Sankhyā : The Indian Journal of Statistics
1979, Volume 41, Series A, Pts. 3 and 4, pp. 271–277.

REPRESENTATION OF QUANTILE PROCESSES WITH NON-UNIFORM BOUNDS

By KESAR SINGH
Indian Statistical Institute

SUMMARY. A representation of quantile processes is established in which the order of the remainder gets sharper and sharper as one moves towards sample extremes.

1. Introduction

Let $\{X_i\}$ be a stationary sequence of random variables (r.v.s.) with marginals as $U[0, 1]$. At the n-th stage we define empirical distribution function (e.d.f.) as

$$F_n(x) = (\# X_i \leqslant x : \ 1 \leqslant i \leqslant n)/n$$

and the t-th sample quantile as

$$F_n^{-1}(t) = \inf\{x : F_n(x) \geqslant t\} \qquad \text{for} \quad t > 0$$

and

$$F_n^{-1}\, 0^+) \qquad\qquad \text{for} \quad t = 0.$$

Bahadur (1966) proved that in the independent case

$$R_n(t) = |F_n^{-1}(t) - t + F_n(t) - t|$$
$$= O(n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}) \text{ a.s.}$$

Later, Kiefer (1967, 1970) concluded that in the independent case

$$R_n(t) = O(n^{-3/4}(\log\log n)^{3/4}) \text{ a.s.}$$

$$\sup_{0 \leqslant t \leqslant 1} R_n(t) = O(n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}) \text{ a.s.} \qquad \dots \quad (1.1)$$

and these orders are exact. Recently, Babu and Singh (1978), Singh (1978a, 1978b) extended some of the above mentioned results for mixing random variables.

It appears quite natural to expect that the order of $R_n(t)$ would be sharper when t is near the extremes of the interval $[0, 1]$ than the uniform order given by (1.1).

The phenomenon was also noted in Kiefer (1970) but it was left there as an open problem. The purpose of this paper is to provide an a.s. asymptotic bound for the process $R_n(t)$ in which the smallness of $t(1-t)$ is reflected clearly. To be precise, this paper establishes the following theorem :

Theorem 1 : *If $\{X_t\}$ is a stationary m-dependent process with marginals* $U[0, 1]$, *then*

$$\sup_{n^{-1}\log n < t < 1-n^{-1}\log n} [t(1-t)]^{-1/4} |F_n^{-1}(t)-t+F_n(t)-t|$$

$$= O(n^{-3/4}(\log n)^{3/4}) \quad a.s. \qquad \ldots \quad (1.2)$$

This bound appears to be helpful in studying the extreme sample quantiles, weighted quantile processes and linear functions of order statistics. These applications are presented in Singh (1978a).

In the next section l_n stands for log n.

2. Proof of theorem 1

We shall prove (1.2) with the supremum over $[n^{-1} l_n, \frac{1}{2}]$ only and a similar argument is applied in the case of the interval $[\frac{1}{2}, 1-n^{-1} l_n]$ to conclude the theorem.

Define, for $0 \leqslant \alpha, \beta \leqslant 1$,

$$x_t(\alpha, \beta) = I(\min(\alpha, \beta) \leqslant X_t \leqslant \max(\alpha, \beta)) - |\alpha-\beta|,$$

($I(A)$ denotes the indicator function of the set A.) We start with the exponential inequality which is the main tool in the proof.

Lemma 1 : *Let $\{X_t\}$ be as in the theorem. Then, there exists a $d > 0$ such that whenever $0 \leqslant \alpha$, $\alpha+\beta \leqslant 1$, $|\beta| \leqslant b > 0$, $1 \leqslant u \leqslant N$, $H \geqslant 0$ and $0 < D \leqslant Nb^{\theta}$ for some $0 \leqslant \theta \leqslant \frac{1}{2}$, one has*

$$P\left(\left|\sum_{t=H+1}^{H+u} x_t(\alpha, \alpha+\beta)\right| > 2dDb^{\theta}\right) \leqslant 2m\, exp(-8D^2N^{-1}). \qquad \ldots \quad (2.1)$$

(In this paper, the lemma is used in the special case $H = 0$, $u = N = n$, and $\theta = \frac{1}{2}$ only.)

Proof : Let us take, without loss of generality, $H = 0$ and $b \geqslant \beta > 0$. Define $z = b^{-\theta}N^{-1}D$ so that $0 \leqslant z \leqslant 1$. Using the stationarity, the m-dependence and Markov's inequality, it follows that

$$P\left(\sum_{i=1}^{u} x_i(\alpha, \alpha+\beta) > dDb^{\theta}\right)$$

$$\leqslant m\left[P\left(\sum_{i=0}^{[u/m]-1} x_{mi+1}(\alpha, \alpha+\beta) > \frac{dDb^{\theta}}{m}\right)\right.$$

$$\left.+P\left(\sum_{i=0}^{[u/m]} x_{mi+1}(\alpha, \alpha+\beta) > \frac{dDb^{\theta}}{m}\right)\right]$$

$$\leqslant 2m \exp(-zm^{-1}dDb^{\theta}) \max_{v=[u/m],[u/m]+1} [E(\exp(zx_1(\alpha, \alpha+\beta))]^{v}.$$

(In the above expressions, we take $\sum_{0}^{[u/m]} = 0$ if $[u/m] = 0$)

$$\leqslant 2m \exp(-m^{-1}dD^2N^{-1}+(u/m+1)\log(1+O(z^2\beta))$$

$$\leqslant 2m \exp(-8D^2N^{-1})$$

$$(\log(1+x) \leqslant x, \ \forall\, x > -1) \quad \ldots \quad (2.2)$$

by choosing d appropriately.

We obtain a similar inequality for $-\sum_{i=1}^{u} x_i(\alpha, \alpha+\beta)$ to complete the proof.

Lemma 2 : *Let $F_n(t)$ be as in Theorem* 1.

$$\sup_{n^{-1}l_n < t < 3/4} t^{-1/2}|F_n(t)-t| = O(n^{-1/2}l_n^{1/2}) \quad a.s.$$

Proof : We divide the interval $[n^{-1}l_n, 3/4]$ into subintervals of length n^{-1} and note that

$$\sup_{n^{-1}l_n \leqslant t \leqslant 3/4} |F_n(t)-t|t^{-1/2} \leqslant \max\{|F_n(s)-s|s^{-1/2} : s \in S_n\}+n^{-1/2} \quad \ldots \quad (2.3)$$

where

$$S_n = \{n^{-1}l_n, n^{-1}l_n+n^{-1}, n^{-1}l_n+2n^{-1}, \ldots, n^{-1}l_n+(3n/4+1)n^{-1}\}.$$

Applying Lemma 1 with $H = 0$, $\theta = \frac{1}{2}$, $u = N = n$, $b = 1$ and $D = n^{\frac{1}{2}} l_n^{\frac{1}{2}}$, it follows that, uniformly in $s \in S_n$,

$$P(n\,|F_n(s) - s| > 2d\, s^{1/2}\, n^{1/2}\, l_n^{1/2}) = O(n^{-3}).$$

Therefore, using Bonferroni inequality and the Borel-Cantelli lemma, it follows that the r.h.s. of (2.3) $= O(n^{-\frac{1}{2}} l_n^{\frac{1}{2}})$ a.s. This yields our desired result.

Lemma 3 : *Under the same set up as in Theorem* 1 *there exists a constant* $K > 0$ *such that*

$$\limsup_{n \to \infty} \sup_{n^{-1} l_n < t < 1/2} n^{1/2}\, l_n^{-1/2}\, t^{-1/2} |F_n^{-1}(t) - t| \leqslant K \quad a.s.$$

Proof : Lemma 2 guarantees the existence of a constant $c > 0$ such that

$$\limsup_{n \to \infty} \sup_{n^{-1} l_n < t < 3/4} n^{1/2}\, l_n^{-1/2}\, t^{-1/2} |F_n(t) - t| < c \quad \text{a.s.}$$

Therefore, for all sufficiently large n and $s \in [n^{-1} l_n, 3/4]$,

$$t - c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2} < F_n(t) < t + c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2} \quad \text{a.s.}$$

(the null set is same for all t). Hence, for $t \in [(c+1)^2 n^{-1} l_n, 1/2]$ and n sufficiently large,

$$F_n(t - c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2}) < t - c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2} + c(t - c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2})^{1/2}\, n^{-1/2}\, l_n^{1/2} < t. \quad \ldots \quad (2.4)$$

and

$$F_n(t + 2c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2}) > t + 2c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2} - c(t + 2c\, t^{1/2}\, n^{-1/2}\, l_n^{1/2})^{1/2}\, n^{-1/2}\, l_n^{1/2} > t \quad \ldots \quad (2.5)$$

with probability one (the null set being free of t).

Obviously, (2.4) and (2.5) imply that

$$\sup_{(c+1)^2 n^{-1} l_n < t < 1/2} t^{-\frac{1}{2}} |F_n^{-1}(t) - t| = O(n^{-1/2}\, l_n^{1/2}) \quad \text{a.s.} \quad \ldots \quad (2.6)$$

Further,

$$\begin{aligned} &\sup_{n^{-1} l_n < t < (c+1)^2 n^{-1} l_n} t^{-\frac{1}{2}} |F_n^{-1}(t) - t| \\ &\leqslant [F_n^{-1}((c+1)^2 n^{-1} l_n) + (c+1)^2 n^{-1} l_n](n^{-1} l_n)^{-1/2} \\ &= O(n^{-\frac{1}{2}}\, l_n^{1/2}) \quad \text{a.s.} \end{aligned} \quad \ldots \quad (2.7)$$

since (2.6) implies that $F_n^{-1}((c+1)^2 n^{-1} l_n) = O(n^{-1} l_n)$ a.s. Now, (2.6) and (2.7) complete the proof of this lemma.

Proof of Theorem 1 : In view of the fact that

$$|F_n F_n^{-1}(t)-t| \leqslant |F_n F_n^{-1}(t)-F_n(F_n^{-1}(t)-0)|,$$

we have,

$$|F_n^{-1}(t)-t+F_n(t)-t|\,t^{-1/4}$$

$$\leqslant |F_n F_n^{-1}(t)-F_n^{-1}(t)-F_n(t)+t|\,t^{-1/4}$$

$$+|F_n F_n^{-1}(t)-(F_n F_n^{-1}(t)-0)|\,t^{-1/4}$$

$$\leqslant 2|F_n F_n^{-1}(t)-F_n^{-1}(t)-F_n(t)+t|\,t^{-1/4}$$

$$+|F_n(F_n^{-1}(t)-0)-F_n^{-1}(t)-F_n(t)+t|\,t^{-1/4}.$$

Combining this with Lemma 3, we conclude that, for all sufficiently large n,

$$\sup_{n^{-1}l_n\leqslant t\leqslant 1/2} |F_n^{-1}(t)-t+F_n(t)-t|\,t^{-1/4}$$

$$\leqslant 3 \sup_{n^{-1}l_n\leqslant t\leqslant 1/2}\ \sup_{|s-t|\leqslant 2Kt^{1/2}n^{-1/2}l_n^{1/2}} |F_n(s)-F_n(t)-s+t|\,t^{-1/4} \text{ a.s.} \quad \ldots \quad (2.8)$$

(K is same as the constant appearing in Lemma 3.)

Let us fix a $t \in [n^{-1}l_n, 1/2]$ and divide the interval $[t-2Kt^{1/2}n^{-1/2}l_n^{1/2}, t+2Kt^{1/2}n^{-1/2}l_n^{1/2}]$ into subintervals of length n^{-3}. Then, due to the usual kind of approximation,

$$\sup_{|s-t|\leqslant 2Kt^{1/2}n^{-1/2}l_n^{1/2}} |F_n(s)-F_n(t)-s+t|\,t^{-1/4}$$

$$\leqslant \max_{|j|\leqslant v_n(t)} |F_n(t+jn^{-3})-F_n(t)-jn^{-3}|\,t^{-1/4}+n^{-1} \quad \ldots \quad (2.9)$$

where

$$v_n(t) = [2Kt^{1/2}n^{5/2}l_n^{1/2}]+1.$$

Further, let us divide the interval $[n^{-1}l_n, \frac{1}{2}]$ into subintervals of length n^{-3} and note that if

$$s \in [n^{-1}l_n+in^{-3},\ n^{-1}l_n+(i+1)n^{-3}],$$

then

$$\max_{|j|\leqslant v_n(s)} |F_n(s+jn^{-3})-F_n(s)-jn^{-3}|\,s^{-1/4}$$

$$\leqslant 2 \max_{|j|\leqslant v_n(a)+1} |F_n(a+jn^{-3})-F_n(a)-jn^{-3}|\,a^{-1/4}+O(n^{-1}), \quad \ldots \quad (2.10)$$

where

$$a = n^{-1}l_n+(i+1)n^{-3}.$$

In view of (2.9) and (2.10)

$$\text{l.h.s. of (2.8)} \leqslant O(n^{-1})+3 \max_{0<r<[n^{3/2}]} \max_{|j|<v_n(n^{-1}l_n+rn^{-3})+1} |F_n(n^{-1}l_n+(r+j)n^{-3})-F_n(n^{-1}l_n+rn^{-3})-jn^{-3}| (n^{-1}l_n+rn^{-3})^{-1/4}. \qquad \ldots \quad (2.11)$$

Let us fix some r and j (j depending upon r) as in the r.h.s. of (2.11). We shall estimate

$$P(n\,|F_n(n^{-1}l_n+(r+j)n^{-3})-F_n(n^{-1}l_n+rn^{-3})-jn^{-3}| \geqslant 2d\sqrt{3}K(n^{-1}l_n+rn^{-3})^{1/4}n^{1/4}l_n^{3/4} \qquad \ldots \quad (2.12)$$

using Lemma 1. We take $\beta = jn^{-3}$, $u = N = n$, $H = 0$, $\theta = 1/2$, $b = 3K\,(n^{-1}l_n+rn^{-3})^{1/2}n^{-1/2}l_n^{1/2}$ and $D = n^{1/2}l_n^{1/2}$ in Lemma 1, for the present use. We have to check the two conditions, namely, $|\beta| \leqslant b$ and $D \leqslant Nb^{1/2}$.

Now

$$|\beta| \leqslant (v_n(n^{-1}l_n+rn^{-3})+1)n^{-3}$$

$$\leqslant 2K(n^{-1}l_n+rn^{-3})^{1/2}n^{-1/2}l_n^{1/2}+2n^{-3}$$

$$\leqslant b \quad \text{(for all large } n\text{)}.$$

The second condition is immediate assuming, without loss of generality, that $K > 1$. Hence, Lemma 1 gives that the l.h.s. of (2.12) $= O(n^{-8})$.

Now, a simple application of Bonferroni inequality and Borel-Cantelli lemma shows that the

$$\text{r.h.s. of (2.11)} = O(n^{-3/4}l_n^{3/4}) \quad \text{a.s.}$$

This completes the proof of Theorem 1.

Remark : As a consequence of Theorem 1

$$\frac{\sqrt{n}(F_n^{-1}(t_n)-t_n)}{[t_n(1-t_n)]^{1/2}} = -\frac{\sqrt{n}(F_n(t_n)-t_n)}{[t_n(1-t_n)]^{1/2}} + o(1) \quad \text{a.s.}$$

whenever $t_n(1-t_n) \to 0$ and $t_n(1-t_n)n\,l_n^{-8} \to \infty$ as $n \to \infty$.

Some extensions of Theorem 1 for mixing random variables are stated below. The proofs of these results are outlined in Singh (1978a). We omit them here since they follow essentially the same lines.

Theorem 2: (i) *If* $\{X_i\}$ *is a sequence of* ϕ*-mixing* $U[0, 1]$ *r.v.s. with* $\phi(n) = O(n^{-\gamma})$ *(if* $\gamma = 2$*, we assume further that* $\Sigma\phi^{\frac{1}{2}}(i) < \infty$*) and* $0 \leqslant 2\varepsilon < 1/2 - \frac{1}{2(\gamma+1)}$*, then*

$$\sup_{n^{-1} < t < 1-n^{-1}} (t(1-t))^{-\varepsilon} |F_n^{-1}(t) - t + F_n(t) - t| = O(n^{-3/4}\, l_n^{3/4}) \quad a.s.$$

(ii) *If* $\{X_i\}$ *is a sequence of* ϕ*-mixing* $U[0, 1]$ *r.v.s. with* $\phi(n) = O(e^{-\gamma n})$*, for some* $\gamma > 0$*, then*

$$\sup_{n^{-1}l_n < t < 1-n^{-1}l_n} (t(1-t))^{-1/4} |F_n^{-1}(t) - t + F_n(t) - t| = O(n^{-3/4}\, l_n^{7/4}) \quad a.s.$$

(iii) *If* $\{X_i\}$ *is a sequence of strong-mixing r.v.s. with* $\alpha(n) = O(e^{-\gamma n})$*, for some* $\gamma > 0$ *then, for all* $0 \leqslant 2\varepsilon \leqslant 1/4$

$$\sup_{n^{-1} < t < 1-n^{-1}} (t(1-t))^{-\varepsilon} |F_n^{-1}(t) - t + F_n(t) - t| = O(n^{-3/4}\, l_n^{3/4}) \quad a.s.$$

References

Babu, G. J. and Singh, K. (1978): On deviations between empirical and quantile processes for mixing random variables. *Jour. Multi. Anal.* **8**, 532–549.

Bahadur, R. R. (1966): A note on quantiles in large samples. *Ann. Math. Statist.*, **37**, 577-580.

Kiefer, J. (1967): On Bahadur representation of sample quantiles. *Ann. Math. Statist.*, **38**, 1323-1341.

——— (1970): *Deviations Between the Sample Quantile Process and the Sample DF. Non-parametric Techniques in Statistical Inference.* (Ed. M. L. Puri) C.U.P., 299-319.

Singh, K. (1978a): On the asymptotic theory of quantiles and L-statistics. Thesis submitted to the Indian Statistical Institute.

——— (1978b): A note on the representation of quantiles for mixing random variables. Submitted for publications.

Paper received: September, 1978.