PROBABILITY

AND MATHEMATICAL STATISTICS

Vol. 4, Fasc. 2 (1984), p. 221-236

1

TWO APPROACHES TO CONSTRUCTING SIMULTANEOUS CONFIDENCE BOUNDS FOR QUANTILES

BY

M. CSÖRGŐ* (OTTAWA) AND P. RÉVÉSZ (BUDAPEST)

Abstract. Given some regularity conditions on the distribution function F of a random sample $X_1, X_2, ..., X_n$, the sequence of quantile processes $[n^{1/2}f(Q(y))(Q_n(y)-Q(y)); 0 < y < 1]$ behaves like a sequence of Brownian bridges $\{B_n(y); 0 < y < 1\}$, where Q(y) $:= F^{-1}(y)$, the inverse of $F(\cdot)$, and $Q_n(y) = X_{k,n}$ if $(k-1)/n < y \le k/n$ (k = 1, 2, ..., n) with the order statistics $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ of the above sample. First, a sequence of consistent direct estimators is proposed for the quantile-density function Q'(y) = 1/f(Q(y)). The latter then also enables us to construct simultaneous confidence bounds for an unknown quantile function Q(y). The second approach makes frequently misused heuristic steps like

$$\begin{aligned} -\alpha &= P\left\{F(x) - n^{-1/2} c(\alpha) \leqslant F_n(x) \leqslant F(x) + n^{-1/2} c(\alpha); \\ &-\infty < x < \infty\right\} \\ &= P\left\{y - n^{-1/2} c(\alpha) \leqslant F_n(F^{-1}(y)) \leqslant y + n^{-1/2} c(\alpha); \\ &F^{-1}(0) < F^{-1}(y) < F^{-1}(1)\right\} \\ &= P\left\{F_n^{-1} \left(y - n^{-1/2} c(\alpha)\right) \leqslant F^{-1}(y) \leqslant F_n^{-1} \left(y + n^{-1/2} c(\alpha)\right); \\ &0 < y < 1\right\}\end{aligned}$$

precise for large *n*, where F_n is the empirical distribution function of the above random sample, and for $\alpha \in (0, 1)$, $c(\alpha)$ is defined by

$$P\left\{\sup_{0\leq y\leq 1}|B(y)|\leq c(\alpha)\right\}=1-\alpha$$

for a Brownian bridge $B(\cdot)$.

1. Introduction. Let $X_1, X_2,...$ be a sequence of i.i.d. r.v. with an absolutely continuous distribution function $F(\cdot)$ and let $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ denote the order statistics of the random sample $X_1, X_2,..., X_n$. Define the

^{*} This research was supported by a Canada Council Killam Senior Research Fellowship and by a Natural Sciences and Engineering Research Council Canada Grant, both held at Carleton University.

empirical distribution function $F_n(x)$ and the empirical quantile function $Q_n(y)$ as follows:

$$F_n(x) = \begin{cases} 0 & \text{if } X_{1:n} > x, \\ k/n & \text{if } X_{k:n} \le x < X_{k+1:n}, \ k = 1, 2, \dots, n-1, \\ 1 & \text{if } X_{n:n} \le x, \end{cases}$$
$$Q_n(y) = X_{k:n} & \text{if } (k-1)/n < y \le k/n, \ k = 1, 2, \dots, n.$$

It is a natural idea to use the empirical quantile functions as an estimator of the quantile function

 $Q(y) := F^{-1}(y)$, where $F^{-1}(y) = \inf \{x: F(x) \ge y\}$ (0 < y < 1).

Properties of the empirical quantile function $Q_n(y)$ were studied in a number of papers. Here we refer only to our earlier papers and references therein (see [3] and [4]) where we proved the following

THEOREM A. Let $X_1, X_2, ...$ be i.i.d. r.v. with a continuous distribution function F. Assume that the following conditions hold:

(i) F(x) is twice differentiable on (a, b), where

$$-\infty \leqslant a = \sup \{x: F(x) = 0\}, \quad \infty \ge b = \inf \{x: F(x) = 1\};$$

(ii) F' = f > 0 on (a, b);

(iii) for some $\gamma > 0$ we have

$$\sup_{x < x < b} F(x) \left(1 - F(x)\right) \frac{|f'(x)|}{f^2(x)} \leq \gamma;$$

(iv)
$$A = \lim_{x \downarrow a} f(x) < \infty, \quad B = \lim_{x \uparrow b} f(x) < \infty;$$

- (v) one of the following conditions holds:
- $(\mathbf{v}_{\alpha}) \quad \min(A, B) > 0,$
- (v_{β}) if A = 0 (B = 0), then f is nondecreasing (nonincreasing) on an interval to the right of a (to the left of b).

One can then define a Brownian bridge $\{B_n(y); 0 \le y \le 1\}$ for each n such that

$$\sup_{0 < y < 1} \left| f(Q(y)) n^{1/2} (Q_n(y) - Q(y)) - B_n(y) \right|$$

$$\stackrel{a.s.}{=} \begin{cases} O(n^{-1/2} \log n) & \text{if } \gamma < 2, \\ O(n^{-1/2} (\log \log n)^{\gamma} (\log n)^{(1+\varepsilon)(\gamma-1)}) & \text{if } \gamma \ge 2. \end{cases}$$

where γ is as in (iii) and $\varepsilon > 0$ is arbitrary.

222

Theorem A is an analogue of Kolmogorov's classical theorem on the empirical distribution function, which gives a confidence bound for the distribution function F. However, Theorem A does not give an immediate confidence bound for the quantile function Q(y). In fact, a direct application of Theorem A produces only the "confidence bound"

$$Q_n(y) - n^{-1/2} \frac{c}{f(Q(y))} \le Q(y) \le Q_n(y) + n^{-1/2} \frac{c}{f(Q(y))} \quad (0 < y < 1)$$

(where the constant c depends on the significance level), which depends on the unknown function 1/f(Q(y)). In our book (cf. [5], Theorem 5.5.2) we proposed to estimate the function 1/f(Q(y)) from the sample X_1, X_2, \ldots, X_n by $1/f_n(Q_n(y))$, where f_n is any appropriate estimator of f and Q_n is the empirical quantile function. This idea can certainly produce a confidence bound, but the use of two estimators (that of f and that of Q) seems to be a somewhat artificial solution of the problem at hand. One of the aims of the present paper is to give a direct estimator of 1/f(Q(y)), and we first prove

THEOREM 1. Let X_1, X_2, \ldots be i.i.d. r.v. with a continuous distribution function $F(\cdot)$ satisfying the conditions of Theorem A. Assume also that the condition

(vi)
$$\sup_{a \le x \le b} \frac{\left[F(x)(1-F(x))\right]^2}{f(x)} \le C$$

holds with some C > 0. Then

$$\lim_{n\to\infty}\omega(n)\sup_{\varepsilon_n\leqslant y\leqslant 1-\varepsilon_n}\left|\frac{Q_n(y+a_n)-Q_n(y-a_n)}{2a_n}-\frac{1}{f(Q(y))}\right|\stackrel{a.s.}{=}0,$$

where $a_n = n^{-\alpha}$, $\varepsilon_n = n^{-\beta}$, $\omega(n) = n^{\delta}$, and $3\beta + \delta < \alpha < \frac{1}{2}$, $2\delta + 4\beta + \alpha < 1$.

The estimator $(2a_n)^{-1}(Q_n(y+a_n)-Q_n(y-a_n))$ proposed here for 1/f(Q(y)) is the natural analogue of the estimator

$$f_n^* = (2a_n)^{-1} \left(F_n(x+a_n) - F_n(x-a_n) \right),$$

which is frequently used to estimate a density function f. However, when studying the problem of density estimation in general, one quickly learns that instead of the estimator f_n^* the more sophisticated estimator

$$f_n(x) = a_n^{-1} \int_{-\infty}^{\infty} \lambda(a_n^{-1}(x-u)) dF_n(u)$$

should be used, where $\lambda(\cdot)$ is an arbitrary density function. Given this experience, we also propose the more general estimator

(1.1)
$$\varphi_n(y) = a_n^{-1} \int_0^1 \lambda\left(\frac{y-u}{a_n}\right) dQ_n(u)$$

for the function 1/f(Q(y)). As to the properties of $\varphi_n(y)$ we prove our main result on quantile-density estimation here, namely

THEOREM 2. Let X_1, X_2, \ldots be i.i.d. r.v. with a continuous distribution function $F(\cdot)$ satisfying the conditions of Theorem 1. Assume also that the condition

$$\sup_{a < x < b} |f''(x)| \le C$$

holds with some C > 0. Let $\lambda(\cdot)$ be a density function which is absolutely continuous on $(-\infty, \infty)$, vanishing outside of the interval $(-\frac{1}{2}, \frac{1}{2})$, and such that

$$\int_{-1/2}^{1/2} x\lambda(x) \, dx = 0, \quad \int_{-1/2}^{1/2} x^2 \, \lambda(x) \, dx < \infty, \quad \sup_{x} |\lambda'(x)| < \infty$$

00.

Then

$$\lim_{n\to\infty}\omega(n)\sup_{\varepsilon_n\leqslant y\leqslant 1-\varepsilon_n}\left|\varphi_n(y)-\frac{1}{f(Q(y))}\right|\stackrel{a.s.}{=}0,$$

where $\varphi_n(y)$ is defined as in (1.1), $a_n = n^{-\alpha}$, $\varepsilon_n = n^{-\beta}$, $\omega(n) = n^{\delta}$, and $\delta + 10\beta < 2\alpha$, $\alpha + 2\beta + \delta < \frac{1}{2}$.

An indirect application of Theorem A provides us an alternative, direct route to constructing simultaneous confidence bounds for quantiles. Namely, the following heuristic steps can be made precise. Let $c_1(\alpha)$ be defined by

(1.2)
$$P\left\{\sup_{0\leq y\leq 1}B(y)\leq c_1(\alpha)\right\}=1-\alpha,$$

where $\{B(y); 0 \le y \le 1\}$ is a Brownian bridge and $\alpha \in (0, 1)$ is fixed. Then, given a random sample of size *n* on a continuous distribution function $F(\cdot)$, in terms of the empirical distribution function $F_n(\cdot)$ of our random sample on $F(\cdot)$ we have

(1.3)
$$P\left\{F_n(x) \leq F(x) + n^{-1/2}c_1(\alpha); -\infty < x < \infty\right\} = 1 - \alpha$$
 as $n \to \infty$,

and letting y = F(x), we get

(1.4)
$$P\left\{F_n(F^{-1}(y)) \le y + n^{-1/2} c_1(\alpha); F^{-1}(0) < F^{-1}(y) < F^{-1}(1)\right\}$$
$$= P\left\{F^{-1}(y) \le F_n^{-1}(y + n^{-1/2} c_1(\alpha)); 0 < y < 1\right\} = 1 - \alpha \quad \text{as } n \to \infty,$$

i.e., in terms of our notation $Q(y) := F^{-1}(y)$ and $Q_n(y) := F_n^{-1}(y)$, for large n we have

(1.5)
$$P\left\{Q(y) \leq Q_n(y+n^{-1/2}c_1(\alpha)); 0 < y < 1\right\} = 1-\alpha.$$

On similar heuristic grounds we have

(1.6)
$$P\{Q_n(y-n^{-1/2}c_1(\alpha)) \le Q(y); 0 < y < 1\} = 1-\alpha \quad \text{as } n \to \infty,$$

and if $c(\alpha)$ is defined by (cf. (1.2))

(1.7)
$$P\left\{\sup_{0\leq y\leq 1}|B(y)|\leq c(\alpha)\right\}=1-\alpha,$$

then for large n we have

(1.8) $P\{Q_n(y-n^{-1/2}c(\alpha)) \leq Q(y) \leq Q_n(y+n^{-1/2}c(\alpha)); 0 < y < 1\} = 1-\alpha.$

It is clear from Theorem A that further conditions on F (other than its continuity) will have to be assumed for the validity of statements like (1.5), (1.6) and (1.8), although (1.3) and similar statements for the empirical distribution function, which lead us to considering (1.5), (1.6) and (1.8), are of course true with the stipulated numbers $c_1(\alpha)$ and $c(\alpha)$ provided only that F is assumed to be continuous. Thus the statements (1.5), (1.6), and (1.8) with $c_1(\alpha)$ and $c(\alpha)$ as above are *candidates* for being simultaneous $1-\alpha$ confidence bounds for Q(y). The question is the following: under what further conditions on F will they hold true as $n \to \infty$ and for which quantiles then? The answers to this question are given in Section 5 under the conditions of Theorem B (cf. Section 5).

2. Proof of Theorem 1. The proof of Theorem 1 is based on two lemmas (cf. Lemmas 1 and 2 below). For the proof of Lemma 2 we need also

LEMMA A (Csörgö and Révész [3], Theorem 6, and [5], Theorem 1.15.2). Let

$$\varrho_n(y) = n^{1/2} f(Q(y)) (Q_n(y) - Q(y)) \quad (0 < y < 1).$$

Then

 $\lim_{n \to \infty} \sup_{1/n < y - a_n < y + a_n < 1 - 1/n} (4a_n \log a_n^{-1})^{-1/2} |\varrho_n(y + a_n) - \varrho_n(y - a_n)| \stackrel{a.s.}{=} 1$

provided that $a_n = n^{-\alpha}$, $0 < \alpha < 1/2$.

LEMMA 1. Assume the conditions of Theorem 1 are satisfied. Then

$$\lim_{n\to\infty}\omega(n)\sup_{\varepsilon_n\leqslant y\leqslant 1-\varepsilon_n}\left|\frac{Q(y+a_n)-Q(y-a_n)}{2a_n}-\frac{1}{f(Q(y))}\right|=0.$$

Proof. By the mean value theorem we obtain

$$\frac{F^{-1}(y+a_n)-F^{-1}(y-a_n)}{2a_n}-\frac{1}{f(F^{-1}(y))}=-2a_n\frac{f'(F^{-1}(\eta))}{f^3(F^{-1}(\eta))},$$

where $y - a_n \leq \eta \leq y + a_n$. By conditions (iii) and (vi), we have

$$a_n \frac{\left|f'\left(F^{-1}\left(\eta\right)\right)\right|}{f^3\left(F^{-1}\left(\eta\right)\right)} \leq O(1) a_n (\varepsilon_n - a_n)^{-3}.$$

This implies our statement if we recall that $Q(y) := F^{-1}(y)$.

LEMMA 2. Assume the conditions of Theorem 1 are satisfied. Then

$$\lim_{n\to\infty}\omega(n)\sup_{\varepsilon_n\leqslant y\leqslant 1-\varepsilon_n}\left|\frac{Q_n(y+a_n)-Q_n(y-a_n)}{2a_n}-\frac{Q(y+a_n)-Q(y-a_n)}{2a_n}\right|\stackrel{a.s.}{=}0.$$

Proof. By the mean value theorem, we have

$$\frac{Q_n(y+a_n)-Q_n(y-a_n)}{2a_n} \frac{F^{-1}(y+a_n)-F^{-1}(y-a_n)}{2a_n}$$

= $\frac{\varrho_n(y+a_n)-\varrho_n(y-a_n)}{2a_n\sqrt{n}f(F^{-1}(y-a_n))} \frac{\varrho_n(y+a_n)}{\sqrt{n}f^2(F^{-1}(y+a_n))} \frac{f(F^{-1}(y+a_n))-f(F^{-1}(y-a_n))}{2a_n}$

$$=\frac{\varrho_n(y+a_n)-\varrho_n(y-a_n)}{2a_n\sqrt{n}f(F^{-1}(y-a_n))}-\frac{\varrho_n(y+a_n)}{\sqrt{n}f^2(F^{-1}(y+a_n))}\frac{f'(F^{-1}(\eta))}{f(F^{-1}(\eta))},$$

where $y - a_n \le \eta \le y + a_n$. Applying again conditions (iii) and (vi), by Lemma A we obtain our statement.

Clearly, Lemmas 1 and 2 imply Theorem 1.

3. **Proof of Theorem 2.** The proof of Theorem 2 is based on two lemmas (cf. Lemmas 3 and 4 below). For the proof of Lemma 4 we need also

LEMMA B (Csörgö and Révész [3], [4]). Let $X_1, X_2, ...$ be i.i.d. r.v. with a continuous distribution function $F(\cdot)$ satisfying the conditions of Theorem A. Then

$$\limsup_{n \to \infty} \sup_{0 < y < 1} \frac{|\varrho_n(y)|}{(\log \log n)^{1/2}} \stackrel{a.s.}{=} 2^{-1/2}.$$

LEMMA 3. Under the conditions of Theorem 2 we have

$$\lim_{n \to \infty} \omega(n) \sup_{\varepsilon_n \leq y \leq 1-\varepsilon_n} \left| \frac{1}{a_n} \int_0^1 \lambda\left(\frac{y-u}{a_n}\right) dQ(u) - \frac{1}{f(Q(y))} \right| = 0.$$

Proof. For n large enough we have

$$\begin{split} &\frac{1}{a_n}\int_{0}^{1}\lambda\left(\frac{y-u}{a_n}\right)dQ\left(u\right) = \frac{1}{a_n}\int_{0}^{1}\lambda\left(\frac{y-u}{a_n}\right)\frac{1}{f\left(F^{-1}\left(u\right)\right)}du\\ &= \int_{(y-1)/a_n}^{y/a_n}\lambda(v)\frac{1}{f\left(F^{-1}\left(y-a_nv\right)\right)}dv = \int_{-1/2}^{1/2}\lambda(v)\frac{1}{f\left(F^{-1}\left(y-a_nv\right)\right)}dv\\ &= \frac{1}{f\left(F^{-1}\left(y\right)\right)} - \int_{-1/2}^{1/2}\frac{a_n^2v^2}{2}\frac{f''\left(F^{-1}\left(\eta\right)\right)f\left(F^{-1}\left(\eta\right)\right) - 3\left(f'\left(F^{-1}\left(\eta\right)\right)\right)^2}{f^5\left(F^{-1}\left(\eta\right)\right)}dv, \end{split}$$

and our statement follows.

LEMMA 4. Under the conditions of Theorem 2 we have

$$\lim_{n\to\infty}\omega(n)\sup_{\varepsilon_n\leqslant y\leqslant 1-\varepsilon_n}\left|\frac{1}{a_n}\int\limits_0^1\lambda\left(\frac{y-u}{a_n}\right)d\left(Q_n(u)-Q(u)\right)\right|\stackrel{a.s.}{=}0.$$

Proof. Integrating by parts we obtain

$$\begin{aligned} \left| \frac{1}{a_n} \int_{0}^{1} \lambda\left(\frac{y-u}{a_n}\right) d\left(Q_n(u) - Q(u)\right) \right| &= \left| \frac{1}{a_n} \int_{0}^{1} \left(Q_n(u) - Q(u)\right) d\lambda\left(\frac{y-u}{a_n}\right) \right| \\ &= \left| \frac{1}{a_n^2 \sqrt{n}} \int_{0}^{1} \frac{\varrho_n(u)}{f\left(F^{-1}(u)\right)} \lambda'\left(\frac{y-u}{a_n}\right) du \right| \\ &= \frac{1}{a_n^2 \sqrt{n}} \left| \int_{y-a_2/2}^{y+a_n/2} \frac{\varrho_n(u)}{f\left(F^{-1}(u)\right)} \lambda'\left(\frac{y-u}{a_n}\right) du \right| \\ &\leqslant \frac{O(1)}{a_n^2 \sqrt{n}} \int_{y-a_n/2}^{y+a_n/2} \left| \frac{\varrho_n(u)}{f\left(F^{-1}(u)\right)} \right| du \\ &\leqslant \frac{O(1)}{a_n \sqrt{n}} \sup_{0 \le u \le 1} |\varrho_n(u)| \sup_{\varepsilon_n/2 \le u \le 1-\varepsilon_n/2} \frac{1}{|f\left(F^{-1}(u)\right)|} \end{aligned}$$

which, by Lemma B, proves our statement.

Clearly, Lemmas 3 and 4 imply Theorem 2.

4. Simultaneous confidence bounds for the quantile function via quantiledensity estimation. Theorems 1 and A together imply

CONSEQUENCE 1. Let $X_1, X_2, ...$ be i.i.d. r.v. with a continuous distribution function $F(\cdot)$ satisfying the conditions of Theorem 1. Then one can define a Brownian bridge $\{B_n(y); 0 \le y \le 1\}$ for each n such that

(4.1)
$$\sup_{\varepsilon_n \leq y \leq 1-\varepsilon_n} \left| \frac{1}{\varphi_n^{(0)}(y)} n^{1/2} (Q_n(y) - Q(y)) - B_n(y) \right| \stackrel{a.s.}{=} O\left(\frac{1}{\omega(n)}\right)$$

and

(4.2)
$$P\left\{Q_n(y) - n^{-1/2} \varphi_n^{(0)}(y) c \leq Q(y) \\ \leq Q_n(y) + n^{-1/2} \varphi_n^{(0)}(y) c; \varepsilon_n < y < 1 - \varepsilon_n\right\} \to K(c) \quad \text{as } n \to \infty,$$

where

$$\varphi_n^{(0)}(y) = \frac{Q_n(y+a_n) - Q_n(y-a_n)}{2a_n}$$

$$\begin{split} \omega(n) &= n^{\delta}, \ a_n = n^{-\alpha}, \ \varepsilon_n = n^{-\beta}, \ 3\beta + \delta < \alpha < 1/2, \ 2\delta + 4\beta + \alpha < 1, \ and \\ K(z) &= 1 - \sum_{k \neq 0} \ (-1)^{k+1} \exp(-2k^2 \, c^2), \quad c \ge 0, \end{split}$$

is the distribution function of $\sup_{0 \le y \le 1} |B(y)|$.

CONSEQUENCE 2. If the conditions of Theorem 2 are satisfied, then Consequence 1 will be true with the corresponding $\omega(n)$ of Theorem 2 when we replace $\varphi_n^{(0)}(y)$ by $\varphi_n(y)$ defined in (1.1).

Proof. Theorems 1 and A combined give (4.1), which implies (4.2) if we observe also that $\sup_{0 \le y \le \varepsilon_n} |B(y)|$ and $\sup_{1-\varepsilon_n \le y \le 1} |B(y)|$ tend to zero almost surely as $n \to \infty$ by the P. Lévy modulus of continuity for a Brownian bridge. Similarly, Theorems 2 and A imply Consequence 2.

The two-sided simultaneous confidence interval for Q(y) is of course not the only consequence of (4.1) with $\varphi_n^{(0)}$ (resp. with φ_n). Other statements of interest, such as various one-sided intervals, can be also based on it. For example, with $\varphi_n^{(0)}$ as in (4.2) we have

(4.3)
$$\lim_{n \to \infty} P\{Q_n(y) - n^{-1/2} \varphi_n^{(0)}(y) c \leq Q(y); \varepsilon_n < y < 1 - \varepsilon_n\} = \lim_{n \to \infty} P\{Q(y) \leq Q_n(y) + n^{-1/2} \varphi_n^{(0)}(y) c; \varepsilon_n < y < 1 - \varepsilon_n\} = P\{\sup_{0 \leq y \leq 1} B(y) \leq c\} = 1 - \exp(-2c^2), \quad c > 0,$$

and similar statements hold with φ_n of (1.1) replacing $\varphi_n^{(0)}$. Some further oneand two-sided confidence intervals for Q are discussed in a somewhat different approach of the next section.

While it is true that in this paper we estimated 1/f(Q(y)) mainly for the sake of gaining confidence intervals like (4.2) and (4.3), it may also appear that the using of $dQ_n(u)$ in (1.1) in order to get a one-step estimator for dQ(y) is more important than the gained confidence intervals themselves, for estimating dQ, the so-called quantile-density function, is of independent interest in statistics. Thus it appears that estimators like φ_n of (1.1) should be further studied. Also, the idea of using $dQ_n(y)$ in (1.1) suggests that the estimator of type $\varphi_n(y)$ can, in turn, be similarly used to get an estimate for J(y) = -d(f(Q(y))), the so-called score function of the density f. Namely, for d(f(Q(y))) one may wish to consider an estimate in terms of $d(1/\varphi_n(y))$ with $\varphi_n(y)$ as in (1.1), i.e., an estimator like, say,

4.4)
$$J_n(y) = -a_n^{-1} \int_0^1 \lambda\left(\frac{y-u}{a_n}\right) d\left(\frac{1}{\varphi_n(u)}\right)$$

for J(y). For further references and considerations of the quantile-density function and the score function of a density function and their estimation we refer to [6].

228

5. Direct simultaneous confidence bounds for the quantile function. First we recall the following version of Theorem A (cf. (5.1) of Theorem 6 in [3]):

THEOREM B. Let X_1, X_2, \ldots be i.i.d. r.v. with a continuous distribution function F and assume that F satisfies conditions (i), (ii), and (iii) of Theorem A. Let $\delta_n := 25n^{-1} \log \log n$. Then with the sequence of Brownian bridges $\{B_n(y); 0 \le y \le 1\}$ of Theorem A we have also

$$\sup_{\delta_n \leq y \leq 1-\delta_n} \left| f(Q(y)) n^{1/2} (Q_n(y) - Q(y)) - B_n(y) \right| \stackrel{a.s.}{=} O(n^{-1/2} \log n).$$

Next we prove two corollaries to Theorem B.

COROLLARY 1. Let $X_1, X_2, ...$ be i.i.d. r.v. with a continuous distribution function F, which is assumed to satisfy all the conditions of Theorem B. Let $\{B(y); 0 \le y \le 1\}$ be a Brownian bridge. Then

(5.1)
$$\lim_{n \to \infty} P\left\{Q(y) \leq Q_n(y+n^{-1/2}c); \varepsilon_n \leq y \leq 1-\varepsilon_n\right\}$$
$$= \lim_{n \to \infty} P\left\{Q_n(y-n^{-1/2}c) \leq Q(y); \varepsilon_n \leq y \leq 1-\varepsilon_n\right\}$$
$$= P\left\{\sup_{0 \leq y \leq 1} B(y) \leq c\right\} = 1 - \exp(-2c^2), \quad c > 0,$$

where $\varepsilon_n := n^{-1/2+\delta}$ with any $\delta \in (0, 1/2)$.

Proof. When calculating below, we use the more suggestive notation F^{-1} (resp. F_n^{-1}) for Q (resp. Q_n). For n = 1, 2, ... we consider the sequence of events

(5.2)
$$\{F_n^{-1}(y+n^{-1/2}c)-F^{-1}(y) \ge 0; \varepsilon_n \le y \le 1-\varepsilon_n\}$$

=
$$\bigcup_{n\varepsilon_n+1 \le k \le n(1-\varepsilon_n)} \{F_n^{-1}(y+n^{-1/2}c)-F^{-1}(y) \ge 0; (k-1)/n < y \le k/n\}.$$

Next, for each $k \in [n\varepsilon_n + 1, n(1 - \varepsilon_n)]$ and *n*, we consider

$$(5.3) \quad \left\{ F_n^{-1} (y + n^{-1/2} c) - F^{-1} (y) \ge 0; \frac{k-1}{n} < y \le \frac{k}{n} \right\}$$
$$= \left\{ F_n^{-1} (y + n^{-1/2} c) - F^{-1} (y + n^{-1/2} c) + F^{-1} (y + n^{-1/2} c) - F^{-1} (y) \ge 0; \frac{k-1}{n} < y \le \frac{k}{n} \right\}$$
$$= \left\{ f \left(F^{-1} \left(\frac{k}{n} + n^{-1/2} c \right) \right) \left(F_n^{-1} (y + n^{-1/2} c) - F^{-1} (y + n^{-1/2} c) \right) + n^{-1/2} c \frac{f \left(F^{-1} (k/n + n^{-1/2} c) \right)}{f \left(F^{-1} (\xi_n) \right)} \ge 0; \frac{k-1}{n} < y \le \frac{k}{n}, y < \xi_n < y + n^{-1/2} c \right\}$$

$$= \left\{ \varrho_n(y+n^{-1/2}c) + \left(\frac{f\left(F^{-1}\left(k/n+n^{-1/2}c\right)\right)}{f\left(F^{-1}\left(y+n^{-1/2}c\right)\right)} - 1 \right) \varrho_n(y+n^{-1/2}c) - c\left(\frac{f\left(F^{-1}\left(k/n+n^{-1/2}c\right)\right)}{f\left(F^{-1}\left(\xi_n\right)\right)} - 1 \right) \leqslant c; \\ \frac{k-1}{n} < y \leqslant \frac{k}{n}, \ y < \xi_n < y+n^{-1/2}c \right\} \\ := \left\{ G_n\left(y, \frac{k}{n}, \xi_n, c\right) \leqslant c; \ y < \xi_n < y+n^{-1/2}c, \ \frac{k-1}{n} < y < \frac{k}{n} \right\},$$

where

$$\begin{split} \varrho_n(y+n^{-1/2}c) &:= n^{1/2} f\left(F^{-1}\left(y+n^{-1/2}c\right)\right) \left(F_n^{-1}\left(y+n^{-1/2}c\right)-F^{-1}\left(y+n^{-1/2}c\right)\right). \end{split}$$
Putting (5.2) and (5.3) together, we get
(5.4) $\{F_n^{-1}\left(y+n^{-1/2}c\right)-F^{-1}\left(y\right) \ge 0; \ \varepsilon_n \le y \le 1-\varepsilon_n\}$
 $= \bigcup_{n\varepsilon_n+1 \le k \le n(1-\varepsilon_n)} \{G_n(y, k/n, \ \xi_n, c) \le c; \ y < \xi_n < y+n^{-1/2}c, (k-1)/n < y \le k/n\}$
 $= \{G_n(y, k/n, \ \xi_n, c) \le c; \ y < \xi_n < y+n^{-1/2}c, (k-1)/n < y < k/n, \ \varepsilon_n+1/n \le k/n \le 1-\varepsilon_n\}. \end{split}$

Hence the problem of verifying the first limit statement of (5.1) is equivalent to showing that

(5.5)
$$\lim_{n \to \infty} P\left\{G_n(y, k/n, \xi_n, c) \le c; y < \xi_n < y + n^{-1/2} c, (k-1)/n < y \le k/n, \varepsilon_n + 1/n \le k/n \le 1 - \varepsilon_n\right\} = P\left\{\sup_{0 \le y \le 1} B(y) \le c\right\}, \quad c > 0.$$

Let now $\{B_n(y); 0 \le y \le 1\}$ be that sequence of Brownian bridges for which Theorem B holds true. We are going to show now that

(5.6)
$$\sup_{\{\xi_n, y, k/n\}} |G_n(y, k/n, \xi_n, c) - B_n(y)| = o_P(1) \quad \text{as } n \to \infty,$$

where

$$\{\xi_n, y, k/n\} \\ := \{y < \xi_n < y + n^{-1/2} c, (k-1)/n < y \le k/n, \varepsilon_n + 1/n \le k/n \le 1 - \varepsilon_n\}.$$

First, by Theorem B we have

(5.7)
$$\sup_{\varepsilon_n \leq y \leq 1-\varepsilon_n} |\varrho_n(y+n^{-1/2}c) - B_n(y+n^{-1/2}c)| \stackrel{\text{a.s.}}{=} o(1) \quad \text{as } n \to \infty.$$

(5.8)
$$\sup_{0 \le y \le 1-n^{-1/2}c} |B(y+n^{-1/2}c) - B(y)| \stackrel{a.s.}{=} O(n^{-1/2}\log n),$$

and since

(5.9)
$$\{B_n(y); 0 \le y \le 1\} \stackrel{\mathcal{D}}{=} \{B(y); 0 \le y \le 1\}$$
 for $n = 1, 2, ...,$

by (5.8) we also obtain

(5.10)
$$\sup_{0 \le y \le 1-n^{-1/2}c} |B_n(y+n^{-1/2}c)-B_n(y)| = o_P(1) \quad \text{as } n \to \infty.$$

Combining now (5.7) and (5.10), we get

(5.11)
$$\sup_{e_n \leq y \leq 1-e_n} |\varrho_n(y+n^{-1/2}c) - B_n(y)| = o_P(1).$$

Next, we show that

(5.12)

$$\sup_{\{y,k/n\}} \left| \left(\frac{f(F^{-1}(k/n + n^{-1/2}c))}{f(F^{-1}(y + n^{-1/2}c))} - 1 \right) \varrho_n(y + n^{-1/2}c) \right| = o_P(1) \quad \text{as } n \to \infty,$$

where $\{y, k/n\} := \{(k-1)/n < y \le k/n, \varepsilon_n + 1/n \le k/n \le 1 - \varepsilon_n\}$. By (5.9) and (5.11) we obtain

(5.13)
$$\sup_{\varepsilon_n \leq y \leq 1-\varepsilon_n} |\varrho_n(y+n^{-1/2}c)| \xrightarrow{\mathscr{D}} \text{an r.v.} \quad \text{as } n \to \infty.$$

(In fact, the latter r.v. is $\sup_{0 \le y \le 1} |B(y)|$, but now this is of no interest to us). Consequently, in order to verify (5.12), it suffices to show that

(5.14)
$$\sup_{\{y,k/n\}} \left| \frac{f\left(F^{-1}\left(k/n+n^{-1/2}c\right)\right)}{f\left(F^{-1}\left(y+n^{-1/2}c\right)\right)} - 1 \right| = o(1) \quad \text{as } n \to \infty.$$

Now, by Lemma 1 in [3], for every pair $y_1, y_2 \in (0, 1)$ and γ as in (iii) of Theorem A we have

(5.15)
$$\frac{f(F^{-1}(y_1))}{f(F^{-1}(y_2))} \leq \left\{ \frac{y_1 \vee y_2}{y_1 \wedge y_2} \frac{1 - (y_1 \wedge y_2)}{1 - (y_1 \vee y_2)} \right\}^{\gamma}.$$

Hence

(5.16)
$$\sup_{\{y,k/n\}} \left| \frac{f\left(F^{-1}\left(k/n+n^{-1/2}c\right)\right)}{f\left(F^{-1}\left(y+n^{-1/2}c\right)\right)} - 1 \right| \\ \leq \sup_{\{y,k/n\}} \left\{ \left(\frac{f\left(F^{-1}\left(k/n+n^{-1/2}c\right)\right)}{f\left(F^{-1}\left(y+n^{-1/2}c\right)\right)} - 1 \right) \vee \left(\frac{f\left(F^{-1}\left(y+n^{-1/2}c\right)\right)}{f\left(F^{-1}\left(k/n+n^{-1/2}c\right)\right)} - 1 \right) \right\}$$

$$\leq 2 \sup_{\{y,k/n\}} \left\{ \left(\frac{k/n + n^{-1/2}c}{y + n^{-1/2}c} \frac{1 - y - n^{-1/2}c}{1 - k/n - n^{-1/2}c} \right)^{\gamma} - 1 \right\}$$

$$\leq 2 \sup_{\varepsilon_n \leq y \leq 1 - \varepsilon_n} \left\{ \left(\frac{y + 1/n + n^{-1/2}c}{y} \frac{1 - y}{1 - y - n^{-1/2}c} \right)^{\gamma} - 1 \right\}$$

$$\leq 2 \left\{ \left[(1 + 1/n\varepsilon_n + n^{-1/2}c/\varepsilon_n)(1 - n^{-1/2}c/\varepsilon_n)^{-1} \right]^{\gamma} - 1 \right\}$$

$$= 2 \left\{ \left[(1 + 1/n^{1/2 + \delta} + c/n^{\delta})(1 - c/n^{\delta})^{-1} \right]^{\gamma} - 1 \right\} \rightarrow 0 \quad \text{as } n \to \infty.$$

Hence, combining (5.16) and (5.13), we get (5.12).

Along the lines of (5.16) one can also show that

(5.17)
$$\sup_{\{\xi_n, y, k/n\}} \left| \frac{f\left(F^{-1}\left(k/n + n^{-1/2} c\right)\right)}{f\left(F^{-1}\left(\xi_n\right)\right)} - 1 \right| = o(1) \quad \text{as } n \to \infty,$$

and (5.12), (5.13), (5.17) together imply (5.6). Now (5.5) follows immediately from the latter after using (5.9) and observing also that sup |B(y)| and

 $\sup_{\substack{1-\varepsilon_n \leq y \leq 1 \\ \text{of continuity for a Brownian bridge. This also completes the proof of the first limit statement of (5.1).} 0 \le y \le \varepsilon_n$

Mutatis mutandis, the above proof also gives the second limit statement of (5.1) and Corollary 1 is proved.

A combination of lines like those of the proof of Corollary 1 yields

COROLLARY 2. Let X_1, X_2, \ldots be i.i.d. r.v. with a continuous distribution function F, which is assumed to satisfy all the conditions of Theorem B. Then

(5.18)
$$\lim_{n \to \infty} P\left\{Q_n(y - n^{-1/2}c) \le Q(y) \le Q_n(y + n^{-1/2}c); \, \varepsilon_n \le y \le 1 - \varepsilon_n\right\} = P\left\{\sup_{0 \le y \le 1} |B(y)| \le c\right\} = K(c), \quad c > 0,$$

where $\varepsilon_n := n^{-1/2+\delta}$ with any $\delta \in (0, 1/2)$ and $K(\cdot)$ is the distribution function as given in (4.2).

Given Corollaries 1 and 2, we can now get correct versions of statements like (1.5), (1.6), and (1.8) under the conditions of Theorem B. For example, with $c = c_1(\alpha)$ as in (1.2) it follows from Corollary 1 that for large *n* a correct version of (1.5) is

(5.19)
$$P\left\{Q\left(y\right) \leqslant Q_n\left(y+n^{-1/2}c_1\left(\alpha\right)\right); \varepsilon_n \leqslant y \leqslant 1-\varepsilon_n\right\} = 1-\alpha,$$

i.e., under the conditions of Theorem B, when estimating the y-th quantile of F by the $(y+n^{-1/2}c_1(\alpha))$ -th quantile of F_n , the probability that no true quantile of F (for quantiles above $n^{-1/2+\delta}$ and below $1-n^{-1/2+\delta}$ with any $\delta \in (0, 1/2)$) exceeds the estimated quantile is $1-\alpha$.

A similar correct version and interpretation of (1.6) follows again from Corollary 1, and that of (1.8) from Corollary 2. The statement of (5.19) also corrects (2) of Alexander [1].

The next three corollaries to Theorem B give some further useful simultaneous confidence bounds for theoretical quantiles.

COROLLARY 3. Let X_1, X_2, \ldots be i.i.d. r.v. with a continuous distribution function F, which is assumed to satisfy all the conditions of Theorem B. Let a_1 and a_2 ($0 < a_1 < a_2 < 1$) be fixed. Then

(5.20)
$$\lim_{n \to \infty} P\left\{Q_n(y - n^{-1/2}c) \le Q(y); a_1 \le y \le a_2\right\}$$
$$= \lim_{n \to \infty} P\left\{Q(y) \le Q_n(y + n^{-1/2}c); 1 - a_2 \le y \le 1 - a_1\right\}$$
$$= P\left\{\sup_{a_1 \le y \le a_2} B(y) \le c\right\}, \quad c > 0,$$

where

(5.21)

$$P\{\sup_{a_1 \leq y \leq a_2} B(y) \leq c\} = (2\pi)^{-1/2} \int_{-c(a_1(1-a_1))^{-1/2}}^{\infty} \exp(-x^2/2) H(x; a_1, a_2, c) dx$$

with

$$\begin{split} H(x; a_1, a_2, c) \\ &= \varPhi \left[c \left(\frac{a_2 - a_1}{(1 - a_1)(1 - a_2)} \right)^{1/2} + \left(x a_1^{1/2} + \frac{c}{(1 - a_1)^{1/2}} \right) \left(\frac{1 - a_2}{a_2 - a_1} \right)^{1/2} \right] - \\ &- \exp \left[- 2c \left(x a_1^{1/2} + c \left(1 - a_1 \right)^{-1/2} \right) (1 - a_1)^{-1/2} \right] \varPhi \left[c \left(\frac{a_2 - a_1}{(1 - a_1)(1 - a_2)} \right)^{1/2} - \\ &- \left(x a_1^{1/2} + \frac{c}{(1 - a_1)^{1/2}} \right) \left(\frac{1 - a_2}{a_2 - a_1} \right)^{1/2} \right]. \end{split}$$

Here, and also in the sequel, $\Phi(\cdot)$ is the unit normal distribution function. The proof of (5.20) goes along the lines of the proof of Corollary 1. Formula (5.21) is taken from [2], where further references are also given. Since

$$\sup_{a_1 \leq y \leq a_2} B(y) \leq \sup_{0 \leq y \leq 1} B(y)$$

for any fixed a_1 and a_2 ($0 < a_1 < a_2 < 1$), given $1 - \alpha$, $c(\alpha)$ of (5.20) is smaller than that of (5.1). Consequently, if we are interested in constructing lower or upper simultaneous confidence intervals for quantiles in the "middle" only, then (5.20) gives a more economical handle than (5.1). If $a_1 = \varepsilon_n$ and $a_2 = 1 - \varepsilon_n$ with ε_n as in Corollary 1, then Corollary 3 gives back the former.

If instead of shorter intervals in the "middle" only, we were interested in constructing shorter lower or upper bounds for quantiles on the tails only, then combining Corollaries 3 and 1 we get (with ε_n as in Corollary 1)

COROLLARY 4. With $a_1 = \varepsilon_n$ and a_2 fixed as before, under the conditions of Theorem B we have

(5.22)
$$\lim_{n \to \infty} P\left\{Q_n(y - n^{-1/2}c) \leq Q(y); \varepsilon_n \leq y \leq a_2\right\}$$
$$= \lim_{n \to \infty} P\left\{Q(y) \leq Q_n(y + n^{-1/2}c); (1 - a_2 \leq y \leq 1 - \varepsilon_n\right\}$$
$$= P\left\{\sup_{0 \leq y \leq a_2} B(y) \leq c\right\}, \quad c > 0,$$

and if $a_2 = 1 - \varepsilon_n$ and a_1 is fixed as before, then

(5.23)
$$\lim_{n \to \infty} P\{Q_n(y - n^{-1/2} c) \leq Q(y); a_1 \leq y \leq 1 - \varepsilon_n\} \\ = \lim_{n \to \infty} P\{Q(y) \leq Q_n(y + n^{-1/2} c); \varepsilon_n \leq y \leq 1 - a_1\} \\ = P\{\sup_{a_1 \leq y \leq 1} B(y) \leq c\}, \quad c > 0,$$

where

(5.24)
$$P\left\{\sup_{0 \le y \le a_2} B(y) \le c\right\}$$
$$= \Phi\left[c\left(a_2(1-a_2)\right)^{-1/2}\right] - \exp(-2c^2)\Phi\left[\frac{(2a_2-1)c}{(a_2(1-a_2))^{1/2}}\right], \quad c > 0,$$

and

(5.25)
$$P\left\{\sup_{a_1 \leq y \leq 1} B(y) \leq c\right\}$$
$$= \Phi\left[c\left(a_1(1-a_1)\right)^{-1/2}\right] - \exp(-2c^2) \Phi\left[\frac{(1-2a_1)c}{(a_1(1-a_1))^{1/2}}\right], \quad c > 0.$$

Here (5.24) results from (5.21) after putting $a_1 = 0$, and (5.25) is (5.21) with $a_2 = 1$. The second line of (5.22) with $c = c(\alpha)$ defined by

$$P\left\{\sup_{0\leq y\leq a_2}B(y)\leq c(\alpha)\right\}=1-\alpha$$

also corrects (4) of [1], while (5.24) is a correct version of (7) in [1].

In case we were interested in shorter simultaneous upper or lower confidence bounds for quantiles on the lower and upper tails at the same time, the next Corollary to Theorem B is useful.

COROLLARY 5. Let a_1 and a_2 ($0 < a_1 \le a_2 < 1$) be fixed. Then under the conditions of Theorem B and with ε_n as in Corollary 1 we have

Simultaneous confidence bounds

(5.26)
$$\lim_{n \to \infty} P \{Q_n(y - n^{-1/2}c) \le Q(y); \varepsilon_n \le y \le a_1, a_2 \le y \le 1 - \varepsilon_n\} \\ = \lim_{n \to \infty} P \{Q(y) \le Q_n(y + n^{-1/2}c); \varepsilon_n \le y \le 1 - a_2, 1 - a_1 \le y \le 1 - \varepsilon_n\} \\ = P \{\sup_{0 \le y \le a_1, a_2 \le y \le 1} B(y) < c\}, \quad c > 0,$$

where

(5.27)
$$P\left\{\sup_{0 \le y \le a_1, a_2 \le y \le 1} B(y) < c\right\} = (2\pi)^{-1/2} \int_{-c(a_2(1-a_2))^{1/2}} \exp(-x^2/2) \times \left\{1 - \exp\left[-2c\left(x(a_2(1-a_2))^{1/2} + c\right)(1-a_2)^{-1}\right]\right\} H(x; a_1, a_2, c) dx\right\}$$

with

$$H(x; a_1, a_2, c) = \Phi\left(\frac{xa_1(1-a_2)^{1/2} + ca_2}{(a_1 a_2(a_2-a_1))^{1/2}}\right) - \exp\left[-2c\left(x(a_2(1-a_2))^{1/2} + c(a_2-a_1)^{1/2}\right) + c(a_2-a_2)^{1/2} + c(a_2-a_2)^{1/2}\right)\right]$$

The proof of (5.26) of Corollary 5 is like that of Corollary 1 combined with that of Corollary 4. Formula (5.27) is taken from [2].

Our reason for reproducing formulae (5.21), (5.24), (5.25), and (5.27) here is due to inaccessibility of Csáki's Hungarian paper [2] to many readers and to the fact that on occasions one sees new attempts to rederive them. Csáki's quoted paper contains an aboundance of further useful formulae. As to the ones quoted here, (5.24) and (5.25) are immediately calculable. It would be of some interest to tabulate also (5.21) and (5.27). For example, a statistical interpretation of (5.26) with $c = c(\alpha)$ defined by

$$P\left\{\sup_{0 \le y \le a_1, a_2 \le y \le 1} B(y) < c(\alpha)\right\} = 1 - \alpha$$

in (5.27) results in

(5.28)
$$P\left\{Q_n\left(y-n^{-1/2}c\left(\alpha\right)\right) \leqslant Q\left(y\right); \varepsilon_n \leqslant y \leqslant a, a_2 \leqslant y \leqslant 1-\varepsilon_n\right\}$$
$$= P\left\{Q\left(y\right) \leqslant Q_n\left(y+n^{-1/2}c\left(\alpha\right)\right); \varepsilon_n \leqslant y \leqslant 1-a_2, 1-a_1 \leqslant y \leqslant 1-\varepsilon_n\right\}$$
$$= 1-\alpha \quad \text{for large } n,$$

i.e., the first line of (5.28) states that, given the conditions of Theorem B and when estimating the y-th quantile of F by the $(y-n^{-1/2}c(\alpha))$ -th quantile of F_n , the probability that no true y-th quantile of F descends the estimated quantile for $y \in [\varepsilon_n, a_1]$ and for $y \in [a_2, 1-\varepsilon_n]$ is $1-\alpha$, while the second line of (5.28) states that when estimating the y-th quantile of F by the $(y + n^{-1/2}c(\alpha))$ -th quantile of F_n , the probability that no true y-th quantile of F exceeds the estimated quantile for $y \in [\varepsilon_n, 1-a_2]$ and for $y \in [1-a, 1-\varepsilon_n]$ is also $1-\alpha$.

8 - Prob. Math. Statist. 4 (2)

On purely mathematical grounds the confidence intervals of this section, which hold under the conditions of Theorem B, are preferable to those of Section 4, due to the fact that for the validity of Theorems 1 and 2 we had to assume further conditions in addition to those of Theorem B. On the other hand, as already mentioned before, Theorems 1 and 2 are hoped to be of independent interest.

REFERENCES

- [1] Ch. Alexander, Simultaneous confidence bounds for the tail of an inverse distribution, Ann. Statist. 8 (1980), p. 1391-1394.
- [2] E. Csáki, On tests based on empirical distribution functions (in Hungarian), Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. 23 (1974), p. 239-327.
- [3] M. Csörgő and P. Révész, Strong approximations of the quantile process, Ann. Statist. 6 (1978), p. 882-894.
- [4] Quantile processes and sums of weighted spacings for composite goodness-of-fit, II, Carleton Mathematical Lecture Note, No. 27 (1980).
- [5] Strong approximations in probability and statistics, Academic Press, New York 1981.
- [6] E. Parzen, Nonparametric statistical data modeling, J. Amer. Statist. Assoc. 74 (1979), p. 105-131.

Department of Mathematics and Statistics Carleton University Ottawa, Canada K1S 5B6 Mathematical Institute of the Hungarian Academy of Sciences H-1053 Budapest Reáltanoda u. 13-15 Hungary

Received on 25. 1. 1982