MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J — Fall 2018

Problem Set 9

Readings:
Notes from Lectures 16-19.
[GS], Section 7.1-7.6
[Cinlar], Chapter III

Exercise 1. We study convergence of algebraic operations:

(a) Show that

$$X_n \stackrel{\text{i.p.}}{\to} X, Y_n \stackrel{\text{i.p.}}{\to} Y \quad \Rightarrow \quad X_n Y_n \stackrel{\text{i.p.}}{\to} XY$$

(*Hint:* reduce to $\stackrel{\text{a.s.}}{\to}$.)

(b) Show, however, that

$$X_n \stackrel{\text{d}}{\to} X, Y_n \stackrel{\text{d}}{\to} Y \quad \not\Rightarrow \quad X_n Y_n \stackrel{\text{d}}{\to} XY$$

(c) Assume $X_n \perp\!\!\!\perp Y_n$ and $X \perp\!\!\!\perp Y$. Show that then

$$X_n \stackrel{\text{d}}{\to} X, Y_n \stackrel{\text{d}}{\to} Y \quad \Rightarrow \quad X_n Y_n \stackrel{\text{d}}{\to} XY$$

(*Hint:* reduce to $\stackrel{\text{a.s.}}{\to}$.)

Solution:

(a) Writing the product as a sum of squares

$$X_n Y_n = \frac{(X_n + Y_n)^2 - (X_n - Y_n)^2}{4}.$$

Therefore, by the triangle inequality,

$$|X_n Y_n| \leq \frac{|X_n + Y_n|^2 + |X_n - Y_n|^2}{4}.$$

Thus

$$\mathbb{P}\left(|X_n Y_n| > \varepsilon\right) \leq \left(\frac{|X_n + Y_n|^2}{4} > \frac{\varepsilon}{2}\right) + \left(\frac{|X_n - Y_n|^2}{4} > \frac{\varepsilon}{2}\right),$$

and it suffices to show that convergence in probability is closed under scalar multiplication, addition and squares. Addition is given in the lecture notes, scalar multiplication follows since, for all $c \geq 0$,

$$\mathbb{P}\left(c|X| > \varepsilon\right)\mathbb{P}\left(|X| > \varepsilon/c\right),$$

and squaring follows since

$$\mathbb{P}\ \ |X|^2 > \varepsilon\ \ = \mathbb{P}\ \ |X| > \sqrt{\varepsilon}\ \ ,$$

i.e. by choose ε appropriately.

(b) Consider the probability space $([0,1], \mathcal{B}, \lambda)$. Let

$$X_n = \begin{cases} \mathbb{1}_{[0,1/2]} & n \text{ odd} \\ \mathbb{1}_{(1/2,1]} & n \text{ even} \end{cases} \qquad Y_n = \begin{cases} \mathbb{1}_{[0,1/2]} & n \text{ even} \\ \mathbb{1}_{(1/2,1]} & n \text{ odd} \end{cases}.$$

Then for all n, X_n and Y_n are Bernoulli $1/2$ random variables and thusly converge in distribution to a Bernoulli $1/2$ random variable. However, $X_n Y_n \equiv 0$ and the product of Bernoulli random variables is not identically zero.

(c) As $X_n \to X$ and $Y_n \to Y$

$$\begin{aligned} \lim_{n\to\infty} F_{X_n,Y_n}(x,y) &= \lim_{n\to\infty} F_{X_n}(x) F_{Y_n}(y) \quad \text{(Independence)} \\ &= \left(\lim_{n\to\infty} F_{X_n}(x)\right)\left(\lim_{n\to\infty} F_{Y_n}(y)\right) \quad \text{(Independence)} \\ &=^{(a)} F_X(x) F_Y(y) \\ &= F_{X,Y}(x,y), \end{aligned}$$

where (a) holds at all continuity points of F_X and F_Y or equivalent of $F_{X,Y}$. Therefore, $(X_n, Y_n) \to (X, Y)$ in distribution. Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded continuous function and q_n, q and p_n, p be the quantile functions for the $\{X_n\}$, X, the $\{Y_n\}$ and Y respectively, i.e. $q_n \to q$ almost everywhere $p_n \to p$ almost everywhere and

$$X_n \sim q_n(U) \quad X \sim q(U) \qquad Y_n \sim p_n(V) \quad Y \sim p(V)$$

for independent random variables U and V. Therefore, $q_n(U)p_n(V) \to$

$q(U)p(V)$ almost everywhere and by the BCT

$$\begin{aligned}\lim_{n\to\infty} E\left[f(X_nY_n)\right] &= \lim_{n\to\infty} E\left[f(q_n(U)p_n(V))\right] \\ &= E\left[\lim_{n\to\infty} f(q_n(U)p_n(V))\right] \quad \text{(BCT)} \\ &= E\left[f(q(U)p(V))\right] \quad (f \text{ continuous }) \\ &= E\left[f(XY)\right].\end{aligned}$$

Hence $X_nY_n \to XY$ in distribution.

Exercise 2 (Metrization of convergence in probability)**.** Define a pseudo-metric on the space of random-variables:

$$d(X,Y) \triangleq \mathbb{E} \quad \frac{|X-Y|}{1+|X-Y|} \quad .$$

Show $X_n \overset{\text{i.p.}}{\to} X$ iff $d(X_n, X) \to 0$.

Solution: Let $\varepsilon > 0$ and WLOG assume $\varepsilon < 1$. Then,

$$\{|X-Y| \geq \varepsilon\} = \left\{\frac{|X-Y|}{1+|X-Y|} \geq \frac{\varepsilon}{1+\varepsilon}\right\},$$

and

$$\left\{\frac{|X-Y|}{1+|X-Y|} \geq \varepsilon\right\} = \left\{|X-Y| \geq \frac{\varepsilon}{1-\varepsilon}\right\}$$

By Markov's inequality

$$\begin{aligned}\mathbb{P}\left(|X-Y| \geq \varepsilon\right) &= \mathbb{P}\left(\frac{|X-Y|}{1+|X-Y|} \geq \frac{\varepsilon}{1+\varepsilon}\right) \\ &\leq \frac{1+\varepsilon}{\varepsilon} E\left[\frac{|X-Y|}{1+|X-Y|}\right] \\ &= \frac{1+\varepsilon}{\varepsilon} d(X,Y).\end{aligned}$$

Therefore, convergence in the metric implies convergence in probability. For the other direction let

$$Z = \frac{|X-Y|}{1+|X-Y|}$$

and

$$Z_\varepsilon = \varepsilon \mathbb{1}\{|Z| < \varepsilon\} + \mathbb{1}\{|Z| \geq \varepsilon\}.$$

Since $Z \le 1$, then $Z \le Z_\varepsilon$ and

$$\begin{aligned} d(X,Y) &= E\,[Z] \\ &\le E\,[Z_\varepsilon] \\ &= E\,[\varepsilon \quad \{|Z| < \varepsilon\} + \quad \{|Z| \ge \varepsilon\}] \\ &\le \varepsilon + \mathbb{P}\,(|Z| \ge \varepsilon) \\ &= \varepsilon + \mathbb{P}\left(|X-Y| \ge \frac{\varepsilon}{1-\varepsilon}\right). \end{aligned}$$

Hence convergence in probability implies convergence in the metric.

Exercise 3 (20 pts)**.** Prove Cauchy criterions for convergence a.s. and i.P.:

(i) Show that X_n converges almost surely iff

$$\forall \epsilon > 0 \quad \mathbb{P}[\sup_{k\ge 0} |X_{n+k} - X_n| > \epsilon] \to 0 \quad n \to \infty$$

(ii) Show that X_n converges in probability iff

$$\forall \epsilon > 0 \quad \sup_{k\ge 0} \mathbb{P}[|X_{n+k} - X_n| > \epsilon] \to 0 \quad n \to \infty$$

Solution:

(i) Equivalently, for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ so that for all $n \ge N$

$$\mathbb{P}\left(\sup_{k\ge 0} |X_{n+k} - X_n| > \varepsilon\right) < \varepsilon.$$

Let $\varepsilon > 0$. Suppose X_n converges almost surely to some random variable X. Then, there exists a measurable set E with $\mathbb{P}(E) < \varepsilon$ so that $X_n \Rightarrow X$ uniformly on E^c. Therefore, there exists $N \in \mathbb{N}$ so that for all $n \ge N$ and for all $\omega \in E^c$

$$|X_n(\omega) - X(\omega)| \le \varepsilon.$$

Hence,

$$\mathbb{P}\left(\sup_{k\ge 0} |X_{n+k} - X_n| > \varepsilon\right) = \mathbb{P}(E) < \varepsilon.$$

Conversely suppose the Cauchy criterion is satisfied. For all k, $\sup_{k\ge 0}|X_{n+k} - X_n| \ge |X_{n+k} - X_n|$. Therefore, for all k,

$$\mathbb{P}\,(|X_{n+k} - X_n| > \varepsilon) \le \mathbb{P}\left(\sup |X_{n+k} - X_n| > \varepsilon\right).$$

Hence

$$\sup_{k\geq 0}\mathbb{P}\left(|X_{n+k}-X_n|>\varepsilon\right)\leq \mathbb{P}\left(\sup_{k\geq 0}|X_{n+k}-X_n|>\varepsilon\right),$$

and thusly this condition is stronger than the condition in part (ii). Therefore, by part (ii) $\{X_n\}$ converges in probability to a random variable X. The set of points where X_n does not converge to X is

$$\{X_n\not\to X\}=\bigcup_{\varepsilon>0}\bigcap_{n=1}^{\infty}\{\sup_{k\geq n}|X_k-X|>\varepsilon\},$$

and by continuity of probability

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty}\left\{\sup_{k\geq n}|X_k-X|>\varepsilon\right\}\right)=\lim_{n\to\infty}\mathbb{P}\left(\sup_{k\geq n}|X_k-X|>\varepsilon\right).$$

Thus it suffices to show that for all $\varepsilon>0$ there exists an $N\in\mathbb{N}$ so for all $n\geq N$

$$\mathbb{P}\left(\sup_{k\geq n}|X_k-X|>\varepsilon\right)<\varepsilon.$$

By the triangle inequality

$$|X_k-X|\leq|X_k-X_n|+|X_n-X|,$$

and thus

$$\begin{aligned}\sup_{k\geq n}|X_k-X| &\leq \sup_{k\geq n}|X_k-X_n|+|X_n-X|\\ &\leq \sup_{k\geq 0}|X_{n+k}-X_n|+|X_n-X|.\end{aligned}$$

Choose N_1 to satisfy the Cauchy criterion for $\varepsilon/2$ and N_2 for the convergence in measure for $\varepsilon/2$. Let $N=\max\{N_1,N_2\}$, then for all $n\geq N$

$$\mathbb{P}\left(\sup_{k\geq n}|X_k-X|>\varepsilon\right)\leq \mathbb{P}\left(\sup_{k\geq 0}|X_{n+k}-X_n|>\frac{\varepsilon}{2}\right)+\mathbb{P}\left(|X_n-X|>\frac{\varepsilon}{2}\right)<\varepsilon$$

as desired.

(ii) Suppose that the sequence satisfies the Cauchy criterion. Then, there exists a subsequence $\{X_{n_k}\}$ such that, if $E_k = \{|X_{n_k} - X_{n_{k+1}}| \geq 2^{-k}\}$, then $\mathbb{P}(E_k) \leq 2^{-k}$. Let $F_k = \cup_{j=k}^{\infty} E_j$. For $\omega \notin F_k$, and $i \geq j \geq k$, we have

$$|X_{n_j}(\omega) - X_{n_i}(\omega)| \leq \sum_{l=j}^{i-1} |X_{n_{l+1}}(\omega) - X_{n_l}(\omega)| \leq 2^{1-j}$$

by the definition of the subsequence. This means that $\{X_{n_k}\}$ is pointwise Cauchy on F_k^c. Let $F = \cap_{k=1}^{\infty} F_k = \limsup E_k$, and note that $\mathbb{P}(F) = 0$. Let us define a random variable X such that on $X(\omega) = 0$ for all $\omega \in F$, and $X(\omega) = \lim X_{n_k}(\omega)$ for all $\omega \notin F$. Then $X_{n_k} \to X$ a.s. and thus $X_{n_k} \to X$ i.p. Finally, we have

$$\{|X_n - X| \geq \epsilon\} \subset \{|X_{n_k} - X_n| \geq \epsilon/2\} \cup \{|X_{n_k} - X| \geq \epsilon/2\},$$

and the right-hand side can be made arbitrarily small with large k and n, and thus we have convergence in probability of X_n to X.

Conversely, for $a > 0$ and $m, n \in \mathbb{N}$, let $E_n(a) = \{\omega : |X_n(\omega) - X(\omega)| \geq a\}$ and $F_{m,n}(a) = \{\omega : |X_m(\omega) - X_n(\omega)| \geq a\}$. Then, for every $a > 0$ and every m, n, we have

$$F_{m,n}(a) \subset E_m(a/2) \cup E_n(a/2).$$

In fact, if ω is neither in $E_m(a/2)$ nor in $E_n(a/2)$, then $|X_m(\omega) - X(\omega)| < a/2$ and $|X_n(\omega) - X(\omega)| < a/2$, so that $|X_m(\omega) - X_n(\omega)| \leq |X_m(\omega) - X(\omega)| + |X_n(\omega) - X(\omega)| < a/2 + a/2 = a$ so $|X_m(\omega) - X(\omega)| < a$, which shows $\omega \notin F_{m,n}(a)$. In virtue of the monotonicity and the subadditivity of the measure, we also have

$$0 \leq \mathbb{P}(F_{m,n}(a)) \leq \mathbb{P}(E_m(a/2) \cup E_n(a/2)) \leq \mathbb{P}(E_m(a/2)) + \mathbb{P}(E_n(a/2)). \quad (1)$$

Fix $a > 0$. Since $X_n \to X$ in probability, for every $\epsilon > 0$ there exists k such that

$$n > k \Rightarrow \mathbb{P}(E_n(a/2)) < \epsilon/2. \quad (2)$$

Then, for $m > k$ and $n > k$, it follows from (1) and (2) that

$$\mathbb{P}(F_{m,n}(a)) < \epsilon/2 + \epsilon/2 = \epsilon$$

which means $lim_{m,n \to \infty} \mathbb{P}(F_{m,n}(a)) = 0$. Since this holds for every $a > 0$, it follows that the Cauchy criterion is satisfied.

Exercise 4. Let $\{X_n\}$ be a sequence of random variables defined on the same probability space.

(a) Show $\mathbb{E}[|X_n - X|] \to 0$ implies $X_n \overset{\text{i.p.}}{\to} X$.

(b) Suppose that $X_n \overset{\text{i.p.}}{\to} 0$ and that for some constant c, we have $|X_n| \leq c$, for all n, with probability 1. Show that

$$\lim_{n\to\infty} \mathbb{E}[|X_n|] = 0.$$

(c) Suppose that each X_n can only take the values 0 and 1 and, that $\mathbb{P}(X_n = 1) = 1/n$.

 (i) Give an example in which we **have** almost sure convergence of X_n to 0.

 (ii) Give an example in which we **do not have** almost sure convergence of X_n to 0.

Solution:

(a) Follows from Markov's inequality

$$\mathbb{P}(|X_n - 0| \geq \epsilon) \leq \frac{E[|X_n|]}{\epsilon},$$

Therefore, if $E[|X_n|]$ approaches 0, then X_n approaches 0 in probability.

(b) Fix $\epsilon > 0$ and define a new random variable X_n^ϵ as follows. We have $X_n^\epsilon = \epsilon$ whenever $|X_n| \leq \epsilon$, and $X_n^\epsilon = c$ whenever $|X_n| > \epsilon$. Then, it is always true that $|X_n| \leq X_n^\epsilon$ and therefore

$$E[|X_n|] \leq E[X_n^\epsilon] = \epsilon P(|X_n| \leq \epsilon) + cP(|X_n| > \epsilon)$$

Taking limits as n goes to infinity, we get

$$\lim_n E[|X_n|] \leq \epsilon,$$

and since this holds for all $\epsilon > 0$, we get $\lim_n E[|X_n|] = 0$.

(c) (i) Consider the Lebesgue probability space $([0,1], \mathcal{B}, \lambda)$. Let $X_n = \mathbb{1}_{[0,\frac{1}{n}]}$. For all $\omega \in (0,1]$ there exists $n \in N$ such that $1/n < \omega$. Thus $X_n \overset{\text{a.s.}}{\to} 0$ and for all n $\mathbb{P}(X_n = 1) = E[X_n] = 1/n$.

(ii) Let $\{X_n\}$ be independent. Then $\{X_n = 1\}$ occurs infinitely often by the Borel-Cantelli Lemma, as

$$\sum_{i=1}^{\infty} P(X_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

and the events $\{X_n = 1\}$ are independent. Hence X_n cannot converge to 0.

Exercise 5. Let $X_1, X_2, \ldots$ be i.i.d. exponential random variables with parameter $\lambda = 1$. Let $S_n = X_1 + \cdots + X_n$. Let $a > 1$. What is the Chernoff upper bound for $\mathbb{P}(S_n \geq na)$?

Solution: To use Chernoff's bound as stated in the lecture notes, we need to work with random variables that have mean 0. Let $Y_i = X_i - 1$, and $S'_n = Y_1 + \cdots + Y_n$. Then

$$P(S_n \geq na) = P(S'_n \geq n(a-1)).$$

Now the moment generating function of the X_i is $1/(1-s)$, so Y_i has moment generating function $e^{-s}/(1-s)$. We must optimize

$$\sup_{s \geq 0} s(a-1) - \log \frac{e^{-s}}{1-s},$$

The optimum occurs at $s = (a-1)/a$, and equals $a - 1 - \log(a)$. So,

$$P(S_n \geq na) \leq e^{-n(a-1-\log(a))} = a^n e^{-n(a-1)}.$$

Exercise 6. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables, uniformly distributed on the interval $[0, 1]$. For n odd, let M_n be the median of $X_1, X_2, \ldots, X_n$, i.e. the $(\frac{n+1}{2})$ order statistic $X^{(\frac{n+1}{2})}$. Show that M_n converges to 1/2, in probability.

Solution: Fix some $\epsilon > 0$. We will show that $\mathbb{P}(M_n > 1/2 + \epsilon)$ converges to zero. By symmetry, this will also imply that $\mathbb{P}(M_n < 1/2 - \epsilon)$ also converges to zero, and will establish the desired convergence.

Let N_n be the number of X_is ($i = 1, \ldots, n$) for which $X_i > 1/2 + \epsilon$. If $M_n > 1/2 + \epsilon$, then $N_n/n > 1/2$. But $\mathbb{E}[N_n/n] = 1/2 - \epsilon$, so that $\mathbb{P}(N_n/n > 1/2) \to 0$, by the weak law of large numbers.

Exercise 7. [Optional, not to be graded] Show that for every $\mathbb{P}_X$ on $(\mathbb{R}, \mathcal{B})$ there exist a sequence $\mathbb{P}_{X_n} \xrightarrow{\text{d}} \mathbb{P}_X$ such that every $\mathbb{P}_{X_n}$ has a continuous, bounded, infinitely-differentiable PDF. Steps:

(i) Show $X_\epsilon = X + \epsilon Z \xrightarrow{\text{d}} X$ as $\epsilon \to 0$.

(ii) Let $X \perp\!\!\!\perp Z$ and $Z \sim \mathcal{N}(0,1)$. Show that CDF of X_ϵ is continuous (*Hint:* BCT) and differentiable (*Hint: Fubini*) with derivative

$$f_{X_\epsilon}(a) = \mathbb{E}\left[f_Z\left(\frac{a - X}{\epsilon}\right)\frac{1}{\epsilon}\right].$$

(iii) Show that $a \mapsto f_{X_\epsilon}(a)$ is continuous.

(iv) [Optional] Conclude the proof (*Hint:* derivatives of f_Z are uniformly bounded on $\mathbb{R}$).

Solution:

(i) Let Z be a random variable defined on the same probability space as X. Let $\delta > 0$ and WLOG assume $\delta < 1$, as $\{\varepsilon|Z| \geq x\} \subset \{\varepsilon|Z| \geq y\}$ for $x \geq y$, Therefore,

$$\begin{aligned}
\mathbb{P}\left(\varepsilon|Z| \geq \delta\right) &= \mathbb{P}\left(Z \leq -\delta/\varepsilon\right) + \mathbb{P}\left(Z \geq \delta/\varepsilon\right) \\
&\leq^{(a)} \mathbb{P}\left(Z \leq -\delta/\varepsilon\right) + \mathbb{P}\left(Z < \delta^2/\varepsilon\right) \\
&= F_Z\left(-\frac{\delta}{\varepsilon}\right) + 1 - F_Z\left(\frac{\delta^2}{\varepsilon}\right) \\
&\to 0 + 1 - 1 = 0,
\end{aligned}$$

where (a) follows since $\delta^2 < \delta$ for $\delta < 1$. Therefore, $\varepsilon Z \to 0$ in probability. Thus, as convergence in probability is closed under addition, $X + \varepsilon Z \to X$ in probability and thusly $X + \varepsilon Z \to X$ in distribution.

(ii) For any measurable function g, as $X \perp\!\!\!\perp Z$,

$$\int g(\gamma)\, \mathbb{P}_{X_\varepsilon}(d\gamma) = \int\int g(\alpha + \beta)\, \mathbb{P}_X(d\alpha)\, \mathbb{P}_{\varepsilon Z}(d\beta),$$

where

$$\mathbb{P}_{\varepsilon Z}(d\beta) = \frac{d}{d\beta} F_Z\left(\frac{\beta}{\varepsilon}\right) = \frac{1}{\varepsilon} f_Z\left(\frac{\beta}{\varepsilon}\right)\lambda(d\beta)$$

and this integration makes sense and can be interchanged by Fubini's Theorem. Letting $g = \mathbb{1}_{(-\infty,z]}$

$$
\begin{aligned}
F_{X_\varepsilon}(z) &= \int \quad _{(-\infty,z]}\, d\mathbb{P}_{X_\varepsilon} \\
&= \int\int \mathbb{1}_{(-\infty,z]}(\alpha+\beta) f_Z\left(\frac{\beta}{\varepsilon}\right)\frac{1}{\varepsilon}\,\lambda(d\beta)\,\mathbb{P}_X(d\alpha) \\
&= \int\int \quad _{(-\infty,z]}(\gamma) f_Z\left(\frac{\gamma-\alpha}{\varepsilon}\right)\frac{1}{\varepsilon}\,\lambda(d(\gamma-\alpha))\,\mathbb{P}_X(d\alpha) \\
&= \int\int \quad _{(-\infty,z]}(\gamma) f_Z\left(\frac{\gamma-\alpha}{\varepsilon}\right)\frac{1}{\varepsilon}\,\lambda(d\gamma)\,\mathbb{P}_X(d\alpha) \quad \text{(Shift invariance)} \\
&= \int_{(-\infty,z]}\int f_Z\left(\frac{\gamma-\alpha}{\varepsilon}\right)\frac{1}{\varepsilon}\,\mathbb{P}_X(d\alpha)\,\lambda(d\gamma) \\
&= \int_{(-\infty,z]} E\left[f_Z\left(\frac{\gamma-X}{\varepsilon}\right)\frac{1}{\varepsilon}\right]\,\lambda(d\gamma).
\end{aligned}
$$

Let

$$f_{X_\varepsilon}(a) := E\left[f_Z\left(\frac{a-X}{\varepsilon}\right)\frac{1}{\varepsilon}\right].$$

Thus

$$F_{X_\varepsilon}(z) = \int_{-\infty}^{z} f_{X_\varepsilon}(\gamma)\,\lambda(d\gamma).$$

From part (iii) f_{X_ε} is continuous and therefore this integral agrees with the Riemann integral. Hence, by the fundamental theorem of calculus, $F_{X_\varepsilon}(z)$ is differential with derivative

$$\frac{d}{dz}F_{X_\varepsilon}(z) = f_{X_\varepsilon}(z).$$

(iii) Limits and integration can be interchanged using the bounded convergence theorem.

(iv) Same as part (iii).