Strongly-consistent, distribution-free confidence intervals for quantiles

David Gilat^{a,1,*}, T.P. Hill^{b,1,2}

^a *School of Mathematical Sciences, Tel Aviv University, Israel*

^b *School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA*

Abstract

Strongly-consistent, distribution-free confidence intervals are derived to estimate the fixed quantiles of an arbitrary unknown distribution, based on order statistics of an iid sequence from that distribution. This new method, unlike classical estimates, works for totally arbitrary (including discontinuous) distributions, and is based on recent one-sided strong laws of large numbers.

Keywords: Quantiles; Quantile intervals; Quantile estimators; One-sided strong laws; Strongly-consistent estimators; Confidence intervals; Order statistics; Empirical distribution function

1. Introduction

The subject of this note is the basic statistical problem (e.g., David, 1981, Section 2.5) of estimating a fixed quantile of an unknown distribution F based on a random sample of iid observations $X_1, X_2, \dots$ from that distribution. More precisely, a *random closed interval* $I_{\beta,n} = I_{\beta,n}(X_1, \dots, X_n)$ is sought to estimate the β -quantile interval for F :

$$I_{\beta}(F) = [q_F(\beta), Q_F(\beta)],$$

where q_F and Q_F are the lower and upper inverses of F given by

$$q_F(\beta) = \inf\{x \in \mathbb{R} : F(x) \geq \beta\}, \quad \beta \in (0, 1),$$

$$Q_F(\beta) = \sup\{x \in \mathbb{R} : F(x) \leq \beta\}, \quad \beta \in (0, 1),$$

* Corresponding author.

¹ Partially supported by U.S.–Israel Binational Science Foundation Grant 88-00005.

² Partially supported by National Science Foundation Grants DMS 89-01267 and DMS-92-03524 and Dutch National Science Foundation (NWO) Dossier B-61-281.

and $q_F(0) = Q_F(0) = \text{ess inf } F$, $q_F(1) = Q_F(1) = \text{ess sup } F$ (where $\text{ess inf } F$ and $\text{ess sup } F$ are the infimum and supremum of the support of the measure F). (That is, $q_F(\beta)$ is the smallest β -quantile of F , and $Q_F(\beta)$ is the largest.)

Among the desirable properties of such quantile confidence intervals $I_{\beta,n}$ are that they be (where here and throughout, $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ are the *order statistics* of $\{X_1, \dots, X_n\}$):

permutation-invariant (i.e., functions only of $X_{1:n}, \dots, X_{n:n}$);
linear ($I_{\beta,n}(aX_1 + b, \dots, aX_n + b) = aI_{\beta,n}(X_1, \dots, X_n) + b$);
distribution-free (not depending on assumptions or parameters of F);
sample-ranged ($I_{\beta,n}(X_1, \dots, X_n) \subset [X_{1:n}, X_{n:n}]$).

Remarks. Note that there are no implications among these properties, and that each property is defined in terms of a *single* β . There are also other desirable properties such as those relating different β 's (e.g., *monotone*: if $I_{\beta_1,n} = [a_1, b_1]$ and $I_{\beta_2,n} = [a_2, b_2]$ for some $\beta_1 < \beta_2$, then $a_1 \leq a_2$ and $b_1 \leq b_2$), but these will not be addressed here.

Letting F_n denote the *empirical distribution function*

$$F_n(x) = \frac{\#\{i \leq n : X_i \leq x\}}{n}$$

and $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denote the integer-valued functions

$$\lceil x \rceil = \min\{j \in \mathbb{Z} : j \geq x\} \quad \text{and} \quad \lfloor x \rfloor = \max\{j \in \mathbb{Z} : j \leq x\},$$

the *smallest sample- β -quantile* of the random sample $X_1, \dots, X_n$ is

$$q_{F_n}(\beta) = X_{\lceil n\beta \rceil : n},$$

and the *largest sample- β -quantile* is

$$Q_{F_n}(\beta) = X_{\lfloor n\beta \rfloor + 1 : n}$$

(note that $\lceil n\beta \rceil = \lfloor n\beta \rfloor$ iff $n\beta \in \mathbb{Z}$, and otherwise $\lceil n\beta \rceil = \lfloor n\beta \rfloor + 1$).

In these terms, the classical estimators of the unknown quantile interval $I_\beta(F)$ are the random closed intervals $I_{\beta,n}$ of the forms

$$[q_{F_n}(\beta), Q_{F_n}(\beta)], \tag{1}$$

$$[q_{F_n}(\beta) - d_n, Q_{F_n}(\beta) + d_n], \tag{2}$$

$$\left[q_{F_n} \left(\beta - \frac{c}{\sqrt{n}} \right), Q_{F_n} \left(\beta + \frac{c}{\sqrt{n}} \right) \right]. \tag{3}$$

The estimator (1) is simply the sample β -quantile interval, that is, the quantile interval for the random probability distribution F_n . Even for an underlying F which is Uniform $[0, 1]$, the intervals in (1) do not satisfy

$$P(I_{\beta,n} \supset I_\beta(F)) \rightarrow 1 \tag{4}$$

and in fact for all irrational β , $P(I_{\beta,n} \supset I_\beta(F)) = 0$ for all n , since $q_{F_n}(\beta) = Q_{F_n}(\beta)$, so $I_{\beta,n}$ as well as $I_\beta(F)$ are singletons. (For this Uniform $[0, 1]$ F , even though (4) fails, both $q_{F_n}(\beta)$ and $Q_{F_n}(\beta)$ converge to $I_\beta(F) = \beta$ a.s.; see Example 3.1 of Gilat and Hill (1992) for a continuous F where both $q_{F_n}(\frac{1}{2})$ and $Q_{F_n}(\frac{1}{2})$ diverge a.s.).

The problem with the estimators in (1) is that the intervals are “too narrow,” and two standard ways of correcting for this are using larger confidence intervals such as those in (2) and (3), both of which *assume continuity* of the underlying distribution F . Those in (2) however are neither sample-ranged nor linear, and in practice (cf. Csörgő (1983) and Shorack and Wellner (1986) where (2) is used to obtain confidence bands for the quantile function) the factors d_n often depend on F (e.g., $d_n = c_x/\sqrt{n}f(q_F(\beta))$) and rely on further assumptions about F (for example, that F has a density, namely f). Even if F is of this form (e.g., $U[0, 1]$), it may be checked using the Hewitt–Savage zero–one law that for *any* $c > 0$, $q_{F_n}(\frac{1}{2}) - (c/\sqrt{n}) > \frac{1}{2} = q_F(\frac{1}{2}) = Q_F(\frac{1}{2})$ i.o. almost surely, so with probability one the confidence intervals in (2) will be disjoint from the true median infinitely often. The objective of method (2), however, is simply to cover the quantile with probability $1 - \alpha$, and with regularity assumptions on F and a proper choice of d_n , the weaker form of (4) $P(I_{\beta,n} \supset I_\beta(F)) \geq 1 - \alpha$ will hold.

The objective of (3) also is to cover the quantile with high probability, and assuming the appropriate regularity conditions on F , this method can also be used (cf. Csörgő, 1983; Csörgő and Horváth, 1989) to construct *confidence bands* for the quantile function. These estimators in (3) (where c is a parameter depending on the desired level of significance) correspond in the median case to “the very simple rule of thumb: for $n > 10$ an approximate $1 - \alpha$ confidence interval for the median is obtained by counting off $\frac{1}{2} \sqrt{nu_\alpha}$ observations to the left and right of the sample median and rounding out to the next integer, where u_α is the upper $\alpha/2$ significance point of a unit normal variate,” (David, 1981, p. 16). But as in the case for (2), even when F is $U[0, 1]$ the confidence intervals in (3) will be disjoint from the true median interval infinitely often with probability one, as follows in this case from the law of the iterated logarithm.

In contrast to these classical quantile-estimators, the main purpose of this note is to derive a new and different easily-calculated class of quantile confidence intervals which not only is permutation-invariant, linear, sample-ranged, and completely distribution-free (even continuity is not assumed), but which also is *strongly-consistent* in the following sense.

Definition. Fix $\beta \in [0, 1]$. Confidence intervals $I_{\beta,n}$ are *strongly-consistent* estimators of the quantile interval $I_\beta(F)$ if

$$\lim_{n \rightarrow \infty} I_{\beta,n} = I_\beta(F) \quad \text{a.s.} \quad (5)$$

Remarks. The limit in (5) is the strong set-theoretic one; it says that $I_{\beta,n}$ converges to $I_\beta(F)$ as *sets*, i.e., $\limsup I_{\beta,n} = \liminf I_{\beta,n} = I_\beta(F)$. Observe that for compact intervals $\{I_n\}$ and I , $\lim I_n = I$ implies that for some N , $I_n \supset I$ for all $n \geq N$. Thus, (5) also implies that $I_{\beta,n} \supset I_\beta(F)$ eventually a.s. (i.e., for F -almost all ω , there exists $N = N_\omega$ so that $I_{\beta,n}(X_1(\omega), \dots, X_n(\omega)) \supset I_\beta(F)$ for all $n \geq N$), so (5) also implies (4). In general, the opposite inclusion $I_{\beta,n} \subset I_\beta(F)$ is *not* possible with positive probability; consider again the case F is $U[0, 1]$ and $\beta = \frac{1}{2}$, so $I_\beta(F) = \{\frac{1}{2}\}$. In general, strong consistency at the endpoints $\beta = 0, 1$ is impossible since $\text{ess inf } F$ and $\text{ess sup } F$ may have zero probability of being observed but as will be shown below, strong consistency at all other β is possible.

2. Main theorem and examples

Recall that the classical quantile-estimators (1), (2), (3), although defined in terms of the empirical distribution (sample) quantiles, may also be expressed in terms of the order statistics since, e.g., $q_{F_n}(\beta) = X_{[n\beta]:n}$. In the following definition new quantile estimators are defined directly in terms of the order statistics, and later it will be seen ((7) below) that they are in fact exactly quantiles for certain alternative empirical distribution functions.

Definition. Fix $\gamma \in (2, \frac{5}{2})$. For each $\beta \in (0, 1)$ and positive integer n , let (suppressing the dependence on γ)

$$\begin{aligned} a &= n^{\gamma-1} + \frac{n}{2} - \left[\left(n^{\gamma-1} + \frac{n}{2} \right)^2 - 2\beta n^\gamma \right]^{1/2}, \\ b &= \left[\left(n^{\gamma-1} - \frac{n}{2} \right)^2 + 2\beta n^\gamma \right]^{1/2} - n^{\gamma-1} + \frac{n}{2}, \\ I_{\beta,n}^* &= [X_{[a]:n}, X_{[b]+1:n}]. \end{aligned} \tag{6}$$

Remark. Although it may be checked directly that $1 \leq [a] \leq [b] + 1 \leq n$, this will also follow easily from the alternate representation (7) of $I_{\beta,n}^*$ below.

Example 2.1. For $\gamma = \frac{9}{4}$, $\beta = \frac{1}{2}$, $n = 100$, then $a \cong 46.07$, $b \cong 53.93$ and $I_{1/2,100}^* = [X_{47:100}, X_{54:100}]$.

The following theorem is the main result of this article; Φ is the df for $N(0, 1)$, $p_n \approx q_n$ means $(1 - p_n)/(1 - q_n) \rightarrow 1$, and recall that λ is Lebesgue measure.

Theorem 1. Fix $\gamma \in (2, \frac{5}{2})$. Then the random intervals $I_{\beta,n}^*$ are permutation-invariant, linear, sample-ranged, distribution-free and strongly-consistent estimators of the unknown quantile intervals $I_\beta(F)$ for all probability distributions F and for all $\beta \in (0, 1)$. Moreover, the consistency $I_{\beta,n} \rightarrow I_\beta$ is such that

- (i) $\lambda(I_{\beta,n}^* \Delta I_\beta) = O(n^{2-\gamma})$ a.s. for (Lebesgue) almost all $\beta \in [0, 1]$;
- (ii) $P(I_{\beta,n}^* \supset I_\beta) \geq \sum_{j=[a]}^{[b]} \binom{n}{j} \beta^j (1 - \beta)^{n-j}$ for all $\beta \in (0, 1)$; and, asymptotically,
- (ii)' $P(I_{\beta,n}^* \supset I_\beta) \approx 2\Phi((\beta^2/\sqrt{\beta(1-\beta)})n^{5/2-\gamma}) - 1$ for all $\beta \in (0, 1)$.

The rates in (i) and (ii)' and the bound in (ii) are sharp and attained.

Remarks. From (i) and (ii)' it can be seen that the interval $(2, \frac{5}{2})$ for γ is sharp. Choice of γ near 2 will increase the rate in (ii)' at which $P(I_{\beta,n}^* \supset I_\beta(F))$ goes to 1, but will do so at the expense of using large intervals (in the sense of (i)), whereas for γ near $\frac{5}{2}$, the opposite occurs. Observe that the rate in (i) is scale-invariant, since replacing a sample $X_1, \dots, X_n$ by $cX_1, \dots, cX_n$ results, by linearity of $I_{\beta,n}^*$ and $I_\beta(F)$, in a new difference in intervals exactly c times the old. That conclusion (i) may fail for some F and β is shown by the next example.

Example 2.2. Let F be the continuous df given by $F(x) = (x^3 + 1)/2$ for $-1 \leq x \leq +1$, and let $\beta = \frac{1}{2}$, so $Q_F(\beta) = 0$. By Lemma 3.3 below, $\hat{F}_n(x) \leq F(x) - (1/16n^{\gamma-2})$ infinitely often a.s. for all x near zero, so infinitely often with probability one,

$$\begin{aligned} Q_{\hat{F}_n} \left(\frac{1}{2} \right) &= \sup \left\{ x : \hat{F}_n(x) \leq \frac{1}{2} \right\} \geq \sup \left\{ x : F(x) - \frac{1}{16n^{\gamma-2}} \leq \frac{1}{2} \right\} \\ &= \sup \left\{ x : (1 + x^3)/2 \leq \frac{1}{2} + \frac{1}{16n^{\gamma-2}} \right\} = \frac{1}{2n^{(\gamma-2)/3}}, \end{aligned}$$

which is not $O(1/n^{\gamma-2})$.

The proof of Theorem 1 will be given in the following section; first a number of examples will be given.

Tables 1–3 allow comparisons for various choices of γ , β and n . (Dependence on n in the order statistics is suppressed, so $X_{(47)} = X_{47:100}$ if $n = 100$.)

Table 1
($\gamma = 2.25$)

n	$\beta = 0.5$		$\beta = 0.55$		$\beta = 0.9$	
	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$
100	$[X_{(47)}, X_{(54)}]$	0.8860	$[X_{(52)}, X_{(59)}]$	0.9464	$[X_{(89)}, X_{(92)}]$	1.0000
250	$[X_{(118)}, X_{(133)}]$	0.9532	$[X_{(130)}, X_{(146)}]$	0.9844	$[X_{(222)}, X_{(228)}]$	1.0000
1000	$[X_{(478)}, X_{(523)}]$	0.9948	$[X_{(528)}, X_{(572)}]$	0.9994	$[X_{(892)}, X_{(908)}]$	1.0000

Table 2
($\gamma = 2.05$)

n	$\beta = 0.5$		$\beta = 0.55$		$\beta = 0.9$	
	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$
100	$[X_{(41)}, X_{(60)}]$	1.0000	$[X_{(46)}, X_{(65)}]$	1.0000	$[X_{(85)}, X_{(93)}]$	1.0000
250	$[X_{(103)}, X_{(148)}]$	1.0000	$[X_{(114)}, X_{(160)}]$	1.0000	$[X_{(214)}, X_{(232)}]$	1.0000
1000	$[X_{(415)}, X_{(586)}]$	1.0000	$[X_{(463)}, X_{(633)}]$	1.0000	$[X_{(857)}, X_{(925)}]$	1.0000

Table 3
($\gamma = 2.45$)

n	$\beta = 0.5$		$\beta = 0.55$		multicolumn2 $\beta = 0.9$	
	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$	$I_{\beta,n}^*$	$P(I_{\beta,n}^* \supset I_\beta)$
100	$[X_{(49)}, X_{(52)}]$	0.4714	$[X_{(54)}, X_{(57)}]$	0.5560	$[X_{(90)}, X_{(91)}]$	0.9994
250	$[X_{(123)}, X_{(128)}]$	0.4900	$[X_{(135)}, X_{(141)}]$	0.5838	$[X_{(225)}, X_{(226)}]$	0.9999
1000	$[X_{(495)}, X_{(506)}]$	0.5200	$[X_{(545)}, X_{(556)}]$	0.6098	$[X_{(898)}, X_{(902)}]$	1.0000

The entries in Tables 1–3 result from simple direct calculations from the definitions of $I_{\beta,n}^*$ and the probability in Theorem 1(ii) (or its approximation in (ii)'). The next example is a typical “inverse” problem; in this case γ is taken in the midpoint of its range; the selection of a “good” γ is another type of numerical inverse problem left to the interested reader.

Example 2.3. Let $\gamma = \frac{9}{4}$. To find the smallest-sample sized 95% confidence interval for the median (interval) of an unknown distribution (i.e., to find the smallest n , and $I_{1/2,n}^*$ so that $P(I_{1/2,n}^* \supset I_{1/2}(F)) \geq 0.95$ for all probability distributions F), there are two approaches. To solve the problem *exactly*, it is necessary (since the inverse df for binomial distributions are not available in closed form) to calculate the interval $I_{1/2,n}^*$ and resulting probability for various choices of n , and then to zero in on the smallest n . On the other hand, to find a close approximation, use Theorem 1(ii)' (with $\beta = \frac{1}{2}$, $\gamma = \frac{9}{4}$) to set $\Phi(n^{1/4}/2) \geq 0.975$, and from the normal tables find $n \geq 238$. Then direct calculations yield $I_{1/2,238}^* = [X_{(111)}, X_{(126)}]$.

Although Theorem 1(i) yields orders of magnitude of how much larger $I_{\beta,n}^*$ is than $I_\beta(F)$, an alternate more explicit comparison can be made by comparing the confidence intervals for a given quantile β with those for a nearby quantile $\hat{\beta}$. For example, from Table 1 with $n = 1000$ it is seen that (with $\gamma = \frac{9}{4}$) the 95% confidence intervals for $\beta = 0.5$ and for $\beta = 0.55$ are disjoint, which is an additional measure of how “tight” $I_{\beta,n}^*$ is.

The bounds given in Theorem 1 are significantly sharper than those in Gilat and Hill (1993), which were obtained via moments using Chebyshev's inequality. For example, the bound corresponding to Theorems 1(ii) and (ii)' was

$$1 - [1 + \frac{1}{9} (1 - 1/n)^2 n^{5-2\gamma} \beta (1 - \beta)]^{-1}$$

which for $\gamma = \frac{9}{4}$, $\beta = \frac{1}{2}$ and $n = 1000$ yields a confidence of 0.467, as compared to the confidence $2\Phi(2.81) - 1 = 0.995$ from Theorem 1(ii)'.

3. Proof of Theorem 1

Recall the definitions of the constants $a = a(\beta, n, \gamma)$ and $b = b(\beta, n, \gamma)$ underlying the definition of the quantile-estimators $I_{\beta,n}^*$ in (6).

Lemma 3.1. (i) $a(\beta, n, \gamma) = n\beta - \frac{1}{2}\beta^2 n^{3-\gamma} + \frac{1}{2}\beta^3 n^{5-2\gamma} + o(n^{5-2\gamma})$;
(ii) $b(\beta, n, \gamma) = n\beta + \frac{1}{2}\beta^2 n^{3-\gamma} - \frac{1}{2}\beta^3 n^{5-2\gamma} + o(n^{5-2\gamma})$.

Proof. Letting $a = c(1 - \sqrt{1-x})$, where $c = n^{\gamma-1} + n/2$ and $x = 2\beta n^\gamma/c^2$, expand a as a Taylor's series in x about zero. Similarly for b . $\square$

Remarks. From these expansions it can also be seen why the interval $(2, \frac{5}{2})$ for γ is sharp. Both endpoints a and b defining the quantile-estimator $I_{\beta,n}^*$ should be of order of magnitude $n\beta$, hence $\gamma > 2$ from the second-order term, and both need to have third-order terms going to infinity in order to keep the endpoints in $\{1, 2, \dots, n\}$, so $5 > 2\gamma$. For the midpoint case $\gamma = \frac{9}{4}$, the expansions for a and b are $n\beta \mp \frac{1}{2}\beta^2 n^{3/4} \pm \frac{1}{2}\beta^3 n^{1/2} + o(n^{1/2})$. The second-order term $n^{3/4}$ overrides the law of iterated logarithm oscillations to allow consistency, and without higher-order terms the estimators $I_{\beta,n}^*$ would not be well defined for all β , as was the case for the classical estimator in (3). Using the expansions in Lemma 3.1, it is easy to check, for example, that for $n = 100$, as $\gamma \searrow 2$ the limiting estimator is $[X_{(44)}, X_{(57)}]$ and as $\gamma \nearrow \frac{5}{2}$ is $[X_{(49)}, X_{(52)}]$, which may be compared with the values in Tables 1–3. In addition, these expansions in Lemma 3.1 may be used to solve inverse “overlap” problems as in the following example.

Example 3.2. Let $\gamma = \frac{9}{4}$. To find the smallest n so that the median and 0.55 quantile estimators do not overlap (i.e., $I_{0.5,n}^* \cap I_{0.55,n}^*$ consists of at most one point), set

$$b(\beta, n, \gamma) = b(0.5, n, \frac{9}{4}) < a(0.55, n, \frac{9}{4}) = a(\beta, n, \gamma)$$

and use the expansions in Lemma 3.1 to deduce $n \geq 929$.

The next definition recalls the *refined empirical distribution functions* introduced in Gilat and Hill (1992), based on non-uniformly weighted convex combinations of the order statistics (cf. Stigler, 1974). Recall that F_n is the classical empirical df.

Definition. For each $\gamma \in (2, \frac{5}{2})$ and each random sample $X_1, X_2, \dots, X_n$ from F , define random probability distributions $\check{F}_n = \check{F}_{n,\gamma}$ and $\hat{F}_n = \hat{F}_{n,\gamma}$ by

$$\check{F}_n = \sum_{i=1}^{nF_n} \left(\frac{1}{n} + \frac{(n+1)}{2n^\gamma} - \frac{i}{n^\gamma} \right) = F_n + \frac{1}{2n^{\gamma-2}} F_n(1 - F_n)$$

and

$$\hat{F}_n = \sum_{i=1}^{nF_n} \left(\frac{1}{n} - \frac{(n+1)}{2n^\gamma} + \frac{i}{n^\gamma} \right) = F_n - \frac{1}{2n^{\gamma-2}} F_n(1 - F_n).$$

Note that the empirical measure defined by $\check{F}_n$ is $\sum_{i=1}^n ((1/n) + ((n+1)/2n^\gamma) - (i/n^\gamma))\delta(X_{i:n})$ and that of $\hat{F}_n$ is $\sum_{i=1}^n ((1/n) - ((n+1)/2n^\gamma) + (i/n^\gamma))\delta(X_{i:n})$, where $\delta(x)$ is the Dirac point-mass at $\{x\}$. From this it is easy to see that $\check{F}_n$ and $\hat{F}_n$ are convex combinations of the Dirac masses at the order statistics, and thus are permutation-invariant, linear, sample-ranged and distribution-free empirical distribution functions.

These new empirical distribution functions not only converge to the true distribution F uniformly almost surely (i.e., Glivenko–Cantelli-wise), but also have the properties that their inverses, i.e., their *quantile* functions also converge almost surely (to the lower and upper quantiles of F , respectively), whereas the inverse of the classical empirical function does not converge in general. For these and other basic properties, see Gilat and Hill (1992, Proposition 3.2, Theorem 3.3). Important for this note are that $\check{F}_n$ and $\hat{F}_n$ are true (random) probability distributions and the following lemma whose proof is given for completeness.

Lemma 3.3. *For all $\gamma \in (2, \frac{5}{2})$ and all $x \in \mathbb{R}$,*

- (i) $\check{F}_n(x) = F(x) + (1/2n^{\gamma-2})F(x)(1 - F(x)) + o(1/n^{\gamma-2})$ a.s.;
- (ii) $\hat{F}_n(x) = F(x) - (1/2n^{\gamma-2})F(x)(1 - F(x)) + o(1/n^{\gamma-2})$ a.s.

Proof. Observe that

$$\begin{aligned} \check{F}_n(x) &= F_n(x) + \frac{1}{2n^{\gamma-2}} F_n(x)(1 - F_n(x)) \\ &= F(x) + o\left(\sqrt{\frac{\log n}{n}}\right) + \frac{1}{2n^{\gamma-2}} [F(x)(1 - F(x)) + o(1)] \quad \text{a.s.} \\ &= F(x) + \frac{1}{2n^{\gamma-2}} F(x)(1 - F(x)) + o\left(\frac{1}{n^{\gamma-2}}\right) \quad \text{a.s.,} \end{aligned}$$

where the first equality follows by definition of $\check{F}_n$, the second equality by the law of iterated logarithm for the first term and the strong law of large numbers for the second term, and the last equality since $\gamma - 2 < \frac{1}{2}$.

The argument for $\hat{F}_n$ is analogous. $\square$

The next lemma relates these refined empirical distribution functions $\check{F}$ and $\hat{F}$ to the new quantile-estimators $I_{\beta,n}^*$ in (6).

Lemma 3.4. *For each $\gamma \in (2, \frac{5}{2})$ and each $\beta \in (0, 1)$, $X_{[a]:n}$ is the lower β -quantile for the refined empirical distribution $\check{F}_n$, and $X_{[b]+1:n}$ is the upper β -quantile for $\hat{F}_n$. In other words,*

$$I_{\beta,n}^* = [q_{\check{F}_n}(\beta), Q_{\hat{F}_n}(\beta)]. \quad (7)$$

Proof. From the definition of $\check{F}_n$, $q_{\check{F}_n}(\beta) = X_{k:n}$, where k is the smallest positive integer satisfying

$$\sum_{i=1}^k \left(\frac{1}{n} + \frac{(n+1)}{2n^\gamma} - \frac{i}{n^\gamma} \right) \geq \beta,$$

and solving the resulting quadratic equation

$$k \left(\frac{1}{n} + \frac{(n+1)}{2n^\gamma} \right) - \frac{k(k+1)}{2n^\gamma} = \beta$$

yields $k = \lceil a \rceil$ as in (6). The calculation for $\hat{F}_n$ is analogous. $\square$

Proof of Theorem 1. Fix $\gamma \in (2, \frac{5}{2})$ and $\beta \in (0, 1)$, and let F be any probability distribution. Since $1 \leq \lceil a \rceil \leq \lfloor b \rfloor + 1 \leq n$, the intervals $I_{\beta, n}^*$ are well-defined, and by their representation as order statistics (in (6)) it follows easily that these intervals are permutation-invariant, linear, sample-ranged and distribution-free estimators of $I_\beta(F)$.

From Lemma 3.3 it easily follows that $q_{\hat{F}_n}(\beta) \rightarrow q_F(\beta)$ and $Q_{\hat{F}_n}(\beta) \rightarrow Q_F(\beta)$ a.s. (in fact these convergences are even eventually from below and from above, respectively), which by Lemma 3.4 implies that the estimators $I_{\beta, n}^*$ are strongly consistent.

To show (i), let $\phi_n : [0, 1] \rightarrow [0, 1]$ be the continuous strictly increasing function $\phi_n(x) = x + x(1-x)/2n^{\gamma-2}$. Then, using essentially the same calculation as for Lemma 3.1, the inverse ϕ_n^{-1} has Taylor's series expansion $\phi_n^{-1}(\beta) = \beta - \frac{1}{2}\beta^2(1/n^{\gamma-2}) + o(1/n^{\gamma-2})$. By Lemma 3.3, $\check{F}_n(x) = \phi_n(F(x)) + o(1/n^{\gamma-2})$ a.s., so

$$q_{\check{F}_n}(\beta) = \check{F}_n^{-1}(\beta) = q_F \left(\phi_n^{-1} \left(\beta + o\left(\frac{1}{n^{\gamma-2}}\right) \right) \right) = q_F \left(\beta + O\left(\frac{1}{n^{\gamma-2}}\right) \right)$$

a.s. for Lebesgue a.e. β , where q_F is again the left inverse of F . Since q_F is monotonically increasing, its derivative exists for Lebesgue almost-all $\beta \in (0, 1)$, and for such β ,

$$\left| q_F \left(\beta + O\left(\frac{1}{n^{\gamma-2}}\right) \right) - q_F(\beta) \right| = O\left(\frac{1}{n^{\gamma-2}}\right) \quad \text{a.s.,}$$

so

$$|q_{\check{F}_n}(\beta) - q_F(\beta)| = O\left(\frac{1}{n^{\gamma-2}}\right) \quad \text{a.s.}$$

The analogous argument gives $|Q_{\check{F}_n}(\beta) - Q_F(\beta)| = O(1/n^{\gamma-2})$ a.s., and together these imply (i) via (7).

Conclusion (ii) follows from (6) and the well known inequality (e.g., David, 1981, (2.5.4) and (2.5.6), p. 14) that for general order statistics $X_{r:n}$ and $X_{s:n}$, and arbitrary F ,

$$P([X_{r:n}, X_{s:n}] \supset I_\beta(F)) \geq \sum_{i=r}^{s-1} \binom{n}{i} \beta^i (1-\beta)^{n-i},$$

and equality is attained for continuous F .

Conclusion (ii)' follows routinely from Lemma 3.1, (ii), and the Central Limit Theorem normal approximation to the binomial distribution. $\square$

Acknowledgements

The authors are indebted to Dr. Amir Ben-Amram for the calculations appearing in Tables 1–3, and to the referee for several suggestions.

References

- Csörgő, M. (1983), *Quantile Processes with Statistical Applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 42 (SIAM, Philadelphia, PA).

Csörgő, M. and L. Horváth (1989), On confidence bands for the quantile function of a continuous distribution function, *Coll. Math. Soc. J. Bolyai* **57**; *Limit Theorems in Probability and Statistics* (North-Holland, Amsterdam) pp. 95–106.

David, H. (1981), *Order Statistics* (Wiley, New York, 2nd ed.).

Gilat, D. and T. Hill (1992), One-sided refinements of the strong law of large numbers and the Glivenko–Cantelli Theorem, *Ann. Probab.* **20**, 1213–1221.

Gilat, D. and T. Hill (1993), Rates of convergence of one-sided estimators for means and quantiles, Technical Report 93-017, School of Mathematics, Georgia Institute of Technology.

Shorack, G. and J. Wellner (1986), *Empirical Processes with Applications to Statistics* (Wiley, New York).

Stigler, S. (1974), Linear functions of order statistics with smooth weight functions, *Ann. Statist.* **2**, 676–693.