Simultaneous concentration of order statistics

Daniel Fresen

To my parents John Fresen and Jill Fresen

ABSTRACT. Let μ be a probability measure on $\mathbb{R}$ with cumulative distribution function F, $(x_i)_1^n$ a large i.i.d. sample from μ, and F_n the associated empirical distribution function. The Glivenko-Cantelli theorem states that with probability 1, F_n converges uniformly to F. In so doing it describes the macroscopic structure of $\{x_i\}_1^n$, however it is insensitive to the position of individual points. Indeed any subset of $o(n)$ points can be perturbed at will without disturbing the convergence.

We provide several refinements of the Glivenko-Cantelli theorem which are sensitive not only to the global structure of the sample but also to individual points. Our main result provides conditions that guarantee simultaneous concentration of all order statistics. The example of main interest is the normal distribution.

1. Introduction

Let μ be a probability measure on $\mathbb{R}$ with cumulative distribution function F and let $(x_i)_1^\infty$ denote an i.i.d. sequence of random variables with distribution μ. For each $n \in \mathbb{N}$ let F_n denote the empirical cumulative distribution function

$$F_n(t) = \frac{1}{n}|\{i \in \mathbb{N} : i \leq n, \, x_i \leq t\}|$$

where $|A|$ denotes the cardinality of a set A. The Glivenko-Cantelli theorem (see e.g. [**8**]) states that with probability 1,

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} |F(t) - F_n(t)| = 0$$

The Dvoretzky-Kiefer-Wolfowitz inequality ([**9**] and [**17**]) provides a quantitative formulation of this and states that for all $n \in \mathbb{N}$ and all $\lambda > 0$, with probability at least $1 - 2\exp(-2\lambda^2)$,

$$\sup_{t\in\mathbb{R}} \sqrt{n}|F(t) - F_n(t)| \leq \lambda$$

2000 *Mathematics Subject Classification.* Primary 62G30; Secondary 60G55.

Key words and phrases. Glivenko-Cantelli theorem, order statistics, log-concave, Lipschitz.

I am grateful to John Fresen and Jill Fresen for my education and for many interesting mathematical discussions throughout the years. Many thanks to Joel Zinn as well as my advisors Alexander Koldobsky and Mark Rudelson for their comments and suggestions.

This titanic theorem would be well deserving of the name '*the fundamental theorem of statistics*' as it is the theoretical foundation behind the idea that a large independent sample is representative of the population. There is, however, a certain crudeness in this noble theorem. Asymptotically, individual points play a negligible role and we learn very little about the finer structure of the sample $\{x_i\}_1^n$. For instance, it gives us almost no information about either the maximum or the minimum. We could take any subset of $o(n)$ points and perturb them as we please without affecting the convergence.

Donsker's theorem (see e.g. [**7**], [**14**] and [**16**]) gives more insight into the structure of the sample. Consider the stochastic process X_n defined on $\mathbb{R}$ by

$$X_n(t) = \sqrt{n}(F_n(t) - F(t))$$

Provided that F is strictly increasing and continuous, X_n converges to a re-scaled Brownian bridge (more precisely, $X_n \circ F^{-1}$ converges to a Brownian bridge on $[0, 1]$). However Donsker's theorem is plagued by a similar insensitivity to the cries of the minority. Through the eyes of Donsker's theorem, we can 'see' subsets as small as $\sqrt{n}$ but are blind to anything smaller such as subsets of size $\log(n)$.

In this paper we provide refined forms of the Glivenko-Cantelli theorem which, under certain conditions, guarantee tight control over all or most points in the sample, not only individually but *simultaneously.* Super-exponential decay of the distribution provides simultaneous concentration of *all* order statistics (see theorem 1) while exponential decay provides simultaneous concentration of *most* order statistics and slightly weaker control over the rest (see theorems 2 and 3). We provide quantitative bounds for log-concave distributions (see theorem 4).

Our results extend the Gnedenko law of large numbers, which guarantees concentration of $\max\{x_i\}_1^n$. They may be compared to the results in [**10**] where the Gnedenko law of large numbers is extended to the multi-dimensional setting, to the paper [**13**] that provides estimates of order statistics in terms of Orlicz functions and to the article [**1**] that concerns optimal matchings of random points uniformly distributed within the unit square. We refer the reader to [**11**] and [**19**] for an extensive treatment of empirical process theory and to [**2**], [**4**] and [**18**] for information on order statistics. Interesting papers on the Glivenko-Cantelli theorem include [**5**], [**20**], [**21**] and [**22**].

Theorem 1. *Let μ be any probability measure on $\mathbb{R}$ with a continuous strictly increasing cumulative distribution function F such that for all $\varepsilon > 0$*

$$\lim_{t\to\infty} \frac{1 - F(t + \varepsilon)}{1 - F(t)} = \lim_{t\to-\infty} \frac{F(t)}{F(t + \varepsilon)} = 0 \tag{1.1}$$

Then there exists a sequence $(\delta_n)_1^\infty$ with $\lim_{n\to\infty} \delta_n = 0$ such that for all $n \in \mathbb{N}$, if $(x_i)_1^n$ is an i.i.d. sample from μ with corresponding order statistics $(x_{(i)})_1^n$, then with probability at least $1 - \delta_n$,

$$\sup_{1\leq i\leq n} |x_{(i)} - x^*_{(i)}| \leq \delta_n \tag{1.2}$$

where $x^*_{(i)} = F^{-1}(i/(n + 1))$.

Theorem 2. *Let μ be any probability measure on $\mathbb{R}$ with a continuous strictly increasing cumulative distribution function F such that for all $\varepsilon > 0$*

$$\limsup_{t\to\infty} \frac{1-F(t+\varepsilon)}{1-F(t)} \quad < \quad 1 \tag{1.3}$$

$$\limsup_{t\to-\infty} \frac{F(t)}{F(t+\varepsilon)} \quad < \quad 1 \tag{1.4}$$

Let $(\omega_n)_1^\infty$ be any sequence in $\mathbb{N}$ with $\lim_{n\to\infty} \omega_n = \infty$. Then there exists a sequence $(\delta_n)_1^\infty$ with $\lim_{n\to\infty} \delta_n = 0$, such that for all $n \in \mathbb{N}$, if $(x_i)_1^n$ is an i.i.d. sample from μ with corresponding order statistics $(x_{(i)})_1^n$, then with probability at least $1 - \delta_n$,

$$\sup_{\omega_n \le i \le n-\omega_n} |x_{(i)} - x^*_{(i)}| \le \delta_n$$

where $x^*_{(i)} = F^{-1}(i/(n+1))$.

Theorem 3. *Let μ be any probability measure on $\mathbb{R}$ that obeys the conditions of theorem 2. Then there exists $k > 0$ such that for all $T > 10^6$ and all $n \in \mathbb{N}$, if $(x_i)_1^n$ is an i.i.d. sample from μ with corresponding order statistics $(x_{(i)})_1^n$, then with probability at least $1 - 400T^{-1/2}$,*

$$\sup_{1\le i\le n} |x_{(i)} - x^*_{(i)}| \le kT$$

Note that in theorem 2 we can take $(\omega_n)_1^\infty$ to grow arbitrarily slowly, for example let $\omega_n = \log\log\log n$. We thus have tight control over almost the entire data set with the exception of a *very* small proportion of points. This is substantially better than the $\sqrt{n}$ 'visibility' of Donsker's theorem.

A probability measure μ is called p-log-concave for some $p \in (0, \infty)$ if it has a density function of the form $f(x) = c\exp(-g(x)^p)$ where g is non-negative and convex. The 1-log-concave distributions are simply referred to as log-concave. If μ is p-log-concave then it is also q-log-concave for all $1 \le q \le p$.

Theorem 4. *Let $p > 1$, $q > 0$ and let μ be a p-log-concave probability measure on $\mathbb{R}$ with a continuous strictly increasing cumulative distribution function F. Then there exists $c > 0$ such that for any $n \in \mathbb{N}$ and any i.i.d. sample $(x_i)_1^n$ from μ with order statistics $(x_{(i)})_1^n$, with probability at least $1 - c(\log n)^{-q}$,*

$$\sup_{1\le i\le n} |x_{(i)} - x^*_{(i)}| \le c\frac{\log\log n}{(\log n)^{1-1/p}}$$

where $x^*_{(i)} = F^{-1}(i/(n+1))$.

The main idea behind the proof of these theorems is to first analyze the uniform distribution on $[0, 1]$. We do this using a powerful representation of the empirical point process via independent random variables that allows us to use classical results such as the law of large numbers (in the form of Chebyshev's inequality) and the law of the iterated logarithm. A key step in this analysis is to exploit the inherent regularity of order statistics which allows for control over all points based on an inspection of merely $\log n$ carefully chosen points. We then transform the points under the action of F^{-1} to analyze the general case. We introduce a new class of metrics on $(0, 1)$ defined by

$$\theta_p(x, y) = \max\left\{ \frac{\log(x^{-1}y)}{(\log x^{-1})^{1-1/p}}, \frac{\log((1-y)^{-1}(1-x))}{(\log(1-y)^{-1})^{1-1/p}} \right\} \tag{1.5}$$

for $1 \leq p < \infty$ and $0 < x \leq y < 1$. To see that each θ_p is indeed a metric, note that $\theta_p(x, y)$ is decreasing in x and increasing in y throughout the triangular region $\{(x, y) \in (0, 1)^2 : x < y\}$. We show that F^{-1} is either Lipschitz or uniformly continuous with respect to these metrics (depending on the assumptions imposed on μ). After this, our main results become straightforward to prove.

There are endless variations on the main theme of this paper. Our intention is simply to highlight a phenomenon and introduce methods by which to study it. Note that our results are purely asymptotic in nature and we can (and do) assume throughout the paper that $n > n_0$ for some $n_0 \in \mathbb{N}$.

2. The uniform distribution

Let $(\gamma_i)_1^n$ denote an i.i.d. sample from the uniform distribution on $[0, 1]$ with corresponding order statistics $(\gamma_{(i)})_1^n$ and let $(z_i)_1^{n+1}$ be an i.i.d. sequence of random variables that follow the standard exponential distribution. For $1 \leq i \leq n$ define

$$y_i = \left(\sum_{j=1}^{i} z_j \right) \left(\sum_{j=1}^{n+1} z_j \right)^{-1}$$

It is of great interest to us that $(y_i)_1^n$ and $(\gamma_{(i)})_1^n$ have the same distribution in $\mathbb{R}^n$ (see chapter 5 in [**6**]). This is nothing but an expression of the fact that the empirical point process locally resembles the Poisson point process. Also of interest is the fact that these random vectors have the same distribution as the partial sums of a random vector uniformly distributed (with respect to Lebesgue measure) in the standard simplex $\Delta^n = \{w \in \mathbb{R}^{n+1} : w_i \geq 0 \ \forall i, \sum_i w_i = 1\}$. The power of this representation is that we have an expression for $(\gamma_{(i)})_1^n$ in terms of independent random variables. Note that

$$y_i = \frac{i}{n+1} \left(\frac{1}{i} \sum_{j=1}^{i} z_j \right) \left(\frac{1}{n+1} \sum_{j=1}^{n+1} z_j \right)^{-1} \tag{2.1}$$

Both lemma 1 and lemma 3 below can be compared to the results in [**23**].

Lemma 1. *Let $T > 10^6$ and $n \in \mathbb{N}$. With probability at least $1 - 400T^{-1/2}$ the following inequalities hold simultaneously for all $1 \leq i \leq n$,*

$$T^{-1} \leq \gamma_{(i)} \left(\frac{i}{n+1} \right)^{-1} \leq T \tag{2.2}$$

$$T^{-1} \leq (1 - \gamma_{(i)}) \left(1 - \frac{i}{n+1} \right)^{-1} \leq T \tag{2.3}$$

Proof. Let $Q = 2^{-1}T^{1/2}$ and momentarily fix $1 \leq i \leq n+1$. The random variable $i^{-1} \sum_{j=1}^{i} z_j$ has mean 1 and variance i^{-1}. Using Chebyshev's inequality, with probability at least $1 - i^{-1}Q^{-2}$ we have

$$-Q < 1 - \frac{1}{i} \sum_{j=1}^{i} z_j < Q$$

The random variable

$$U_i = |\{j \in \mathbb{N} : j \leq i, z_j \leq 2Q^{-1}\}|$$

follows a binomial distribution with i trials and success probability $1-\exp(-2Q^{-1}) \leq 2Q^{-1}$. Using Chebyshev's inequality again, with probability at least $1-32i^{-1}Q^{-1}$ we have $U_i < i/2$, which implies that $i^{-1}\sum_{j=1}^{i} z_j > Q^{-1}$. Hence, with probability at least $1-33i^{-1}Q^{-1}$ we have

$$Q^{-1} < \frac{1}{i}\sum_{j=1}^{i} z_j < Q+1 \tag{2.4}$$

Let $M = \lfloor \log_2(n) \rfloor$. With probability at least $1 - 33Q^{-1}\sum_{j=0}^{M} 2^{-j} - 33(n+1)^{-1}Q^{-1} \geq 1-100Q^{-1}$equation (2.4) holds simultaneously for $i = 1, 2, 2^2, 2^3 \ldots 2^M$ and for $i = n+1$. Hence, by (2.1), with probability at least $1-100Q^{-1}$ we have that for all such i

$$\frac{1}{2}Q^{-2}\frac{i}{n+1} \leq y_i \leq 2Q^2\frac{i}{n+1}$$

Since $(y_i)_1^n$ is an increasing sequence, control over the values $(y_{2^j})_{j=1}^M$ leads to control over the entire sequence and, recalling the representation of $(\gamma_{(i)})_1^n$ in terms of $(y_i)_1^n$, the bound (2.2) follows for all $1 \leq i \leq n$. The bound (2.3) then follows by symmetry. □

Lemma 2. *Let* $t \in (0,1)$ *and* $n \in \mathbb{N}$. *With probability at least* $1 - 2\exp(-nt^2/5)$ *the following inequality holds simultaneously for all* $1 \leq i \leq n$*,*

$$\left|\gamma_{(i)} - \frac{i}{n+1}\right| \leq t \tag{2.5}$$

Proof. We can assume without loss of generality that $n^{-1} \leq 2t/3$ (otherwise the probability bound becomes trivial). Note that since our sample is taken from the uniform distribution we have

$$\begin{aligned}\sup_{1\leq i\leq n} |\gamma_{(i)} - i(n+1)^{-1}| &\leq n^{-1} + \sup_{1\leq i\leq n} |\gamma_{(i)} - in^{-1}| \\ &= n^{-1} + \sup_{0\leq t\leq 1} |F_n(t) - F(t)|\end{aligned}$$

where $F(t) = t$ is the cumulative distribution function and F_n is the empirical distribution function. By the Dvoretzky-Kiefer-Wolfowitz inequality (as mentioned in the introduction), with probability at least $1 - 2\exp(-5^{-1}nt^2)$ we have

$$\sup_{0\leq t\leq 1} |F_n(t) - F(t)| \leq t/3$$

and the result follows. □

Note that in the preceding proof one can also use Doob's martingale inequality (in the form of Kolmogorov's inequality) and the representation of $(\gamma_{(i)})_1^n$ in terms of $(y_n)_1^n$, although this approach yields an inferior probability bound.

Lemma 3. *Let* $(\omega_n)_1^\infty$ *be any sequence in* $\mathbb{N}$ *such that* $\lim_{n\to\infty}\omega_n = \infty$. *Then for all* $T > 1$ *and all* $\delta \in (0,1)$ *there exists* $n_0 \in \mathbb{N}$ *such that for all* $n > n_0$*, if* $(\gamma_{(i)})_1^n$ *are the order statistics from an i.i.d. sample from the uniform distribution on* $[0,1]$*, then with probability at least* $1-\delta$*, (2.2) and (2.3) hold for all* $\omega_n \leq i \leq n - \omega_n$*.*

Proof. We use the representation (2.1). Let $T > 1$ and $\delta \in (0,1)$ be given. Without loss of generality we may assume that $T \leq 2$. Let $(\widetilde{z}_i)_1^\infty$ denote any i.i.d.

sequence of random variables that follow the standard exponential distribution. Define the deterministic sequence $(\lambda_j)_1^\infty$ as follows,

$$\lambda_j = \mathbb{P}\{\sup_{i\geq j}(2i\log\log i)^{-1/2}\left|\sum_{k=1}^{i}(\widetilde{z}_k - 1)\right| \leq 2\}$$

Note that $(\lambda_j)_1^\infty$ is an increasing sequence and by the law of the iterated logarithm, $\lim_{j\to\infty}\lambda_j = 1$. Fix $n_0 \in \mathbb{N}$ with $n_0 \geq 64\delta^{-1}(T^{1/2}-1)^{-2}$ such that for all $n > n_0$ we have the following inequalities,

$$\begin{aligned} \lambda_{\omega(n)} &\geq 1-\delta/4 \\ \left(\frac{8\log\log\omega_n}{\omega_n}\right)^{1/2} &\leq T^{1/2}-1 \end{aligned}$$

Now consider any $n > n_0$ and let $(\gamma_{(i)})_1^n$ denote the order statistics mentioned in the statement of the lemma. With probability at least $1-\delta/4$, for all $\omega(n) \leq i \leq n$,

$$\begin{aligned} \left|1-\frac{1}{i}\sum_{j=1}^{i}z_j\right| &\leq \left(\frac{8\log\log\omega_n}{\omega_n}\right)^{1/2} \\ &\leq T^{1/2}-1 \end{aligned}$$

By Chebyshev's inequality and the fact that the function $u \mapsto u^{-1}$ is 4-Lipschitz on $[1/2,\infty)$, with probability at least $1-16n^{-1}(T^{1/2}-1)^{-2} \geq 1-\delta/4$

$$\left|1-\left(\frac{1}{n+1}\sum_{j=1}^{n+1}z_j\right)^{-1}\right| < T^{1/2}-1$$

By (2.1), with probability at least $1-\delta/2$, (2.2) holds for all $\omega(n) \leq i \leq n$. By symmetry, with the same probability (2.3) holds for all $1 \leq i \leq n-\omega(n)$. The lemma is thus proven. □

3. The general case

Lemma 4. *Let F be a continuous strictly increasing cumulative distribution function that satisfies (1.1). Then F^{-1} is continuous and for all $T > 1$ and all $\delta > 0$ there exists $\eta \in (0,1)$ such that for all $x, y \in (0,\eta)$ with $T^{-1} \leq xy^{-1} \leq T$ and all $x, y \in (1-\eta, 1)$ with $T^{-1} \leq (1-x)(1-y)^{-1} \leq T$ we have $|F^{-1}(x) - F^{-1}(y)| \leq \delta$.*

Proof. Consider any $T > 1$ and $\delta > 0$. By (1.1) there exists $t_0 \in \mathbb{R}$ such that for all $t \leq t_0$, $TF(t) < F(t+\delta)$. Let $\eta_1 = F(t_0)$. Consider any $x, y \in (0, \eta_1)$ such that $T^{-1} \leq xy^{-1} \leq T$. Without loss of generality, $x < y$. Let $s = F^{-1}(x)$ and $t = F^{-1}(y)$. Then $s \leq t_0$, hence $F(t) = y \leq Tx = TF(s) < F(s+\delta)$, from which it follows that $t < s+\delta$ and that $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. Analysis of the right hand tail is identical and provides us with $\eta_2 > 0$ such that for all $x, y \in (1-\eta_2, 1)$ with $T^{-1} \leq (1-x)(1-y)^{-1} \leq T$ we have $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. The result follows with $\eta = \min\{\eta_1, \eta_2\}$. □

Lemma 5. *Let F be a continuous strictly increasing cumulative distribution function that satisfies both (1.3) and (1.4). Then F^{-1} is continuous and for all $\delta > 0$ there exists $T > 1$ such that for all $x, y \in (0,1)$ such that $T^{-1} \leq xy^{-1} \leq T$ and $T^{-1} \leq (1-x)(1-y)^{-1} \leq T$ we have $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. In particular, F^{-1} is uniformly continuous with respect to the metric θ_1 (see (1.5)).*

Proof. Consider any $\delta > 0$. By (1.4) there exists $T_1 > 1$ and $t_0 \in \mathbb{R}$ such that for all $t < t_0$, $T_1 F(t) \leq F(t+\delta)$. Let $\eta_1 = \min\{F(t_0), 2^{-1}\}$. As in the proof of the previous lemma, it follows that for all $x, y \in (0, \eta_1)$ with $T_1^{-1} \leq xy^{-1} \leq T_1$ we have $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. Similarly (using (1.3)), there exists $T_2 > 1$ and $\eta_2 \in (2^{-1}, 1)$ such that for all $x, y \in (\eta_2, 1)$ with $T_2^{-1} \leq (1-x)(1-y)^{-1} \leq T_2$ we have $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. By continuity of F^{-1} relative to the standard topology on $(0,1)$, and by compactness of $[2^{-1}\eta_1, 1 - 2^{-1}\eta_2]$ there exists $0 < \delta' < 10^{-1}\min\{\eta_1, \eta_2\}$ such that for all $x, y \in [2^{-1}\eta_1, 1-2^{-1}\eta_2]$ with $|x-y| < \delta'$ we have $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. We leave it to the reader to verify that the result holds with

$$T = \min\{T_1, T_2, 1+\delta'\}$$

□

Proof of theorem 1. We shall construct a function h that takes an arbitrary $\delta \in (0,1)$ and produces an appropriate $n_0 = h(\delta) \in \mathbb{N}$. Then, using this function we shall define the desired sequence $(\delta_n)_1^\infty$ that is mentioned in the statement of the theorem. To this end, let $\delta \in (0,1)$ be given. Define

$$T = 10^6 \delta^{-2} \tag{3.1}$$

By lemma 4 there exists $\eta \in (0,1)$ such that if $x, y \in (0, \eta)$ and $T^{-1} \leq xy^{-1} \leq T$, or $x, y \in (1-\eta, 1)$ and $T^{-1} \leq (1-x)(1-y)^{-1} \leq T$, then $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. By compactness, F^{-1} is uniformly continuous on $[\eta/2, 1-\eta/2]$, which implies the existence of $t \in (0, \eta/2)$ such that if $x, y \in [\eta/2, 1-\eta/2]$ and $|x-y| \leq t$, then $|F^{-1}(x) - F^{-1}(y)| \leq \delta$. Define

$$n_0 = \left\lceil 5t^{-2}\log(4\delta^{-1})\right\rceil \tag{3.2}$$

and consider any $n \geq n_0$. Let $(\gamma_{(i)})_1^n$ denote the order statistics corresponding to an i.i.d. sample from the uniform distribution on $[0,1]$. Note that we have the representation

$$x_{(i)} = F^{-1}(\gamma_{(i)}) \tag{3.3}$$

valid for all $1 \leq i \leq n$. By lemmas 1 and 2, as well as equations (3.1) and (3.2), with probability at least $1-\delta$ inequalities (2.2), (2.3) and (2.5) hold simultaneously for all $1 \leq i \leq n$. Suppose that these inequalities do indeed hold and consider any fixed $1 \leq i \leq n$. Since $t \leq \eta/2$, one of the three sets $[0, \eta]$, $[\eta/2, 1-\eta/2]$ and $[1-\eta, 1]$ contains both $\gamma_{(i)}$ and $i(n+1)^{-1}$, which implies that $|F^{-1}(\gamma_{(i)}) - F^{-1}(i(n+1)^{-1})| \leq \delta$, which is inequality (1.2).

Define the non-decreasing sequence $(\kappa_n)_1^\infty$ by $\kappa_n = \max\{h(e^{-i}) : 1 \leq i \leq n\}$ and set

$$\delta_n = \exp(-\max\{i \in \mathbb{N} : \kappa_i \leq n\})$$

where we define $\max\emptyset = 0$. It is clear that $\lim_{n\to\infty}\delta_n = 0$. Consider any fixed $n \in \mathbb{N}$. If $\{i \in \mathbb{N} : \kappa_i \leq n\} = \emptyset$ then the probability bound is trivial, otherwise let $j = \max\{i \in \mathbb{N} : \kappa_i \leq n\}$. The result follows by the inequality $h(\delta_n) = h(e^{-j}) \leq \kappa_j \leq n$ and by definition of the function h. □

Proof of theorems 2 and 3. The proof is very similar to that of theorem 1. We use the representation (3.3). The main difference is that we use lemmas 3 and 5 instead of lemmas 1 and 4. The details are left to the reader. □

4. Log-concave distributions

The following two lemmas are modifications of lemmas 6 and 9 in [**10**].

Lemma 6. *Let μ be a log-concave probability measure on $\mathbb{R}$ with a continuous strictly increasing cumulative distribution function F. Then there exists $c > 0$ such that for all $0 < x < y < 1$,*

$$|F^{-1}(y) - F^{-1}(x)| \leq c \max\left\{ \left|F^{-1}(y)\right| \frac{\log(x^{-1}y)}{\log y^{-1}}, \left|F^{-1}(x)\right| \frac{\log((1-x)/(1-y))}{\log(1-x)^{-1}} \right\} \tag{4.1}$$

Proof. By theorem 5.1 in [**15**] (see lemma 5 in [**10**] for a proof) F is log-concave. Hence the function $u(t) = -\log F(t)$ is convex (and strictly decreasing). Let $\mathbb{E}\mu$ denote the centroid of μ (the expected value of a random variable with distribution μ). By lemma 5.12 in [**15**] (see also lemma 3.3 in [**3**]) $F(\mathbb{E}\mu) \geq e^{-1}$, hence $u(\mathbb{E}\mu) \leq 1$. By convexity of u we have the inequality $(t-s)^{-1}(u(t)-u(s)) \leq (\mathbb{E}\mu - t)^{-1}(u(\mathbb{E}\mu) - u(t))$, which is valid for all $s < t < \mathbb{E}\mu$. Let $0 < x < y < \min\{e^{-2}, F(0), F(-2\mathbb{E}\mu)\}$ and define $s = F^{-1}(x)$ and $t = F^{-1}(y)$. Then we have

$$F^{-1}(y) - F^{-1}(x) \leq (\mathbb{E}\mu - F^{-1}(y)) \frac{\log(x^{-1}y)}{\log y^{-1} - u(\mathbb{E}\mu)}$$

It follows from the restrictions on y that $F^{-1}(y) < 0$ and that $\left|F^{-1}(y)\right| \geq 2\,|\mathbb{E}\mu|$. Since $y < F(\mathbb{E}\mu)^2$, it follows that $\log y^{-1} > 2u(\mathbb{E}\mu)$ and (4.1) follows for such x and y with $c = 4$. For other values of x and y, inequality (4.1) follows by compactness, continuity and symmetry. □

Lemma 7. *Let $p \geq 1$ and let μ be a p-log-concave probability measure on $\mathbb{R}$ with cumulative distribution function F. Then there exists $c > 0$ such that for all $x \in (0,1)$,*

$$|F^{-1}(x)| \leq c \max\{(\log x^{-1})^{1/p}, (\log(1-x)^{-1})^{1/p}\} \tag{4.2}$$

As a consequence of (4.2) and (4.1), F^{-1} is Lipschitz with respect to the metric θ_p (see (1.5)).

Proof. By lemma 9 in [**10**] (which holds for $p \geq 1$) there exists $c_1, c_2 > 0$ and $t_0 > 1$ such that for all $t < -t_0$, $F(t) \leq c_1|t|^{1-p}\exp(-c_2|t|^p)$. Let $\eta_1 = \min\{F(-t_0), c_1^{-1}\}$ and consider any $x \in (0, \eta_1)$. Let $t = F^{-1}(x)$. Hence $x = F(t) \leq c_1|t|^{1-p}\exp(-c_2|t|^p)$, which implies that

$$\begin{aligned} |F^{-1}(x)| &= -t \\ &\leq (c_2^{-1}(\log c_1 + \log x^{-1}))^{1/p} \\ &\leq 2^{1/p} c_2^{-1/p} (\log x^{-1})^{1/p} \end{aligned}$$

The result now follows by symmetry, compactness and continuity. □

Lemma 8. *Let F be a continuous strictly increasing cumulative distribution function associated to a log-concave probability measure. Then there exists $c > 0$ such that for all $\varepsilon \in (0, 1/2)$ and all $x, y \in [\varepsilon, 1 - \varepsilon]$,*

$$|F^{-1}(x) - F^{-1}(y)| \leq c\varepsilon^{-1}|x - y|$$

PROOF. This follows from lemmas 6 and 7 with $p = 1$ and the inequality $\log t \leq t - 1$. □

PROOF OF THEOREM 4. By lemmas 1, 6 and 7, with probability at least $1 - 400(\log n)^{-q}$, for all $i \leq n^{3/4}$ and all $i \geq n - n^{3/4}$ we have

$$|x_{(i)} - x^*_{(i)}| \leq c\frac{\log \log n}{(\log n)^{1-1/p}}$$

Let $I = [2^{-1}n^{-1/4}, 1 - 2^{-1}n^{-1/4}]$. By lemma 8, for all $x, y \in I$ we have

$$|F^{-1}(x) - F^{-1}(y)| \leq cn^{1/4}|x - y|$$

By lemma 2, with probability at least $1 - 2\exp(-5n^{1/4})$, for all $1 \leq i \leq n$ we have

$$|\gamma_{(i)} - i(n + 1)^{-1}| \leq n^{-3/8}$$

Hence for all $n^{3/4} \leq i \leq n - n^{3/4}$ both $\gamma_{(i)}$ and $i(n + 1)^{-1}$ are elements of I and the result follows. □

References

[1] Ajtai, M., Komlòs, J., Tusnàdy, G.: On optimal matchings. Combinatorica **4,** 259-264 (1984)

[2] Balakrishnan, N., Clifford Cohen, A.: Order Statistics and Inference. Statistical Modeling and Descision Science. Academic Press (1991)

[3] Bobkov, S.: On concentration of distributions of random weighted sums. Ann. Probab. **31** (1), 195-215 (2003)

[4] David, H. A.: Order Statistics. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons (1970)

[5] Dehardt, J.: Generalizations of the Glivenko-Cantelli theorem. Ann. Math. Statist. **42** (6), 2050-2055 (1971)

[6] Devroye, L.: Non-Uniform Random Variate Generation. Originally published with Springer-Verlag, New York (1986)

[7] Donsker, M. D.: Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. Ann. Math. Statist. **23**, 277–281 (1952)

[8] Dudley, R. M.: Real Analysis and Probability. Wadsworth & Brooks/Cole (1989)

[9] Dvoretzky, A., Kiefer, J., Wolfowitz, J.: Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. Ann. Math. Statist. **27** (3), 642–669 (1956)

[10] Fresen, D.: A multivariate Gnedenko law of large numbers. arXiv:1101.4887v1

[11] Gaensller, P., Stute, W.: Empirical processes: a survey of results for independent and identically distributed random variables. Ann. Probab. **7** (2) 193-243 (1979)

[12] Gnedenko, B.: Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math. **44,** 423-453 (1943)

[13] Gordon, Y., Litvak, A., Schütt, C., Werner, E.: Uniform estimates for order statistics and Orlicz functions. arXiv:0809.2989v1

[14] Komlòs, J., Major, P., Tusnàdy, G.: An approximation of partial sums of independent RV'-s and the sample DF. I. Z. Wahrscheinlichkeitstheorie verw. Gebiete **32**, 111–131. (1975)

[15] Lovász, L., Vempala, S.: The geometry of logconcave functions and sampling algorithms. Random Structures Algorithms **30** (3), 307-358 (2007)

[16] Mason, D. M., van Zwet, W.: A refinement of the KMT inequality for the uniform empirical process. Ann. Probab. **15**, 871-884 (1987)

[17] Massart, P.: The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. Ann. Probab. **18** (3), 1269–1283 (1990)
[18] Sarhan, A. E., Greenberg, B. G. (eds.): Contributions to Order Statistics. Wiley Publications in Statistics. John Wiley & Sons (1962)
[19] Shorack, G., Wellner, J.: Empirical Processes with Applications to Statistics. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons (1986)
[20] Talagrand, M.: The Glivenko-Cantelli problem. Ann. Probab. **15** (3), 837-870 (1987)
[21] Talagrand, M.: The Glivenko-Cantelli problem, ten years later. J. Theoret. Probab. **9** (2), 371-384 (1996)
[22] Wellner, J.: A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. Ann. Statist. **5** (3), 473-480 (1977)
[23] Wellner, J.: Limit theorems for the ratio of the empirical distribution function to the true distribution function. Probab. Theory Relat. Fields **45** (1) 73-88 (1978)

Department of Mathematics, University of Missouri
E-mail address: djfb6b@mail.missouri.edu