ON ASYMPTOTIC INDEPENDENCE OF ORDER STATISTICS

SADAO IKEDA AND TADASHI MATSUNAWA

(Received Feb. 20, 1970)

Summary and Introduction

Type $(\boldsymbol{B})_d$ asymptotic independence property of order statistics as the size of sample increases is investigated in both non-censored and censored cases. The results [1], [2], [3], [4], [5] hitherto obtained in non-censored cases are improved. So far as the authors are awared, the property has not been investigated in literature for censored samples.

Let $X_1 < X_2 < \cdots < X_N$ be order statistics based on a random sample of size N drawn from a univariate continuous distribution with cdf. $F(x)$ and pdf. $f(x)$. The results [1], [2], [3] have been improved by Homma [4] in the form that X_n and X_{N-m+1} are asymptotically independent in somewhat weaker sense than type $(\boldsymbol{B})_d$ as N increases indefinitely, provided that $n/N \to \mu_1$ (<1) and $m/N \to 0$, or that $n/N \to 0$ and $m/N \to \mu_2$ (<1). On the other hand, Ikeda [5] showed that $X_{(n)} = (X_1, \cdots, X_n)$ and $Y_{(m)} = (X_{N-m+1}, \cdots, X_N)$ are mutually asymptotically independent $(\boldsymbol{B})_d$ as $N \to \infty$, if $n+m = o(\sqrt{N})$.

In the present article, it is shown that under the same conditions as in [4] $X_{(n)}$ and $Y_{(m)}$ are asymptotically independent $(\boldsymbol{B})_d$ as $N \to \infty$, and as a consequence it is also shown that the set $\{X_{(n)}, Z_{k(h)}, Y_{(m)}\}$ is asymptotically independent $(\boldsymbol{B})_d$ as $N \to \infty$, if $n/N \to 0$, $m/N \to 0$, $k/N \to \lambda$ and $h/N \to \mu$ $(0 < \lambda,\ \mu < 1;\ \lambda + \mu < 1)$, where $Z_{k(h)} = (X_k, X_{k+1}, \cdots, X_{k+h-1})$. Similar results are obtained in singly and doubly censored cases, too.

In Section 1, some preliminary results are stated, which are useful in the subsequent sections. In Sections 2 and 3, two cases of type I censoring are treated. In Section 2, a fundamental result is shown on asymptotic independence $(\boldsymbol{B})_d$ of a set of lower and upper extremes in a doubly censored case from the top and the bottom; results in singly and non-censored cases are obtained as limiting cases. In Section 3, we consider the case of a singly censored sample of type I where the middle part of the sample is censored. Section 4 is devoted to discussion of the asymptotic independence $(\boldsymbol{B})_d$ property of censored sample of type II.

1. Preliminaries

In this section, we list up some well-known results on integral calculus and infinite series in Lemma 1.1, and results on asymptotic independence $(B)_d$ in Lemmas 1.2 and 1.3, which are useful in later sections.

LEMMA 1.1. (i) *For positive integers p and q,*

$$\int_0^1 z^{p-1}(1-z)^{q-1}\log z dz = -\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\sum_{i=1}^{q}\frac{1}{p-1+i}. \tag{1.1}$$

(ii) *For integer* $p \geqq 2$,

$$\sum_{i=1}^{p}\frac{1}{i} = C + \log p + \frac{1}{2p} - \sum_{i=1}^{\infty}\frac{a_{i+1}}{p(p+1)\cdots(p+i)}, \tag{1.2}$$

where C denotes the Euler constant and a_{i+1} *are defined by*

$$a_r = \frac{1}{r}\int_0^1 z(1-z)(2-z)\cdots(r-1-z)dz, \quad (r \geqq 2). \tag{1.3}$$

(iii) *For* a_r *defined above,*

$$\Gamma(r-1)/6r \leqq a_r \leqq \Gamma(r)/6r, \quad (r \geqq 2). \tag{1.4}$$

(iv) *For positive integer* $p \geqq 2$,

$$\sum_{i=0}^{\infty}\frac{\Gamma(i+1)}{\Gamma(p+1+i)} = \frac{1}{(p-1)\Gamma(p)}. \tag{1.5}$$

Using the formula (1.2), we get for positive integers $p \geqq 2$ and q

$$p\left(\sum_{i=1}^{q}\frac{1}{p+i} + \log\frac{p}{p+q}\right) = -\frac{q}{2(p+q)} - \frac{p}{p+q}T(p+q) + T(p), \tag{1.6}$$

where $T(p)$ is defined by

$$T(p) = \sum_{i=1}^{\infty}\frac{a_{i+1}}{(p+1)\cdots(p+i)}, \quad (p \geqq 2). \tag{1.7}$$

Since it holds by (1.4) and (1.5) that

$$\begin{aligned} 0 < T(p) &\leqq \frac{1}{6}\sum_{i=1}^{\infty}\frac{\Gamma(i+1)}{(p+1)\cdots(p+i)(i+1)} \\ &\leqq \frac{\Gamma(p+1)}{6}\sum_{i=1}^{\infty}\frac{\Gamma(i+1)}{\Gamma(p+1+i)} = \frac{p}{6(p-1)}, \end{aligned}$$

the series $T(p)$ is absolutely convergent for any integer $p \geqq 2$. More-over, the sequence $\{T(p);\ p=2, 3, \cdots\}$ is monotone decreasing, and hence, is convergent.

Now, some results on asymptotic independence $(\boldsymbol{B})_d$ of a set of random variables are stated. As for the notions of asymptotic equivalence $(\boldsymbol{B})_d$ and asymptotic independence $(\boldsymbol{B})_d$, reference should be made to Ikeda [6].

For each positive integer s, let $(X^s_{(n_1)}, Y^s_{(n_2)}, Z^s_{(n_3)})$ be a real random variable defined over the $(n_1+n_2+n_3)$-dimensional Euclidean space, $R_{(n_1+n_2+n_3)}=R_{(n_1)}\times R_{(n_2)}\times R_{(n_3)}$, where n_1, n_2 and n_3 may vary depending on s. Then, it is shown [5] that the asymptotic independence $(\boldsymbol{B})_d$ of the set $\{X^s_{(n_1)}, Y^s_{(n_2)}, Z^s_{(n_3)}\}$ as $s\to\infty$ implies the same type of asymptotic independence of the set $\{\bar{X}^s_{(m_1)}, \bar{Y}^s_{(m_2)}, \bar{Z}^s_{(m_3)}\}$, where the variables in the latter set are marginal random variables of those of the former set, respectively, and m_1, m_2 and m_3 may depend on s. Furthermore, it is not so difficult to show that, if $\{X^s_{(n_1)}, (Y^s_{(n_2)}, Z^s_{(n_3)})\}$ and $\{Y^s_{(n_2)}, Z^s_{(n_3)}\}$ are both asymptotically independent $(\boldsymbol{B})_d$, then the set $\{X^s_{(n_1)}, Y^s_{(n_2)}, Z^s_{(n_3)}\}$ is asymptotically independent $(\boldsymbol{B})_d$ as $s\to\infty$.

We shall consider a slightly different situation. Let $I_{(n)}$ be a countable set of discrete points of $R_{(n)}$, and let $W^s_{(n)}$ be a random variable defined over $I_{(n)}$, the σ-field of subsets of $I_{(n)}$ being defined to be the class of all subsets of $I_{(n)}$, where n may be assumed to be dependent on s. Given $W^s_{(n)}=w_{(n)}$, a conditional random variable, $(X^{s,w}_{(n_1)}, Y^{s,w}_{(n_2)}, Z^{s,w}_{(n_3)})$, is distributed over the Euclidean space $R_{(n_1+n_2+n_3)}$, where n_1, n_2 and n_3 may be dependent on s, but not on $w_{(n)}$. Let us designate the joint variable of these random variables by $(W^s_{(n)}:(X^{s,W}_{(n_1)}, Y^{s,W}_{(n_2)}, Z^{s,W}_{(n_3)}))$, which we denote, for the notational simplicity, by $(W^s:(X^s_W, Y^s_W, Z^s_W))$.

Under this set-up, the asymptotic independence $(\boldsymbol{B})_d$ of the set $\{X^s_W, Y^s_W, Z^s_W\}$, or more precisely, $\{(W^s:X^s_W), (W^s:Y^s_W), (W^s:Z^s_W)\}$, should be understood to mean the validity of the condition

$$\sup_{C,E}\left|\sum_{w\in C} \mathrm{P}(W^s=w)(P^s_w(E)-Q^s_w(E))\right|\to 0\ , \qquad (s\to\infty)\ , \tag{1.8}$$

where the supremum is taken over all subsets, C, of $I_{(n)}$ and all subsets, E, of $R_{(n_1+n_2+n_3)}$, belonging to the usual Borel field, and P^s_w and Q^s_w designate probability measures corresponding to the conditional variables (X^s_w, Y^s_w, Z^s_w) and $(X^s_w)(Y^s_w)(Z^s_w)$, respectively, the latter being an $(n_1+n_2+n_3)$-dimensional real random variable, say $(\tilde{X}^s, \tilde{Y}^s, \tilde{Z}^s)$, such that $\{\tilde{X}^s, \tilde{Y}^s, \tilde{Z}^s\}$ form an independent set and the marginals are identically distributed with those of (X^s_w, Y^s_w, Z^s_w), respectively.

First, we state the following

LEMMA 1.2. (i) *If $\{X^s_W, Y^s_W, Z^s_W\}$ is asymptotically independent $(\boldsymbol{B})_d$ as $s\to\infty$, so is $\{\bar{X}^s_W, \bar{Y}^s_W, \bar{Z}^s_W\}$ too, where the latter is a set of mar-*

ginals of the former in the sense that the conditional random variables $\bar{X}^s_w$, $\bar{Y}^s_w$ *and* $\bar{Z}^s_w$, *given* $W^s=w$, *are marginal variables, respectively, of* X^s_w, Y^s_w *and* Z^s_w.

(ii) *If both of the sets,* $\{X^s_W, (Y^s_W, Z^s_W)\}$ *and* $\{Y^s_W, Z^s_W\}$, *are asymptotically independent* $(\boldsymbol{B})_d$ *as* $s\to\infty$, *so is the set* $\{X^s_W, Y^s_W, Z^s_W\}$ *too.*

PROOF. The first result (i) is straightforward.

We prove (ii). It is easily seen, in the first place, that the condition (1.8) is equivalent to the condition

$$\sum_{w\in I_{(n)}} \mathrm{P}(W^s=w)\delta_d(P^s_w, Q^s_w: \boldsymbol{B})\to 0\ , \qquad (s\to\infty)\ , \tag{1.9}$$

where

$$\delta_d(P^s_w, Q^s_w: \boldsymbol{B})=\sup_{E\in \boldsymbol{B}} |P^s_w(E)-Q^s_w(E)|\ ,$$

$\boldsymbol{B}$ being the Borel field of subsets of $R_{(n_1+n_2+n_3)}$. In fact, the left-hand members of (1.8) and (1.9) are equal to each other.

As was shown in [6], it holds, in general, that

$$\delta_d(P, Q: \boldsymbol{B})=\delta_d(P, Q: \boldsymbol{A})$$

for any probability measures P and Q, where $\boldsymbol{A}$ is the class of all finite disjoint unions of $(n_1+n_2+n_3)$-dimensional rectangles, left-closed and right-opened. Since any set A belonging to $\boldsymbol{A}$ can be expressed in the form $A=\sum_{i=1}^{N}(A_{1i}\times A_{2i})$, where A_{1i}, $i=1,\cdots,N$, are mutually disjoint subsets belonging to $\boldsymbol{A}_1=\boldsymbol{A}_{(n_1)}$ and $A_{2i}\in\boldsymbol{A}_2=\boldsymbol{A}_{(n_2+n_3)}$, $i=1,\cdots,N$, we have

$$\begin{aligned}&|P^{(X^s_w)(Y^s_w, Z^s_w)}(A)-P^{(X^s_w)(Y^s_w)(Z^s_w)}(A)|\\&\leqq\sum_{i=1}^{N}P^{X^s_w}(A_{1i})\,|P^{(Y^s_w, Z^s_w)}(A_{2i})-P^{(Y^s_w)(Z^s_w)}(A_{2i})|\ ,\end{aligned}$$

from which it follows that

$$\delta_d(P^{(X^s_w)(Y^s_w, Z^s_w)}, P^{(X^s_w)(Y^s_w)(Z^s_w)}: \boldsymbol{A})=\delta_d(P^{(Y^s_w, Z^s_w)}, P^{(Y^s_w)(Z^s_w)}: \boldsymbol{A}_2)\ .$$

Now, since

$$\begin{aligned}&\sum_{w\in I_{(n)}}\mathrm{P}(W^s=w)\delta_d(P^s_w, Q^s_w: \boldsymbol{B})\\&\leqq\sum_{w\in I_{(n)}}\mathrm{P}(W^s=w)\delta_d(P^s_w, P^{(X^s_w, Y^s_w, Z^s_w)}: \boldsymbol{B})\\&\quad+\sum_{w\in I_{(n)}}\mathrm{P}(W^s=w)\delta_d(P^{(Y^s_w, Z^s_w)}, P^{(Y^s_w)(Z^s_w)}: \boldsymbol{A}_2)\ ,\end{aligned}$$

and the conditions of the lemma imply that the two members of the right-hand side of this inequality tend to zero as $s\to\infty$, we have (1.9), which completes the proof of the lemma.

Suppose in the next place that (X_w^s, Y_w^s, Z_w^s) is absolutely continuous with respect to the Euclid-Lebesgue measure, μ, over $R=R_{(n_1+n_2+n_3)}$, for any given s and w. Let us denote by $f_w^s(x, y, z)$, $f_{w1}^s(x)$, $f_{w2}^s(y)$, and $f_{w3}^s(z)$ the pdf.'s of (X_w^s, Y_w^s, Z_w^s), X_w^s, Y_w^s and Z_w^s, respectively. Then, the Kullback-Leibler mean information is defined by

$$I(P_w^s : Q_w^s)=\int_R f_w^s(x, y, z) \log \frac{f_w^s(x, y, z)}{f_{w1}^s(x) f_{w2}^s(y) f_{w3}^s(z)} d\mu \ , \tag{1.10}$$

for which it holds [5] that

$$\delta_d(P_w^s, Q_w^s : \boldsymbol{B})^2 \leqq I(P_w^s : Q_w^s) \ .$$

Hence, by applying the Cauchy-Schwarz inequality to the left-hand member of (1.9), we obtain

$$\sum_{w \in I_{(n)}} \mathrm{P}(W^s=w)\delta_d(P_w^s, Q_w^s : \boldsymbol{B}) \leqq \left[\sum_{w \in I_{(n)}} \mathrm{P}(W^s=w) I(P_w^s : Q_w^s)\right]^{1/2} .$$

Thus we can state the following lemma, which gives us a criterion for the asymptotic independence $(\boldsymbol{B})_d$ of $\{X_W^s, Y_W^s, Z_W^s\}$ as $s \to \infty$.

LEMMA 1.3. *The set* $\{X_W^s, Y_W^s, Z_W^s\}$ *is asymptotically independent* $(\boldsymbol{B})_d$ *as* $s \to \infty$, *if it holds that*

$$\sum_{w \in I_{(n)}} \mathrm{P}(W^s=w) I(P_w^s : Q_w^s) \to 0 \ , \qquad (s \to \infty) \ , \tag{1.11}$$

2. The case of type I censoring (1)

We consider two cases of doubly censoring of type I. Let $X_1 < X_2 < \cdots < X_N$ be order statistics based on a random sample of size N drawn from a univariate distribution whose cdf. and pdf. are given by $F(x)$ and $f(x)$. Let α and β be preassigned extended real numbers such that $F(\alpha) < F(\beta)$.

Case 1. The variables X_i's such that $X_i < \alpha$ or $\beta \leqq X_i$ are censored. Thus, if $F(\alpha)=0$ and $F(\beta)<1$, we have the case of singly censoring from the top; if $0<F(\alpha)$ and $F(\beta)=1$, we have the case from the bottom; and if $F(\alpha)=0$ and $F(\beta)=1$, it turns out to be the case of non-censoring.

Case 2. The variables X_i's such that $\alpha \leqq X_i < \beta$ are censored.

According to Sarhan-Greenberg [8], these types of censoring are referred to as 'type I censoring'.

In the present section, we consider the case 1.

Suppose, in case 1, that $X_S < X_{S+1} < \cdots < X_{S+L-1}$ are observed, where the initial number, S, and the number of variables observed, L, are regarded as random variables, whose probabilities are given by

$$p_N(s, l)=\frac{N!}{(s-1)!l!(N-s-l+1)!}\xi^{s-1}\eta^l\zeta^{N-s-l+1}, \quad (1\leqq s\leqq N;\ 0\leqq l\leqq N-s+1). \tag{2.1}$$

Here, we have put $\xi=F(\alpha)$, $\eta=F(\beta)-F(\alpha)$ and $\zeta=1-F(\beta)$. Under the condition $(S, L)=(s, l)$, $X_s<X_{s+1}<\cdots<X_{s+l-1}$ are order statistics based on a random sample of size l drawn from a doubly truncated distribution, whose pdf. and cdf. are given by

$$g(x)=\begin{cases} f(x)/\eta, & \text{if } \alpha\leqq x<\beta, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$G(x)=\begin{cases} 0, & \text{if } x<\alpha, \\ (F(x)-F(\alpha))/\eta, & \text{if } \alpha\leqq x<\beta, \\ 1, & \text{if } \beta\leqq x. \end{cases}$$

We are now interested in the asymptotic independence property of the set of size 2

$$\{X_{(n)}^{S,L}=(X_S, \cdots, X_{S+n-1}),\ Y_{(m)}^{S,L}=(X_{S+L-m}, \cdots, X_{S+L-1})\}, \tag{2.2}$$

where n and m are positive integers which may depend only on N and not on the realizations of (S, L). When $L<n+m$, the above set is considered to be a set of size 1, i.e., the joint distribution of the whole sample, and we shall make a convention that a set of size 1 is always asymptotically independent $(\boldsymbol{B})_d$.

As was stated in the preceding section, asymptotic independence $(\boldsymbol{B})_d$ of the set (2.2) is defined to be

$$\sup_{\substack{K\in I^2\\ E\in \boldsymbol{B}}}\left|\sum_{(s,l)\in K} p_N(s, l)P^{U^{s,l}}(E)-\sum_{(s,l)\in K} p_N(s, l)P^{V^{s,l}}(E)\right|\to 0, \tag{2.3}$$

as $N\to\infty$, where $I^2=\{(i, j): i\geqq 0, j\geqq 0\}$, $\boldsymbol{B}$ denotes the usual Borel field of subsets of the $(n+m)$-dimensional Euclidean space, and $U^{s,l}=(X_{(n)}^{s,l}, Y_{(m)}^{s,l})$ and $V^{s,l}=(X_{(n)}^{s,l})(Y_{(m)}^{s,l})$. Then, by Lemma 1.3, it is sufficient for (2.3) to hold that

$$\sum_{\substack{s,l\\ l\geqq n+m}} p_N(s, l)I(U^{s,l}: V^{s,l})\to 0, \quad (N\to\infty). \tag{2.4}$$

Now, we shall prove the following

THEOREM 2.1. *The set* (2.2) *is asymptotically independent* $(\boldsymbol{B})_d$, *if*

$n/N \to \mu$ and $m/N \to 0$, or if $n/N \to 0$ and $m/N \to \mu$, as $N \to \infty$, where μ is any given number such that $0 \leqq \mu < \eta$.

PROOF. Since the conditional variables, $G(X_s) < G(X_{s+1}) < \cdots < G(X_{s+l-1})$, given $(S, L) = (s, l)$, are order statistics from a uniform distribution over $(0, 1)$, it is easy to see that the Kullback-Leibler mean information in (2.4) becomes

$$I(U^{s,l} : V^{s,l}) = I_{1N}(l) + I_{2N}(l) , \qquad (l \geqq n+m) , \tag{2.5}$$

where

$$I_{1N}(l) = \log \frac{\Gamma(l-n+1)\Gamma(l-m+1)}{\Gamma(l+1)\Gamma(l-n-m+1)} \tag{2.6}$$

and

$$\begin{aligned} I_{2N}(l) = (l-n) \sum_{i=1}^{n} \frac{1}{l-n+i} + (l-m) \sum_{i=1}^{m} \frac{1}{l-m+i} \\ -(l-n-m) \sum_{i=1}^{n+m} \frac{1}{l-n-m+i} . \end{aligned} \tag{2.7}$$

Since these values are independent of s, the condition (2.4) is implied by the condition

$$\sum_{l \geqq n+m} p_N(l)(I_{1N}(l) + I_{2N}(l)) \to 0 , \qquad (N \to \infty) , \tag{2.8}$$

where $p_N(l)$ are the probabilities of L given by

$$p_N(l) = \binom{N}{l} \eta^l (1-\eta)^{N-l} , \qquad (0 \leqq l \leqq N) . \tag{2.9}$$

Let us put

$$W_{N,\varepsilon} = \{l : N(\eta - \varepsilon) \leqq l \leqq N(\eta + \varepsilon)\} ,$$

and let $W_{N,\varepsilon}^c$ be the complementary set of $W_{N,\varepsilon}$ with respect to the set of all non-negative integers, where ε is an arbitrarily fixed number such that $0 < \varepsilon < \eta - \mu$. Then, by [7], it holds that

$$P_L(W_{N,\varepsilon}^c) \leqq 2e^{-2N\varepsilon^2} , \tag{2.10}$$

for any given N.

Suppose N be large. Then, for each l belonging to $W_{N,\varepsilon}$, we have, by the results in Lemma 1.1 and the Stirling formula,

$$|(I_{1N}(l) + I_{2N}(l)) - (J_{1N}(l) + J_{2N}(l))| \leqq 1/3(l-n-m) , \tag{2.11}$$

where

$$J_{1N}(l) = (1/2)\{\log(1-n/l) + \log(1-m/l) - \log(1-(n+m)/l)\} \tag{2.12}$$

and

(2.13) $$J_{2N}(l)=T(l-n)+T(l-m)-T(l-n-m)-T(l)\ ,$$

$T(p)$ being the same as that defined by (1.7) in general. From (2.12) it is readily seen that

$$\sup_{l\in W_{N,\varepsilon}}|J_{1N}(l)|\leqq\sup_{l\in W_{N,\varepsilon}} nm/l(l-n-m)$$
$$\leqq nm/2(\eta-\varepsilon)(\eta-\varepsilon-(n+m)/N)N^2\ ,$$

from which we obtain

(2.14) $$\sup_{l\in W_{N,\varepsilon}}|J_{1N}(l)|\to 0\ ,\qquad (N\to\infty)\ .$$

Since $T(p)$ is monotone decreasing with increasing p, we have

$$\sup_{l\in W_{N,\varepsilon}}|J_{2N}(l)|\leqq 2\{T([N(\eta-\varepsilon)]-n-m)-T([N(\eta+\varepsilon)])\}\ ,$$

where [ ] designates the ordinary Gauss symbol. Hence, the convergence of the sequence $\{T(p);\ p=2,3,\cdots\}$ and the conditions of the theorem imply that

(2.15) $$\sup_{l\in W_{N,\varepsilon}}|J_{2N}(l)|\to 0\ ,\qquad (N\to\infty)\ .$$

It then follows from (2.11), (2.14) and (2.15) that

(2.16) $$\sum_{l\in W_{N,\varepsilon}} p_N(l)(I_{1N}(l)+I_{2N}(l))\to 0\ ,\qquad (N\to\infty)\ .$$

Now, for $l\in W^c_{N,\varepsilon}$, since $I_{1N}(l)\leqq 0$ and $|I_{2N}(l)|\leqq 3N^2$ by (2.7), we have

$$\sum_{l\in W^c_{N,\varepsilon}} p_N(l)(I_{1N}(l)+I_{2N}(l))\leqq 3N^2P_L(W^c_{N,\varepsilon})\ .$$

Hence, by (2.10) it holds that

(2.17) $$\sum_{l\in W^c_{N,\varepsilon}} p_N(l)(I_{1N}(l)+I_{2N}(l))\to 0\ ,\qquad (N\to\infty)\ .$$

Thus, from (2.16) and (2.17) we get (2.8). This completes the proof of the theorem.

By using Lemma 1.2, the following is an immediate consequence of the above theorem.

COROLLARY 2.1. *Let n, k, h and m be positive integers depending on N in such a way that $n<k<k+h-1$, $n/N\to 0$, $m/N\to 0$, $k/N\to\lambda$ and $h/N\to\mu$, where λ and μ are any given numbers such that $\xi<\lambda<1-\zeta$ and $0\leqq\mu<\eta+\xi-\lambda$. Then the set of size* 3

(2.18) $$\{X_{(n)}^{S,L}=(X_S,\cdots,X_{S+n-1}),\ Z_{k(h)}^{S,L}=(X_k,\cdots,X_{k+h-1}),\ Y_{(m)}^{S,L}=(X_{S+L-m},\cdots,X_{S+L-1})\}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$.

The following results in singly censored cases of type I are also direct consequences of the above theorem.

COROLLARY 2.2. (i) *Suppose that the sample are censored from the top by* β. *Then, the set*

(2.19) $$\{X_{(n)}^{L}=(X_1,\cdots,X_n),\ Y_{(m)}^{L}=(X_{L-m+1},\cdots,X_L)\}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *provided that* $n/N\to\mu$ *and* $m/N\to 0$, *or that* $n/N\to 0$ *and* $m/N\to\mu$, *where* μ *is any given number such that* $0\leqq\mu<F(\beta)$.

(ii) *Under the same situation as in* (i), *the set*

(2.20) $$\{X_{(n)}^{L}=(X_1,\cdots,X_n),\ Z_{k(h)}^{L}=(X_k,\cdots,X_{k+h-1}),\ Y_{(m)}^{L}=(X_{L-m+1},\cdots,X_L)\}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *if* $n/N\to 0$, $m/N\to 0$, $k/N\to\lambda$ *and* $h/N\to\mu$ *for any fixed number* λ *and* μ *such that* $0<\lambda<F(\beta)$ *and* $0\leqq\mu<F(\beta)-\lambda$.

(iii) *Suppose the sample be censored from the bottom by* α. *Then, the set*

(2.21) $$\{X_{(n)}^{S}=(X_S,\cdots,X_{S+n-1}),\ Y_{(m)}^{S}=(X_{N-m+1},\cdots,X_N)\}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *if* $n/N\to\mu$ *and* $m/N\to 0$, *or if* $n/N\to 0$ *and* $m/N\to\mu$, *where* μ *is any given number such that* $0\leqq\mu<1-F(\alpha)$.

(iv) *In the same situation as in* (iii), *the set*

(2.22) $$\{X_{(n)}^{S}=(X_S,\cdots,X_{S+n-1}),\ Z_{k(h)}^{S}=(X_k,\cdots,X_{k+h-1}),\ Y_{(m)}^{S}=(X_{N-m+1},\cdots,X_N)\}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *if* $n/N\to 0$, $m/N\to 0$, $k/N\to\lambda$ *and* $h/N\to\mu$ *for some* λ *and* μ *such that* $\xi<\lambda<1$ *and* $0\leqq\mu<1-\lambda$.

The following theorem gives the results in non-censored cases, and is an immediate consequence of Theorem 2.1.

THEOREM 2.2. (i) *Let* $X_1<X_2<\cdots<X_N$ *be order statistics based on a radom sample of size* N *from a univariate continuous distribution. Then, the set*

(2.23) $$\{X_{(n)}=(X_1,\cdots,X_n),\ Y_{(m)}=(X_{N-m+1},\cdots,X_N)\}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *if* $n/N\to\mu$ *and* $m/N\to 0$, *or if* $n/N\to 0$ *and* $m/N\to\mu$, *where* μ *is any given number such that* $0\leqq\mu<1$.

(ii) *In the same situation as in* (i), *the set*

$$\{X_{(n)}=(X_1,\cdots,X_n),\ Z_{k(h)}=(X_k,\cdots,X_{k+h-1}),\ Y_{(m)}=(X_{N-m+1},\cdots,X_n)\} \tag{2.24}$$

is asymptotically independent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *if* $n/N\to 0$, $m/N\to 0$, $k/N\to\lambda$ *and* $h/N\to\mu$, *where* λ *and* μ *are any given number such that* $0<\lambda<1$ *and* $0\leqq\mu<1-\lambda$.

The result (i) of this theorem implies all the results hitherto obtained [1], [2], [3], [4], [5]. For (i), evaluation of approximation error is given by

$$\delta_d((X_{(n)},Y_{(m)}),(X_{(n)})(Y_{(m)}))\leqq[I((X_{(n)},Y_{(m)}):(X_{(n)})(Y_{(m)}))]^{1/2}. \tag{2.25}$$

Some of the values of the right-hand member of this inequality are tabulated below.

Table 1. Error evaluation in non-censored cases ($\mu=n/N$)

N	m \ μ	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80
10	1	0.07583							
50	1	0.03344	0.05017	0.06572	0.08201				
	2	0.04782	0.07179	0.09411					
	3	0.05922	0.08898						
	5	0.07822							
100	1	0.02362	0.03536	0.04635	0.05784	0.07086	0.08680		
	2	0.03359	0.05038	0.06600	0.08242				
	4	0.04800	0.07210	0.09455					
	7	0.06456	0.09712						
	10	0.07850							
150	1	0.01926	0.02887	0.03784	0.04725	0.05777	0.07078	0.08839	
	2	0.02733	0.04100	0.05369	0.06704	0.08214			
	4	0.03895	0.05847	0.07664	0.05971				
	6	0.04807	0.07228	0.09467					
	11	0.06638	0.09976						
	15	0.07873							
200	1	0.01654	0.02482	0.03261	0.04097	0.05000	0.06117	0.07634	
	2	0.02355	0.03542	0.04645	0.05810	0.07106	0.08711		
	3	0.02901	0.04357	0.05713	0.07140	0.08745			
	5	0.03769	0.05665	0.07427	0.09281				
	8	0.04811	0.07232	0.09483					
	14	0.06474	0.09746						
	20	0.07873							
300	1	0.01353	0.02041	0.02661	0.03345	0.04088	0.04993	0.06248	0.08162
	2	0.01932	0.02877	0.03785	0.04755	0.05787	0.07093	0.08867	
	3	0.02374	0.03547	0.04674	0.05840	0.07125	0.08724		
	5	0.03071	0.04598	0.06045	0.07560	0.09248			
	8	0.03929	0.05867	0.07696	0.09621				
	13	0.05055	0.07564	0.09921					
	21	0.06507	0.09768						
	30	0.07905							

3. The case of type I censoring (2)

In this section, we consider the case 2 of type I censoring stated in the beginning part of the preceding section.

Let α and β be any preassigned real numbers such that $0<F(\alpha)<F(\beta)<1$, and the variables X_i such that $\alpha\leqq X_i<\beta$ be censored. The observed variables are then grouped into two parts,

$$O_1 : X_1<\cdots<X_L \ (<\alpha)\ , \qquad \text{and} \qquad O_2 : (\beta\leqq)\ X_{N-T+1}<\cdots<X_N\ ,$$

where L and T are random variables whose probabilities are given by

$$(3.1) \qquad p_N(l,t)=\frac{N!}{l!(N-l-t)!t!}\xi^l\eta^{N-l-t}\zeta^t\ , \qquad (0\leqq l,\ t,\ l+t\leqq N)\ ,$$

with $\xi=F(\alpha)$, $\eta=F(\beta)-F(\alpha)$ and $\zeta=1-F(\beta)$ as before.

Under the condition that $(L,T)=(l,t)$, two groups of conditional variables, O_1 with $L=l$ and O_2 with $T=t$, are stochastically independent of each other, and they can be regarded as order statistics based on random samples of size l and t drawn from truncated distributions, whose cdf.'s are given by

$$G_1(x)=\begin{cases} F(x)/\xi\ , & \text{if } x<\alpha\ ,\\ 1\ , & \text{otherwise}\ ,\end{cases}$$

and

$$G_2(x)=\begin{cases} (F(x)-F(\beta))/\zeta\ , & \text{if } \beta\leqq x\ ,\\ 0\ , & \text{otherwise}\ ,\end{cases}$$

respectively.

Now, let us consider a set of size 4 of extremes

$$(3.2) \qquad \left\{\begin{array}{l} X_{(n)}^{L,T}=(X_1,\cdots,X_n),\ Y_{(m)}^{L,T}=(X_{L-m+1},\cdots,X_L)\\ U_{(p)}^{L,T}=(X_{N-T+1},\cdots,X_{N-T+p}),\ V_{(q)}^{L,T}=(X_{N-q+1},\cdots,X_N)\end{array}\right\}.$$

As in the preceding section, type $(\boldsymbol{B})_d$ asymptotic independence of this set is guaranteed by the vanishing, as $N\to\infty$, of the quantity

$$(3.3) \qquad \sum_{\substack{l\geqq n+m\\ t\geqq p+q}} p_N(l,t)\{I((X_{(n)}^{l,t},Y_{(m)}^{l,t}):(X_{(n)}^{l,t})(Y_{(m)}^{l,t})) \\ +I((U_{(p)}^{l,t},V_{(q)}^{l,t}):(U_{(p)}^{l,t})(V_{(q)}^{l,t}))\}\ ,$$

where we have used the fact that the variables, $O_1(L=l)$ and $O_2(T=t)$, are mutually independent. Furthermore, for the vanishing of this quantity, it is sufficient that

$$(3.4) \qquad \sum_{l \geq n+m} p_N(l) I((X^l_{(n)}, Y^l_{(m)}) : (X^l_{(n)})(Y^l_{(m)})) \\ + \sum_{t \geq p+q} p_N(t) I((U^t_{(p)}, V^t_{(q)}) : (U^t_{(p)})(V^t_{(q)})) \to 0 , \qquad (N \to \infty) ,$$

where $p_N(l)$ and $p_N(t)$ are the probabilities of marginals, L and T, respectively, which are binomial $B(N, \xi)$ and $B(N, \zeta)$, and we have deleted the unnecessary suffices l and t from the conditional variables in (3.3).

Analogous calculation to that of the proof of Theorem 2.1 shows that the two members of (3.4) tend to zero as $N \to \infty$, if it holds that $n/N \to \mu_1$, $m/N \to 0$, $p/N \to \mu_2$ and $q/N \to 0$ for some μ_1 and μ_2.

Thus, we get the following

THEOREM 3.1. *Suppose that one of the following conditions is satisfied*: (a) $n/N \to \mu_1$, $m/N \to 0$, $p/N \to \mu_2$ *and* $q/N \to 0$, (b) $n/N \to 0$, $m/N \to \mu_1$, $p/N \to \mu_2$ *and* $q/N \to 0$, (c) $n/N \to \mu_1$, $m/N \to 0$, $p/N \to 0$ *and* $q/N \to \mu_2$, *and* (d) $n/N \to 0$, $m/N \to \mu_1$, $p/N \to 0$ *and* $q/N \to \mu_2$, *as* $N \to \infty$, *where* μ_1 *and* μ_2 *are any given number such that* $0 \leqq \mu_1 < \xi$ *and* $0 \leqq \mu_2 < \zeta$. *Then, the set* (3.2) *is asymptotically independent* $(\boldsymbol{B})_d$ *as* $N \to \infty$.

By Lemma 1.2, the following is an immediate consequence of the above theorem.

COROLLARY 3.1. *The set of size* 6,

$$(3.5) \qquad \{X^{L,T}_{(n)}, Z^{L,T}_{k(h)}, Y^{L,T}_{(m)}, U^{L,T}_{(p)}, W^{L,T}_{v(u)}, V^{L,T}_{(q)}\} ,$$

where $Z^{L,T}_{k(h)} = (X_k, \cdots, X_{k+h-1})$ *and* $W^{L,T}_{v(u)} = (X_{N-T+v}, \cdots, X_{N-T+v+u-1})$, *is asymptotically independent* $(\boldsymbol{B})_d$ *as* $N \to \infty$, *provided that* $n/N \to 0$, $k/N \to \lambda_1$, $h/N \to \mu_1$, $m/N \to 0$, $p/N \to 0$, $v/N \to \lambda_2$, $u/N \to \mu_2$ *and* $q/N \to 0$ *for any fixed* λ_1, λ_2, μ_1 *and* μ_2 *such that* $0 < \lambda_1 < \xi$, $0 \leqq \mu_1 < \xi - \lambda_1$, $1 - \xi - \eta < \lambda_2 < 1$, *and* $0 \leqq \mu_2 < 1 - \lambda_2$.

4. The case of type II censoring

Suppose that p-fraction (in number) of the whole N variables are censored from the top, or from the bottom. Such a censoring is called the type II censoring. It may also be the cases of type II censoring that the variables of p-fraction from the top and of q-fraction from the bottom $(p+q<1)$ are censored.

In the present section, we shall consider the asymptotic independence $(\boldsymbol{B})_d$ property of censored samples of type II, for which the fractions of censoring depend on N in general.

Let, as before, $X_1 < X_2 < \cdots < X_N$ be order statistics based on a random sample of size N drawn from a univariate continuous distribution, whose cdf. and pdf. are given by $F(x)$ and $f(x)$, respectively. The first P and the last Q variables are supposed to be censored, where P and Q may depend on N. Observed variables are thus $X_{P+1}<X_{P+2}<\cdots<X_{N-Q}$.

In this situation, we consider the set

$$\{X_{(n)}=(X_{P+1},\cdots,X_{P+n}),\ Y_{(m)}=(X_{N-Q-m+1},\cdots,X_{N-Q})\}, \tag{4.1}$$

where n and m may depend on N. The Kullback-Leibler mean information for criticizing the asymptotic independence $(\boldsymbol{B})_d$ of this set is seen to be

$$\begin{aligned} &I((X_{(n)},Y_{(m)}):(X_{(n)})(Y_{(m)})) \\ &=\log\frac{\Gamma(N-P-n+1)\Gamma(N-Q-m+1)}{\Gamma(N+1)\Gamma(N-P-Q-n-m+1)} \\ &\quad+(N-P-n)\sum_{i=1}^{P+n}\frac{1}{N-P-n+i}+(N-Q-m)\sum_{i=1}^{Q+m}\frac{1}{N-Q-m+i} \\ &\quad-(N-P-Q-n-m)\sum_{i=1}^{P+Q+n+m}\frac{1}{N-P-Q-n-m+i}. \end{aligned} \tag{4.2}$$

Through the above calculation, we notice that, so far as the Kullback-Leibler mean information is used as a criterion, the asymptotic independence $(\boldsymbol{B})_d$ properties of the original (non-censored) sample are preserved through type II censoring, or, in other words, type II censoring does not affect the property of the original sample, except for the numbers of variables under consideration. In fact, the quantity (4.2) tends to zero as $N\to\infty$, if $(P+n)/N\to\mu$ $(\geqq 0)$ and $(Q+m)/N\to 0$, or if $(P+n)/N\to 0$ and $(Q+m)/N\to\mu$ $(\geqq 0)$, and hence, these conditions imply the asymptotic independence $(\boldsymbol{B})_d$ of the set (4.1). This result is a restatement of Theorem 2.2 (i), in which n and m are replaced by $P+n$ and $Q+m$, respectively.

This consideration leads us to the following

THEOREM 4.1. (i) *The set* (4.1) *is asymptotically independent* $(\boldsymbol{B})_d$ *as* $N\to\infty$, *if* $(P+n)/N\to\mu$ *and* $(Q+m)/N\to 0$, *or if* $(P+n)/N\to 0$ *and* $(Q+m)/N\to\mu$, *where* μ *is any given number such that* $0\leqq\mu<1$.

(ii) *Suppose that* $(P+n)/N\to 0$, $k/N\to\lambda$, $h/N\to\mu$ *and* $(Q+m)/N\to 0$, *where* λ *and* μ *are any given numbers such that* $0<\lambda<1$ *and* $0\leqq\mu<1-\lambda$. *Then, the set*

$$\{X_{(n)},Z_{k(h)},Y_{(m)}\} \tag{4.3}$$

is asymptotically indendent $(\boldsymbol{B})_d$ *as* $N\to\infty$, *where* $Z_{k(h)}=(X_k,\cdots,X_{k+h-1})$.

In the case of type II censoring, where the variables X_i with $P<i\leqq N-Q$ are censored, we can state the following

THEOREM 4.2. *Suppose that* $P/N\to\rho_1$ *and* $Q/N\to\rho_2$ *as* $N\to\infty$, *where*

ρ_1 and ρ_2 are preassigned positive numbers such that $0<\rho_1, \rho_2, \rho_1+\rho_2<1$.

(i) *The set*

$$\{X_{(n)}=(X_1, \cdots, X_n), Y_{(m)}=(X_{N-m+1}, \cdots, X_N)\} \tag{4.4}$$

is asymptotically independent $(\boldsymbol{B})_d$ as $N\to\infty$, if $n/N\to\mu_1$ and $m/N\to 0$, or if $n/N\to 0$ and $m/N\to\mu_2$, where μ_1 and μ_2 are any given numbers such that $0\leqq\mu_1\leqq\rho_1$ and $0\leqq\mu_2\leqq\rho_2$.

(ii) *The set*

$$\{X_{(n)}, Z_{(h_1+h_2)}, Y_{(m)}\} \tag{4.5}$$

is asymptotically independent $(\boldsymbol{B})_d$ as $N\to\infty$, if $n/N\to 0$, $h_1/N\to\mu_1$, $h_2/N\to\mu_2$ and $m/N\to 0$, where $Z_{(h_1+h_2)}=(X_{P-h_1+1}, \cdots, X_P, X_{N-Q+1}, \cdots, X_{N-Q+h_2})$, and μ_1 and μ_2 are any given numbers such that $0\leqq\mu_1<\rho_1$ and $0\leqq\mu_2<\rho_2$.

Finally, we give some remarks.

Remark 1. All the results obtained in this paper are still valid if we replace such a condition as $p/N\to\mu$ $(\geqq 0)$ by $\limsup_{N\to\infty} p/N=\mu$. For instance, Theorem 2.2 (i) holds true if the condition $n/N\to\mu$, or $m/N\to\mu$, is replaced by the condition $\limsup_{N\to\infty} n/N=\mu$, or $\limsup_{N\to\infty} m/N=\mu$, respectively.

Remark 2. All results of this paper seem to be optimum in the sense that the conditions there can not be weakened any more. For example, in Theorem 2.2 (i), if $n/N\to\mu_1$ (>0) and $m/N\to\mu_2$ (>0), then the set (2.23) is not asymptotically independent $(\boldsymbol{B})_d$ as $N\to\infty$, which will be noticed by considering the fact that the joint distribution of $a_N X_n+b_N$ and $c_N X_m+d_N$ converges in law to a dependent normal distribution of two dimensions for some sequence of real numbers, a_N, b_N, c_N and d_N.

Remark 3. Since for any distributions X and Y it holds that

$$\delta_d(X, Y; \boldsymbol{B})\leqq(1-\rho^2(X, Y))^{1/2}\leqq I(X:Y)^{1/2},$$

it is more desirable to use the Matusita affinity $\rho(X, Y)$ for criticizing, and for estimating the error of, asymptotic independence $(\boldsymbol{B})_d$ of a set of variables. The values of Table 1 of Section 2 are expected to be reduced if we use, instead of (2.25), the inequality

$$\delta_d(U_{(n,m)}, V_{(n,m)}: \boldsymbol{B})\leqq[1-\rho^2(U_{(n,m)}, V_{(n,m)})]^{1/2},$$

where $U_{(n,m)}=(X_{(n)}, Y_{(m)})$ and $V_{(n,m)}=(X_{(n)})(Y_{(m)})$, although certain difficulties exist in calculating Matusita's affinity.

Remark 4. The investigations we have done in this paper would contribute to the asymptotic distribution theory of order statistics, and also to statistical inference by using the theory, which will be worked out in the forthcoming papers.

References

[1] Cramér, H. (1946). *Mathematical Methods of Statistics*, Princeton Univ. Press.
[2] Gumbel, E. J. (1946). On the independence of the extremes in a sample, *Ann. Math. Statist.*, **17**, 78-81.
[3] Gumbel, E. J. and Keeney, R. D. (1950). The extremal quotient, *Ann. Math. Statist.*, **21**, 523-538.
[4] Homma, T. (1951). On the asymptotic independence of order statistics, *Rep. Stat. Appl. Res. JUSE*, Tokyo, **1**, 1-8.
[5] Ikeda, S. (1963). Asymptotic equivalence of probability distributions with applications to some problems of asymptotic independence, *Ann. Inst. Statist. Math.*, **15**, 87-116.
[6] Ikeda, S. (1968). Asymptotic equivalence of real probability distributions, *Ann. Inst. Statist. Math.*, **20**, 339-362.
[7] Okamoto, M. (1958). Some inequalities relating to the partial sum of binomial probabilities, *Ann. Inst. Statist. Math.*, **10**, 29-35.
[8] Sarhan, A. E. and Greenberg, B. G. (ed.) (1962). *Contributions to Order Statistics*, John Wiley & Sons, Inc.