ESTIMATION OF QUANTILES FROM DATA WITH ADDITIONAL MEASUREMENT ERRORS

Technische Universität Darmstadt

Supplementary Material

The following supplementary material contains detailed proofs of the Theorems 1 to 5.

In three of the proofs we use the following lemma, which relates the plug-in estimate with data containing additional measurement errors to plug-in estimates with i.i.d. data without additional measurement errors.

Lemma 1. *Let $a > 0$ be a (possibly random) finite constant and set*

$$\delta_n = \frac{1}{n} \sum_{i=1}^{n} I_{\left\{ \left| X_i - \bar{X}_{i,n} \right| > a \right\}}.$$

Then it holds for $\alpha \in \mathbb{R}$ and the plug-in estimates defined above that

$$\hat{q}_{X,n,\alpha-\delta_n} - a \leq \hat{q}_{\bar{X},n,\alpha} \leq \hat{q}_{X,n,\alpha+\delta_n} + a$$

Proof. Consider

$$\bar{F}_n(x) - F_n(x+a) = \frac{1}{n} \sum_{i=1}^{n} \left(I_{\left\{ \bar{X}_{i,n} \leq x \right\}} - I_{\{X_i \leq x+a\}} \right).$$

The i-th summand becomes one, if

$$\bar{X}_{i,n} \leq x \quad \text{and } X_i > x + a.$$

In this case $\left|X_i - \bar{X}_{i,n}\right| > a$ also holds true. So we can conclude

$$\bar{F}_n(x) - F_n(x+a) \leq \frac{1}{n} \sum_{i=1}^{n} I_{\{|X_i - \bar{X}_{i,n}| > a\}} = \delta_n.$$

Analogously we can also show

$$\bar{F}_n(x) - F_n(x-a) \geq -\frac{1}{n} \sum_{i=1}^{n} I_{\{|X_i - \bar{X}_{i,n}| > a\}} = -\delta_n.$$

Hence we get

$$\begin{aligned}
\hat{q}_{\bar{X},n,\alpha} &= \min\left\{z \in \mathbb{R} : \bar{F}_n\left(z\right) \geq \alpha\right\} \\
&= \min\left\{z \in \mathbb{R} : \bar{F}_n\left(z\right) - F_n\left(z+a\right) + F_n\left(z+a\right) \geq \alpha\right\} \\
&\geq \min\left\{z \in \mathbb{R} : \delta_n + F_n\left(z+a\right) \geq \alpha\right\} \\
&= \min\left\{z \in \mathbb{R} : F_n\left(z\right) \geq \alpha - \delta_n\right\} - a \\
&= \hat{q}_{X,n,\alpha-\delta_n} - a
\end{aligned}$$

and

$$\begin{aligned}\hat{q}_{\bar{X},n,\alpha} &= \min\left\{z \in \mathbb{R} : \bar{F}_n(z) \geq \alpha\right\} \\ &= \min\left\{z \in \mathbb{R} : \bar{F}_n(z) - F_n(z-a) + F_n(z-a) \geq \alpha\right\} \\ &\leq \min\{z \in \mathbb{R} : -\delta_n + F_n(z-a) \geq \alpha\} \\ &= \min\{z \in \mathbb{R} : F_n(z) \geq \alpha + \delta_n\} + a \\ &= \hat{q}_{X,n,\alpha+\delta_n} + a,\end{aligned}$$

which yields the assertion. □

S1 Proof of Theorem 1

Let $\alpha_n \in (0,1)$ be such that

$$\alpha_n \to \alpha \quad a.s.$$

We divide the proof into three steps:

In the first step of the proof we show that

$$dist\left(\hat{q}_{X,n,\alpha_n}, Q_{X,\alpha}\right) \to 0 \quad a.s. \tag{S1.1}$$

Therefore set

$$N := \left\{\alpha_n \to \alpha \ (n \to \infty) \quad \text{and} \quad \sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \to 0 \ (n \to \infty)\right\}.$$

Notice that

$$\mathbf{P}(N) = 1$$

because of the Glivenko-Catelli theorem (cf., e.g., Theorem 12.4 in Devroye, Györfi and Lugosi (1996)) and $\alpha_n \to \alpha$ $a.s.$ Let $\epsilon > 0$ be arbitrary. We know

$$F\left(q_{X,\alpha}^{[low]} - \epsilon\right) < \alpha < F\left(q_{X,\alpha}^{[up]} + \epsilon\right). \tag{S1.2}$$

Setting

$$\rho_1 = \min\left(\alpha - F\left(q_{X,\alpha}^{[low]} - \epsilon\right), F\left(q_{X,\alpha}^{[up]} + \epsilon\right) - \alpha\right),$$

we can conclude

$$F\left(q_{X,\alpha}^{[low]} - \epsilon\right) + \frac{\rho_1}{2} < \alpha < F\left(q_{X,\alpha}^{[up]} + \epsilon\right) - \frac{\rho_1}{2}.$$

Assume N to hold in the following. Then we can (for all $\omega \in N$) find n_0, such that for all $n \geq n_0$ we have

$$|\alpha_n - \alpha| < \frac{\rho_1}{4} \quad \text{and} \quad \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| < \frac{\rho_1}{4},$$

which implies

$$F_n\left(q_{X,\alpha}^{[low]} - \epsilon\right) < \alpha_n < F_n\left(q_{X,\alpha}^{[up]} + \epsilon\right)$$

and consequently

$$q_{X,\alpha}^{[low]} - \epsilon \leq \hat{q}_{X,n,\alpha_n} \leq q_{X,\alpha}^{[up]} + \epsilon.$$

Hence,

$$\mathbf{P}\left(\limsup_{n\to\infty} dist\left(\hat{q}_{X,n,\alpha_n}, Q_{X,\alpha}\right) \leq \epsilon\right) \geq \mathbf{P}(N) = 1.$$

Since $\epsilon > 0$ was arbitrary this implies the assertion.

Let $\epsilon > 0$ again be arbitrary and set

$$\delta_n = \frac{1}{n}\sum_{i=1}^{n} I_{\{|X_i - \bar{X}_{i,n}| > \epsilon\}}.$$

In the second step of the proof we show

$$\delta_n \to 0 \quad a.s. \tag{S1.3}$$

Therefore we observe

$$\frac{1}{n}\sum_{i=1}^{n} I_{\{|X_i - \bar{X}_{i,n}| > \epsilon\}} \leq \frac{1}{\epsilon}\frac{1}{n}\sum_{i=1}^{n} |X_i - \bar{X}_{i,n}|,$$

which yields the assertion by (2.1).

Furthermore, we know by Lemma 1

$$\hat{q}_{X,n,\alpha-\delta_n} - \epsilon \leq \hat{q}_{\bar{X},n,\alpha} \leq \hat{q}_{X,n,\alpha+\delta_n} + \epsilon \tag{S1.4}$$

In the third step of the proof we finally show the assertion. By the second step, we know $\alpha - \delta_n \to \alpha$ $a.s.$ and $\alpha + \delta_n \to \alpha$ $a.s.$, so by choosing $\alpha_n = \alpha - \delta_n$ or $\alpha_n = \alpha + \delta_n$, resp., we conclude by (S1.4) and by the first step for arbitrary $\epsilon > 0$

$$\begin{aligned} & dist\left(\hat{q}_{\bar{X},n,\alpha}, Q_{X,\alpha}\right) \\ & \leq dist\left(\hat{q}_{X,n,\alpha-\delta_n}, Q_{X,\alpha}\right) + \epsilon + dist\left(\hat{q}_{X,n,\alpha+\delta_n}, Q_{X,\alpha}\right) + \epsilon \longrightarrow 2\epsilon \quad a.s. \end{aligned} \tag{S1.5}$$

Since $\epsilon > 0$ was arbitrary this implies the assertion. □

S2 Proof of Theorem 2

In order to proof Theorem 2, we need the following lemma, which is a straightforward extension of ideas in Theorem 4 in Feldman and Tucker (1966) to random sequences.

Lemma 2. *Let $\alpha \in (0,1)$ be arbitrary and $X, X_1, X_2, \dots$ be independent and identically distributed real valued random variables with cdf. F.*

(a) Let $\gamma_{n,l}$ be a (possibly random) sequence, that satisfies

$$\gamma_{n,l} + (1+\nu)\sqrt{\frac{2\log(\log(n/2))}{n}} < \alpha \quad \text{and} \quad \gamma_{n,l} \to \alpha \quad a.s.$$

for some $\nu > 0$. Then it holds

$$\hat{q}_{X,n,\gamma_{n,l}} \to q_{X,\alpha}^{[low]} \quad a.s. \tag{S2.6}$$

(b) Let $\gamma_{n,r}$ be a (possibly random) sequence, that satisfies

$$\gamma_{n,r} - (1+\nu)\sqrt{\frac{2\log(\log(n/2))}{n}} > \alpha \quad \text{and} \quad \gamma_{n,l} \to \alpha \quad a.s.$$

for some $\nu > 0$. Then it holds

$$\hat{q}_{X,n,\gamma_{n,r}} \to q_{X,\alpha}^{[up]} \quad a.s. \tag{S2.7}$$

Proof of Lemma 2. (a) It suffices to show

(i) $\mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,l}} \leq q_{X,\alpha}^{[low]} - \epsilon \quad i.o.\right) = 0$ for any $\epsilon > 0$, and

(ii) $\mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,l}} > q_{X,\alpha}^{[low]} \quad i.o.\right) = 0$,

where $i.o.$ means infinitely often. First of all we show (i). Therefore let $\epsilon > 0$ be arbitrary. We know

$$F\left(q_{X,\alpha}^{[low]} - \epsilon\right) < \alpha.$$

Setting

$$\rho_2 = \alpha - F\left(q_{X,\alpha}^{[low]} - \epsilon\right),$$

we can conclude

$$F\left(q_{X,\alpha}^{[low]} - \epsilon\right) + \frac{\rho_2}{2} < \alpha.$$

Choose

$$N := \left\{\gamma_{n,l} \to \alpha \ (n \to \infty) \quad \text{and} \quad \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \to 0 \ (n \to \infty)\right\}.$$

As in the proof of Theorem 1 we have $\mathbf{P}(N) = 1$. We can (for every $\omega \in N$) find n_0 such that for all $n \geq n_0$ it holds

$$|\gamma_{n,l} - \alpha| \leq \frac{\rho_2}{4} \quad \text{and} \quad \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \leq \frac{\rho_2}{4}.$$

This implies (for every $\omega \in N$)

$$F_n\left(q_{X,\alpha}^{[low]} - \epsilon\right) < \gamma_{n,l}$$

and hence

$$\hat{q}_{X,n,\gamma_{n,l}} > q_{X,\alpha}^{[low]} - \epsilon$$

for n large enough. So we actually have shown

$$1 - \mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,l}} \leq q_{X,\alpha}^{[low]} - \epsilon \quad i.o.\right) \geq \mathbf{P}\left(N\right) = 1,$$

which proves (i).

It remains to show (ii). Therefore set

$$U_i = 1 - 2 \cdot I_{\left\{X_i \leq q_{X,\alpha}^{[low]}\right\}} \quad \text{for } i = 1, ..., n$$

and

$$p_1 = \mathbf{P}\left(X \leq q_{X,\alpha}^{[low]}\right) \geq \alpha.$$

We know

$$\mathbf{E}\left\{U_i\right\} = 1 - 2p_1 \leq 1 - 2\alpha \quad \text{and} \quad s = \mathbf{V}\left\{U_i\right\} = 4p_1 \cdot (1 - p_1)$$

and

$$\sum_{i=1}^{n} U_i = n - 2n \cdot F_n\left(q_{X,\alpha}^{[low]}\right).$$

Thus,

$$\begin{aligned} \left\{\hat{q}_{X,n,\gamma_{n,l}} > q_{X,\alpha}^{[low]}\right\} &= \left\{F_n\left(q_{X,\alpha}^{[low]}\right) < \gamma_{n,l}\right\} \\ &= \left\{-2n \cdot F_n\left(q_{X,\alpha}^{[low]}\right) > -2\gamma_{n,l} \cdot n\right\} \\ &\subseteq \left\{\sum_{i=1}^{n} U_i \geq n - 2\gamma_{n,l} \cdot n\right). \end{aligned} \tag{S2.8}$$

It is only necessary to consider the nontrivial case where $s > 0$. Set $\psi_n = (2ns \cdot \log(\log(ns)))^{1/2}$, which we will need in the subsequent application of Kolmogorov's law of the iterated logarithm. Observe that ψ_n is well-defined for n large enough. Since $0 \leq x \cdot (1-x) \leq \frac{1}{4}$ for $x \in [0,1]$, we have $0 \leq s \leq 1$ and thus $(2n \cdot \log(\log(n)))^{1/2} \geq \psi_n$. Because of

$$\alpha - \gamma_{n,l} > (1+\nu) \cdot \sqrt{\frac{2\log(\log(n/2))}{n}},$$

we can conclude

$$\alpha - \gamma_{n,l} \geq \frac{1+\nu}{2} \cdot \sqrt{\frac{2\log(\log(n))}{n}}$$

for all n large enough. Combining this with

$$1 - 2p_1 \leq 1 - 2\alpha,$$

we get by (S2.8)

$$\begin{aligned}
&\mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,l}} > q_{X,\alpha}^{[low]} \;\; i.o.\right) \\
&\leq \mathbf{P}\left(\sum_{i=1}^{n} U_i \geq n - 2\gamma_{n,l} \cdot n \;\; i.o.\right) \\
&\leq \mathbf{P}\left(\sum_{i=1}^{n} U_i \geq n \cdot (1-2\alpha) + 2 \cdot (\alpha \cdot n - \gamma_{n,l} \cdot n) \;\; i.o.\right) \\
&\leq \mathbf{P}\left(\sum_{i=1}^{n} U_i \geq n \cdot (1-2p_1) + (1+\nu) \cdot \psi_n \;\; i.o.\right).
\end{aligned}$$

We know by Kolmogorov's law of the iterated logarithm (cf., e.g., Theorem

1 on page 140 in Tucker (1967))

$$\mathbf{P}\left(\limsup_{n\to\infty} \frac{\sum_{i=1}^{n} U_i - n\cdot(1-2p_1)}{\psi_n} = 1\right) = 1,$$

from which we can conclude

$$\mathbf{P}\left(\sum_{i=1}^{n} U_i \geq n\cdot(1-2p_1) + (1+\nu)\cdot\psi_n \;\; i.o.\right) = 0.$$

This completes the proof of (a).

(b) It suffices to show

(i) $\mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,r}} > q_{X,\alpha}^{[up]} + \epsilon \quad i.o.\right) = 0$ for any $\epsilon > 0$, and

(ii) $\mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,r}} < q_{X,\alpha}^{[up]} \quad i.o.\right) = 0.$

The proof of (i) is analogously to (i) in part (a). It remains to show (ii). Therefore set

$$V_i = 2\cdot I_{\left\{X_i < q_{X,\alpha}^{[up]}\right\}} - 1 \quad \text{for } i = 1, ..., n$$

and

$$p_2 = \mathbf{P}\left(X < q_{X,\alpha}^{[up]}\right) \leq \alpha.$$

We have $\mathbf{E}\{V_i\} = 2p_2 - 1 \leq 2\alpha - 1$ and $\tilde{s} = \mathbf{V}\{V_i\} = 4p_2\cdot(1-p_2)$. Observe that if

$$\hat{q}_{X,n,\gamma_{n,r}} < q_{X,\alpha}^{[up]},$$

then

$$\frac{1}{n}\sum_{i=1}^{n} I_{\left\{X_i<q_{X,\alpha}^{[up]}\right\}} \geq \frac{1}{n}\sum_{i=1}^{n} I_{\left\{X_i\leq \hat{q}_{X,n,\gamma_{n,r}}\right\}} = F_n\left(\hat{q}_{X,n,\gamma_{n,r}}\right) \geq \gamma_{n,r}.$$

Thereby, we can analogously to (ii) in part (a) conclude

$$\left\{\hat{q}_{X,n,\gamma_{n,r}} < q_{X,\alpha}^{[up]}\right\} \subseteq \left\{\sum_{i=1}^{n} V_i \geq 2\gamma_{n,r}\cdot n - n\right\}.$$

Again, we only need to consider the nontrivial case $\tilde{s} > 0$ and set $\tilde{\psi}_n = (2n\tilde{s}\cdot\log(\log(n\tilde{s})))^{1/2}$. Since $0 \leq x\cdot(1-x) \leq \frac{1}{4}$ for $x\in[0,1]$, we have $(2n\cdot\log(\log(n)))^{1/2} \geq \tilde{\psi}_n$. The assumption on $\gamma_{n,r}$ implies

$$\gamma_{n,r} - \alpha \geq \frac{1+\nu}{2}\cdot\sqrt{\frac{2\log(\log(n))}{n}}$$

for all n large enough. Thus, using $2\alpha - 1 \geq 2p_2 - 1$, we can conclude

$$\mathbf{P}\left(\hat{q}_{X,n,\gamma_{n,r}} < q_{X,\alpha}^{[up]}\ \ i.o.\right) \leq \mathbf{P}\left(\sum_{i=1}^{n} V_i \geq n\cdot(2p_2-1) + (1+\nu)\cdot\tilde{\psi}_n\ \ i.o.\right)$$

Again, by Kolmogorov's law of the iterated logarithm, we get

$$\mathbf{P}\left(\sum_{i=1}^{n} V_i \geq n\cdot(2p_2-1) + (1+\nu)\cdot\tilde{\psi}_n\ \ i.o.\right) = 0,$$

which completes the proof. □

Proof of Theorem 2. Set

$$\delta_n = \frac{1}{n}\sum_{i=1}^{n} I_{\{|X_i-\bar{X}_{i,n}|>\sqrt{\eta_n}\}}$$

and observe that (2.2) implies

$$\delta_n = \frac{1}{n}\sum_{i=1}^{n} I_{\{|X_i-\bar{X}_{i,n}|>\sqrt{\eta_n}\}} \leq \frac{1}{\sqrt{\eta_n}}\frac{1}{n}\sum_{i=1}^{n}|X_i-\bar{X}_{i,n}| \leq \frac{\eta_n}{\sqrt{\eta_n}} = \sqrt{\eta_n}\quad a.s. \tag{S2.9}$$

Using Lemma 1 and (S2.9), we can conclude that for any (random) sequence γ_n holds

$$\hat{q}_{X,n,\gamma_n-\sqrt{\eta_n}} - \sqrt{\eta_n} \leq \hat{q}_{\bar{X},n,\gamma_n} \leq \hat{q}_{X,n,\gamma_n+\sqrt{\eta_n}} + \sqrt{\eta_n} \tag{S2.10}$$

for every $n \in \mathbb{N}$. By setting $\gamma_n = \alpha_n$ in (S2.10) we know

$$\hat{q}_{X,n,\alpha_n-\sqrt{\eta_n}} - \sqrt{\eta_n} \leq \hat{q}_{\bar{X},n,\alpha_n} \leq \hat{q}_{X,n,\alpha_n+\sqrt{\eta_n}} + \sqrt{\eta_n} \tag{S2.11}$$

for all $n \in \mathbb{N}$. Having regard to

$$\alpha_n + (1+\nu) \cdot \sqrt{\frac{2\log(\log(n/2))}{n}} + \sqrt{\eta_n} < \alpha$$

for all $0 < \nu < 1$, as well as $\alpha_n \to \alpha$ $a.s.$, we also know that $\gamma_{n,l} = \alpha_n + \sqrt{\eta_n}$ and $\gamma_{n,l} = \alpha_n - \sqrt{\eta_n}$ fullfill the assumptions of Lemma 2a). So we get

$$\hat{q}_{X,n,\alpha_n-\sqrt{\eta_n}} - \sqrt{\eta_n} \to q_{X,\alpha}^{[low]} \quad a.s. \quad \text{and} \quad \hat{q}_{X,n,\alpha_n+\sqrt{\eta_n}} + \sqrt{\eta_n} \to q_{X,\alpha}^{[low]} \quad a.s.,$$

which yields

$$\hat{q}_{\bar{X},n,\alpha_n} \to q_{X,\alpha}^{[low]} \quad a.s.$$

Analogously we can show

$$\hat{q}_{\bar{X},n,\beta_n} \to q_{X,\alpha}^{[up]} \quad a.s.$$

by using Lemma 2b), which completes the proof. □

S3 Proof of Theorem 3

Let $\alpha \in (0,1)$ be arbitrary. Assume to the contrary that there exists a sequence $(\hat{q}_{n,\alpha})_{n\in\mathbb{N}}$ of quantile estimates statisfying

$$\hat{q}_{n,\alpha}\left(\bar{X}_1,\ldots,\bar{X}_n\right) \to^{\mathbf{P}} q_{X,\alpha}^{[low]} \tag{S3.12}$$

whenever $\bar{X}_1, \bar{X}_2, \ldots$ are such that for some independent and identically as X distributed $X_1, X_2, \ldots$ we have

$$\frac{1}{n}\sum_{i=1}^{n} |X_i - \bar{X}_i| \to 0 \quad a.s. \tag{S3.13}$$

Let $X, X_1, X_2, \ldots$ be independent and indentically distributed with cdf.

$$F(x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ x & \text{if} \quad 0 \le x < \alpha \\ \alpha & \text{if} \quad \alpha \le x < 1+\alpha \\ x-1 & \text{if} \quad 1+\alpha \le x < 2 \\ 1 & \text{if} \quad 2 \le x \end{cases}$$

and α-quantile $q_{X,\alpha}^{[low]} = \alpha$. For $k \in \mathbb{N}$ set

$$F_k(x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ x & \text{if} \quad 0 \le x < \alpha - \frac{\alpha}{k} \\ \alpha - \frac{\alpha}{k} & \text{if} \quad \alpha - \frac{\alpha}{k} \le x < 1+\alpha - \frac{\alpha}{k} \\ x-1 & \text{if} \quad 1+\alpha - \frac{\alpha}{k} \le x < 2 \\ 1 & \text{if} \quad 2 \le x \end{cases}$$

and

$$X_i^{(k)} = \begin{cases} X_i & \text{if} \quad X_i \notin \left[\alpha - \frac{\alpha}{k}, \alpha\right] \\ \\ X_i + 1 & \text{if} \quad X_i \in \left[\alpha - \frac{\alpha}{k}, \alpha\right] \end{cases} .$$

Then $X_1^{(k)}, X_2^{(k)}, \dots$ are independent and identically distributed random variables with cdf. F_k and α-quantile $q_{k,\alpha}^{[low]} = 1 + \alpha$. So if we set $\bar{X}_i = X_i^{(k)}$ for all $i \geq N$ with $N \in \mathbb{N}$ arbitrary, (S3.13) is fullfilled (with X_i replaced by $X_i^{(k)}$) and we know by (S3.12) that

$$\hat{q}_{n,\alpha}\left(\bar{X}_1, \dots, \bar{X}_n\right) \to^{\mathbf{P}} q_{k,\alpha}^{[low]} \tag{S3.14}$$

Next we define for suitably chosen deterministic $n_0 := 0 < n_1 < n_2 < \dots$ (where $n_i \in \mathbb{N}$ for all $i \in N$) our data with measurement error by

$$\bar{X}_i = X_i^{(k)} \quad \text{if } n_{k-1} < i \leq n_k \quad (k \in \mathbb{N}).$$

For all $i \in \mathbb{N}$ we have

$$\mathbf{P}\left(|X_i - \bar{X}_i| = 0\right) \geq 1 - \alpha \quad \text{and} \quad \mathbf{P}\left(|X_i - \bar{X}_i| = 1\right) \leq \alpha$$

and hence

$$0 \leq \mathbf{E}\left\{|X_i - \bar{X}_i|\right\} \leq \alpha \quad \text{and} \quad \mathbf{V}\left\{|X_i - \bar{X}_i|\right\} \leq \mathbf{E}\left\{|X_i - \bar{X}_i|^2\right\} \leq \alpha.$$

So

$$\sum_{i=1}^{\infty} \frac{\mathbf{V}\{|X_i - \bar{X}_i|\}}{i^2} \leq \sum_{i=1}^{\infty} \frac{\alpha}{i^2} < \infty.$$

By a criterion which is sometimes called the Kolmogorov criterion (cf., e.g., Theorem 14.5 in Burckel and Bauer (1996)), we get

$$\frac{1}{n}\sum_{i=1}^{n}\left(|X_i - \bar{X}_i| - \mathbf{E}\{|X_i - \bar{X}_i|\}\right) \to 0 \quad a.s. \tag{S3.15}$$

But since $|X_i - X_i^{(k)}| \geq |X_i - X_i^{(l)}|$ for all $l \geq k$ and $i \in \mathbb{N}$, we can conclude

$$\begin{aligned}
0 \leq \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}\{|X_i - \bar{X}_i|\} &= \frac{1}{n}\sum_{i=1}^{n_k}\mathbf{E}\{|X_i - \bar{X}_i|\} + \frac{1}{n}\sum_{i=n_k+1}^{n}\mathbf{E}\{|X_i - \bar{X}_i|\} \\
&\leq \frac{1}{n}\sum_{i=1}^{n_k}\alpha + \frac{1}{n}\sum_{i=n_k+1}^{n}\mathbf{E}\{|X_i - X_i^{(k)}|\} \\
&= \frac{n_k}{n}\cdot\alpha + \frac{1}{n}\sum_{i=n_k+1}^{n}\frac{\alpha}{k} \\
&\leq \frac{n_k}{n}\cdot\alpha + \frac{\alpha}{k} \quad \longrightarrow \frac{\alpha}{k} \quad (n\to\infty),
\end{aligned}$$

for every $k \in \mathbb{N}$, which implies

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}\{|X_i - \bar{X}_i|\} \to 0$$

and finally by (S3.15)

$$\frac{1}{n}\sum_{i=1}^{n}\left|X_i - \bar{X}_i\right| \to 0 \quad a.s.$$

So it suffies to show, that for some $\epsilon > 0$

$$\limsup_{n\to\infty}\mathbf{P}\left(\left|\hat{q}_{n,\alpha}\left(\bar{X}_1,\ldots,\bar{X}_n\right) - q_{X,\alpha}^{[low]}\right| > \epsilon\right) > 0. \tag{S3.16}$$

To do this we will choose n_k such that (S3.16) holds. Let $0 < \epsilon < 1$ be fixed and choose n_1 such that

$$\mathbf{P}\left(\left|\hat{q}_{n_1,\alpha}\left(\bar{X}_1^{(1)},\ldots,\bar{X}_{n_1}^{(1)}\right) - q_{1,\alpha}^{[low]}\right| > \epsilon\right) < \frac{1}{2}.$$

This is possible because of (S3.14). Given $n_1, \dots, n_{k-1}$, we choose $n_k > n_{k-1}$ such that

$$\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_{k-1}}, \bar{X}^{(k)}_{n_{k-1}+1}, \dots, \bar{X}^{(k)}_{n_k}\right) - q^{[low]}_{k,\alpha}\right| > \epsilon\right) < \frac{1}{2},$$

which is again possible because of (S3.14). The choice of $n_1, n_2, \dots$ implies

$$\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{k,\alpha}\right| > \epsilon\right) < \frac{1}{2}$$

and accordingly

$$\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{k,\alpha}\right| \leq \epsilon\right) \geq \frac{1}{2}$$

for $k \in \mathbb{N}$. Using the triangle inequality, we know

$$1 = \left|q^{[low]}_{k,\alpha} - q^{[low]}_{X,\alpha}\right| \leq \left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{k,\alpha}\right| + \left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{X,\alpha}\right|.$$

Thereby, we can conclude for any $k \in \mathbb{N}$

$$\begin{aligned}
&\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{X,\alpha}\right| > 1 - \epsilon\right) \\
&\geq \mathbf{P}\left(1 - \left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{k,\alpha}\right| > 1 - \epsilon\right) \\
&= \mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_1, \dots, \bar{X}_{n_k}\right) - q^{[low]}_{k,\alpha}\right| < \epsilon\right) \\
&\geq \frac{1}{2},
\end{aligned} \tag{S3.17}$$

which completes the proof. □

S4 Proof of Theorem 4

For the sake of simplicity we write $q_{X,\alpha}$ for the lower α-quantile of X instead of $q^{[low]}_{X,\alpha}$.

We divide the proof into two steps:

In the first step of the proof we show that if α_n is a (possibly random) sequence with

$$\alpha_n \to \alpha \quad a.s.$$

it holds

$$\left|\hat{q}_{X,n,\alpha_n} - q_{X,\alpha}\right| = O_{\mathbf{P}}\left(\left(\frac{1}{\sqrt{n}} + |\alpha_n - \alpha|\right)^{1/\gamma}\right). \tag{S4.18}$$

Therefore it suffices to show

$$\limsup_{n\to\infty} \mathbf{P}\left(\left|\hat{q}_{X,n,\alpha_n} - q_{X,\alpha}\right| \leq \frac{2c_1}{c_2^{1/\gamma}} \cdot \left(\frac{1}{\sqrt{n}} + |\alpha_n - \alpha|\right)^{1/\gamma}\right) \geq 1-2\exp\left(-2c_1^2\right)$$

for every $c_1 \geq 1$, with the finite constant $c_2 > 0$ of (2.5).

Now set

$$B_n := \left\{\frac{2c_1}{c_2}\left|\alpha_n - \alpha\right| \leq \frac{\zeta^\gamma}{2}\right\}$$

and

$$C_n := \left\{\sup_{t\in\mathbb{R}} \left|F(t) - F_n(t)\right| \leq \frac{c_1}{\sqrt{n}}\right\}.$$

We know

$$\mathbf{P}\left(B_n^c\right) \to 0 \quad (n\to\infty) \quad \text{and} \quad \mathbf{P}\left(C_n^c\right) \leq 2\exp\left(-2c_1^2\right)$$

because of $\alpha_n \to \alpha$ $a.s.$ and the Dvoretzky-Kiefer-Wolfowitz inequality (cf., Dvoretzky, Kiefer and Wolfowitz (1956)) in combination with Corollary 1 in Massart (1990). Choose $n_0 \in \mathbb{N}$, such that $0 < \frac{2}{c_2} \cdot \frac{c_1}{\sqrt{n}} \leq \frac{\zeta^\gamma}{2}$ is fullfilled for all $n \geq n_0$. Assume in the following, that the events B_n and C_n hold and consider $n \geq n_0$. Set $\theta_n = 2c_1 \cdot |\alpha_n - \alpha| + 2 \cdot \frac{c_1}{\sqrt{n}}$. The assumptions imply

$$0 < \left(\frac{1}{c_2} \cdot \theta_n\right)^{1/\gamma} = \left(\frac{2c_1}{c_2} \cdot |\alpha_n - \alpha| + \frac{2}{c_2} \cdot \frac{c_1}{\sqrt{n}}\right)^{1/\gamma} \leq \left(\frac{\zeta^\gamma}{2} + \frac{\zeta^\gamma}{2}\right)^{1/\gamma} = \zeta$$

so we can conclude by the assumption in (2.5) and $F(q_{X,\alpha}) = \alpha$

$$\theta_n = c_2 \left| q_{X,\alpha} - q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right|^\gamma \leq \left|\alpha - F\left(q_{X,\alpha} + \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)\right| \tag{S4.19}$$

and

$$\theta_n = c_2 \left| q_{X,\alpha} - q_{X,\alpha} + \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right|^\gamma \leq \left|\alpha - F\left(q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)\right|. \tag{S4.20}$$

Since $\theta_n > 0$ for all n, (S4.19) and (S4.20) imply

$$F\left(q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right) < \alpha - \frac{\theta_n}{2} < \alpha < \alpha + \frac{\theta_n}{2} < F\left(q_{X,\alpha} + \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right). \tag{S4.21}$$

Since the event C_n holds, we know

$$F_n\left(q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right) - \frac{c_1}{\sqrt{n}} \leq F\left(q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)$$

and

$$F\left(q_{X,\alpha}+\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)\leq F_n\left(q_{X,\alpha}+\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)+\frac{c_1}{\sqrt{n}}.$$

Combining this with (S4.21) and the definition of θ_n leads to

$$F_n\left(q_{X,\alpha}-\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)<\alpha-c_1\cdot|\alpha-\alpha_n|\leq\alpha+c_1\cdot|\alpha-\alpha_n|<F_n\left(q_{X,\alpha}+\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right).$$

Since $c_1\geq 1$ we have

$$\alpha-c_1\cdot|\alpha-\alpha_n|\leq\alpha_n\leq\alpha+c_1\cdot|\alpha-\alpha_n|\,,$$

which implies

$$F_n\left(q_{X,\alpha}-\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)<\alpha_n<F_n\left(q_{X,\alpha}+\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right).$$

So finally we have shown

$$\mathbf{P}\left(B_n\cap C_n\right)\leq\mathbf{P}\left(F_n\left(q_{X,\alpha}-\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)<\alpha_n<F_n\left(q_{X,\alpha}+\left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)\right),$$

which by the definition of $\hat{q}_{X,n,\alpha_n}$ and for $n \geq n_0$ leads to

$$\begin{aligned}
&\mathbf{P}\left(|\hat{q}_{X,n,\alpha_n} - q_{X,\alpha}| \leq \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right) \\
&= \mathbf{P}\left(q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma} \leq \hat{q}_{X,n,\alpha_n} \leq q_{X,\alpha} + \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right) \\
&\geq \mathbf{P}\left(F_n\left(q_{X,\alpha} - \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right) < \alpha_n < F_n\left(q_{X,\alpha} + \left(\frac{1}{c_2}\theta_n\right)^{1/\gamma}\right)\right) \\
&\geq \mathbf{P}\left(B_n \cap C_n\right) \\
&= 1 - \mathbf{P}\left(B_n^c \cup C_n^c\right) \\
&\geq 1 - \mathbf{P}\left(B_n^c\right) - \mathbf{P}\left(C_n^c\right) \\
&\geq 1 - \mathbf{P}\left(B_n^c\right) - 2\exp\left(-2c_1^2\right) \to 1 - 2\exp\left(-2c_1^2\right) \quad (n \to \infty).
\end{aligned}$$

This was the assertion.

Furthermore, we know (see proof of Theorem 2 in combination with (2.4))

$$\delta_n = \frac{1}{n}\sum_{i=1}^{n} I_{\{|X_i - \hat{X}_{i,n}| > \sqrt{\eta_n}\}} \leq \frac{\eta_n}{\sqrt{\eta_n}} = \sqrt{\eta_n} \to 0 \quad a.s. \tag{S4.22}$$

Using (S4.22), application of Lemma 1 yields

$$\hat{q}_{X,n,\alpha-\sqrt{\eta_n}} - \sqrt{\eta_n} \leq \hat{q}_{\bar{X},n,\alpha} \leq \hat{q}_{X,n,\alpha+\sqrt{\eta_n}} + \sqrt{\eta_n} \tag{S4.23}$$

for all $n \in \mathbb{N}$.

In the second step of the proof we finally show the assertion. By the first step we can conclude

$$\left|\hat{q}_{X,n,\alpha-\sqrt{\eta_n}} - q_{X,\alpha}\right| = O_{\mathbf{P}}\left(\left(\frac{1}{\sqrt{n}} + \sqrt{\eta_n}\right)^{1/\gamma}\right)$$

and

$$\left|\hat{q}_{X,n,\alpha+\sqrt{\eta_n}} - q_{X,\alpha}\right| = O_{\mathbf{P}}\left(\left(\frac{1}{\sqrt{n}} + \sqrt{\eta_n}\right)^{1/\gamma}\right).$$

By (S4.23) we know

$$\begin{aligned}\left|\hat{q}_{\bar{X},n,\alpha} - q_{X,\alpha}\right| &\leq \left|\hat{q}_{X,n,\alpha-\sqrt{\eta_n}} - \sqrt{\eta_n} - q_{X,\alpha}\right| + \left|\hat{q}_{X,n,\alpha+\sqrt{\eta_n}} + \sqrt{\eta_n} - q_{X,\alpha}\right| \\ &\leq \left|\hat{q}_{X,n,\alpha-\sqrt{\eta_n}} - q_{X,\alpha}\right| + \left|\hat{q}_{X,n,\alpha+\sqrt{\eta_n}} - q_{X,\alpha}\right| + 2\sqrt{\eta_n},\end{aligned}$$

which completes the proof. □

S5 Proof of Theorem 5

Let $\alpha \in (0,1)$ be arbitrary. For the sake of simplicity we write $q_{X,\alpha}$ for the lower α-quantile of X instead of $q_{X,\alpha}^{[low]}$. Assume to the contrary that there exists an estimator $(\hat{q}_{n,\alpha})_{n\in N}$ such that for all random variables $\bar{X}_{1,n}, \bar{X}_{2,n}, \ldots,$ which are such that for some independent and identically as X distributed $X_1, X_2, \ldots$ it holds

$$\eta_n = \frac{1}{n}\sum_{i=1}^{n} |X_i - \bar{X}_{i,n}| \to 0 \quad a.s., \tag{S5.24}$$

we have

$$\lim_{c\to\infty} \limsup_{n\to\infty} \mathbf{P}\left(\left|\hat{q}_{n,\alpha}\left(\bar{X}_{1,n}, \ldots, \bar{X}_{n,n}\right) - q_{X,\alpha}\right| > c \cdot \left(\frac{1}{\sqrt{n}} + \tilde{\eta}_n\right)\right) = 0, \tag{S5.25}$$

with a sequence $\tilde{\eta}_n$ that fullfills

$$\frac{\tilde{\eta}_n}{\sqrt{\eta_n}} \to^{\mathbf{P}} 0. \tag{S5.26}$$

Let $X, X_1, X_2, \ldots$ be independent and identically uniformly on $(0,1)$ distributed, i.e., with cdf.

$$F(x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ x & \text{if} \quad 0 \leq x < 1 \\ 1 & \text{if} \quad x \geq 1 \end{cases}$$

and lower α-quantile $q_{X,\alpha} = \alpha$. Set $\beta = \min(\alpha, 1-\alpha)/2$ and for $k \in \mathbb{N}$ let $Y^{(k)}$ have the distribution function

$$F_k(x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ x & \text{if} \quad 0 \leq x < \alpha - \beta\sqrt{\frac{1}{k}} \\ \alpha - \beta\sqrt{\frac{1}{k}} & \text{if} \quad \alpha - \beta\sqrt{\frac{1}{k}} \leq x < \alpha \\ 2(x-\alpha) + \alpha - \beta\sqrt{\frac{1}{k}} & \text{if} \quad \alpha \leq x < \alpha + \beta\sqrt{\frac{1}{k}} \\ x & \text{if} \quad \alpha + \beta\sqrt{\frac{1}{k}} \leq x < 1 \\ 1 & \text{if} \quad 1 \leq x. \end{cases}$$

In other words the distribution of the random variable $Y^{(k)}$ is obtained by shifting all mass, that is contained in the interval $\left[\alpha - \beta\sqrt{\frac{1}{k}}, \alpha\right]$, by $\beta\sqrt{\frac{1}{k}}$ to the right. This distribution has the lower α-quantile $q_{Y^{(k)},\alpha} = \alpha + \frac{\beta}{2}\sqrt{\frac{1}{k}}$.

Furthermore, we set

$$X_{i,n}^{(k)} = \begin{cases} X_i + \beta\sqrt{\frac{1}{k}} & \text{if} \quad X_i \in \left[\alpha - \beta\sqrt{\frac{1}{k}}, \alpha\right] \text{ and } X_i \text{ is one of the } \lfloor \beta\sqrt{\frac{1}{k}} \cdot n \rfloor \\ & \qquad \text{largest samples of } (X_j)_{j=1,\ldots,n} \text{ in } \left[\alpha - \beta\sqrt{\frac{1}{k}}, \alpha\right] \\ X_i, & \text{otherwise} \end{cases}$$

and notice that this is almost surely well defined, since ties occur only with probability zero because F is continuous. Now let $Y_1^{(k)}, Y_2^{(k)}, \ldots$ be independet and identically as $Y^{(k)}$ distributed. Then we know by (S5.25) that for every $k \in \mathbb{N}$

$$\limsup_{n\to\infty} \mathbf{P}\left(\left|\hat{q}_{n,\alpha}\left(Y_1^{(k)}, \ldots, Y_n^{(k)}\right) - q_{Y^{(k)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{k}}\right) = 0. \qquad \text{(S5.27)}$$

Denote by $A_n^{(k)}$ the event, that there are not more than $\lfloor \beta\sqrt{\frac{1}{k}} \cdot n \rfloor$ of the samples $(X_i)_{i=1,\ldots,n}$ in thegalleys interval $\left[\alpha - \beta\sqrt{\frac{1}{k}}, \alpha\right]$. Then the de Moivre-Laplace theorem (cf., e.g., Theorem 1 and Corollary 1 on pp. 47-48 in Chow and Teicher (1978)), which is a special case of the central limit theorem for binomially-distributed random variables, implies for a $B\left(n, \beta\sqrt{\frac{1}{k}}\right)$-

distributed random variable Z, and $p = \beta\sqrt{\frac{1}{k}}$

$$\begin{aligned}
\mathbf{P}\left(A_n^{(k)}\right) &= \sum_{l=0}^{\lfloor pn \rfloor} \binom{n}{l} \cdot \mathbf{P}\left(X \in [\alpha - p, \alpha]\right)^l \cdot \mathbf{P}\left(X \notin [\alpha - p, \alpha]\right)^{n-l} \\
&= \sum_{l=0}^{\lfloor pn \rfloor} \binom{n}{l} \cdot p^l \cdot (1-p)^{n-l} \\
&= \mathbf{P}\left(Z \leq \lfloor pn \rfloor\right) \\
&= \mathbf{P}\left(\frac{Z - \lfloor pn \rfloor}{\sqrt{np(1-p)}} \leq 0\right) \to \frac{1}{2} \quad (n \to \infty)
\end{aligned}$$

and

$$\mathbf{P}\left(\left(A_n^{(k)}\right)^c\right) \to \frac{1}{2} \quad (n \to \infty)$$

for every $k \in \mathbb{N}$. So we can conclude by (S5.27) that for every $k \in \mathbb{N}$

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbf{P}\left(\left|\hat{q}_{n,\alpha}\left(X_{1,n}^{(k)}, \dots, X_{n,n}^{(k)}\right) - q_{Y^{(k)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{k}}\right) \\
&\leq \limsup_{n\to\infty} \left[\mathbf{P}\left(\left\{\left|\hat{q}_{n,\alpha}\left(X_{1,n}^{(k)}, \dots, X_{n,n}^{(k)}\right) - q_{Y^{(k)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{k}}\right\} \cap A_n^{(k)}\right) + \mathbf{P}\left(\left(A_n^{(k)}\right)^c\right)\right] \\
&= 0 + \frac{1}{2} = \frac{1}{2},
\end{aligned} \tag{S5.28}$$

because if we intersect with the event $A_n^{(k)}$ the samples $X_{1,n}^{(k)}, \dots, X_{n,n}^{(k)}$ are in fact samples drawn from the distribution of the random variable $Y^{(k)}$. So for every $k \in \mathbb{N}$ we get in particular for n large enough

$$\mathbf{P}\left(\left|\hat{q}_{n,\alpha}\left(X_{1,n}^{(k)}, \dots, X_{n,n}^{(k)}\right) - q_{Y^{(k)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{k}}\right) \leq \frac{3}{4}. \tag{S5.29}$$

It suffices to show, that there exists a strictly increasing sequence $(n_k)_{k\in\mathbb{N}}$ and data with measurement error $\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}$, fullfilling (S5.24), and $\tilde{\eta}_n$ satisfying (S5.26), such that for every $c_3 > 0$

$$\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}\right) - q_{X,\alpha}\right| > c_3 \cdot \left(\frac{1}{\sqrt{n_k}} + \tilde{\eta}_{n_k}\right)\right) \geq \frac{1}{8} \quad \text{(S5.30)}$$

for k large enough.

We will now sequentially construct such a sequence n_k and the data $\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}$ and show that (S5.30) holds. Choose $n_1 \geq 1$ such that

$$\mathbf{P}\left(\left|\hat{q}_{n_1,\alpha}\left(X^{(1)}_{1,n_1}, \ldots, X^{(1)}_{n_1,n_1}\right) - q_{Y^{(1)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{1}}\right) \leq \frac{3}{4}$$

holds. This is possible because of (S5.29). Given n_{k-1}, choose $n_k > n_{k-1}$ such that $n_k \geq k^2$ and

$$\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(X^{(k)}_{1,n_k}, \ldots, X^{(k)}_{n_k,n_k}\right) - q_{Y^{(k)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{k}}\right) \leq \frac{3}{4}.$$

hold. This is again possible because of (S5.29). Setting

$$\begin{aligned} \bar{X}_{i,n} &= X^{(1)}_{i,n} \quad \text{for} \quad 0 < n \leq n_1 \quad &&\text{and } i = 1, ..., n \quad \text{and} \\ \bar{X}_{i,n} &= X^{(k)}_{i,n} \quad \text{for} \quad n_{k-1} < n \leq n_k \quad &&\text{and } i = 1, ..., n, \end{aligned} \quad \text{(S5.31)}$$

we can conclude for $n_{k-1} < n \leq n_k$

$$\eta_n = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \bar{X}_{i,n}\right| = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - X^{(k)}_{i,n}\right| \leq \frac{1}{n} \cdot \left\lfloor \beta\sqrt{\frac{1}{k}} \cdot n \right\rfloor \cdot \beta\sqrt{\frac{1}{k}} \leq \frac{\beta^2}{k}$$

and in particular

$$\eta_{n_k} \leq \frac{\beta^2}{k} \quad \text{for all } k \in \mathbb{N}$$

and

$$\eta_n \to 0 \quad a.s.$$

In this way we have constructed a strictly increasing sequence $(n_k)_{k\in\mathbb{N}}$ and data with measurement error $\bar{X}_{1,n_k}, ..., \bar{X}_{n_k,n_k}$ such that for all $k \in \mathbb{N}$

$$\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k},\ldots,\bar{X}_{n_k,n_k}\right) - q_{Y^{(k)},\alpha}\right| \geq \frac{\beta}{4}\sqrt{\frac{1}{k}}\right) \leq \frac{3}{4}. \tag{S5.32}$$

By the triangle inequality, we know

$$\begin{aligned}\frac{\beta}{2}\sqrt{\frac{1}{k}} &= \left|q_{Y^{(k)},\alpha} - q_{X,\alpha}\right| \\ &\leq \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k},\ldots,\bar{X}_{n_k,n_k}\right)\right| + \left|\hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k},\ldots,\bar{X}_{n_k,n_k}\right) - q_{X,\alpha}\right|.\end{aligned} \tag{S5.33}$$

Thereby, we can conclude for all $k \in \mathbb{N}$

$$\begin{aligned}&\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k},\ldots,\bar{X}_{n_k,n_k}\right) - q_{X,\alpha}\right| > c_3\cdot\left(\frac{1}{\sqrt{n_k}} + \tilde{\eta}_{n_k}\right)\right) \\ &\geq \mathbf{P}\left(\frac{\beta}{2}\sqrt{\frac{1}{k}} - \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k},\ldots,\bar{X}_{n_k,n_k}\right)\right| > c_3\cdot\left(\frac{1}{\sqrt{n_k}} + \tilde{\eta}_{n_k}\right)\right) \\ &= \mathbf{P}\left(\frac{\beta}{2}\sqrt{\frac{1}{k}} - c_3\cdot\left(\frac{1}{\sqrt{n_k}} + \tilde{\eta}_{n_k}\right) > \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k},\ldots,\bar{X}_{n_k,n_k}\right)\right|\right).\end{aligned}$$

Since $\eta_{n_k} \leq \frac{\beta^2}{k}$, we know by (S5.26)

$$\frac{\tilde{\eta}_{n_k}}{\frac{\beta}{4}\sqrt{\frac{1}{k}}} \leq \frac{4\tilde{\eta}_{n_k}}{\sqrt{\eta_{n_k}}} \to^{\mathbf{P}} 0 \quad (k\to\infty).$$

Furthermore, since $n_k \geq k^2$ for all $k \in \mathbb{N}$ by construction, we have

$$\frac{\frac{1}{\sqrt{n_k}}}{\frac{\beta}{4}\sqrt{\frac{1}{k}}} \leq \frac{\frac{1}{\sqrt{k^2}}}{\frac{\beta}{4}\sqrt{\frac{1}{k}}} \to 0 \quad (k\to\infty),$$

which implies for every $c_3 > 0$

$$\frac{c_3\left(\tilde{\eta}_{n_k} + \frac{1}{\sqrt{n_k}}\right)}{\frac{\beta}{4}\sqrt{\frac{1}{k}}} \to^{\mathbf{P}} 0 \quad (k \to \infty).$$

So setting

$$B_k = \left\{ c_3 \cdot \left(\tilde{\eta}_{n_k} + \frac{1}{\sqrt{n_k}}\right) \leq \frac{\beta}{4}\sqrt{\frac{1}{k}} \right\}$$

yields

$$\mathbf{P}(B_k) \to 1 \;\; (k \to \infty)$$

and thus

$$\mathbf{P}(B_k) \geq \frac{7}{8}$$

for k large enough. Thereby, we finally get for every $c_3 > 0$ and k large enough

$$\begin{aligned}
&\mathbf{P}\left(\left|\hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}\right) - q_{X,\alpha}\right| > c_3 \cdot \left(\tilde{\eta}_{n_k} + \frac{1}{\sqrt{n_k}}\right)\right) \\
&\geq \mathbf{P}\left(\frac{\beta}{2}\sqrt{\frac{1}{k}} - c_3 \cdot \left(\tilde{\eta}_{n_k} + \frac{1}{\sqrt{n_k}}\right) > \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}\right)\right|\right) \\
&\geq \mathbf{P}\left(\left\{\frac{\beta}{2}\sqrt{\frac{1}{k}} - c_3 \cdot \left(\tilde{\eta}_{n_k} + \frac{1}{\sqrt{n_k}}\right) > \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}\right)\right|\right\} \cap B_k\right) \\
&\geq \mathbf{P}\left(\left\{\frac{\beta}{4}\sqrt{\frac{1}{k}} > \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}\right)\right|\right\} \cap B_k\right) \\
&\geq \mathbf{P}\left(\frac{\beta}{4}\sqrt{\frac{1}{k}} > \left|q_{Y^{(k)},\alpha} - \hat{q}_{n_k,\alpha}\left(\bar{X}_{1,n_k}, \ldots, \bar{X}_{n_k,n_k}\right)\right|\right) - \mathbf{P}\left(B_k^c\right) \\
&\geq \frac{1}{4} - \frac{1}{8} = \frac{1}{8},
\end{aligned}$$

where we have used (S5.32) in the last inequality. This yields the assertion.

□

References

Burckel, R. and Bauer, H. (1996). *Probability Theory.* De Gruyter, Berlin.

Chow, Y. S. and Teicher, H. (1978). *Probability Theory: Independence, interchangeability, martingales.* Springer, New York.

Devroye, L., Györfi, L. and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition.* Springer, New York.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.*, **27**, 642-669.

Feldman, D. and Tucker, H. G. (1966). Estimation of non-unique quantiles. *Ann. Math. Statist.*, **37**, 451-457.

Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.*, **18**, 1269-1283.

Tucker, H. G. (1967). *A Graduate Course in Probability.* Academic Press, New York.