ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVII, 1 (1980)

KRYSTYNA KĘDZIORA (Wrocław)

LIMIT DISTRIBUTIONS OF QUANTILES FOR A RANDOM SAMPLE OF RANDOM LENGTH

Consider independent random variables $X_1, X_2, \ldots$ with a common distribution function $F(x) = \Pr\{X_1 < x\}$. Let Z_K^N with $K = [Np]+1$ $(0 < p < 1)$ be a quantile for the random sample $X_1, X_2, \ldots, X_N$, where N is an integer-valued random variable, and let z_p be the quantile of order p of F. In the present paper the limit distributions of quantiles Z_K^N are studied.

The following lemma (see [2] and [3]) will be used in the sequel.

LEMMA. *Let I_{nk} $(k, n = 1, 2, \ldots)$ be random variables, independent for every fixed n, with distribution*

$$\Pr\{I_{nk} = 1\} = p_n, \qquad \Pr\{I_{nk} = 0\} = 1-p_n = q_n.$$

For any sequence of positive integers $\{k(n)\}$ tending to infinity, if

$$\lim_{n\to\infty} k(n)p_n q_n = \infty,$$

then

$$\lim_{n\to\infty}\Pr\left\{\sum_{k=1}^{k(n)} \frac{I_{nk}-p_n}{\sqrt{k(n)p_n q_n}} < x\right\} = \varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt.$$

THEOREM 1. *Let N_n $(n = 1, 2, \ldots)$ be positive integer-valued random variables satisfying*

$$N_n/n \xrightarrow{\mathrm{P}} c > 0, \qquad n \to \infty,$$

and let I_{nk} $(k, n = 1, 2, \ldots)$ be such as in the Lemma with $p_n \in (\delta, 1-\delta)$ for $0 < \delta < 1/2$ and sufficiently large n. Then

$$\lim_{n\to\infty}\Pr\{J_{N_n}/\sqrt{N_n} < x\} = \varphi(x),$$

where

$$J_r = \sum_{k=1}^{r} (I_{nk}-p_n)/\sqrt{p_n q_n}.$$

Proof. Let $0 < \varepsilon < 1/5$. Choose n_0 such that for every $n \geqslant n_0$

$$\Pr\{|N_n - cn| \geqslant cn\varepsilon\} \leqslant \varepsilon.$$

Since, obviously,

$$\Pr\{J_{N_n}/\sqrt{N_n} < x\} = \sum_{|r-cn| \geqslant cn\varepsilon} \Pr\{J_r/\sqrt{r} < x, N_n = r\} + \sum_{|r-cn| < cn\varepsilon} \Pr\{J_r/\sqrt{r} < x, N_n = r\}$$

and

$$\sum_{|r-cn| \geqslant cn\varepsilon} \Pr\{J_r/\sqrt{r} < x, N_n = r\} \leqslant \varepsilon \quad \text{for } n \geqslant n_0,$$

we have

$$\Big|\Pr\{J_{N_n}/\sqrt{N_n} < x\} - \sum_{|r-cn| < cn\varepsilon} \Pr\{J_r/\sqrt{r} < x, N_n = r\}\Big| \leqslant \varepsilon.$$

Put $N_1^* = N_1^*(n) = [cn(1-\varepsilon)]$, $N_2^* = N_2^*(n) = [cn(1+\varepsilon)]$ and let x be any real positive number. Proceeding as in [4] (see the proof of Theorem 1) we obtain

$$\Pr\{J_{N_1^*}/\sqrt{N_1^*} < x - \varepsilon^{1/3}\} - 7\varepsilon^{1/3} \leqslant \Pr\{J_{N_n}/\sqrt{N_n} < x\}$$
$$\leqslant \Pr\{J_{N_1^*}/\sqrt{N_1^*} < x\sqrt{(1+2\varepsilon)/(1-2\varepsilon)} + \varepsilon^{1/3}\} + 6\varepsilon^{1/3},$$

which together with the Lemma gives the assertion for $x \geqslant 0$. For $x < 0$ we prove the theorem in a similar way.

THEOREM 2. *Let the sequence of random variables $\{N_n\}$ satisfy the conditions of Theorem 1 and let the sequence $y_n = a_n x + b_n$, where $0 < a_n$, b_n, $x \in R$ $(n = 1, 2, \ldots)$, be such that*

$$V_K^n(x) \xrightarrow{P} V(x), \quad n \to \infty,$$

where

$$V_K^n(x) = \frac{N_n F(y_n) - K}{\sqrt{N_n F(y_n)\big(1 - F(y_n)\big)}}.$$

If there exists a δ $(0 < \delta < 1/2)$ such that $F(y_n) \in (\delta, 1-\delta)$ for sufficiently large n, then

$$\lim_{n\to\infty} \Pr\{(Z_K^{N_n} - b_n)/a_n < x\} = \varphi\big(V(x)\big).$$

Proof. Put

$$I_{nk} = \begin{cases} 1 & \text{if } X_k < y_n \\ 0 & \text{if } X_k \geqslant y_n \end{cases} \quad (k, n = 1, 2, \ldots).$$

For every fixed n the random variables I_{nk} are independent and, obviously,

$$\Pr\{Z_K^{N_n} < y_n\} = \Pr\Big\{\sum_{k=1}^{N_n} I_{nk} \geqslant K\Big\}.$$

Now, applying Theorem 1, we obtain

$$\Pr\{(Z_K^{N_n} - b_n)/a_n < x\} = \Pr\{Z_K^{N_n} < y_n\}$$

$$= \Pr\Big\{\sum_{k=1}^{N_n} I_{nk} \geqslant K\Big\} = \Pr\{J_{N_n}/\sqrt{N_n} \geqslant -V_K^n(x)\} \to \varphi\big(V(x)\big), \qquad n \to \infty,$$

which completes the proof of Theorem 2.

COROLLARY. *If the distribution function F is continuously differentiable in the neighbourhood of z_p and if $f(z_p) = F'(z_p) > 0$, if the sequence $\{N_n\}$ satisfies the conditions of Theorem 2 with $c = 1$ and*

$$y_n = z_p + x\sqrt{\frac{p(1-p)}{n}}\,\frac{1}{f(z_p)} \qquad (n = 1, 2, \ldots),$$

then $Z_K^{N_n}$ is asymptotically normal with parameters z_p and

$$\sqrt{\frac{p(1-p)}{n}}\,\frac{1}{f(z_p)}$$

(see [1]).

It is sufficient to prove that

$$V_K^n(x) \xrightarrow{P} x, \qquad n \to \infty. \tag{1}$$

Using the mean value theorem under the continuity condition of f we have

$$F\left(z_p + x\sqrt{\frac{p(1-p)}{n}}\,\frac{1}{f(z_p)}\right) = p + x\frac{a_n}{\sqrt{n}},$$

where

$$\lim_{n\to\infty} a_n = \sqrt{p(1-p)}.$$

Thus we can write

$$V_K^n(x) = \frac{\sqrt{N_n/n}\,a_n x - \sqrt{N_n}(K/N_n - p)}{\sqrt{\big(p + (a_n/\sqrt{n})x\big)\big(1 - p - (a_n/\sqrt{n})x\big)}}.$$

Since

$$0 \leqslant \sqrt{N_n}\left(\frac{K}{N_n} - p\right) \leqslant \frac{1}{\sqrt{N_n}},$$

we have

$$\sqrt{N_n}\left(\frac{K}{N_n}-p\right)\xrightarrow{P}0, \quad n\to\infty. \tag{2}$$

We have also

$$\sqrt{\frac{N_n}{n}}\xrightarrow{P}1, \quad n\to\infty. \tag{3}$$

Now, combining (2) and (3), we get (1).

References

[1] F. I. Anscombe, *Large-sample theory of sequential estimation*, Proc. Cambridge Phil. Soc. 48 (1952), p. 600-607.

[2] K. C. Chanda, *Asymptotic properties of order statistics attracted to normal law*, Calcutta Statistical Association Bulletin 24 (1975), p. 13-22.

[3] M. Loève, *Ranking limit problems*, Proc. Third Berkeley Symp. 2 (1956), p. 174-194.

[4] A. Rényi, *On the asymptotic distribution of the sum of the random number of independent random variables*, Acta Math. Acad. Sci. Hung. 8 (1957), p. 193-197.

MATHEMATICAL INSTITUTE
UNIVERSITY OF WROCŁAW
50-384 WROCŁAW

Received on 30. 3. 1978

KRYSTYNA KĘDZIORA (Wrocław)

GRANICZNE ROZKŁADY KWANTYLI Z PRÓBY PROSTEJ O LOSOWEJ DŁUGOŚCI

STRESZCZENIE

Przedmiotem pracy są twierdzenia graniczne dla kwantyli z próby prostej o losowej długości. Podaje się warunki dostateczne na zbieżność tych kwantyli do rozkładu normalnego z odpowiednimi parametrami.
