq 1-(\varepsilon_2+\varepsilon_2')$. If ε_2' is so chosen that $\varepsilon_2'=\varepsilon_2$, then $P^Y(A_{\varepsilon_2}) \geqq 1-2\varepsilon_2$. Interchanging X and Y we have the dual result $P^X(A_{\varepsilon_2}) \geqq 1-2\varepsilon_2$. Thus, we get

$$\min(P^X(A_{\varepsilon_2}), P^Y(A_{\varepsilon_2})) \geqq 1-2\varepsilon_2\,. \tag{2.11}$$

Next, let us put $B=\{x;\ f^*(x) \geqq g^*(x),\ x \in A_{\varepsilon_2}\}$ and $C=A_{\varepsilon_2}-B$, then

$$\begin{aligned} V^*(X, Y; A_{\varepsilon_2}) &= \int_B (f^*-g^*)d\mu + \int_C (g^*-f^*)d\mu \\ &\leqq 2\sup_{E\in\boldsymbol{B}} |P^X(E)-P^Y(E)| = 2D(X, Y; \boldsymbol{B}) \leqq 2\varepsilon_2\,. \end{aligned}$$

Thus, we have

$$\rho^*(X, Y; A_{\varepsilon_2}) \geqq \frac{1}{2}\{P^X(A_{\varepsilon_2})+P^Y(A_{\varepsilon_2})-V^*(X, Y; A_{\varepsilon_2})\}$$
$$\geqq 1-2\varepsilon_2-\varepsilon_2=1-3\varepsilon_2 ,$$

which completes the proof of the lemma.

On the way of the proof to the lemma we have obtained the following result which is of independent interest and will be used later.

COROLLARY 2.1. (i) *If for any given* ε_3 $[0 \leqq \varepsilon_3 \leqq 3-2\sqrt{2}=0.17157\cdots]$ *there exist a measurable set* A_{ε_3} $(\in \boldsymbol{B})$ *and a non-negative small number* $\eta_1=\eta_1(\varepsilon_3)$ *such that*

$$\min(P^X(A_{\varepsilon_3}), P^Y(A_{\varepsilon_3})) \geqq 1-\varepsilon_3 , \tag{2.12}$$

and that

$$V^*(X, Y; A_{\varepsilon_3}) \leqq \eta_1(\varepsilon_3) \leqq 2\sqrt{\varepsilon_3} , \tag{2.13}$$

then

$$X \overset{\phi_2}{\sim} Y, \qquad (\boldsymbol{B})_d , \tag{2.14}$$

where

$$\phi_2=\phi_2(\varepsilon_3; V^*) \leqq \eta_1(\varepsilon_3)+\varepsilon_3 \leqq 2\sqrt{\varepsilon_3}+\varepsilon_3 . \tag{2.15}$$

(ii) *If for any given* ε_4 $[0 \leqq \varepsilon_4 \leqq 0.5]$

$$D(X, Y; \boldsymbol{B}) \leqq \varepsilon_4 , \tag{2.16}$$

then there exists a non-empty measurable set $A_{\varepsilon_4} \in \boldsymbol{B}$ *such that*

$$V^*(X, Y; A_{\varepsilon_4}) \leqq 2\varepsilon_4 . \tag{2.17}$$

The following lemma plays an important role in the subsequent discussions.

LEMMA 2.2. *Let* ε *be any given non-negative number and define the set*

$$B_\varepsilon=\left\{x;\ \left|\ln\frac{f^*(x)}{g^*(x)}\right| \leqq \varepsilon,\ x \in A \subset R\right\} . \tag{2.18}$$

Then, it holds that

$$\min\{P^X(B_\varepsilon), P^Y(B_\varepsilon)\} \geqq \min\{P^X(A), P^Y(A)\}-\frac{1}{c(\varepsilon)}V^*(X, Y; A) , \tag{2.19}$$

where $c(\varepsilon)=\min(\varepsilon, 1)$.

PROOF. Let us define the set

$$S_\varepsilon=\left\{x\,;\ \max\left(\left|\frac{f^*(x)}{g^*(x)}-1\right|,\ \left|\frac{g^*(x)}{f^*(x)}-1\right|\right)\leqq\varepsilon,\ x\in A\subset R\right\}, \tag{2.20}$$

then

$$\begin{aligned}V^*(X, Y; A)&=\int_A |f^*-g^*|d\mu\geqq\int_{A-S_\varepsilon} f^*\left|\frac{g^*}{f^*}-1\right|d\mu+\left|\int_{S_\varepsilon}(f^*-g^*)d\mu\right|\\ &\geqq\varepsilon\int_{A-S_\varepsilon} f^*d\mu+|P^X(S_\varepsilon)-P^Y(S_\varepsilon)|\\ &\geqq c(\varepsilon)[\{P^X(A)-P^X(S_\varepsilon)\}+|P^X(S_\varepsilon)-P^Y(S_\varepsilon)|]\,.\end{aligned}$$

We have also the dual inequality:

$$V^*(X, Y; A)\geqq c(\varepsilon)[\{P^Y(A)-P^Y(S_\varepsilon)\}+|P^X(S_\varepsilon)-P^Y(S_\varepsilon)|]\,.$$

Then,

$$\begin{aligned}V^*(X, Y; A)&\geqq c(\varepsilon)[\min(P^X(A), P^Y(A))\\ &\qquad -\{\max(P^X(S_\varepsilon), P^Y(S_\varepsilon))-|P^X(S_\varepsilon)-P^Y(S_\varepsilon)|\}]\\ &=c(\varepsilon)[\min(P^X(A), P^Y(A))-\min(P^X(S_\varepsilon), P^Y(S_\varepsilon))]\,,\end{aligned}$$

from which

$$\min(P^X(S_\varepsilon), P^Y(S_\varepsilon))\geqq\min\{P^X(A), P^Y(A)\}-\frac{1}{c(\varepsilon)}V^*(X, Y; A)\,. \tag{2.21}$$

Incidentally, since $|\ln t|\leqq\max(|t-1|, |1/t-1|)$ $(t>0)$, then $S_\varepsilon\subseteqq B_\varepsilon$ and hence we immediately obtain the desired inequality (2.19).

Remark 2.1. The lemma becomes a generalization to Pinsker's result (1.13). In fact, taking $A=R$ we have an improved inequality corresponding to (1.13).

Now we are in a position to state the following

THEOREM 2.1. (i) *If for any given* ε_5 $[0\leqq\varepsilon_5\leqq1-\sqrt{2(\sqrt{2}-1)}=0.35640\cdots]$ *there exist a measurable set* $A_{\varepsilon_5}\in\boldsymbol{B}$ *and a non-negative small number* $\eta_2=\eta_2(\varepsilon_5)$ *such that*

$$P^X(A_{\varepsilon_5})\geqq1-\varepsilon_5\,, \tag{2.22}$$

and

$$|I^*(X, Y; A_{\varepsilon_5})|\leqq\eta_2(\varepsilon_5)\leqq(1-\varepsilon_5)\ln\left[\frac{(1-\varepsilon)^2}{2(\sqrt{2}-1)}\right], \tag{2.23}$$

simultaneously, then it holds that

(2.24) $$X \overset{\phi_3}{\sim} Y\,, \qquad (\boldsymbol{B})_d\,,$$

with

(2.25) $$\phi_3=\phi_3(\varepsilon_5\,;\,I^*)=\phi_1(\varepsilon_5^*\,;\,I^*)\leqq 2\sqrt{\varepsilon_5^*}+\varepsilon_5^*\,,$$

where

(2.26) $$\varepsilon_5^*\leqq 1-(1-\varepsilon_5)^2\exp\{-\eta_2/(1-\varepsilon_5)\}\,.$$

(ii) *If for any given* ε_6 $[0\leqq\varepsilon_6<0.25]$

(2.27) $$D(X, Y; \boldsymbol{B})\leqq\varepsilon_6\,,$$

then there exist a constant α *and a measurable set* $A_{\varepsilon_6}\in\boldsymbol{B}$ *such that the set defined by*

(2.28) $$B_{\varepsilon_6}^{\alpha}\equiv\left\{x\,;\,\left|\ln\frac{f^*(x)}{g^*(x)}\right|\leqq\varepsilon_6^{\alpha},\right.$$
$$\left.0<\alpha<1-\ln(0.5-\varepsilon_6)/\ln\varepsilon_6 \text{ and } x\in A_{\varepsilon_6}\subset R\right\}\,,$$

satisfies the conditions

(2.29) $$P^X(B_{\varepsilon_6}^{\alpha})\geqq 1-\varepsilon_{7\alpha}\,,$$

and

(2.30) $$-\varepsilon_6\leqq I^*(X, Y; B_{\varepsilon_6}^{\alpha})\leqq\min\{\varepsilon_6^{\alpha},\ \varepsilon_6\exp(\varepsilon_6^{\alpha}/2)\}\,,$$

simultaneously.

PROOF. (i) Since $f^*/P^X(A_{\varepsilon_5})$ is a generalized probability density function with respect to the measure μ over the set A_{ε_5}, then we can apply the Jensen inequality to get

$$I^*(X, Y; A_{\varepsilon_5})\geqq -2P^X(A_{\varepsilon_5})[-\ln P^X(A_{\varepsilon_5})+\ln\rho^*(X, Y; A_{\varepsilon_5})]\,.$$

Using the conditions (2.22) and (2.23) we obtain

(2.31) $$[\rho^*(X, Y; A_{\varepsilon_5})]^2\geqq(1-\varepsilon_5)^2\exp\left\{-\frac{\eta_2(\varepsilon_5)}{1-\varepsilon_5}\right\}\equiv 1-\varepsilon_5^*\,,$$

from which it is seen that $0\leqq\varepsilon_5^*\leqq 3-2\sqrt{2}$ for $0\leqq\varepsilon_5\leqq 1-\sqrt{2(\sqrt{2}-1)}$. Thus, by virtue of (i) in Lemma 2.1 we get (2.24) with (2.25).

(ii) In view of Lemma 2.2 and Corollary 2.1 (ii), we have

(2.32) $$P^X(B_{\varepsilon_6}^{\alpha})\geqq P^X(A_{\varepsilon_6})-\frac{1}{\varepsilon_6^{\alpha}}V^*(X, Y; A_{\varepsilon_6})$$
$$\geqq 1-2\varepsilon_6(1+\varepsilon_6^{-\alpha})\equiv 1-\varepsilon_{7\alpha}>0\,,$$

where we have used the same evaluations on $P^X(A_{\epsilon_6})$ and $V^*(X, Y; A_{\epsilon_6})$ as those derived in (2.11) and (2.17), respectively. As for the R.H.S. inequality in (2.30) we can easily get

$$I^*(X, Y; B^\alpha_{\epsilon_6}) \leqq \int_{B^\alpha_{\epsilon_6}} f^* \left| \ln \frac{f^*}{g^*} \right| d\mu \leqq \varepsilon_6^\alpha P^X(B^\alpha_{\epsilon_6}) \leqq \varepsilon_6^\alpha ,$$

or we can get with the aid of the inequality $\sqrt{t} \ln t \leqq t-1$ $(t \geqq 1)$ as

$$\begin{aligned} I^*(X, Y; B^\alpha_{\epsilon_6}) &\leqq \int_{B^\alpha_{\epsilon_6} \cap \{x;\, f^*(x) \geqq g^*(x)\}} \sqrt{f^*/g^*}\,(f^*-g^*) d\mu \\ &\leqq \exp(\varepsilon_6^\alpha) D(X, Y; \boldsymbol{B}) \leqq \varepsilon_6 \exp(\varepsilon_6^\alpha/2) . \end{aligned}$$

On the contrary, using the inequalities $\sqrt{t}\,|\ln t| \leqq |t-1| \leqq (t+1)|\ln t|/2$ $(t>0)$ or a less precise inequality $\ln t > 1-1/t$ $(t>0)$, we have

$$I^*(X, Y; B^\alpha_{\epsilon_6}) \geqq \int_{B^\alpha_{\epsilon_6}} (f^*-g^*) d\mu \geqq -D(X, Y; \boldsymbol{B}) \geqq -\varepsilon_6 ,$$

which completes the proof of the theorem.

Remark 2.2. From Lemma 2.2 $P^Y(B^\alpha_{\epsilon_6}) \geqq 1-\varepsilon_{7\alpha}$ automatically holds under the same condition of the part (ii) in the above theorem.

For $W^*(X, Y; \cdot)$ we have the following

THEOREM 2.2. (i) *If for any given* ε_8 $[0 \leqq \varepsilon_8 \leqq 3-2\sqrt{2}]$ *there exist a measurable set* A_{ϵ_8} $(\in \boldsymbol{B})$ *and a non-negative small number* $\eta_3 = \eta_3(\varepsilon_8)$ *such that*

$$P^Y(A_{\epsilon_8}) \geqq 1-\varepsilon_8 , \tag{2.33}$$

and

$$W^*(X, Y; A_{\epsilon_8}) \leqq \eta_3(\varepsilon_8) \leqq 4\varepsilon_8 , \tag{2.34}$$

simultaneously, then it holds that

$$X \overset{\phi_4}{\sim} Y \qquad (\boldsymbol{B})_d , \tag{2.35}$$

where

$$\phi_4 = \phi_4(\varepsilon_8; W^*) \leqq 2\sqrt{\varepsilon_8} + \varepsilon_8 . \tag{2.36}$$

(ii) *If for any given* ε_9 $[0 \leqq \varepsilon_9 < 0.5]$

$$D(X, Y; \boldsymbol{B}) \leqq \varepsilon_9 , \tag{2.37}$$

then there exist a constant β *and a measurable set* $A_{\epsilon_9} \in \boldsymbol{B}$ *such that the set defined by*

(2.38) $$B_{\varepsilon_9}^{\beta}\equiv\left\{x;\ \left|\frac{f^*(x)}{g^*(x)}-1\right|\leqq\varepsilon_9^{\beta},\right.$$
$$\left.-1<\beta<1-\ln(0.5-\varepsilon_9)/\ln\varepsilon_9,\ x\in A_{\varepsilon_9}\subset R\right\}$$

satisfies the conditions

(2.39) $$P^Y(B_{\varepsilon_9}^{\beta})\geqq 1-2\varepsilon_9-2\varepsilon_9^{1-\beta}>0\ ,$$

and

(2.40) $$W^*(X, Y; B_{\varepsilon_9}^{\beta})\leqq\min\{\varepsilon_9^{2\beta}, 2\varepsilon_9^{\beta+1}\}\ .$$

PROOF. (i) Since

$$\begin{aligned}V^*(X, Y; A_{\varepsilon_8})&=\int_{A_{\varepsilon_8}}|f^*-g^*|d\mu=\int_{A_{\varepsilon_8}}\left|\frac{f^*}{g^*}-1\right|\sqrt{g^*}\sqrt{g^*}d\mu\\&\leqq\left[\int_{A_{\varepsilon_8}}\left(\frac{f^*}{g^*}-1\right)^2g^*d\mu\cdot\int_{A_{\varepsilon_8}}g^*d\mu\right]^{1/2}\\&=[W^*(X, Y; A_{\varepsilon_8})\cdot P^Y(A_{\varepsilon_8})]^{1/2}\ ,\end{aligned}$$

we get from the condition (2.34)

$$\leqq\sqrt{\eta_3(\varepsilon_8)P^Y(A_{\varepsilon_8})}\leqq 2\sqrt{\varepsilon_8}\ .$$

Thus, noticing the condition (2.33) and Corollary 2.1 we get the desired result (2.35) with (2.36).

(ii) By Lemma 2.2 and Corollary 2.1 (ii), we have

$$P^Y(B_{\varepsilon_9}^{\beta})\geqq P^Y(A_{\varepsilon_9})-\varepsilon_9^{-\beta}V^*(X, Y; A_{\varepsilon_9})\geqq 1-2\varepsilon_9-2\varepsilon_9^{1-\beta}\equiv 1-\varepsilon_{10\beta}>0\ .$$

In respect to (2.40) it is clear that

$$W^*(X, Y; B_{\varepsilon_9}^{\beta})=\int_{B_{\varepsilon_9}^{\beta}}\left(\frac{f^*}{g^*}-1\right)^2g^*d\mu\leqq\varepsilon_9^{2\beta}P^Y(B_{\varepsilon_9}^{\beta})\leqq\varepsilon_9^{2\beta}\ ,$$

and that

$$W^*(X, Y; B_{\varepsilon_9}^{\beta})=\int_{B_{\varepsilon_9}^{\beta}}\left|\frac{f^*}{g^*}-1\right||f^*-g^*|d\mu\leqq\varepsilon_9^{\beta}V^*(X, Y; B_{\varepsilon_9}^{\beta})\leqq 2\varepsilon_9^{\beta+1}\ ,$$

which completes the proof of the theorem.

Now we shall state a relation between $I^*(X, Y; \cdot)$ and $W^*(X, Y; \cdot)$ in the following

THEOREM 2.3. *Assume that for any given* ε_{11} $[0\leqq\varepsilon_{11}<1]$ *there exists a measurable set* $A_{\varepsilon_{11}}$ $(\in \boldsymbol{B})$ *such that*

(2.41) $$\min(P^X(A_{\varepsilon_{11}}), P^Y(A_{\varepsilon_{11}}))\geqq 1-\varepsilon_{11}\ .$$

Then, (i) *it holds that*

$$I^*(X, Y; A_{\varepsilon_{11}}) \leqq \ln \left[1+\frac{1}{1-\varepsilon_{11}}\{W^*(X, Y; A_{\varepsilon_{11}})+\varepsilon_{11}\}\right]. \tag{2.42}$$

(ii) *In addition to* (2.41), *if there exist a non-negative number* $\varepsilon_{12} = \varepsilon_{12}(\varepsilon_{11})$ $[0 \leqq \varepsilon_{11} \leqq \varepsilon_{12} < 1]$ *and a measurable set* $B_{\varepsilon_{12}} \subset A_{\varepsilon_{11}}$ *such that*

$$\min(P^X(B_{\varepsilon_{12}}), P^Y(B_{\varepsilon_{12}})) \geqq 1-\varepsilon_{12}, \tag{2.43}$$

then it holds that

$$I^*(X, Y; B_{\varepsilon_{12}}) \geqq \frac{1}{2}\left(1-\frac{1}{3}\varepsilon_{12}\right)W^*(X, Y; B_{\varepsilon_{12}})-\varepsilon_{12} \tag{2.44}$$

and

$$I^*(X, Y; B_{\varepsilon_{12}}) \leqq \frac{1}{2}\left(1+\frac{1}{3}\varepsilon_{12}\right)W^*(X, Y; B_{\varepsilon_{12}})+\varepsilon_{12}. \tag{2.45}$$

PROOF. (i) From the definition

$$\begin{aligned} W^*(X, Y; A_{\varepsilon_{11}}) &= \int_{A_{\varepsilon_{11}}} \frac{f^{*2}}{g^*} d\mu - 2\int_{A_{\varepsilon_{11}}} f^* d\mu + \int_{A_{\varepsilon_{11}}} g^* d\mu \\ &= \int_{A_{\varepsilon_{11}}} \exp\left(\ln \frac{f^*}{g^*}\right) \cdot f^* d\mu - 2P^X(A_{\varepsilon_{11}}) + P^Y(A_{\varepsilon_{11}}), \end{aligned}$$

using Jensen's inequality

$$\geqq P^X(A_{\varepsilon_{11}}) \exp\left(\int_{A_{\varepsilon_{11}}} \frac{f^*}{P^X(A_{\varepsilon_{11}})} \ln \frac{f^*}{g^*} d\mu\right) - 2P^X(A_{\varepsilon_{11}}) + P^Y(A_{\varepsilon_{11}}). \tag{2.46}$$

Thus, from (2.41) and (2.46), we have

$$\begin{aligned} &I^*(X, Y; A_{\varepsilon_{11}}) \\ &\leqq P^X(A_{\varepsilon_{11}}) \ln \left[1+\frac{1}{P^X(A_{\varepsilon_{11}})}\{P^X(A_{\varepsilon_{11}})-P^Y(A_{\varepsilon_{11}})+W^*(X, Y; A_{\varepsilon_{11}})\}\right] \\ &\leqq \ln \left[1+\frac{1}{1-\varepsilon_{11}}\{W^*(X, Y; A_{\varepsilon_{11}})+\varepsilon_{11}\}\right]. \end{aligned}$$

(ii) Under the condition (2.41), we can easily find a measurable set satisfying (2.43). For example, take the set of the form

$$B_{\varepsilon_{12}} = \left\{x;\ \max\left[\left|\frac{f^*}{g^*}-1\right|, \left|\frac{g^*}{f^*}-1\right|\right] < \varepsilon_{12},\right. \\ \left. 0 \leqq \varepsilon_{11} \leqq \varepsilon_{12} < 1,\ x \in A_{\varepsilon_{11}} \subset R\right\}, \tag{2.47}$$

then

$$\min (P^X(B_{\varepsilon_{12}}), P^Y(B_{\varepsilon_{12}})) \geqq 1-\varepsilon_{12} . \tag{2.48}$$

Incidentally, since

$$f^* \ln \frac{f^*}{g^*} \geqq (f^*-g^*)+\frac{(f^*-g^*)^2}{2g^*}-\frac{(f^*-g^*)^3}{6g^{*2}} ,$$

(cf. [5], [6]), we have

$$\begin{aligned} I^*(X, Y; B_{\varepsilon_{12}}) &\geqq \int_{B_{\varepsilon_{12}}} (f^*-g^*)d\mu+\frac{1}{2}\int_{B_{\varepsilon_{12}}} \left(\frac{f^*}{g^*}-1\right)^2 g^* d\mu \\ &\quad -\frac{1}{6}\int_{B_{\varepsilon_{12}}} \left|\frac{f^*}{g^*}-1\right| \left(\frac{f^*}{g^*}-1\right)^2 g^* d\mu \\ &\geqq P^X(B_{\varepsilon_{12}})-P^Y(A_{\varepsilon_{11}})+\frac{1}{2}\left(1-\frac{1}{3}\varepsilon_{12}\right)W^*(X, Y; B_{\varepsilon_{12}}) \\ &\geqq \frac{1}{2}\left(1-\frac{1}{3}\varepsilon_{12}\right)W^*(X, Y; B_{\varepsilon_{12}})-\varepsilon_{12} . \end{aligned}$$

Conversely, the inequality

$$f^* \ln \frac{f^*}{g^*} \leqq (f^*-g^*)+\frac{(f^*-g^*)^2}{2g^*}-\frac{(f^*-g^*)^3}{6f^*g^*} ,$$

implies that

$$\begin{aligned} I^*(X, Y; B_{\varepsilon_{12}}) &\leqq P^X(A_{\varepsilon_{11}})-P^Y(B_{\varepsilon_{12}})+\frac{1}{2}W^*(X, Y; B_{\varepsilon_{12}}) \\ &\quad +\frac{1}{6}\int_{B_{\varepsilon_{12}}} \left|\frac{g^*}{f^*}-1\right| \left(\frac{f^*}{g^*}-1\right)^2 g^* d\mu \\ &\leqq \frac{1}{2}\left(1+\frac{1}{3}\varepsilon_{12}\right)W^*(X, Y; B_{\varepsilon_{12}})+\varepsilon_{12} . \end{aligned}$$

Remark 2.3. If we define the quantity

$$\tau^*(X, Y; A) \equiv \int_A \frac{f^{*2}}{g^*} d\mu , \tag{2.49}$$

then $W^*(X, Y; A)=\tau^*(X, Y; A)-1$. So we have other versions of the uniform ϕ-equivalence rewritten in terms of $\tau^*(X, Y; A)$, which may be sometimes tractable in practical calculations.

Remark 2.4. $W^*(X, Y; A)$ is closely related to the K. Pearson's chi-square statistic X_P^2 under a multinomial schemes. In this case the basic measure space $(R, \boldsymbol{B}, \mu)$ is taken such that R is the set of all non-negative integers, $\boldsymbol{B}$ is the σ-field consisting of all subsets of R, and μ is the counting measure on R. Suppose that we have n independent observations whose possible outcomes fall in either one of the

k (fixed and $\geqq 2$) mutually exclusive cells C_i $(i=1, 2, \cdots, k)$. Let n_i be the observed cell frequency in C_i after n independent trials and let p_i $(0<p_i<1)$ be the probability that any one of the outcomes falls in C_i, for each i. Then, it must be $\sum_{i=1}^{k} n_i = n$ $(0 \leqq n_i \leqq n)$ and $\sum_{i=1}^{k} p_i = 1$. Define the $(k-1)$-dimensional simplex as

$$A=\left\{(x_1, \cdots, x_k) \mid x_i \geqq 0 \ \ (i=1, \cdots, k), \ \sum_{i=1}^{k} x_i = 1\right\} .$$

Further, let P^X and P^Y be two k-term probability distributions whose respective discrete densities are given by

$$(n_1/n, n_2/n, \cdots, n_k/n) \qquad \text{and} \qquad (p_1, p_2, \cdots, p_k) .$$

Then both of them lie in the simplex A, and

$$W^*(X, Y; A)=\int_A \left(\frac{f^*}{g^*}-1\right)^2 g^* d\mu = \sum_{i=1}^{k} \left(\frac{n_i/n}{p_i}-1\right)^2 p_i = \frac{1}{n} \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} ,$$

that is to say

$$n \cdot W^*(X, Y; A) = X_P^2 , \tag{2.50}$$

which is K. Pearson's χ^2-statistic. Rényi [8] called $W(X, Y; R)$ the χ^2 divergence. Now, assume that p_i's and n_i's fulfill the conditions: for any given $\varepsilon>0$ $\max_{1 \leqq i \leqq k} \left\{\left|\frac{p_i}{n_i/n}-1\right|, \left|\frac{n_i/n}{p_i}-1\right|\right\} < \varepsilon$.

In that case, it should be noted that by the similar manner obtained (2.44), (2.45) and (2.50) we have

$$\left(1-\frac{1}{3}\varepsilon\right) X_P^2 - \varepsilon \leqq 2n \cdot I^*(X, Y; B_\varepsilon) \leqq \left(1+\frac{1}{3}\varepsilon\right) X_P^2 + \varepsilon . \tag{2.51}$$

Since ε is arbitrary and $X_P^2 \xrightarrow{w} \chi^2(k-1)$ (the chi-square distribution with $k-1$ degrees of freedom) as $n \to \infty$, then

$$2n \cdot I^*(X, Y; B_\varepsilon) \xrightarrow{w} \chi^2(k-1) , \qquad (n \to \infty) . \tag{2.52}$$

Further, choosing ε so small that $P^X(B_\varepsilon) \doteqdot 1 = P^X(R)$, we can roughly speak that

$$2n \cdot I(X, Y; R) \xrightarrow{w} \chi^2(k-1) , \qquad (n \to \infty) . \tag{2.53}$$

This property is very often referred, but its simple derivations are seldom stated in the standard text books.

3. Applications to the type $(\boldsymbol{B})_d$ uniform asymptotic equivalence

In this section some criteria will be given on the type $(\boldsymbol{B})_d$ asymptotic equivalence of probability distributions with the help of the results obtained in the previous section. Notations in the present section are almost the same as in the previous section: Let $\{X_s\}$ $(s=1, 2, \cdots)$ and $\{Y_s\}$ $(s=1, 2, \cdots)$ be two sequences of random variables defined on an abstract measurable space $(R_s, \boldsymbol{B}_s)$, for each s. Then, according to Ikeda [2], the two sequences are said to be *asymptotically equivalent in the sense of type* $(\boldsymbol{B})_d$ if

$$D(X_s, Y_s; \boldsymbol{B}_s) \equiv \sup_{E \in \boldsymbol{B}_s} |P^{X_s}(E) - P^{Y_s}(E)| \to 0 \tag{3.1}$$

as $s \to \infty$, and is denoted by

$$X_s \sim Y_s \qquad (\boldsymbol{B})_d,\ (s \to \infty)\,. \tag{3.2}$$

As before, we shall consider the case where both X_s and Y_s are absolutely continuous with respect to a σ-finite measure μ_s over a non-empty sub-σ-field of $\boldsymbol{B}_s$ generated by a measurable set A_s belonging to the σ-field $\boldsymbol{B}_s$, for each s. The set A_s is not necessarily identical with R_s, in which case X_s and Y_s may or may not be dominated by μ_s outside the set A_s. Now the following results are straightforwardly obtained. The statements (i)–(iv) below are the corresponding results to Lemma 2.1, Corollary 2.1, Theorem 2.1 and Theorem 2.2, respectively, and their proofs can be done by modifying slightly the corresponding parts in the previous section by assuming $\varepsilon_s \to 0$ as $s \to \infty$.

THEOREM 3.1. *Each one of the following conditions* (i)–(iv) *are necessary and sufficient for* $\{X_s\}$ $(s=1, 2, \cdots)$ *and* $\{Y_s\}$ $(s=1, 2, \cdots)$ *to be asymptotically equivalent in the sense of type* $(\boldsymbol{B})_d$:

There exists a sequence of measurable sets $\{A_s\,(\in \boldsymbol{B}_s)\}$ $(s=1, 2, \cdots)$ *such that*

(i) $\rho^*(X_s, Y_s; A_s) \to 1$,

(ii) $P^{X_s}(A_s) \to 1$ *and* $V^*(X_s, Y_s; A_s) \to 0$,

(iii) $P^{X_s}(A_s) \to 1$ *and* $I^*(X_s, Y_s; A_s) \to 0$,

(iv) $P^{X_s}(A_s) \to 1$ *and* $W^*(X_s, Y_s; A_s) \to 0$,

as $s \to \infty$.

Remark 3.1. In the necessary parts of (iii) and (iv), we may take, as the sets A_s, the corresponding sets of the analogous form to those in (2.28) and (2.38), respectively. In the sufficient parts of the theorem we can take $\{R_s\}$ $(s=1, 2, \cdots)$ (whole spaces) as $\{A_s\}$ $(s=1, 2, \cdots)$ and the reduced results are well known. But, it is not always possible for us to do so in the necessary parts of (iii) and (iv). This is shown by the following counter example.

Example. Let $\{X_s\}$ $(s=1, 2, \cdots)$ and $\{Y_s\}$ $(s=1, 2, \cdots)$ be sequences of random variables with the following respective densities, for each s.

$$f_s(x)=\begin{cases} \dfrac{1-1/s^{\alpha+1}}{1-1/s} & \text{for } 0\leqq x<1-1/s\,, \\ 1/s^{\alpha} & \text{for } 1-1/s\leqq x<1\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{3.3}$$

and

$$g_s(x)=\begin{cases} \dfrac{1-e^{-s^{\beta}}/s}{1-1/s} & \text{for } 0\leqq x<1-1/s\,, \\ e^{-s^{\beta}} & \text{for } 1-1/s\leqq x<1\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{3.4}$$

Then, it is seen that when $\alpha, \beta>0$

$$\begin{aligned} D(X_s, Y_s; \boldsymbol{B}_s)&=\frac{1}{2}V(X_s, Y_s; R_s)=\frac{1}{2}\int_{[0,1)}|f_s-g_s|d\mu \\ &\leqq\frac{1}{2}\left\{\left|\frac{1}{s^{\alpha+1}}-\frac{e^{-s^{\beta}}}{s}\right|+\frac{1}{s}\left|\frac{1}{s^{\alpha}}-e^{-s^{\beta}}\right|\right\}\to 0 \qquad (s\to\infty)\,, \end{aligned}$$

where μ is the Lebesgue measure over $[0, 1)$. On the other hand

$$\begin{aligned} I(X_s, Y_s; R_s)&=\int_{[0,1)} f_s \ln\frac{f_s}{g_s}d\mu \\ &=\left(1-\frac{1}{s^{\alpha+1}}\right)\ln\frac{1-1/s^{\alpha+1}}{1-e^{-s^{\beta}}/s}+\frac{1}{s^{\alpha+1}}\ln\frac{1/s^{\alpha}}{e^{-s^{\beta}}}\,, \end{aligned}$$

in which the first member of the last expression tends to zero as $s\to\infty$. The second member can be evaluated as follows:

$$\frac{1}{s^{\alpha+1}}\ln\frac{1/s^{\alpha}}{e^{-s^{\beta}}}=\frac{1}{s^{\alpha+1}}(-\alpha\ln s+s^{\beta})\sim s^{\beta-\alpha-1}\,.$$

Hence, if $\beta>\alpha+1$, then $s^{\beta-\alpha-1}\to\infty$, and therefore

$$I(X_s, Y_s; R_s)\to\infty\,, \qquad (s\to\infty)\,.$$

On the other hand, $I(X_s, Y_s; R_s)\to 0$ $(s\to\infty)$, provided that $\beta<\alpha+1$. However, this is not the case for the quantity

$$W(X_s, Y_s; R_s)=\int_{[0,1)}\left(\frac{f_s}{g_s}-1\right)^2 g_s d\mu$$

$$=\left(\frac{e^{-s\beta}}{s}-\frac{1}{s^{\alpha+1}}\right)^2\frac{1}{1-e^{-s\beta}/s}+\frac{1}{se^{-2s\beta}}\left(\frac{1}{s^\alpha}-e^{-s\beta}\right)^2$$

which does not always vanish. Because, the first term of the above expression tends to zero as $s\to\infty$, but even if $\beta<\alpha+1$ the second term

$$\frac{1}{se^{-2s\beta}}\left(\frac{1}{s^\alpha}-e^{-s\beta}\right)^2=\frac{1}{s}\left(\frac{e^{s\beta}}{s^\alpha}\right)^2\left(1-\frac{s^\alpha}{e^{s\beta}}\right)^2$$

can not necessarily converge to zero as $s\to\infty$.

Such being the case we can not take as $A_s=R_s$ in the example. However, if we take the set $[0, 1-1/s)$, then

$$P^{X_s}(A_s)\to 1$$

and both quantities $I^*(X_s, Y_s; A_s)$ and $W^*(X_s, Y_s; A_s)$ are simultaneously converge to zero as $s\to\infty$.

Relating to Theorem 3.1 and the example considered above we have also the following result:

THEOREM 3.2. *If there exist two positive numbers m and M, such that*

$$m\leqq\frac{f_s(x)}{g_s(x)}\leqq M \tag{3.5}$$

for almost all (μ_s) x *in* R_s, *then* (a) $W(X_s, Y_s; R_s)\to 0$ $(s\to\infty)$, (b) $I(X_s, Y_s; R_s)\to 0$ $(s\to\infty)$ *and* (c) *the condition* (3.1) *are mutually equivalent.*

PROOF. Since, the inclusion relations (a)$\Rightarrow$(b)$\Rightarrow$(c) clearly hold (without the condition (3.5)), it suffices to show (c)$\Rightarrow$(a) under the condition (3.5). From Lemma 2.2 we get for any given $\varepsilon>0$,

$$P^{X_s}\left(\left\{x;\left|\frac{f_s(x)}{g_s(x)}-1\right|\geqq\varepsilon\right\}\right)\leqq\frac{1}{\varepsilon}V(X_s, Y_s; R_s)\,. \tag{3.6}$$

Now,

$$\begin{aligned} W(X_s, Y_s; R_s)&=\int_{R_s}\left(\frac{f_s}{g_s}-1\right)^2 g_s d\mu_s \\ &=\int_{A_{s,\varepsilon}}\left(\frac{f_s}{g_s}-1\right)^2 g_s d\mu_s+\int_{\bar{A}_{s,\varepsilon}}\left(\frac{f_s}{g_s}-1\right)^2 g_s d\mu_s\,, \end{aligned} \tag{3.7}$$

where

$$A_{s,\varepsilon}=\left\{x;\left|\frac{f_s(x)}{g_s(x)}-1\right|<\varepsilon\right\}, \qquad \bar{A}_{s,\varepsilon}=R_s-A_{s,\varepsilon}\,.$$

It is then seen that

$$(3.8) \qquad \left|\int_{A_{s,\epsilon}}\left(\frac{f_s}{g_s}-1\right)^2 g_s d\mu_s\right| \leqq \epsilon^2 P^{Y_s}(A_{s,\epsilon}) \leqq \epsilon^2 .$$

From the assumption (3.5) of the theorem, it is true that

$$\left|\frac{f_s}{g_s}-1\right| \leqq \max(|M-1|, |m-1|) \equiv C , \qquad \text{a.e. } (\mu_s) \text{ over } R_s ,$$

and therefore

$$(3.9) \qquad \left|\int_{\bar{A}_{s,\epsilon}}\left(\frac{f_s}{g_s}-1\right)^2 g_s d\mu_s\right| \leqq C^2 \cdot P^{Y_s}(\bar{A}_{s,\epsilon})$$

which tends to zero as $s \to \infty$ by (3.6). Thus,

$$W(X_s, Y_s; R_s) \leqq 2\epsilon^2 \qquad \text{for sufficiently large } s .$$

Since ϵ is arbitrary, the proof of the theorem is completed.

Acknowledgment

The author wishes to thank the referee for his careful reading and useful comments.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Ikeda, S. (1963). Asymptotic equivalence of probability distributions to some problems of asymptotic independence, *Ann. Inst. Statist. Math.*, **15**, 87-116.

[2] Ikeda, S. (1968). Asymptotic equivalence of real probability distributions, *Ann. Inst. Statist. Math.*, **20**, 339-362.

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

[4] Kagan, A. M. (1963). On the theory of Fisher's amount of information, *Dokl. Akad. Nauk, SSSR*, **151**, 277-278.

[5] Matsunawa, T. (1976). Some inequalities based on inverse factorial series, *Ann. Inst. Statist. Math.*, **28**, A, 291-305.

[6] Matsunawa, T. (1977). Approximations to the probabilities of binomial and multinomial random variables and chi-square type statistics, *Ann. Inst. Statist. Math.*, **29**, A, 333-358.

[7] Pinsker, M. S. (1964). *Information and Information Stability of Random Variables and Processes*, Holden-Day, Inc.

[8] Rényi, A. (1970). *Foundations of Probability*, Holden-Day, Inc.