ASYMPTOTIC EQUIVALENCE OF PROBABILITY DISTRIBUTIONS WITH APPLICATIONS TO SOME PROBLEMS OF ASYMPTOTIC INDEPENDENCE

BY SADAO IKEDA

(Received Nov. 26, 1962; revised Oct. 17, 1963)

Contents

Summary

The problem of asymptotic approximation is formulated in a generalized form, and the basic concepts of asymptotic equivalence of probability distributions, which are fundamental to the study of asymptotic approximation problems, are established precisely. For the purpose of practical applications, some useful criterions for the asymptotic equivalence are given.

As a special case of asymptotic equivalence, notions of asymptotic independence are introduced, partly with the aim of application to the asymptotic approximation problem, or more precisely, to the construction of asymptotically approximate distributions.

By applying the idea to a problem of asymptotic independence of a set of elementary coverages, to that of extremes of an ordered sample, and to that of a set of multinominal variates, new and interesting results are obtained.

PART I.

BASIC CONCEPTS OF ASYMPTOTIC EQUIVALENCE IN THE PROBLEM OF ASYMPTOTIC APPROXIMATION

1.1. Problem of asymptotic approximation

As for the comparison of two sequences of probability distributions, some notions have been introduced in the literature, for most of which the existence of limiting distributions of the sequences is assumed, or in other words, these notions rely upon those of stochastic convergence (Loéve [13]). Recently, in the study of extreme values, Jeffroy [3] introduced a notion of asymptotic independence of two random variables, based upon a two-dimensional extension of Lévy's distance [12], where he did not assume the existence of limiting distributions.

Now we shall formulate the problem of asymptotic approximation under a sufficiently general situation.

Let us consider a family $\mathscr{T}$ of σ-finite measure spaces, (R, S, μ)'s, where R is an abstract space, S is a σ-field of subsets of R, and μ is a σ-finite measure over S. For any space (R, S, μ) belonging to $\mathscr{T}$, $\mathscr{P}(R, S, \mu)$ designates the family of all probability distributions, or equally, of all random variables, defined on the measurable space (R, S), which are absolutely continuous with respect to μ. Then, as is well-known, for every member X of $\mathscr{P}(R, S, \mu)$ there exists a generalized probability density function with respect to μ ('*gpdf*(μ)', in short), $f(z)$, such that

$$P^X(E)=\int_E f(z)d\mu$$

for any E belonging to S.

Let us consider a limiting process, $t \to t_0$, of a certain 'parameter' depending on t in a topological parameter space, and, throughout the present discussion, the above limiting process is fixed unless otherwise stated. Here, we use the term 'parameter' in a wide sense: It stands for a usual parameter of a probability distribution, or a size of sample, or a time-parameter of time series, etc.

Suppose that, corresponding to the above limiting process, we have a sequence, $\{X_t\}(t \to t_0)$, of probability distributions, where X_t belongs to $\mathscr{P}(R_t, S_t, \mu_t)$. Here, of course, the basic spaces (R_t, S_t, μ_t)'s are not necessarily different from one another. Further, suppose that we are given a sequence of measurable subsets, $\{E_t\}(t \to t_0)$, where E_t belongs to S_t for every $t(\to t_0)$, and that we are required to evaluate the value of $P^{X_t}(E_t)$ asymptotically as $t \to t_0$. This is an *asymptotic evaluation problem.* The problem of *asymptotic approximation* arises when it is difficult to evaluate the probability directly. In such a case, we must seek another sequence of probability distributions, $\{Y_t\}(t\to t_0)$, such that the probability $P^{Y_t}(E_t)$ is easy to evaluate and the relation

$$P^{X_t}(E_t)-P^{Y_t}(E_t)\to 0 \quad (t\to t_0)$$

holds.

In the asymptotic approximation problem, it is required to (a) construct an asymptotically approximate distribution to the one under consideration, and (b) define a notion of asymptotic equivalence appropriate for showing that a probability distribution can be used as an asymptotically approximate distribution.

However, it is difficult to discuss the former problem (a) in general, except for the case when the distribution under consideration is an asymptotically independent (in some sense) system of probability distributions. Notions of asymptotic independence in the asymptotic approximation problem will be introduced later in section 1.5. Section 1.2 and the subsequent two sections will be devoted to discussion of the problem (b).

1.2. Definitions of asymptotic equivalence

In the asymptotic approximation problem introduced in the preceding section, two sorts of error may be considered. The one is the absolute error, $|P^{X_t}(E_t)-P^{Y_t}(E_t)|$, and the other is the relative error, $|P^{X_t}(E_t)/P^{Y_t}(E_t)-1|$.

In order that the absolute error tends to zero for any sequence of measurable subsets, $\{E_t\}(t\to t_0)$, it is sufficient that

(1.2.1) $$\delta(P^{X_t}, P^{Y_t})=\sup|P^{X_t}(E_t)-P^{Y_t}(E_t)|\to 0 \quad (t\to t_0),$$

where the 'sup' is taken with respect to $E_t\in S_t$.

Now, define

(1.2.2) $$V(X_t, Y_t)=\int_{R_t}|f_t(z_t)-g_t(z_t)|d\mu_t,$$

where $f_t(z_t)$ and $g_t(z_t)$ stand for the $gpdf(\mu_t)$'s of X_t and Y_t, respectively. Then, it is evident that

(1.2.3) $$\delta(P^{X_t}, P^{Y_t})\leqq V(X_t, Y_t)\leqq 2\delta(P^{X_t}, P^{Y_t}),$$

from which it follows that the condition (1.2.1) is equivalent to the vanishing of the distance $V(X_t, Y_t)$ as $t\to t_0$.

DEFINITION 1.2.1. Two sequences of probability distributions, $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$, are said to be *asymptotically equivalent in the nse of type I*, or simply *asymptotically equivalent* (I), as $t\to t_0$, if it holds that

$$V(X_t,\ Y_t)\to 0 \quad (t\to t_0)\,. \tag{1.2.4}$$

Briefly we shall call this equivalence the 'asymptotic equivalence (I)'.

It should be remarked that Jeffroy's notion of asymptotic equivalence [3] (asymptotic independence is a special case of asymptotic equivalence) is strictly weaker than the above one. This fact is shown more precisely in the following section.

Now, we shall proceed to the consideration of relative error. Put

$$K(X_t,\ Y_t)=\sup\left|\frac{P^{X_t}(E_t)}{P^{Y_t}(E_t)}-1\right|\,, \tag{1.2.5}$$

where the 'sup' is taken with respect to E_t contained in S_t. Based upon this quantity $K(X_t,\ Y_t)$, we define the *asymptotic equivalence* (*II*) as follows.

DEFINITION 1.2.2. Two sequences of probability distributions, $\{X_t\}$ $(t\to t_0)$ and $\{Y_t\}(t\to t_0)$, are said to be *asymptotically equivalent in the sense of type II*, or simply *asymptotically equivalent* (*II*), as $t\to t_0$, if it holds that

$$K(X_t,\ Y_t)\to 0 \quad (t\to t_0)\,. \tag{1.2.6}$$

Generally speaking, this notion is too strong for the application to a problem of asymptotic approximation. For asymptotic approximation problem, where we are given a sequence $\{E_t'\}(t\to t_0)$, and the probability $P(E_t')$ is to be evaluated asymptotically, it is sufficient to define a weaker notion of asymptotic equivalence based upon the following distance:

$$K(X_t,\ Y_t;\ F_t)=\sup|(P^{X_t}(E_t)/P^{Y_t}(E_t))-1|\,, \tag{1.2.5$'$}$$

where $\{F_t\}(t\to t_0)$ is a suitable sequence of members of S_t such that $E_t'\subseteqq F_t$ for every $t(\to t_0)$, and the 'sup' is taken with respect to E_t belonging to $S_t(F_t)=\{E_t'\cap F_t : E_t'\in S_t\}$.

Asymptotic equivalence (II) is useful, for example, for the following problem of asymptotic approximation. Suppose that the value of

$$A(X_t)=\sum_{i=1}^{n(t)}\lambda_{it}P^{X_t}(E_{it})$$

is to be evaluated asymptotically as $t\to t_0$, where $n(t)$ tends to infinity as t tends to t_0 and the series is known to be absolutely convergent. If we can find a sequence of probability distributions, $\{Y_t\}(t\to t_0)$, which is asymptotically equivalent (II) to $\{X_t\}(t\to t_0)$, then, exchanging X_t for Y_t

in the above definition of $A(X_t)$, we have $|A(Y_t)-A(X_t)|\to 0$ as $t\to t_0$. Here, it is noted that the asymptotic equivalence (I) between $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$ does not necessarily assure the asymptotic equality between $A(X_t)$ and $A(Y_t)$.

1.3. Properties of asymptotic equivalence

In the present section, some properties of two sorts of asymptotic equivalence defined in the previous section will be stated. Throughout the present paper, $D(X)$ denotes the carrier of a gpdf. of X. For example,

$$D(X_t)=\{z_t : f_t(z_t)>0\} .$$

First of all, we can show the following relation between two notions of asymptotic equivalence.

LEMMA 1.3.1. *The notion of asymptotic equivalence (II) is strictly stronger than that of asymptotic equivalence (I).*

PROOF. Put

$$A_t=D(X_t)\frown\{z_t : f_t(z_t)\geqq g_t(z_t)\} ,$$
$$B_t=D(X_t)\frown\{z_t : f_t(z_t)<g_t(z_t)\} ,$$

and

$$C_t=D(Y_t)-D(X_t)\frown D(Y_t) .$$

Then we have

$$V(X_t,\ Y_t)=P^{X_t}(A_t)-P^{Y_t}(A_t)+P^{Y_t}(B_t)-P^{X_t}(B_t)+P^{Y_t}(C_t) ,$$

from which it follows that

$$V(X_t,\ Y_t)\leqq K(X_t,\ Y_t)+P^{Y_t}(C_t) . \tag{1.3.1}$$

Since the condition (1.2.6) implies that $P^{Y_t}(C_t)\to 0$, $(t\to t_0)$, the above inequality shows that the notion of asymptotic equivalence (II) is stronger than that of asymptotic equivalence (I).

The 'strictly' assertion is confirmed by the following example.

EXAMPLE 1.3.1. Let X_n and Y_n be one-dimensional real random variables, whose pdf.'s are

$$f_n(x)=\begin{cases}(n-1)/n^3, & \text{if } 0\leqq x<n^2,\\ 1/n^2, & \text{if } n^2\leqq x<n^2+n,\\ 0, & \text{otherwise,}\end{cases}$$

and

$$g_n(x)=\begin{cases}(n-2)/n^3, & \text{if } 0\leqq x<n^2\,,\\ 2/n^2, & \text{if } n^2\leqq x<n^2+n,\\ 0, & \text{otherwise,}\end{cases}$$

respectively. Then, it is clear that $V(X_n,\ Y_n)\to 0$ $(n\to\infty)$, while it does not hold that $K(X_n,\ Y_n)\to 0$ $(n\to\infty)$.

In the second place, we shall give some conditions for the asymptotic equivalence.

LEMMA 1.3.2. *In order that* $\{X_t\}(t\to t_0)$ *and* $\{Y_t\}(t\to t_0)$ *are asymptotically equivalent (I) as* $t\to t_0$, *it is necessary and sufficient that*

$$P^{X_t}(E_t)-P^{Y_t}(E_t)\to 0\ \ (t\to t_0)\,, \tag{1.3.2}$$

for any sequence $\{E_t\}(t\to t_0)$ *with* $E_t\in S_t$.

PROOF. The result is straightforward from the relations

$$|P^{X_t}(E_t)-P^{Y_t}(E_t)|\leqq V(X_t,\ Y_t)$$

and

$$V(X_t,\ Y_t)=P^{X_t}(A'_t)-P^{Y_t}(A'_t)+P^{Y_t}(B'_t)-P^{X_t}(B'_t)\,,$$

where

$$A'_t=\{z_t : f_t(z_t)\geqq g_t(z_t)\}\ ,$$

and

$$B'_t=\{z_t : f_t(z_t)<g_t(z_t)\}\ .$$

LEMMA 1.3.3. *In order that the condition* (1.2.6) *of asymptotic equivalence (II) holds, it is necessary and sufficient that one of the following conditions* (a) *and* (b) *is satisfied.*

(a) *For any sequence* $\{E_t\}(t\to t_0)$ *with* $E_t\in S_t$, *it holds that*

$$|(P^{X_t}(E_t)/P^{Y_t}(E_t))-1|\to 0\quad (t\to t_0)\,. \tag{1.3.3}$$

(b) *For any sequence* $\{E_t\}(t\to t_0)$ *with* $E_t\in S_t$, *there exists a function of* t, $\varphi(t)$, *independent of any particular choice of the sequence* $\{E_t\}(t\to t_0)$, *such that* $\varphi(t)\to 0$ *as* $t\to t_0$ *and*

$$|P^{X_t}(E_t)-P^{Y_t}(E_t)|\leqq\varphi(t)P^{Y_t}(E_t)\,, \tag{1.3.4}$$

for all values of t *sufficiently close to* t_0.

PROOF. (a) It is clear that the condition (1.2.6) is sufficient for (a).

Conversely, suppose that (1.2.6) is false. Then, there exist a sequence $\{E_t\}(t\to t_0)$ of the members of S_t and a positive constant δ such that a subsequence of $\{E_t\}(t\to t_0)$, say, $\{E_{t'}\}(t'\to t_0)$, satisfies the relation

$$|(P^{X_{t'}}(E_{t'})/P^{Y_{t'}}(E_{t'}))-1|\geqq\delta$$

for all $t'(\to t_0)$. This contradicts (a), which shows that (a) is sufficient for (1.2.6).

(b) Putting $\varphi(t)=K(X_t, Y_t)$, it can be seen that (1.2.6) is sufficient for (b).

Suppose (b) holds. If $P^{Y_t}(E_t)=0$, then it follows from (1. 3. 4) that $P^{X_t}(E_t)=0$. Hence, it holds that $K(X_t, Y_t)\leqq\varphi(t)$, which shows that (b) is sufficient for (1.2.6).

Thus, the proof of our lemma is complete.

In the last half of this section, we shall discuss some other properties of the asymptotic equivalence (I).

As is easily noticed, the situation, under which the notion of asymptotic equivalence (I) is defined, permits some specializations. If, in the definition 1.2.1, all the basic spaces (R_t, S_t, μ_t)'s are independent of t and identical with a certain space (R, S, μ), then the condition (1.2.4) turns out to be

$$\int_R |f_t(z)-g_t(z)|d\mu\to 0 \quad (t\to t_0)\,, \tag{1.3.5}$$

which is equivalent to the condition that

$$P^{X_t}(E)-P^{Y_t}(E)\to 0 \ (t\to t_0)\,, \tag{1.3.5$'$}$$

uniformly in E belonging to S. This is the case of 'equal basic spaces', and it is just in this case that Jeffroy [3] defined his notion of asymptotic equivalence. As was already remarked in the preceding section, Jeffroy's notion of asymptotic equivalence is strictly weaker than ours. This is seen in the following example.

EXAMPLE 1.3.2. Consider the condition (1.3.5) in the case when (R, S, μ) is the one-dimensional Euclidean space with Borel field S and Lebesgue measure μ. Then, to X_t and Y_t there correspond the cumulative distribution functions, $F_t(z)$ and $G_t(z)$, respectively. Jeffroy's definition of asymptotic equivalence (in the strong sense) is based upon the condition that

$$F_t(z)-G_t(z)\to 0 \quad (t\to t_0)\,, \tag{1.3.6}$$

uniformly in z. It is evident that (1.3.5) is stronger than (1.3.6).

The 'strictly' assertion is shown by the following example, which is a slight modification of Robbins' example [15]. Let

$$f_n(z)=\begin{cases} 1/(1-1/n^2), & \text{if } 1/n^2\leqq z\leqq 1\,, \\ 0, & \text{otherwise}\,, \end{cases}$$

and, for any fixed λ with $0<\lambda<1/2$ and $\delta_n=\lambda/(n2^n)$, let

$$g_n(z)=\begin{cases} 1/(n\delta_n), & \text{if } i/n-\delta_n\leqq z\leqq i/n,\ (i=1,\,2,\,\cdots,\,n)\,, \\ 0, & \text{otherwise}\,, \end{cases}$$

for $n=1, 2, \cdots$. Then, it is seen that

$$\sup_z |F_n(z)-G_n(z)|\to 0 \quad (n\to\infty)\,.$$

On the other hand, for $E_n=\{z:g_n(z)>0\}$, it can be seen that

$$\int_{E_n} f_n(z)dz=\frac{1}{1-1/n^2}\,\frac{\lambda}{2^n}<\frac{\lambda}{2^{n-1}}\,,\ (n=2,\,3,\,\cdots)\,,$$

and

$$\int_{E_n} g_n(z)dz=1,\ (n=1,\,2,\,\cdots)\,.$$

This shows that the condition (1.3.5) is not satisfied.

Now, we shall consider a more specialized situation. In addition to the equal basic spaces, if we assume that $g_t(z)$ is independent of t and is identical with $g(z)$ for every $t(\to t_0)$, then (1.2.4) becomes

$$\int_R |f_t(z)-g(z)|d\mu\to 0 \quad (t\to t_0)\,, \tag{1.3.7}$$

which is equivalent to the following.

$$P^{X_t}(E)\to P^Y(E) \quad (t\to t_0)\,, \tag{1.3.7$'$}$$

uniformly in E belonging to S, where Y is a probability distribution whose $gpdf(\mu)$ is $g(z)$. This is an important concept of convergence in statistics, together with that of 'in law' convergence. For the above convergence, a useful criterion has been given by Scheffé [16], or in a general situation of basic spaces, by, for example, Halmos [4] and Loéve [13]. Another criterion was introduced by Ikeda [9], which is a special case of the theorem 1.4.2 in the following section. It should be noted that our results of the present part are applicable to the problems of asymptotic approximation under these special situations, too.

Returning again to the general situation of definition 1.2.1, we shall consider another property of asymptotic equivalence (I).

Let $\{X_t\}(t\to t_0)$ be a sequence as before. Suppose there exists another family, $\bar{\mathscr{T}}$ of σ-finite measure spaces, and let u_t be a measurable transformation from the space R_t onto a space $\bar{R}_t$, of which $(\bar{R}_t, \bar{S}_t, \bar{\mu}_t)$ belongs to $\bar{\mathscr{T}}$ for every $t(\to t_0)$. For this transformation, put

$$\bar{X}_t = u_t(X_t)\,. \tag{1.3.8}$$

Then, the new random variable $\bar{X}_t$ defines a probability distribution over the measurable space $(\bar{R}_t, \bar{S}_t)$:

$$P^{\bar{X}_t}(\bar{E}_t) = P^{X_t}(u^{-1}(\bar{E}_t))\,.$$

Here we assume that $\bar{X}_t$ is absolutely continuous with respect to $\bar{\mu}_t$.

Now, suppose that $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$ are asymptotically equivalent (I) as $t\to t_0$, and that $\{\bar{X}_t\}(t\to t_0)$ and $\{\bar{Y}_t\}(t\to t_0)$ correspond respectively to the above sequences by measurable transformations u_t's in such a manner as stated above, for every $t(\to t_0)$. Under this situation we can get the following:

LEMMA 1.3.4. *If $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$ are asymptotically equivalent (I) as $t\to t_0$, then $\{\bar{X}_t\}(t\to t_0)$ and $\{\bar{Y}_t\}(t\to t_0)$ are also asymptotically equivalent (I) as $t\to t_0$.*

If, moreover, the transformations u_t's are all non-singular (i.e., one-to-one and the inverse is also measurable), then the asymptotic equivalence (I) of $\{\bar{X}_t\}(t\to t_0)$ and $\{\bar{Y}_t\}(t\to t_0)$ implies that of $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$.

PROOF. By lemma 1.3.2, the asymptotic equivalence (I) of $\{\bar{X}_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$ is equivalent to the condition

$$P^{\bar{X}_t}(\bar{E}_t) - P^{\bar{Y}_t}(\bar{E}_t) \to 0 \quad (t\to t_0)\,,$$

for any sequence $\{\bar{E}_t\}(t\to t_0)$ of the members of $\bar{S}_t$'s. Since

$$P^{\bar{X}_t}(\bar{E}_t) - P^{\bar{Y}_t}(\bar{E}_t) = P^{X_t}(u^{-1}(\bar{E}_t)) - P^{Y_t}(u^{-1}(\bar{E}_t))\,,$$

the asymptotic equivalence (I) of $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$ implies that of $\{\bar{X}_t\}(t\to t_0)$ and $\{\bar{Y}_t\}(t\to t_0)$. This proves the first assertion.

If u_t is non-singular for every $t(\to t_0)$, an analogous argument is made in the opposite direction, which shows that, together with the first argument, the asymptotic equivalence (I) of $\{\bar{X}_t\}(t\to t_0)$ and $\{\bar{Y}_t\}(t\to t_0)$ is equivalent to that of $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$. Thus the proof of our lemma is complete.

Lastly, we shall investigate an asymptotic equality of characteristic quantities of asymptotically equivalent (I) distributions.

Let $v_t(z_t)$ be a real-valued, measurable (S_t) and integrable (with respect to P^{X_t} and P^{Y_t}) function, defined on (R_t, S_t, μ_t), for every $t(\to t_0)$, and let

$$\mathscr{E}[v_t(X_t)]=\int_{R_t} v_t(z_t)f_t(z_t)d\mu_t ,$$

and

$$\mathscr{E}[v_t(Y_t)]=\int_{R_t} v_t(z_t)g_t(z_t)d\mu_t .$$

Then, we can show the following:

LEMMA 1.3.5. *Let $\{X_t\}(t\to t_0)$ and $\{Y_t\}(t\to t_0)$ be asymptotically equivalent (I) as $t\to t_0$. Then, in order that the asymptotic equality*

$$\mathscr{E}[v_t(X_t)]-\mathscr{E}[v_t(Y_t)]\to 0 \quad (t\to t_0) , \tag{1.3.9}$$

holds, it is sufficient that both of $\mathscr{E}[v_t(X_t)^2]$ and $\mathscr{E}[v_t(Y_t)^2]$ are bounded from the above by some $K(>0)$ independent of t.

PROOF. By using Schwarz's inequality, it is easy to see that

$$\begin{aligned}&|\mathscr{E}[v_t(X_t)]-\mathscr{E}[v_t(Y_t)]|^2\\ &\leqq\int_{R_t}|v_t(z_t)|^2|f_t(z_t)-g_t(z)_t|d\mu_t\cdot\int_{R_t}|f_t(z_t)-g_t(z_t)|d\mu_t .\end{aligned}$$

Hence, by assumption we obtain

$$|\mathscr{E}[v_t(X_t)]-\mathscr{E}[v_t(Y_t)]|\leqq\sqrt{2KV(X_t, Y_t)} ,$$

from which the lemma easily follows.

1.4. Asymptotic equivalence criterions

In order to make our notions of asymptotic equivalence applicable to practical problems, it is indispensable to give some criterions for proving the asymptotic equivalence of two sequences of probability distributions. In the present section, we shall introduce some of such criterions.

First, we investigate relations between several distances defined on a class of probability distributions. Let (R, S, μ) be a σ-finite measure space, and let $\mathscr{P}(R, S, \mu)$ be a family of probability distributions which are absolutely continuous with respect to μ. In addition to the distance

$V(X, Y)$ defined by (1.2.2), i.e.,

$$V(X, Y)=\int_R |f(z)-g(z)|d\mu \,, \tag{1.4.1}$$

we introduce the distance

$$W(X, Y)=\sqrt{\int_R (\sqrt{f(z)}-\sqrt{g(z)})^2 d\mu} \,, \tag{1.4.2}$$

and the so-called Kullback-Leibler mean information with a general definition [11]

$$I(X:Y)=\int_R f(z)\log\frac{f(z)}{g(z)}d\mu \,, \tag{1.4.3}$$

where $f(z)$ and $g(z)$ are the $gpdf(\mu)$'s of X and Y, respectively.

As for the distance $W(X, Y)$ it is seen that

$$W(X, Y)=\sqrt{2(1-\rho(X, Y))} \,,$$

where $\rho(X, Y)$ is the 'affinity' between X and Y (Matusita [14]), defined by

$$\rho(X, Y)=\int_R \sqrt{f(z)g(z)}\, d\mu \,. \tag{1.4.4}$$

Relations between these distances have been given by some authors, for example, by Hannan [5], Hoeffding and Wolfowitz [6], Ikeda [9], Matusita [14], or implicitly by Kudo [10, 11]. We shall summarize them in the following:

LEMMA 1.4.1. *For any members X and Y of $\mathscr{P}(R, S, \mu)$ it holds that*

$$0\leqq 2(1-\rho(X, Y))\leqq V(X, Y)\leqq 2\sqrt{1-\rho^2(X, Y)}\leqq 2\sqrt{I(X:Y)} \,, \tag{1.4.5}$$

and the last inequality holds true if we replace $I(X:Y)$ by $I(Y:X)$.

The proof is straightforward. For example, as to the last inequality, when $D(X)\not\subseteqq D(Y)(\mu)$, we have $I(X:Y)=\infty$ by definition, and the inequality clearly holds. When $D(X)\subseteqq D(Y)(\mu)$, we have

$$\begin{aligned} I(X:Y)&=\int_{D(X)} f(z)\log(f(z)/g(z))d\mu \\ &\geqq -2\log\int_{D(X)} f(z)\sqrt{g(z)/f(z)}\,d\mu=-2\log\rho(X, Y) \\ &\geqq 2(1-\rho(X, Y))\geqq 1-\rho^2(X, Y) \,, \end{aligned}$$

from which it follows that

$$2\sqrt{1-\rho^2(X,\ Y)}\leqq 2\sqrt{I(X:Y)}\,.$$

Now, let us return to the situation of section 1.2. From the above lemma, the following results are evident, which will be regarded as the asymptotic equivalence (I) criterions.

THEOREM 1.4.1. *In order that* $\{X_t\}\,(t\to t_0)$ *and* $\{Y_t\}\,(t\to t_0)$ *are asymptotically equivalent* (*I*) *as* $t\to t_0$, *it is necessary and sufficient that*

$$\rho(X_t,\ Y_t)=\int_{R_t}\sqrt{f_t(z_t)g_t(z_t)}\,d\mu_t\to 1\quad (t\to t_0)\,, \tag{1.4.6}$$

and the 'error estimation' is given by

$$V(X_t,\ Y_t)\leqq 2\sqrt{1-\rho^2(X_t,\ Y_t)}\,. \tag{1.4.7}$$

THEOREM 1.4.2. *In order that* $\{X_t\}\,(t\to t_0)$ *and* $\{Y_t\}\,(t\to t_0)$ *are asymptotically equivalent* (*I*) *as* $t\to t_0$, *it is sufficient that*

$$I(X_t:Y_t)=\int_{R_t}f_t(z_t)\log\frac{f_t(z_t)}{g_t(z_t)}\,d\mu_t\to 0\quad (t\to t_0)\,, \tag{1.4.8}$$

or

$$I(Y_t:X_t)=\int_{R_t}g_t(z_t)\log\frac{g_t(z_t)}{f_t(z_t)}\,d\mu_t\to 0\quad (t\to t_0)\,, \tag{1.4.9}$$

where the error estimation is given by

$$V(X_t,\ Y_t)\leqq 2\sqrt{I(X_t:Y_t)}\,,\ \text{or}\ \ V(X_t,\ Y_t)\leqq 2\sqrt{I(Y_t:X_t)} \tag{1.4.10}$$

correspondingly to the case (1.4.8) *or* (1.4.9).

As mentioned above, we have

$$V(X_t,\ Y_t)\leqq 2\sqrt{1-\rho^2(X_t,\ Y_t)}\leqq 2\sqrt{I(X_t:Y_t)}\,. \tag{1.4.11}$$

Therefore, when applying these theorems to practical problems, it is desirable to use the first one as much as possible, but there might exist some problems in which $\rho(X_t,\ Y_t)$ is difficult to calculate but $I(X_t,\ Y_t)$ is much easier to handle.

As to the application of the second theorem, it should be remarked that, when $D(X_t)\subseteqq D(Y_t)(\mu_t)$ the condition (1.4.8) must be examined, while when $D(Y_t)\subseteqq D(X_t)(\mu_t)$ the condition (1.4.9) must be used. In cases when these implication relations do not hold, theorem 1.4.2 is useless, and other criterions must be prepared. The following is one of criterions for such cases.

LEMMA 1.4.2. *For the asymptotic equivalence (I) of* $\{X_t\}(t\to t_0)$ *and* $\{Y_t\}(t\to t_0)$ *it is sufficient that*

$$P^{X_t}(D(X_t)-D(Y_t))\to 0,\ P^{Y_t}(D(Y_t)-D(X_t))\to 0\ , \tag{1.4.12}$$

and

$$d_0(X_t,\ Y_t)=ess.\ sup\left|\frac{f_t(z_t)}{g_t(z_t)}-1\right|\to 0\ , \tag{1.4.13}$$

as $t\to t_0$, *where 'ess. sup' is taken for* $z_t\in D(X_t)\cap D(Y_t)$ *with respect to* μ_t.

PROOF. This is straightforward from the inequality

$$V(X_t,\ Y_t)\leqq P^{X_t}(D(X_t)-D(Y_t))+P^{Y_t}(D(Y_t)-D(X_t))+d_0(X_t,\ Y_t)\ .$$

Next, we shall give a sufficient condition for the asymptotic equivalence (II).

THEOREM 1.4.3. *In order that* $\{X_t\}(t\to t_0)$ *and* $\{Y_t\}(t\to t_0)$ *are asymptotically equivalent (II) as* $t\to t_0$, *it is sufficient that*

$$d(X_t,\ Y_t)=ess.\ sup\left|\frac{f_t(z_t)}{g_t(z_t)}-1\right|\to 0,\ (t\to t_0)\ , \tag{1.4.14}$$

where the 'ess. sup' is taken for $z_t\in D(X_t)\cup D(Y_t)$ *with respect to* μ_t.

PROOF. It is clear that the condition (1.4.14) implies that $D(X_t)=D(Y_t)(\mu_t)$ for all values of t, at least for t sufficiently close to t_0. Hence, for such values of t, it holds that

$$|P^{X_t}(E_t)-P^{Y_t}(E_t)|\leqq\int_{E_t}|f_t(z_t)-g_t(z_t)|d\mu_t\leqq d(X_t,\ Y_t)P^{Y_t}(E_t)\ ,$$

for any set E_t belonging to S_t. Thus the result follows from lemma 1.3.3 (b).

1.5. Asymptotic independence

In order to define notions of asymptotic independence, we shall assume that the family $\mathscr{T}$ of basic spaces is closed under the finite-product operation, i.e., if the σ-finite measure spaces $(R_1,\ S_1,\ \mu_1)$, $(R_2,\ S_2,\ \mu_2)$, $\cdots$, $(R_n,\ S_n,\ \mu_n)$ belong to $\mathscr{T}$, then their product space $(R_{(n)},\ S_{(n)},\ \mu_{(n)})=\prod_{i=1}^{n}(R_i,\ S_i,\ \mu_i)$ also belongs to $\mathscr{T}$ for any finite positive integer n.

Now, suppose we are given a set of random variables, $\{X_1,\ X_2,\ \cdots,\ X_n\}$, each X_i belonging to $\mathscr{P}(R_i,\ S_i,\ \mu_i)$, and let $X_{(n)}=(X_1,\ X_2,\ \cdots,\ X_n)$ belong to $\mathscr{P}(R_{(n)},\ S_{(n)},\ \mu_{(n)})$. As is well-known, according to the usual definition, a set of random variables $\{X_i\}$ $(i=1, 2, \cdots, n)$ is said to be *independent* if, for any members E_i's of S_i's, it holds that

$$P^{X_{(n)}}(E_{(n)})=\prod_{i=1}^{n}P^{X_i}(E_i)\,, \tag{1.5.1}$$

where $E_{(n)}=\prod_{i=1}^{n}E_i$ denotes the direct-product of E_i's. By this definition, (a) independence of every proper subset of $\{X_i\}$ $(i=1, 2, \cdots, n)$ does not necessarily imply that of the whole set, and conversely, (b) independence of the whole set $\{X_i\}$ $(i=1, 2, \cdots, n)$ implies that of any subset of it (see, for example, Cramér [1]). From the viewpoint of applications to the problem of asymptotic approximation, it is convenient to define notions of asymptotic independence by specializing those of asymptotic equivalence defined in section 1.2. The concept of asymptotic independence (I) given below by such specialization, implies that of independence defined by (1.5.1) in the limit as $t\to t_0$, provided that n is independent of t.

Let us consider a set of random variables $\{X_i^t\}$ $(i=1, 2, \cdots, n)$, where n may depend on t as $t\to t_0$, and $X_i^t\in\mathscr{P}(R_i^t, S_i^t, \mu_i^t)$ for $i=1, 2, \cdots, n$. We shall call n the 'size' of the set. Let $X_{(n)}^t=(X_1^t, X_2^t, \cdots, X_n^t)$ be their joint variable, and assume that it belongs to $\mathscr{P}(R_{(n)}^t, S_{(n)}^t, \mu_{(n)}^t)$ and that there exists $gpdf\,(\mu_{(n)}^t)f_{(n)}^t(z_{(n)}^t)$ of $X_{(n)}^t$. On the other hand, let $f_i^t(z_i^t)$ be a $gpdf\,(\mu_i^t)$ of X_i^t for $i=1, 2, \cdots, n$, and put

$$g_{(n)}^t(z_{(n)}^t)=\prod_{i=1}^{n}f_i^t(z_i^t)\,, \tag{1.5.2}$$

where $z_{(n)}^t=(z_1^t, z_2^t, \cdots, z_n^t)$. Then, there exists an independent set of random variables, $\{Y_i^t\}$ $(i=1, 2, \cdots, n)$, say, such that each Y_i^t is identically distributed with X_i^t and their joint variable $Y_{(n)}^t=(Y_1^t, Y_2^t, \cdots, Y_n^t)$ necessarily has $g_{(n)}^t(z_{(n)}^t)$ as a $gpdf\,(\mu_{(n)}^t)$. Under this specialized situation, the definitions (1.2.2) and (1.2.5) become in turn

$$V(X_{(n)}^t, Y_{(n)}^t)=\int_{R_{(n)}^t}|f_{(n)}^t(z_{(n)}^t)-g_{(n)}^t(z_{(n)}^t)|d\mu_{(n)}^t\,, \tag{1.5.3}$$

and

$$K(X_{(n)}, Y_{(n)})=\sup\left|\frac{P^{X_{(n)}^t}(E_{(n)}^t)}{P^{Y_{(n)}^t}(E_{(n)}^t)}-1\right|\,, \tag{1.5.4}$$

where the 'sup' is taken for $E_{(n)}^t\in S_{(n)}^t$. It should be noted that, under the present situation, the inclusion relation of the carriers $D(X_{(n)}^t)\subseteqq D(Y_{(n)}^t)$ $(\mu_{(n)}^t)$ holds.

We shall define the asymptotic independence (I) as follows:

DEFINITION 1.5.1. A set of random variables, $\{X_i^t\}(i=1, 2, \cdots, n)$, is said to be *asymptotically independent in the sense of type I*, or simply *asymptotically independent (I)*, as $t \to t_0$, if $\{X_{(n)}^t\}(t \to t_0)$ and $\{Y_{(n)}^t\}(t \to t_0)$ are asymptotically equivalent (I) as $t \to t_0$, i.e., if it holds that

$$V(X_{(n)}^t, Y_{(n)}^t) \to 0 \quad (t \to t_0) . \tag{1.5.5}$$

In connection with this definition, we define the asymptotic n-independence as follows:

DEFINITION 1.5.2. A set of random variables, $\{X_k^t\}(k=1, 2, \cdots, N)$, is said to be *asymptotically n-independent in the sense of type I*, or briefly *asymptotically n-independent (I)*, as $t \to t_0$, if all subsets of size n of $\{X_{k_i}^t\}(i=1, 2, \cdots, N)$, are asymptotically independent (I) as $t \to t_0$, where N and n may be dependent on t.

Quite analogously, asymptotic independence (II) is defined as follows:

DEFINITION 1.5.3. We shall say that a set of random variables, $\{X_i^t\}(i=1, 2, \cdots, n)$, is *asymptotically independent in the sense of type II*, or simply *asymptotically independent (II)*, as $t \to t_0$, if $\{X_{(n)}^t\}(t \to t_0)$ and $\{Y_{(n)}^t\}(t \to t_0)$ are asymptotically equivalent (II) as $t \to t_0$, where $\{Y_{(n)}^t\}$ have the above stated property, i.e., if it holds that

$$K(X_{(n)}^t, Y_{(n)}^t) \to 0 \quad (t \to t_0) . \tag{1.5.6}$$

DEFINITION 1.5.4. A set of random variables, $\{X_k^t\}(k=1, 2, \cdots, N)$, is said to be *asymptotically n-independent in the sense of type II*, or simply *asymptotically n-independent (II)*, as $t \to t_0$, if all subsets of size n of $\{X_{k_i}^t\}(i=1, 2, \cdots, N)$, are asymptotically independent (II) as $t \to t_0$, where N and n may depend on t.

As a consequence of theorem 1.4.1, we have

COROLLARY 1.5.1. *In order that a set of random variables $\{X_i^t\}(i=1, 2, \cdots, n)$ is asymptotically independent (I) as $t \to t_0$, it is necessary and sufficient that*

$$\rho(X_{(n)}^t, Y_{(n)}^t) \to 1 \quad (t \to t_0) . \tag{1.5.7}$$

From theorem 1.4.2 we have:

COROLLARY 1.5.2. *In order that a set of random variables $\{X_i^t\}(i=1, 2, \cdots, n)$ is asymptotically independent (I) as $t \to t_0$, it is sufficient that*

$$I(X_{(n)}^t : Y_{(n)}^t) \to 0 \quad (t \to t_0) . \tag{1.5.8}$$

By specializing theorem 1.4.3, we obtain

COROLLARY 1.5.3. *In order that a set of random variables* $\{X_i^t\}$ $(i=1, 2, \cdots, n)$ *is asymptotically independent (II) as* $t\to t_0$, *it is sufficient that*

$$d(X_{(n)}^t, Y_{(n)}^t)\to 0 \quad (t\to t_0)\,, \tag{1.5.9}$$

where the distance $d(X_{(n)}^t, Y_{(n)}^t)$ *is defined analogously to that of* (1.4.14).

Now, we shall remark some properties of asymptotic independence (I).

First, we have

LEMMA 1.5.1. *Under the situation of equal basic spaces, if* n *is independent of* t, *and for every* i *the limiting distribution of* X_i^t *exists, then* $\{X_i^t\}$ $(i=1, 2, \cdots, n)$ *is independent (in the sense of* (1.5.1)) *in the limit as* $t\to t_0$.

Secondly, taking a measurable transformation such that

$$u_t(X_{(n)}^t)=X_{(m)}^t\,,$$

where $X_{(m)}^t=(X_{i_1}^t, X_{i_2}^t, \cdots, X_{i_m}^t)$ and $m\leqq n$, and applying the lemma 1.3.4, we have

LEMMA 1.5.2. *If a set of random variables* $\{X_i^t\}$ $(i=1, 2, \cdots, n)$ *is asymptotically independent (I) as* $t\to t_0$, *then every subset of* $\{X_i^t\}$ *are also asymptotically independent (I) as* $t\to t_0$.

This result is extended in a more general form: Partition the set of random variables $\{X_i^t\}$ $(i=1, 2, \cdots, n)$ into s mutually exclusive subsets, $\{X_{k_j}^t\}$ $(j=1, 2, \cdots, n_k\,; k=1, 2, \cdots, s)$, and let

$$X_{(n_k)}^t=(X_{k_1}^t, X_{k_2}^t, \cdots, X_{k_{n_k}}^t), \quad (k=1, 2, \cdots, s)\,.$$

Then, we have the decomposition of $X_{(n)}^t$

$$X_{(n)}^t=(X_{(n_1)}^t, X_{(n_2)}^t, \cdots, X_{(n_s)}^t)\,,$$

and the corresponding decomposition of the basic space

$$(R_{(n)}^t, S_{(n)}^t, \mu_{(n)}^t)=\prod_{k=1}^{s}(R_{(n_k)}^t, S_{(n_k)}^t, \mu_{(n_k)}^t)\,, \tag{1.5.10}$$

where, of course, $X_{(n_k)}^t$ belongs to $\mathscr{P}(R_{(n_k)}^t, S_{(n_k)}^t, \mu_{(n_k)}^t)$.

Now, let $u^t=(u_1^t, u_2^t, \cdots, u_s^t)$ be a measurable transformation from the space (1.5.10) into a certain measure space

$$(\bar{R}_{(s)}^t, \bar{S}_{(s)}^t, \bar{\mu}_{(s)}^t)=\prod_{k=1}^{s}(\bar{R}_k^t, \bar{S}_k^t, \bar{\mu}_k^t)$$

belonging to a family $\bar{\mathscr{T}}$ of σ-finite measure spaces, such that

(1.5.11) $$u^t(z^t_{(n)})=(u^t_1(z^t_{(n_1)}),\ u^t_2(z^t_{(n_2)}),\ \cdots,\ u^t_s(z^t_{(n_s)}))\ ,$$

where $z^t_{(n)}=(z^t_{(n_1)},\ z^t_{(n_2)},\ \cdots,\ z^t_{(n_s)})$ correspondingly to (1.5.10). Here, s may depend on t. This transformation defines that of $X^t_{(n)}$ such that

(1.5.12) $$\bar{X}^t_{(s)}=u^t(X^t_{(n)})\ ,$$

or

(1.5.12)′ $$(\bar{X}^t_1,\ \bar{X}^t_2,\ \cdots,\ \bar{X}^t_s)=(u^t_1(X^t_{(n_1)}),\ u^t_2(X^t_{(n_2)}),\ \cdots,\ u^t_s(X^t_{(n_s)}))\ ,$$

where each $\bar{X}^t_k$ is assumed to belong to $\mathscr{P}(\bar{R}^t_k,\ \bar{S}^t_k,\ \bar{\mu}^t_k)$.

Under the above situation, the following is an immediate consequence of lemma 1.3.4.

LEMMA 1.5.3. *If $\{X^t_i\}(i=1,\ 2,\ \cdots,\ n)$ is asymptotically independent (I) as $t\to t_0$, then $\{\bar{X}^t_k\}(k=1,\ 2,\ \cdots,\ s)$ is also asymptotically independent (I) as $t\to t_0$. If, moreover, $s=n$ and all the transformations u^t_k's are non-singular, then the asymptotic independence (I) of $\{\bar{X}^t_k\}(k=1,\ 2,\ \cdots,\ s)$ implies that of $\{X^t_i\}(i=1,\ 2,\ \cdots,\ n)$.*

PART II

APPLICATION TO CERTAIN PROBLEMS OF ASYMPTOTIC INDEPENDENCE

2.1. Asymptotic independence of a set of coverages

Basic results obtained in the preceding part is applied, in the present section, to the examination of asymptotic independence in the sense of type I, of a set of coverages.

According to the definition by Wilks [17], a set of random variables, $\{C_i\}(i=1,\ 2,\ \cdots,\ N+1)$, is said to be a *set of elementary coverages*, if it is subject to the restraint

$$\sum_{i=1}^{N+1} C_i=1\ ,$$

and the probability element of the joint distribution of C_i's is given by

(2.1.1) $$N!\,dc_1dc_2\cdots dc_N\ ,\ (0\leqq c_i,\ \sum_{i=1}^{N} c_i\leqq 1)\ .$$

As far as the author is aware, no work has been presented in the literature as to the property of asymptotic independence (I) of this set as the size N tends to infinity. In the previous study [8], the author is faced with essentially the same problem as this, where he strongly felt the need of the theory of asymptotic independence.

Now, we shall consider a subset, $\{C_{i_j}\}$ $(j=1, 2, \cdots, n)$, of a set of elementary coverages, where the size of the subset may be dependent on N as N tends to infinity. A question may arise, how large the value of n should be in order that $\{C_{i_j}\}$ $(j=1, 2, \cdots, n)$ are asymptotically independent (I) as N tends to infinity.

In the first place, it is noted that the joint distribution of C_{i_j}'s has a probability density function, independent of any particular choice of the subset, such as

$$f_{(n)}(c_{(n)})=\frac{\Gamma(N+1)}{\Gamma(N-n+1)}(1-\sum_{i=1}^{n} c_i)^{N-n}, \quad (0\leqq c_i, \sum_{i=1}^{n} c_i\leqq 1), \tag{2.1.2}$$

where $c_{(n)}=(c_1, c_2, \cdots, c_n)$. Hence, there is no loss of generality to restrict our attention to a set of first n coverages, $\{C_i\}$ $(i=1, 2, \cdots, n)$, for which we put $C_{(n)}=(C_1, C_2, \cdots, C_n)$.

The probability density function of each marginal C_i is independent of i and is given by

$$f(c)=N(1-c)^{N-1}, \quad (0\leqq c\leqq 1), \tag{2.1.3}$$

and the sum of any r coverages is distributed according to a beta distribution whose probability density function is

$$\frac{\Gamma(N+1)}{\Gamma(r)\Gamma(N-r+1)}u^{r-1}(1-u)^{N-r}, \quad (0\leqq u\leqq 1). \tag{2.1.4}$$

Let $C'_{(n)}$ be a random variable of n dimensions, whose probability density function is

$$g_{(n)}(c_{(n)})=\prod_{i=1}^{n} f(c_i)=N^n \prod_{i=1}^{n} (1-c_i)^{N-1}, \tag{2.1.5}$$

for $0\leqq c_i\leqq 1$, $(i=1, 2, \cdots, n)$.

In order to apply the result of corollary 1.5.2 as an asymptotic independence (I) criterion, we shall calculate the Kullback-Leibler mean information, the vanishing of which as N tends to infinity implies the asymptotic independence (I) of the subset $\{C_i\}$ $(i=1, 2, \cdots, n)$. Now, in our present situation, the Kullback-Leibler mean information of (1.5.8) becomes

$$I(C_{(n)} : C'_{(n)})=\mathscr{E}\left[\log \frac{f_{(n)}(C_{(n)})}{g_{(n)}(C_{(n)})}\right], \tag{2.1.6}$$

where the expectation $\mathscr{E}$ is taken with respect to the distribution of $C_{(n)}$, i.e., with respect to (2.1.2). Thus, from (2.1.2) and (2.1.5) it follows that

$$(2.1.7)\quad I(C_{(n)}:C'_{(n)})=\log\frac{\Gamma(N+1)}{N^n\Gamma(N-n+1)}+(N-n)\mathscr{E}[\log(1-\sum_{i=1}^{n}C_i)]$$
$$-(N-1)\sum_{i=1}^{n}\mathscr{E}[\log(1-C_i)]\,.$$

For the exact calculation of this information, we shall prepare a formula of integral calculus.

LEMMA 2.1.1.

$$(2.1.8)\quad \int_0^1 x^{p-1}(1-x)^{q-1}\log x\,dx=-\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\sum_{i=1}^{q}\frac{1}{p-1+i}\,,$$

for integers p and q.

For the second member of the right-hand side of (2.1.7), by using this lemma, we obtain

$$(2.1.9)\quad \mathscr{E}[\log(1-\sum_{i=1}^{n}C_i)]$$
$$=\frac{\Gamma(N+1)}{\Gamma(n)\Gamma(N-n+1)}\int_0^1 u^{n-1}(1-u)^{N-n}\log(1-u)\,du$$
$$=-\sum_{i=1}^{n}\frac{1}{N-n+i}\,,$$

and for the third term we have

$$(2.1.10)\quad \mathscr{E}[\log(1-C_i)]=N\int_0^1(1-c)^{N-1}\log(1-c)dc=-\frac{1}{N}\,.$$

Thus, from (2.1.7), (2.1.9) and (2.1.10) it follows that

$$(2.1.11)\quad I(C_{(n)}:C'_{(n)})=\log\frac{N(N-1)\cdots(N-n+1)}{N^n}$$
$$-(N-n)\sum_{i=1}^{n}\frac{1}{N-n+i}+\frac{(N-1)n}{N}\,.$$

Then, we obtain

$$(2.1.12)\quad I(C_{(n)}:C'_{(n)})\leqq n(n-1)/N\,,$$

from which we can conclude that, if $n=o(\sqrt{N})$, then $I(C_{(n)}:C'_{(n)})$ tends to zero as N tends to infinity.

Hence, we have the following

THEOREM 2.1.1. *A set of elementary coverages,* $\{C_i\}(i=1, 2, \cdots, N+1)$, *is asymptotically n-independent (I) as* $N\to\infty$, *provided that* $n=o(\sqrt{N})$.

This result will be obtained in a more general form.

Let us take n subsets of elementary coverages, $\{C_{1_i}\}(i=1, 2, \cdots, m_1)$, $\{C_{2_i}\}(i=1, 2, \cdots, m_2)$, $\cdots$, $\{C_{n_i}\}(i=1, 2, \cdots, m_n)$, which are chosen mutually exclusively from a set of elementary coverages $\{C_i\}(i=1, 2, \cdots, N+1)$, where n and m_k's can depend on N according as N tends to infinity. Let $C_{(m_k)}=(C_{k_1}, C_{k_2}, \cdots, C_{k_{m_k}})$, for each $k(=1, 2, \cdots, n)$. Under this situation, as for the asymptotic independence (I) of a set of random variables, $\{C_{(m_k)}\}(k=1, 2, \cdots, n)$, the following corollary is obtained as a direct consequence of lemma 1.5.3 and the above theorem.

COROLLARY 2.1.1. *A set* $\{C_{(m_k)}\}(k=1, 2, \cdots, n)$ *is asymptotically independent (I) as* $N\to\infty$, *provided that* $m(=\sum_{k=1}^{n} m_k)=o(\sqrt{N})$.

By lemma 1.5.3 the following is easily follows from this corollary.

COROLLAPY 2.1.2. *Under the same situation as in the preceding corollary, let*

$$\bar{C}_k=u_k(C_{(m_k)}),\ (k=1, 2, \cdots, n),$$

where u_k*'s are all measurable transformations. Then, in order that* $\{\bar{C}_k\}(k=1, 2, \cdots, n)$ *is asymptotically independent (I) as N tends to infinity, it is sufficient that* $m=o(\sqrt{N})$, *where as before* $m=\sum_{i=1}^{n} m_k$. *Hence, if* $n=0(1)$ *independently of N, then the condition* $\bar{m}=o(\sqrt{N})$ *implies the asymptotic independence (I) of that set, where* $\bar{m}=\max(m_1, m_2, \cdots, m_n)$.

Now, let $X_1\leqq X_2\leqq\cdots\leqq X_N$ be an ordered sample of size N drawn from a uniform distribution on a certain interval, (θ_1, θ_2), say, of unit length. Then, putting

$$C_i=X_i-X_{i-1},\ (i=1, 2, \cdots, N+1), \tag{2.1.13}$$

where $X_0=\theta_1$ and $X_{N+1}=\theta_2$, we have a set of elementary coverages $\{C_i\}$ $(i=1, 2, \cdots, N+1)$. Our results obtained above are applicable to this set.

More generally, as is seen easily, the following is an immediate consequence of lemma 1.5.3 and theorem 2.1.1.

COROLLARY 2.1.3. *Let* $X_1\leqq X_2\leqq\cdots\leqq X_N$ *be an ordered sample from a uniform distribution on a certain interval* (θ_1, θ_2) *with any given length, and let* $\bar{C}_i$*'s be random variables defined analogously to* (2.1.18). *Then, in order that* $\{\bar{C}_i\}(i=1, 2, \cdots, N+1)$ *is asymptotically n-independent (I) as N tends to infinity, it is sufficient that* $n=o(\sqrt{N})$.

Under the situation of lemma 1.3.4, it is not difficult to see that

$$V(\bar{X}_t, \bar{Y}_t) \leqq V(X_t, Y_t),$$

where the equality holds when the transformation u_t is non-singular. Accordingly, if we define, analogously to the distance (1.5.3) in the general theory, the distances for those variables which are considered in corollary 2.1.1, corollary 2.1.2 and corollary 2.1.3 (in the first two results, replacing m by n), those distances are not larger than that in theorem 2.1.1, i.e., not larger than $V(C_{(n)}, C'_{(n)})$, and in particular, the distance in the last corollary 2.1.3, $V(\bar{C}_{(n)}, \bar{C}'_{(n)})$, is equal to $V(C_{(n)}, C'_{(n)})$. In this sense, the value of $V(C_{(n)}, C'_{(n)})$ is regarded as a sort of 'weakness of independence (I)' of a set of elementary coverages, and it will be of some importance in practical applications to check the values for some N and n.

As was already shown in (1.4.10), it holds that

$$V(C_{(n)}, C'_{(n)}) \leqq 2\sqrt{I(C_{(n)} : C'_{(n)})}$$

in our present case, where the exact values of $I(C_{(n)} : C'_{(n)})$ for given N and n may be calculable from (2.1.11). Exact values of the right-hand member of this inequality for some values of N and n are tabulated in the following table.

Table 1.
Values of $2\sqrt{I(C_{(n)}:C'_{(n)})}$.

n \ N	10	20	50	100	500
2	0.1516	0.0732	0.0286	0.0144	0.0049
3		0.1290	0.0593	0.0259	
4			0.0739	0.0397	
6			0.1153	0.0566	0.0187
8				0.0767	
10				0.0990	0.0195
20					0.0894

2.2. Asymptotic independence of extreme values

In the present section we shall be concerned with the asymptotic independence (I) of extreme values of a random sample when the size of the sample increases infinitely.

Let $Z_1 \leqq Z_2 \leqq \cdots \leqq Z_N$ be an ordered sample of size N drawn from a one-dimensional continuous probability distribution defined on the real line. Under some restrictions on the population distribution, Gumbel [2] has shown that the nth lower extreme and the mth upper extreme are mutually 'independently' distributed as N tends to infinity, provided that n, $m=0(1)$ independently of N. The notion of asymptotic independence in Gumbel's sense is not so clear, but, scrutinizing his paper, one will notice that Gumbel's notion is weaker than or equivalent to that of the asymptotic independence of type I in our present case. The problem to be discussed in this section is regarded as an extension of the above result. By applying the so-called probability integral transformation

$$X_i=F(Z_i),\ i=1,\ 2,\ \cdots,\ N\,, \tag{2.2.1}$$

we have a new set of random variables, $X_1 \leqq X_2 \leqq \cdots \leqq X_N$, which is regarded as an ordered sample drawn from a uniform distribution on the interval (0,1). In the first place, we shall be concerned with these variables.

Let us take a set of first n variables, $\{X_1, X_2, \cdots, X_n\}$, and that of last m variables, $\{X_{N-m+1}, X_{N-m+2}, \cdots, X_N\}$, which will be called briefly the *set of lower n extremes* and the *set of upper m extremes*, respectively. For the sake of notational simplicity, we put, for the second set,

$$Y_1=X_N,\ Y_2=X_{N-1},\ \cdots,\ Y_m=X_{N-m+1}\,.$$

Let $X_{(n)}=(X_1, X_2, \cdots, X_n)$ and $Y_{(m)}=(Y_1, Y_2, \cdots, Y_m)$ be their joint variables respectively. Then, since the set $\{C_i\}$, $(i=1, 2, \cdots, N+1)$, defined by

$$C_i=X_i-X_{i-1},\ i=1,\ 2,\ \cdots,\ N+1,\ \text{with } X_0=0 \text{ and } X_{N+1}=1\,,$$

is a set of elementary coverages, and since

$$X_i=\sum_{s=1}^{i} C_s,\ i=1,\ 2,\ \cdots,\ n,\ \text{and } Y_j=1-\sum_{s=N+2-j}^{N+1} C_s,\ j=1,\ 2,\ \cdots,\ m\,,$$

the following theorem follows from corollary 2.1.2.

THEOREM 2.2.1. *A set of two random variables, $\{X_{(n)}, Y_{(m)}\}$, is asymptotically independent (I) as N tends to infinity, provided that $n+m=o(\sqrt{N})$.*

According to the usual definition, we shall say that 'two sets of extremes, i.e., the set of lower n extremes and that of upper m extremes, are mutually asymptotically independent in the sense of type I as N tends to infinity', if $\{X_{(n)}, Y_{(m)}\}$ is asymptotically independent (I) as N tends to infinity. By virtue of lemma 1.5.3, the following is an immediate consequence of the above theorem.

COROLLARY 2.2.1. *Any subset of the set of lower n extremes and any subset of the set of upper m extremes of an ordered sample of size N drawn from a uniform distribution on (0,1) are mutually asymptotically independent in the sense of type I as N tends to infinity, provided that $n+m=o(\sqrt{N})$.*

As a special case of this corollary, we have

COROLLARY 2.2.2. *The nth lower extreme and the mth upper extreme of an ordered sample of size N drawn from a uniform distribution on* (0,1) *are mutually asymptotically independent in the sense of type I as N tends to infinity, provided that* $n+m=o(\sqrt{N})$.

This is an improvement of Gumbel's result stated in the beginning of the present section.

Now, we shall return to the original ordered sample considered in the beginning of this section. As before, let $Z_1 \leqq Z_2 \leqq \cdots \leqq Z_N$ be an ordered sample of size N, drawn from a real, one-dimensional continuous distribution, whose cumulative distribution function and probability density function are $F(z)$ and $f(z)$ respectively. The probability integral transformation (2.2.1) is rewritten as

$$X_{(N)}=u_{(N)}(Z_{(N)})=(F(Z_1),\ F(Z_2),\ \cdots,\ F(Z_N))\,, \tag{2.2.2}$$

where $X_{(N)}=(X_1, X_2, \cdots, X_N)$ and $Z_{(N)}=(Z_1, Z_2, \cdots, Z_N)$,

and we have the inverse transformation

$$Z_{(N)}=u_{(N)}^{-1}(X_{(N)})=(F^{-1}(X_1),\ F^{-1}(X_2),\ \cdots,\ F^{-1}(X_N))\,. \tag{2.2.3}$$

Choose the intervals on which $F(z)$ is strictly increasing, and let R be their union, where we can assume without any loss of generality that R is an open domain on the real line. For our present consideration of the asymptotic independence problem, it will be sufficient to regard R as a basic space. Then, the transformation $x=F(z)$ from R onto the interval (0,1) is one-to-one and the inverse transformation $z=F^{-1}(x)$ is measurable, and hence the inverse transformation (2.2.3) is non-singular. Thus, by applying the lemma 1.5.3, we have

THEOREM 2.2.2. *Let* $Z_1 \leqq Z_2 \leqq \cdots \leqq Z_N$ *be an ordered sample of size N drawn from a real, one-dimensional continuous distribution. Then, the set of lower n extremes,* $\{Z_i\}$ $(i=1, 2, \cdots, n)$, *and that of upper m extremes,* $\{Z_j\}$ $(j=N-m+1, N-m+2, \cdots, N)$, *are mutually asymptotically independent (I) as N tends to infinity, provided that* $n+m=o(\sqrt{N})$.

In particular, correspondingly to corollary 2.2.2, we have

COROLLARY 2.2.3. *Under the same situation as in the above theorem, the nth lower extreme* Z_n *and the mth upper extreme* Z_{N-m+1} *are mutually asymptotically independent (I) as N tends to infinity, provided that* $n+m=o(\sqrt{N})$.

From the standpoint of the asymptotic independence consideration, this is regarded as an extension of Gumbel's result in two directions, that is, in widening the basic distributions and in loosening the condition to be satisfied by n and m.

It is also noted that theorem 2.2.2 is an extension of the preceding theorem 2.2.1, and it states that, in practical applications to the asymptotic evaluation problems which concerns the lower and upper extremes of an ordered sample, we can employ them as if they are mutually independent in the usual sense. For such cases, it will be important that we know how large the 'error' of approximation would be. Since the transformations (2.2.2) and (2.2.3) are non-singular, it is not so difficult to see that the error of approximation in the above sense is less than or equal to $2\sqrt{I(n, m, N)}$ for any given n, m and N in all cases under the situation of theorem 2.2.2, where $I(n, m, N)$ is given by

$$
\begin{aligned}
(2.2.4) \qquad I(n, m, N) = \log \frac{\Gamma(N-n+1)\Gamma(N-m+1)}{\Gamma(N+1)\Gamma(N-n-m+1)} \\
+(N-n-m)\,\mathscr{E}[\log(Y_m - X_n)] \\
-(N-n)\,\mathscr{E}[\log(1-X_n)] \\
-(N-m)\,\mathscr{E}[\log Y_m] .
\end{aligned}
$$

Now, the right-hand members of (2.2.4) can be calculated exactly, as is shown in the following. First, it follows from the joint distribution of X_n and Y_m that

$$(2.2.5) \qquad \mathscr{E}[\log(Y_m - X_n)] = C(n, m, N)\, J(n, m, N) ,$$

where

$$C(n, m, N) = \frac{\Gamma(N+1)}{\Gamma(n)\Gamma(N-n-m+1)\Gamma(m)} ,$$

and

$$J(n, m, N) = \iint_{0<x<y<1} x^{n-1}(y-x)^{N-n-m}(1-y)^{m-1} \log(y-x)\, dx dy .$$

By changing the variable as $x = yz$, we obtain

$$
\begin{aligned}
J(n, m, N) &= \int_0^1 y^{N-m}(1-y)^{m-1} dy \int_0^1 z^{n-1}(1-z)^{N-n-m} \log(y(1-z))\, dz \\
&= B(n, N-n-m+1) \int_0^1 y^{N-m}(1-y)^{m-1} \log y\, dy \\
&\quad + B(m, N-m+1) \int_0^1 z^{n-1}(1-z)^{N-n-m} \log(1-z)\, dz .
\end{aligned}
$$

Here, it holds by lemma 2.1.1 that

$$\int_0^1 y^{N-m}(1-y)^{m-1} \log y \, dy = -B(m, N-m+1) \sum_{i=1}^{m} \frac{1}{N-m+i},$$

and

$$\int_0^1 z^{n-1}(1-z)^{N-n-m} \log (1-z) dz = -B(n, N-n-m+1) \sum_{i=1}^{n} \frac{1}{N-n-m+i},$$

where $B(p, q)$ designates the usual beta function. Thus we have, by (2.2.5),

$$\mathscr{E}[\log (Y_m - X_n)] = -\left[\sum_{i=1}^{m} \frac{1}{N-m+i} + \sum_{i=1}^{n} \frac{1}{N-n-m+i}\right]. \tag{2.2.6}$$

Similarly, we can obtain

$$\mathscr{E}[\log (1-X_n)] = -\sum_{i=1}^{n} \frac{1}{N-n+i}, \tag{2.2.7}$$

and

$$\mathscr{E}[\log Y_m] = -\sum_{i=1}^{m} \frac{1}{N-m+i}. \tag{2.2.8}$$

Substituting (2.2.6), (2.2.7) and (2.2.8) into (2.2.4), we have

$$\begin{aligned} I(n, m, N) = {} & (N-n)U(N-n\,; n) + (N-m)U(N-m\,; m) \\ & -(N-n-m)U(N-n-m\,; n+m) - W(n, m, N), \end{aligned} \tag{2.2.9}$$

where

$$U(p\,; q) = \sum_{i=1}^{q} \frac{1}{p+i} \quad \text{for positive integers } p \text{ and } q, \tag{2.2.10}$$

and

$$W(n, m, N) = \log \frac{\Gamma(N+1)\Gamma(N-n-m+1)}{\Gamma(N-n+1)\Gamma(N-m+1)}. \tag{2.2.11}$$

By virtue of (2.2.9), exact calculation of the values of $2\sqrt{I(n, m, N)}$ is possible for any given n, m and N, and some of them are tabulated in table 2.

2.3. Asymptotic independence of a system of multinomial variates

A set of discrete random variables, $\{X_1, X_2, \cdots, X_r\}$, distributed according to a multinomial distribution with probabilities

Table 2.

Values of $2\sqrt{I(n, m, N)}$

N=10

m \ n	1
1	0.04796

N=50

m \ n	1	2	3	4	5	6
1	0.02867	0.04096	0.05070	0.05918	0.06690	0.07412
2		0.05854	0.07246	0.08446	0.09564	
3			0.08972	0.10474		

N=100

m \ n	1	2	4	6	8	10	12
1	0.01424	0.02023	0.02891	0.03578	0.04167	0.04722	0.05229
2		0.02875	0.04109	0.05088	0.05939	0.06716	
4			0.05874	0.07272	0.08490		
6				0.09004			

N=500

m \ n	4	8	12	16	20
4	0.01139	0.01615	0.01985	0.02309	0.02587
8		0.02303	0.02832	0.03282	
12			0.03473		

$$P\{(X_1, X_2, \cdots, X_r)=(n_1, n_2, \cdots, n_r)\} = \frac{n!}{\prod_{i=1}^{r} n_i!}\prod_{i=1}^{r} p_i^{n_i}, \tag{2.3.1}$$

$$(0 \leqq n_i, \sum_{i=1}^{r} n_i = n),$$

where $0<p_i$, $\sum_{i=1}^{r} p_i=1$, will be called *a system of multinomial variates.* In this section, asymptotic independence (I) properties of this system are investigated.

LEMMA 2.3.1. *Let X_n be a random variable with a probability distribution, P^{X_n}, defined over a measurable space (R, S), where R is a metric space with distance $D(z, z')$, and S is a σ-field of subsets of R, containing all open subsets of R. Here, throughout $n=1, 2, \cdots$, (R, S) is fixed. Furthermore, let $t(z)$ be a real-valued function defined on R. Under this situation, suppose that X_n converges in probability to a given point z_0 of R, as n tends to infinity. Then, if $t(z)$ is continuous at z_0, and if $\mathscr{E}[t^2(X_n)]<K$ (independent of n), we have*

$$\mathscr{E}[t(X_n)] \to t(z_0), \quad (n \to \infty). \tag{2.3.2}$$

The proof of this lemma is easy and is omitted.

Now, we shall consider two subsets of the above system, $\{X_1, \cdots, X_k\}$ and $\{X_{k+1}, \cdots, X_{r-1}\}$, where k is a positive integer such that $1 \leqq k \leqq r-2$.

The limiting process of related parameters, under which our discussions are made, is:

(2.3.3) $$n\to\infty,\ n(p_1+\cdots+p_k)\to K(<\infty),\ \text{and}\ p_r\to p_r^*(>0)\,,$$

where k and r may vary depending on this limiting process.

First, we show the following

THEOREM 2.3.1. *If* $k\geqq 2$, *then the set* $\{X_1,\cdots,X_k\}$ *is asymptotically independent* (*I*) *under the limiting process* (2.3.3).

PROOF. Let $P(n_1,\cdots,n_k)$ and $P(n_i)$ be the probabilities of the distributions of $X_{(k)}=(X_1,\cdots,X_k)$ and X_i, respectively, i.e.,

$$P(n_1,\cdots,n_k)=\frac{n!}{n_1!\cdots n_k!(n-n_1-\cdots-n_k)!}p_1^{n_1}\cdots p_k^{n_k}(1-p_1-\cdots-p_k)^{n-n_1-\cdots-n_k},$$

$$(0\leqq n_i,\ \sum_{i=1}^{k}n_i\leqq n)\,,$$

and

$$P(n_i)=\frac{n!}{n_i!(n-n_i)!}p_i^{n_i}(1-p_i)^{n-n_i}\,,\ (0\leqq n_i\leqq n)\,.$$

To prove the theorem, it is sufficient to show that

$$V_{\text{indep}}(X_1,\cdots,X_k)=\sum_{\substack{0\leq n_i\leq n\\ i=1,\cdots,k}}|P(n_1,\cdots,n_k)-P(n_1)\cdots P(n_k)|\to 0\,,$$

under (2.3.3). By theorem 1.4.2, it holds that

(2.3.4) $$V_{\text{indep}}(X_1,\cdots,X_k)\leqq 2\sqrt{I_{\text{indep}}(X_1,\cdots,X_k)}\,,$$

where

$$I_{\text{indep}}(X_1,\cdots,X_k)=\mathscr{E}\left[\log\frac{P(X_1,\cdots,X_k)}{P(X_1)\cdots P(X_k)}\right],$$

$\mathscr{E}$ being the expectation sign with respect to the distribution of $X_{(k)}$. Then, we can get

$$I_{\text{indep}}(X_1,\cdots,X_k)\leqq n(p_1+\cdots+p_k)^2,$$

from which the theorem follows.

It is noted, concerning this theorem, that if $np_i\to\alpha_i$, $i=1,2,\cdots,k$, under the limiting process (2.3.3), then the limiting distributions of X_i's are Poisson distributions with mean α_i's respectively.

In the second place, we shall prove the following

THEOREM 2.3.2. *Let $X_{(k)}=(X_1, \cdots, X_k)$ and $X_{(r-k-1)}=(X_{k+1}, \cdots, X_{r-1})$. Then, $X_{(k)}$ and $X_{(r-k-1)}$ are asymptotically independent (I) under the limiting process* (2.3.3).

PROOF. Put

$$\bar{X}_k=\sum_{i=1}^{k} X_i, \quad \bar{X}_{r-k-1}=\sum_{i=k+1}^{r-1} X_i ,$$

and

$$\bar{p}_k=\sum_{i=1}^{k} p_i, \quad \bar{p}_{r-k-1}=\sum_{i=k+1}^{r-1} p_i .$$

The Kullback-Leibler mean information for proving the independence under consideration,

$$I_{\text{indep}}(X_{(k)}, X_{(r-k-1)})=\mathscr{E}\left[\log \frac{P(X_1, \cdots, X_{r-1})}{P(X_1, \cdots, X_k)P(X_{k+1}, \cdots, X_{r-1})}\right],$$

is identical with

$$I_{\text{indep}}(\bar{X}_k, \bar{X}_{r-k-1})=\mathscr{E}\left[\log \frac{P(\bar{X}_k, \bar{X}_{r-k-1})}{P(\bar{X}_k)P(\bar{X}_{r-k-1})}\right], \tag{2.3.5}$$

where $\mathscr{E}$ is taken with respect to the joint distribution of $\bar{X}_k$ and $\bar{X}_{r-k-1}$. We shall calculate the right-hand member of (2.3.5). From the distributions of $\bar{X}_k$, $\bar{X}_{r-k-1}$ and $(\bar{X}_k, \bar{X}_{r-k-1})$, it follows that

$$\begin{aligned} I_{\text{indep}}(\bar{X}_k, \bar{X}_{r-k-1})=&\ \mathscr{E}\left[\log \frac{(n-\bar{X}_k)!\,(n-\bar{X}_{r-k-1})!}{n!\,(n-\bar{X}_k-\bar{X}_{r-k-1})!}\right] \\ &+n(1-\bar{p}_k-\bar{p}_{r-k-1})\log(1-\bar{p}_k-\bar{p}_{r-k-1}) \\ &-n(1-\bar{p}_k)\log(1-\bar{p}_k)-n(1-\bar{p}_{r-k-1})\log(1-\bar{p}_{r-k-1}) . \end{aligned} \tag{2.3.6}$$

The first member of the right-hand side of this equality satisfies the inequality

$$\begin{aligned} I_1 &= \mathscr{E}\left[\log \frac{(n-\bar{X}_k)!\,(n-\bar{X}_{r-k-1})!}{n!\,(n-\bar{X}_k-\bar{X}_{r-k-1})!}\right] \leqq \mathscr{E}\left[\bar{X}_k \log\left(1-\frac{\bar{X}_{r-k-1}}{n}\right)\right] \\ &= n\frac{\bar{p}_k}{1-\bar{p}_{r-k-1}}\,\mathscr{E}'\left[\left(1-\frac{\bar{X}_{r-k-1}}{n}\right)\log\left(1-\frac{\bar{X}_{r-k-1}}{n}\right)\right], \end{aligned}$$

where the expectation sign $\mathscr{E}'$ is taken with respect to $\bar{X}_{r-k-1}$. Since $\bar{X}_{r-k-1}/n$ converges to $1-p_r^*$ in probability under (2.3.3), by applying lemma 2.3.1 we obtain

$$I_1 \leqq n\bar{p}_k \log p_r^*$$

for sufficiently large n. Thus we can get

$$I_{\text{indep}}(\bar{X}_k, \bar{X}_{r-k-1}) \leqq n\bar{p}_k \log \frac{p_r^*}{1-\bar{p}_k-\bar{p}_{r-k-1}}$$

for large n, from which the theorem follows.

Acknowledgements

This research originates in the previous studies [8,9] which were suggested by Professors J. Ogawa and S. Yamamoto. The author expresses his gratitude to them for their helpful suggestions and kind encouragements. His gratitude is also due to the referees of Ann. Inst. Statist. Math. for their helpful comments, by which this paper is revised.

DEPT. OF MATH., COLLEGE OF SCI. AND ENG., NIHON UNIVERSITY

REFERENCES

[1] H. Cramér, *Mathematical Methods of Statistics*, Princeton Univ. Press, 1946.

[2] E. J. Gumbel, "On the independence of the extremes in a sample," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 78-81.

[3] J. Geffroy, "Contribution à la théorie des valeurs extrêmes," *Publ. Inst. Univ. Paris*, Vol. 7 (1958), pp. 37-121; Vol. 8 (1959), pp. 123-184.

[4] P. R. Halmos, *Measure Theory*, Van Nost. Comp. Inc., 1950.

[5] J. Hannan, "Consistency of maximum likelihood estimation of discrete distributions," *Contributions to Probability and Statistics*, Essays in Honor of H. Hotelling, Stanford Univ. Press, 1960, pp. 249-257.

[6] W. Hoeffding and J. Wolfowitz, "Distinguishability of sets of distributions," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 700-718.

[7] S. Ikeda, "On an equality by Bouman, Velden and Yamamoto relating to the threshold number of quanta in human vision," *Proc. Inst. Stat. Math.*, Vol. 9, No. 1 (1961), pp. 37-46, (in Japanese).

[8] S. Ikeda, "On an asymptotic independence of time-intervals between the successive occurences of absorptions and their approximate distribution," (a complementary remark to [9]), *Proc. Inst. Stat. Math.*, Vol. 9, No. 2 (1962), pp. 113-126, (in Japanese).

[9] S. Ikeda, "Necessary conditions for the convergence of Kullback-Leibler's mean information," *Ann. Inst. Stat. Math.*, Vol. 14 (1962), pp. 107-118.

[10] H. Kudo, "A theorem of Kakutani on infinite product measures," *Nat. Sci. Rep. Ochanomizu Univ.*, Vol. 3 (1952), pp. 10-22.

[11] H. Kudo, "Dependent experiments and sufficient statistics," *Nat. Sci. Rep. Ochanomizu Univ.*, Vol. 4 (1953), pp. 151-163.

[12] P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, 1937.

[13] M. Loéve, *Probability Theory*, Van Nost. Comp. Inc., 1955.

[14] K. Matusita, "Decision rules, based on the distance, for problems of fit, two samples, and estimation," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 631-640.

[15] H. Robbins, "Convergence of distributions," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 72–76.
[16] H. Scheffé, "A useful convergence theorem for probability distributions," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 434–438.
[17] S. S. Wilks, "Order statistics," *Bull. Amer. Math. Soc.*, Vol. 54 (1947), pp. 6–50.