**Probability 1**
**CEU Budapest, fall semester 2018**
Imre Péter Tóth
**Homework sheet 4 – solutions**

4.1 Show that if $X_n \Rightarrow X$ and $f : \mathbb{R} \to \mathbb{R}$ is continuous, then $f(X_n) \Rightarrow f(X)$.

4.2 Let $F : \mathbb{R} \to [0,1]$ be a probability distribution function, and let $Y$ be a random variable which is uniformly distributed in $[0,1]$. Let $X = \sup\{x|F(x) < Y\}$. Show that the distribution function of $X$ is exactly $F$.

4.3 **(homework)** For a distribution function $F : \mathbb{R} \to [0,1]$, define its generalized inverse $F^{-1} : [0,1] \to \bar{\mathbb{R}}$ as $F^{-1}(y) := \sup\{x \in \mathbb{R}|F(x) < y\}$. Let $F, F_1, F_2, \dots$ be distribution functions such that $F_n \Rightarrow F$. Let $\Omega = [0,1]$, let $\mathbb{P}$ be Lebesgue measure on $\Omega$, and define de random variables $X(\omega) := F^{-1}(\omega)$, $X_n(\omega) := F_n^{-1}(\omega)$ for $\omega \in \Omega$. Show that $X_n \to X$ almost surely.

**Solution:** see Durrett [1], Theorem 3.2.2.

4.4 **(homework)** Durrett [1], Exercise 3.2.6

**Solution:** We first show that $\rho$ is a metric:

a.) $\rho(F,G) \ge 0$ because $F$ is increasing, so $F(x-\varepsilon) - \varepsilon \le F(x+\varepsilon) + \varepsilon$ can not hold for $\varepsilon < 0$.

b.) $F(x-\varepsilon) - \varepsilon \le G(x)$ for every $x$ is the same as $F(y) \le G(y+\varepsilon) + \varepsilon$ for every $y$ (by the substitution $x = y + \varepsilon$). Similarly $G(x) \le F(x+\varepsilon) + \varepsilon$ for every $x$ is the same as $G(y-\varepsilon) - \varepsilon \le F(y)$ for every $y$. So

$$[\forall x(F(x-\varepsilon) - \varepsilon \le G(x) \le F(x+\varepsilon) + \varepsilon)] \Leftrightarrow [\forall y(G(y-\varepsilon) - \varepsilon \le F(y) \le G(y+\varepsilon) + \varepsilon)],$$

which means that $\rho(F,G) = \rho(G,F)$.

c.) If $\rho(F,G) = 0$, then $G(x) \le F(x+\varepsilon) + \varepsilon$ for every $\varepsilon > 0$, so $G(x) \le \lim_{y \searrow x} F(y) = F(x)$, since $F$ is a distribution function, thus continuous from the right. By symmetry $F(x) \le G(x)$ as well, so $F = G$.

d.) If $\varepsilon_1 > \rho(F,G)$ and $\varepsilon_2 > \rho(G,H)$, then $F(x-\varepsilon_1) - \varepsilon_1 \le G(x)$ and $G(y-\varepsilon_2) - \varepsilon_2 \le H(y)$ for every $x$ and $y$, including $x = y - \varepsilon_2$, so

$$F(y - \varepsilon_2 - \varepsilon_1) - \varepsilon_2 - \varepsilon_1 \le G(y - \varepsilon_2) - \varepsilon_2 \le H(y) \quad \text{for every } y.$$

Similarly

$$H(y) \le F(y + \varepsilon_2 + \varepsilon_1) + \varepsilon_2 + \varepsilon_1 \quad \text{for every } y.$$

Since these hold for every $\varepsilon_1 > \rho(F,G)$ and $\varepsilon_2 > \rho(G,H)$, we have that $\rho(F,H) \le \rho(F,G) + \rho(G,H)$.

We have shown that $\rho$ is a metric. Now we show that if $\rho(F,F_n) \to 0$, then $F_n \Rightarrow F$. Indeed, $\rho(F,F_n) \to 0$ implies that for every $x$

$$\lim_{y \nearrow x} F(y) \le \varliminf_{n \to \infty} F_n(x) \le \varlimsup_{n \to \infty} F_n(x) \le \lim_{y \searrow x} F(y).$$

If $F$ is continuous at $x$, then $\lim_{y \nearrow x} F(y) = \lim_{y \searrow x} F(y) = F(x)$, so this means that $\lim_{n \to \infty} F_n(x) = F(x)$, meaning exactly that $F_n \Rightarrow F$.

Eventually, we show that if $F_n \Rightarrow F$, then $\rho(F,F_n) \to 0$. This is the key part of the statement, and this shows that the definition of $\rho$ is celever. The difficulty is that although

$F_n(x) \to F(x)$ for all but countably many $x$, this convergence is not at all uniform, since $F$ may not be continuous. Indeed, if $F_n$ is the indicator function of $\left[\frac{1}{n}, \infty\right)$ and $F$ is the indicator function of $[0, \infty)$, then $F_n \Rightarrow F$, but $|F_n - F| = 1$ on a short interval for every $n$. So, since there is no uniformity, we make use of monotonicity of the distribution functions.

Fix $\varepsilon > 0$. Since $F$ is a distribution function, $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$, so there is an $M < \infty$ such that $F(x) < \frac{\varepsilon}{2}$ for all $x \le -M$ and $F(x) > 1 - \frac{\varepsilon}{2}$ for all $x \ge M$. Furthermore, $F$ is continuous except for at most countably many points, so we can cut up the interval $[-M, M]$ into finitely many subintervals of length at most $\varepsilon$, using only continuity points as endpoints: let $x_0 < -M < x_1 < x_2 < x_3 \cdots < x_{N-2} < x_{N-1} < M < x_N$ such that $F$ is continuous at $x_0, \dots, x_N$ and $x_{k+1} - x_k \le \varepsilon$ for all $k \in \{0, 1, \dots, N-1\}$. By assumption, $F_n(x_k) \to F(x_k)$ as $n \to \infty$ for all $k \in \{0, 1, \dots, N-1\}$, so if $n$ is big enough, then $|F_n(x_k) - F(x_k)| < \frac{\varepsilon}{2}$ simultaneously for all $k$ (there are only finitely many $k$s). Then, by monotonicity of $F_n$ and $F$, if we take $x \in [x_k, x_{k+1}]$ for some $k$, then

$$F(x-\varepsilon) - \varepsilon \le F(x_k) - \varepsilon \le F_n(x_k) \le F_n(x) \le F_n(x_{k+1}) \le F(x_{k+1}) + \varepsilon \le F(x+\varepsilon) + \varepsilon.$$

On the other hand, if $x \le x_0$, then

$$F(x-\varepsilon) - \varepsilon \le F(x_0) - \varepsilon \le \varepsilon - \varepsilon = 0 \le F_n(x) \le F_n(x_0) \le F(x_0) + \frac{\varepsilon}{2} \le \varepsilon \le F(x+\varepsilon) + \varepsilon.$$

Finally, if $x \ge x_N$, then

$$F(x-\varepsilon) - \varepsilon \le 1 - \varepsilon \le F(x_N) - \frac{\varepsilon}{2} \le F_n(x_N) \le F_n(x) \le 1 = 1 - \varepsilon + \varepsilon \le F(x_N) + \varepsilon \le F(x+\varepsilon) + \varepsilon.$$

We have shown for all $x \in \mathbb{R}$ that

$$F(x-\varepsilon) - \varepsilon \le F_n(x) \le F(x+\varepsilon) + \varepsilon,$$

so $\rho(F, F_n) \le \varepsilon$ if $n$ is big enough.

4.5 Durrett [1], Exercise 3.2.9

4.6 Durrett [1], Exercise 3.2.12

4.7 Durrett [1], Exercise 3.2.14

4.8 Durrett [1], Exercise 3.2.15

4.9 **(homework)** *Weak convergence and densities.*

(a) Prove the following

**Theorem 1** *Let $\mu_1, \mu_2, \dots$ and $\mu$ be a sequence of probability distributions on $\mathbb{R}$ which are absolutely continouos w.r.t. Lebesgue measure. Denote their densities by $f_1, f_2, \dots$ and $f$, respectively. Suppose that $f_n(x) \xrightarrow{n\to\infty} f(x)$ for every $x \in \mathbb{R}$. Then $\mu_n \Rightarrow \mu$ (weakly).*

(Hint: denote the cumulative distribution functions by $F_1, F_2, \dots$ and $F$, respectively. Use the Fatou lemma to show that $F(x) \le \liminf_{n\to\infty} F_n(x)$. For the other direction, consider $G(x) := 1 - F(x)$.

(b) Show examples of the following facts:

i. It can happen that the $f_n$ converge pointwise to some $f$, but the sequence $\mu_n$ is not weakly convergent, because $f$ is not a density.

ii. It can happen that the $\mu_n$ are absolutely continuous, $\mu_n \Rightarrow \mu$, but $\mu$ is not absolutely continuous.

iii. It can happen that the $\mu_n$ and also $\mu$ are absolutely continuous, $\mu_n \Rightarrow \mu$, but $f_n(x)$ does not converge to $f(x)$ for any $x$.

**Solution:**

(a) $F_n(x) = \int_{-\infty}^x f_n(x)\,\mathrm{d}x$ and $f_n(x) \to f(x)$ for every $x$, so the Fatou lemma says that

$$F(x) = \int_{-\infty}^x f(x)\,\mathrm{d}x = \int_{-\infty}^x \liminf_{n\to\infty} f_n(x)\,\mathrm{d}x \le \liminf_{n\to\infty} \int_{-\infty}^x f_n(x)\,\mathrm{d}x = \liminf_{n\to\infty} F_n(x).$$

Similarly,

$$\begin{aligned} 1 - F(x) &= \int_x^\infty f(x)\,\mathrm{d}x = \int_x^\infty \liminf_{n\to\infty} f_n(x)\,\mathrm{d}x \\ &\le \liminf_{n\to\infty} \int_x^\infty f_n(x)\,\mathrm{d}x = \liminf_{n\to\infty}(1 - F_n(x)) = 1 - \limsup_{n\to\infty} F_n(x), \end{aligned}$$

which implies $\limsup_{n\to\infty} F_n(x) \le F(x)$, so $F_n(x) \to F(x)$ for every $x$, and we are done.

(b) i. Let $\mu_n$ be uniform on $[n, n+1]$, so $f_n$ is the indicator function of $[n, n+1]$. Then $f_n \to 0$ for all $x$.

ii. Let $\mu_n$ be the uniform distribution on $\left[-\frac{1}{n}, \frac{1}{n}\right]$ and let $\mu$ be the probability measure concentrated on $\{0\}$.

iii. Let $f$ be the uniform density on $[0,1]$ and let $f_n = f + h_n$ where the deviation $h_n$ is constructed to be "small" in the sense of weak convergence, but spoils pointwise convegrence totally. In particular, for $m = 1, 2, 3, \ldots$ and $k = 0, 1, \ldots, 2m^2 - 1$ let

$$h_{m,k} = \mathbf{1}_{\left[-m+\frac{k}{m}, -m+\frac{k+1}{m}\right]} - \mathbf{1}_{\left[0, \frac{1}{m}\right]},$$

where $\mathbf{1}$ denotes indicator function. Now let the sequence $h_n$ contain all the $h_{m,k}$ (in any order). Draw these functions and see that they work.

4.10 **(homework)** Let $X_1, X_2, \ldots$ be independent and uniformly distributed on $[0,1]$. Let $M_n = \max\{X_1, \ldots, X_n\}$ and let $Y_n = n(1 - M_n)$. Find the weak limit of $Y_n$. *(Hint: Calculate the distribution functions.)*

**Solution:** Let $F_X$ be the common distribution function of the $X_i$:

$$F_X(x) = \mathbb{P}(X_i \le x) = \begin{cases} 0 & \text{if } x \le 0 \\ x & \text{if } 0 < x < 1 \\ 1 & \text{if } x \ge 1 \end{cases}.$$

Now $M_n = \max\{X_1, \ldots, X_n\}$, so $M_n \le x$ iff $X_i \le x$ for all $i$. So the distribution function of $M_n$ is

$$\begin{aligned} F_{M_n}(x) &:= \mathbb{P}(M_n \le x) = \mathbb{P}(X_1 \le x, \ldots, X_n \le x) = \mathbb{P}(X_1 \le x) \cdots \mathbb{P}(X_n \le x) = \\ &= (F_X(x))^n = \begin{cases} 0 & \text{if } x \le 0 \\ x^n & \text{if } 0 < x < 1 \\ 1 & \text{if } x \ge 1 \end{cases}. \end{aligned}$$

We have used the independence of the $X_i$. Now the distribution function of $Y_n$ is

$$\begin{aligned} F_n(y) &:= \mathbb{P}(Y_n \le y) = \mathbb{P}(n(1 - M_n) \le y) = \mathbb{P}\left(M_n \ge 1 - \frac{y}{n}\right) = 1 - F_{M_n}\left(1 - \frac{y}{n}\right) = \\ &= \begin{cases} 0 & \text{if } y \le 0 \\ 1 - \left(1 - \frac{y}{n}\right)^n & \text{if } 0 < y < n \\ 1 & \text{if } y \ge n \end{cases}. \end{aligned}$$

Given any $y > 0$, as $n$ grows, we will eventually have $y < n$, so the second case matters, and $F_n(y) \to \lim_{n\to\infty} 1 - \left(1 - \frac{y}{n}\right)^n = 1 - e^{-y}$. All in all, we got that

$$\lim_{n\to\infty} F_n(y) = F(y) := \begin{cases} 0 & \text{if } y \leq 0 \\ 1 - e^{-y} & \text{if } y > 0 \end{cases}$$

for every $y \in \mathbb{R}$, so $F_n \Rightarrow F$. This $F$ is exactly the distribution function of the exponential distribution with parameter 1, so we have shown that $Y_n \Rightarrow Exp(1)$.

4.11 Let $X_1, X_2, \ldots$ be independent and exponentially distributed with parameter $\lambda = 1$. Let $M_n = \max\{X_1, \ldots, X_n\}$ and let $Y_n = M_n - \ln n$. Find the weak limit of $Y_n$. *(Hint: Calculate the distribution functions.)*

4.12 Let $S = \mathbb{Z}$ and let the random variables $X, X_1, X_2, \cdots \in S$.

a.) Show that $X_n \Rightarrow X$ if and only if $\mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$ as $n \to \infty$ for every $k \in S$.

b.) It this also true for some arbitrary countable $S \subset \mathbb{R}$?

# References

[1] Durrett, R. *Probability: Theory and Examples.* **4th** edition, Cambridge University Press (2010)