K-SAMPLE ANALOGUES OF RÉNYI'S KOLMOGOROV-SMIRNOV TYPE THEOREMS

BY MIKLÓS CSÖRGÖ[1]

Communicated by J. L. Doob, March 19, 1965

1. Let ξ_{ji}, $i=1, \cdots, n_j$, $j=1, \cdots, k$, be k mutually independent samples of mutually independent random variables having a common continuous distribution function $F(x)$. Let $F_{n_j}(x)$, $j=1, \cdots, k$, be the corresponding empirical distribution functions, that is $F_{n_j}(x) = (\text{number of } \xi_{ji} \leqq x,\ 1 \leqq i \leqq n_j)/n_j$.

We define $N = n_1/(\sum_{j=1}^{k} n_1/n_j)$ and let $N \to \infty$ mean that $n_j \to \infty$, $j=1, \cdots, k$, so that $n_1/n_j \to \rho_j$, $j=2, \cdots, k$, where ρ_j's are constant for each j.

Under above conditions the following theorems hold.

THEOREM 1.

$$\lim_{N\to\infty} P\left\{N^{1/2} \sup_{a \leqq F(x)} \left(\prod_{j=1}^{k} F_{n_j}(x) - F^k(x)\right) \Big/ F^k(x) < y\right\}$$

$$= (2/\pi)^{1/2} \int_0^{y\{a/(1-a)\}^{1/2}} e^{-t^2/2} dt, \textit{ if } y > 0, \textit{ zero otherwise, and } 0<a<1.$$

THEOREM 2.

$$\liminf_{N\to\infty} P\left\{N^{1/2} \sup_{a \leqq F(x)} \left|\prod_{j=1}^{k} F_{n_j}(x) - F^k(x)\right| \Big/ F^k(x) < y\right\}$$

$$\geqq 4/\pi \sum_{k=0}^{\infty} (-1)^k (2k+1)^{-1} \exp\{-(2k+1)^2\pi^2(1-a)/8ay^2\},$$

$$\textit{if } y > 0, \textit{ zero otherwise, and } 0 < a < 1$$

$$= L(y; a).$$

THEOREM 3.

$$\lim_{N\to\infty} P\left\{N^{1/2} \sup_{a \leqq F(x) \leqq b} \left(\prod_{j=1}^{k} F_{n_j}(x) - F^k(x)\right) \Big/ F^k(x) < y\right\}$$

$$= 1/\pi \int_{-\infty}^{y\{b/(1-b)\}^{1/2}} e^{-u^2/2} \cdot \left[\int_0^{\{y[b/(1-b)]^{1/2}-u\}\cdot\{a(1-b)/(b-a)\}^{1/}} e^{-t^2/2} dt\right] du$$

[1] This work was in part supported by the Office of Naval Research, Project ONR 042-023, Contract Nonr-1858(05) to Princeton University.

where $-\infty < y < +\infty$, $0 < a < b < 1$.

THEOREM 4.

$$\liminf_{N\to\infty} P\left\{N^{1/2} \sup_{a\leqq F(x)\leqq b} \left| \prod_{j=1}^{k} F_{n_j}(x) - F^k(x) \right| \Big/ F^k(x) < y\right\}$$
$$\geqq L(y;a)\cdot\left\{1 - 2/(2\pi)^{1/2}\int_{y\{b/(1-b)\}^{1/2}}^{\infty} e^{-u^2/2}du + A_k\right\}$$

where

$$A_k = (2\pi)^{-1/2}\{a/(1-a)\}^{-1/2}(1/y)2\exp\{-by^2/2(1-b)\}$$
$$\cdot \int_0^{(2k+1)\pi/2} \{\exp(1-b)u^2/2by^2\} \sin u du$$

and $L(y; a)$ *is as defined in Theorem* 2 *above for* $y>0$, *zero otherwise, and* $0<a<b<1$.

When $k=1$, Theorems 1, 2, 3, 4 are equivalent to Rényi's Kolmogorov-Smirnov type theorems in [2] with the additional remark with regard to Theorems 2 and 4 that in this case ($k=1$) we have lim statements instead of lim inf statements with the present lower bounds as limiting forms.

Theorems 1, 2, 3, 4 can be used to construct asymptotic confidence intervals for an unknown continuous distribution function $F(x)$ or to test the goodness-of-fit hypothesis

$$H_0: F_1(x) = F_2(x) = \cdots = F_k(x) = F(x) \tag{1.1}$$

where $F(x)$ is some specified continuous distribution function in this case. The class of alternatives to H_0 can be considered to be all sets $(F_1(x), \cdots, F_k(x))$ which violate (1.1).

Theorem 1 implies

COROLLARY 1.

$$\lim_{N\to\infty} P\left\{\sup_{a\leqq F(x)}\left(\prod_{j=1}^{k} F_{n_j}(x) - F^k(x)\right) < 0\right\} = 0$$

and from this in turn follows that

$$\lim_{N\to\infty}\left\{\sup_{0<F(x)<1}\left(\prod_{j=1}^{k} F_{n_j}(x) - F^k(x)\right) < 0\right\} = 0.$$

On the other hand we get from Theorem 3

COROLLARY 2.

$$\lim_{N\to\infty} P\left\{ \sup_{a\leqq F(x)\leqq b} \left(\prod_{j=1}^{k} F_{n_j}(x) - F^k(x) \right) < 0 \right\}$$

$$= 1/\pi \int_0^\infty e^{-u^2/2} \int_0^{u\{a(1-b)/(b-a)\}^{1/2}} e^{-t^2/2} dt du$$

$$= 1/\pi \arcsin \{a(1-b)/b(1-a)\}^{1/2},$$

a statement which, when $k=1$, *agrees with what has been said by Rényi* [2] *and Gihman* [1] *concerning the asymptotic behavior of an empirical and theoretical distribution function.*

2. We shall not give the proof here. We remark only that to prove these theorems one shows first that the set $\{x: a \leqq F(x)\}$ for which sup of random variables of Theorems 1 and 2 is examined can be replaced by $\bigcap_{j=1}^{k} \{x: a \leqq F_{n_j}(x)\} \cap \{x: a \leqq F_{\Sigma_j n_j}(x)\}$ in the limit (the same is true, mutatis mutandis, about Theorem 3 and 4) where $F_{\Sigma_j n_j}(x)$ is the empirical distribution function of the random sample gained by pooling the k random samples of size n_j, $j=1, \cdots, k$, respectively.

What follows after is a lengthy adaptation of Rényi's method of proof of his original Kolmogorov-Smirnov type theorems in [2].

Some aspects of these theorems in the one sample case are also discussed in a recent paper by the author in Ann. Math. Statist., **36** (1965), 322–326.

REFERENCES

1. T. Gihman, *Ob empiriceskoj funkcii raspredelenija slucaje grouppirovki dannych*, Dokl. Akad. Nauk SSSR **82** (1952), 837–840.

2. Alfréd Rényi, *On the theory of order statistics*, Acta Math. Acad. Sci. Hungar. **4** (1953), 191–213.

PRINCETON UNIVERSITY