Stat 310A/Math 230A Theory of Probability

Final

Andrea Montanari *December 7, 2015*

Solutions should be complete and concisely written. Please, mark clearly the beginning and end of each problem.

You have 3 hours but you are not required to solve all the problems!

Just solve those that you can solve within the time limit. Points assigned to each problem are indicated in parenthesis. I recommend to look at all problems before starting.

For any clarification on the text, one of the TAs will be outside the room.

You can consult textbooks and your notes. You cannot use computers, and in particular you cannot use the web. You can cite theorems (propositions, corollaries, lemmas, etc.) from Amir Dembo's lecture notes by number, and exercises you have done as homework by number as well. Any other non-elementary statement must be proved!

Problem 1 (20 points)

(*a*) Let X be a random variable taking values in $\mathbb{R}$. Prove that, for each $\epsilon > 0$, there exists $x_0 \in \mathbb{R}$ such that

$$\mathbb{P}\big(X \in [x_0 - \epsilon, x_0 + \epsilon]\big) > 0\,. \tag{1}$$

[Notice: The inequality is strict!]

(*b*) Let X, Y be independent and identically distributed. Prove that, for any $\epsilon > 0$,

$$\mathbb{P}\big(|X - Y| \le \epsilon\big) > 0\,. \tag{2}$$

Problem 2 (15 points)

Given two distribution functions F, G on $\mathbb{R}$, define their Wasserstein distance as

$$W_1(F, G) \equiv \int_{\mathbb{R}} |F(t) - G(t)|\, \mathrm{d}t \tag{3}$$

(whenever this integral is finite.)

Let $\{X_i\}_{i\ge 1}$ be a collection of i.i.d. random variable, with common distribution F, such that $\mathbb{E}\{|X_1|\} < \infty$. Define the empirical distribution $F_n : \mathbb{R} \to [0, 1]$ as

$$F_n(t) \equiv \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{t \ge X_i\}}\,. \tag{4}$$

The following result might turn useful (Theorem 2.3.6 in the lecture notes).

Theorem 1 (Glivenko-Cantelli)**.** *The following limit holds almost surely:*

$$\lim_{n\to\infty} \big\{ \sup_{x\in\mathbb{R}} \big|F_n(x) - F(x)\big| \big\} = 0\,. \tag{5}$$

Prove that

$$\lim_{n\to\infty} W_1(F_n, F) = 0\,. \tag{6}$$

[Hint: As a warm-up step, it might be useful to consider initially the case of bounded random variables, i.e. $|X_i| \le M$ almost surely, for some non-random constant M.]

Problem 3 (20 points)

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables, with $X_n \sim \mathsf{N}(a_n, b_n)$, for some $a_n \in \mathbb{R}$, $b_n \in \mathbb{R}_{\geq 0}$ (in words, X_n is Gaussian, with mean $\mathbb{E}\{X_n\} = a_n$ and variance $\mathrm{Var}(X_n) = b_n$). As usual, the case $b_n = 0$ is admitted (a random variable with distribution $\mathsf{N}(a, 0)$ is equal to a with probability one).

(*a*) Assume $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$, for some $a \in \mathbb{R}$, $b \in \mathbb{R}_{\geq 0}$. Prove that X_n converges in distribution to $X \sim \mathsf{N}(a, b)$.

(*b*) Prove the converse statement. In other words, assume $X_n \xrightarrow{\mathrm{d}} X$. Prove that there exist $a \in \mathbb{R}$, $b \in \mathbb{R}_{\geq 0}$ such that $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$, and that $X \sim \mathsf{N}(a, b)$.

Problem 4 (30 points)

Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables, uniformly distributed in the interval $[0, 1]$. For each n, let $Y_1^{(n)} \leq Y_2^{(n)} \leq \cdots \leq Y_n^{(n)}$ be the variables obtained by ordering $\{X_1, X_2, \ldots, X_n\}$.

(*a*) Prove that, for each $n \in \mathbb{N}$, $\ell \in \{1, \ldots, n\}$, $Y_\ell^{(n)}$ is a random variable.

(*b*) Define the rescaled median

$$Z_n \equiv \sqrt{4n}\Big(Y^{(n)}_{\lfloor n/2\rfloor} - \frac{1}{2}\Big)\,. \tag{7}$$

(Here $\lfloor x\rfloor$ denotes the larges integer j such that $j \leq x$.) Prove that Z_n converges in distribution (as $n \to \infty$) to a standard normal random variable $\mathsf{N}(0, 1)$.

(*c*) Assume now that $\{X_i\}_{i\in\mathbb{N}}$ are i.i.d. with common distribution F, which is differentiable with continuous, strictly positive derivative $F'(x) > 0$ for all $x \in \mathbb{R}$. Let x_0 be the unique solution of $F(x) = 1/2$. Define as above $Y_1^{(n)} \leq Y_2^{(n)} \leq \cdots \leq Y_n^{(n)}$ to be the variables obtained by ordering $\{X_1, X_2, \ldots, X_n\}$. Define

$$Z_n \equiv \sqrt{4n}\big(Y^{(n)}_{\lfloor n/2\rfloor} - x_0\big)\,, \tag{8}$$

Prove that, as $n \to \infty$, $Z_n \xrightarrow{\mathrm{d}} Z_\infty \sim \mathsf{N}(0, \sigma^2)$, and compute the variance of the limit σ^2.

Problem 5 (40 points)

This exercise aims at generalizing some results that we proved in the context of real random variables. Let $\mathcal{H}$ be a separable Hilbert space and denote by $\langle \boldsymbol{x}, \boldsymbol{y}\rangle_{\mathcal{H}}$ the scalar product between vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$, and by $\|\boldsymbol{x}\|_{\mathcal{H}} \equiv \langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$ the corresponding norm. It might be useful to recall the following facts:

- $\mathcal{H}$ is a vector space, and is complete for the topology induced by $\|\cdot\|_{\mathcal{H}}$.
- $\mathcal{H}$ has a countable orthonormal basis, i.e. a countable set of vectors $\{\boldsymbol{e}_i\}_{i\in\mathbb{N}}$, with $\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = 0$ for $i \neq j$ and $\langle \boldsymbol{e}_i, \boldsymbol{e}_i\rangle = 1$. Any vector $\boldsymbol{x} \in \mathcal{H}$ is represented as $\boldsymbol{x} = \sum_{i=1}^{\infty} x_i \boldsymbol{e}_i$ (with convergence in the norm $\|\cdot\|_{\mathcal{H}}$).
- Viceversa, for any $\{x_i\}_{i\in\mathbb{N}}$, such that $\sum_{i=1}^{\infty} x_i^2 < \infty$, the series $\sum_{i=1}^{\infty} x_i \boldsymbol{e}_i$ converges to a vector $\boldsymbol{x} \in \mathcal{H}$.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a $\mathcal{H}$-valued random variable is a measurable mapping $\boldsymbol{Z} : (\Omega, \mathcal{F}) \to (\mathcal{H}, \mathcal{B}_{\mathcal{H}})$. We will also call $\boldsymbol{Z}$ a random vector.

(*a*) For a random vector $\boldsymbol{Z}$ such that $\mathbb{E}\{\|\boldsymbol{Z}\|_{\mathcal{H}}\} < \infty$, $\mathbb{E}\{\boldsymbol{Z}\}$ is defined as the unique vector such that $\langle \mathbb{E}\{\boldsymbol{Z}\}, \boldsymbol{v}\rangle = \mathbb{E}\{\langle \boldsymbol{Z}, \boldsymbol{v}\rangle\}$ for all $\boldsymbol{v} \in \mathcal{H}$.

Prove that this definition is well posed. In other words, you need to prove that there exists a vector $\boldsymbol{u} \in \mathcal{H}$ such that $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \mathbb{E}\{\langle \boldsymbol{Z}, \boldsymbol{v}\rangle\}$, and that this vector is unique

(*b*) Let $\boldsymbol{Z}$, $\boldsymbol{Y}$ be two independent random vectors, such that $\mathbb{E}\{\|\boldsymbol{Z}\|_{\mathcal{H}}^2\} < \infty$,$\mathbb{E}\{\|\boldsymbol{Y}\|_{\mathcal{H}}^2\} < \infty$. Prove that

$$\mathbb{E}\{\langle \boldsymbol{Z}, \boldsymbol{Y}\rangle_{\mathcal{H}}\} = \langle \mathbb{E}\{\boldsymbol{Z}\}, \mathbb{E}\{\boldsymbol{Y}\}\rangle_{\mathcal{H}} \,. \tag{9}$$

[As a part of this proof, you are required to prove that the left-hand side is well defined, i.e. that $\langle \boldsymbol{Z}, \boldsymbol{Y}\rangle_{\mathcal{H}}$ is a random variable, and it has a well-defined expectation.]

(*c*) Let $\{\boldsymbol{Z}_\ell\}_{\ell \geq 1}$ be a collection of independent random vectors, with $\mathbb{E}\{\boldsymbol{Z}_\ell\} = 0$, and $\mathbb{E}\{\|\boldsymbol{Z}_\ell\|_{\mathcal{H}}^2\} < \infty$. Define $\boldsymbol{S}_\ell \equiv \boldsymbol{Z}_1 + \boldsymbol{Z}_2 + \cdots + \boldsymbol{Z}_\ell$. Prove that, for all $n \geq 1$, $t > 0$

$$\mathbb{P}\Big\{ \max_{\ell \in \{1,\ldots,n\}} \big\|\boldsymbol{S}_\ell\big\|_{\mathcal{H}} > t \Big\} \leq \frac{1}{t^2} \sum_{\ell=1}^{n} \mathbb{E}\big\{ \|\boldsymbol{Z}_\ell\|_{\mathcal{H}}^2 \big\} \,. \tag{10}$$

(*d*) With the notations in the previous point, prove that, if $\sum_{\ell=1}^{\infty} \mathbb{E}\{\|\boldsymbol{Z}_\ell\|_{\mathcal{H}}^2\} < \infty$, then the sequence $\{\boldsymbol{S}_\ell\}_{\ell \in \mathbb{N}}$ converges almost surely (with respect to the topology defined by $\|\cdot\|_{\mathcal{H}}$).