MATH524 – Spring 2025

Problem Set: Week 1

1. (**Integrability**) Let f and g be integrable functions.

   (a) Show that $\max\{f,g\} = (f+g+|f-g|)/2$ and $\min\{f,g\} = (f+g-|f-g|)/2$. **Solution**: recall that $|z| = z$ whenever $z \geq 0$ and $-z$ otherwise. Thus, when $f \geq g$ we have that $\max\{f,g\} = f$ and $|f-g| = f-g$, which gives $(f+g+|f-g|)/2 = (f+g+f-g)/2 = f$, so the two functions coincide. Similarly, when $f < g$ we have $\max\{f,g\} = g$ and $(f+g+|f-g|)/2 = (f+g-f+g)/2 = g$ which shows that the two functions are equal. The case for $\min\{f,g\}$ is analogous.

   (b) Show that $\max\{f,g\}$ and $\min\{f,g\}$ are measurable and integrable. **Solution**: for the first case, we want to show that $\int |\max\{f,g\}|d\mu < \infty$. From (a) we have that $\int |\max\{f,g\}|d\mu = \int |(f+g+|f-g|)/2|d\mu$. Now, applying the triangle inequality twice, we have that $|f+g+|f-g|| \leq |f+g|+|f-g| \leq 2|f|+2|g|$ and thus by properties (4) and (3) in the slides we get $\int |(f+g+|f-g|)/2|d\mu \leq \int (|f|+|g|)\,d\mu = \int |f|d\mu + \int |g|d\mu < \infty$ by integrability of f and g. The case for the minimum is analogous.

2. (**Mixed Random Variable**) Let $X \sim \text{Uniform}(0,1)$.

   (a) Show that $g_c(x) = x\mathbf{1}(x > c)$ for $c \in (0,1)$ is measurable. **Solution**: Note that $g_c(x) = f_1(x) \cdot f_2(x)$ where $f_1(x) = x$ and $f_2(x) = \mathbf{1}(x > c)$. Now, $f_2(x)$ is an indicator function on the Borel set $(c, +\infty)$, so it is measurable, and $f_2(x)$ is a continuous function so it is also measurable. Finally, the product of measurable functions is measurable, from which we conclude that $g_c(x)$ is measurable.

   (b) Derive the d.f. of $Y(c) = g_c(X)$, where $c \in (0,1)$. **Solution**: clearly $0 \leq g_c(X) \leq 1$ so $\mathbb{P}(Y \leq y) = 0$ for $y < 0$ and $\mathbb{P}(Y \leq y) = 1$ for $y > 1$. Then, for $y \in [0,1]$,

$$\begin{aligned}
\mathbb{P}(Y \leq y) &= \mathbb{P}(X\mathbf{1}(X > c) \leq y) = \mathbb{P}(X\mathbf{1}(X > c) \leq y, X > c) + \mathbb{P}(X\mathbf{1}(X > c) \leq y, X \leq c) \\
&= \mathbb{P}(X \leq y, X > c) + \mathbb{P}(X \leq c) = (F_X(y) - F_X(c))\mathbf{1}(y \geq c) + F_X(c) \\
&= (y-c)\mathbf{1}(y \geq c) + c
\end{aligned}$$

   This yields:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \begin{cases} 0 & \text{if } y < 0 \\ c & \text{if } 0 \leq y \leq c \\ y & \text{if } c < y \leq 1 \\ 1 & \text{if } y > 1 \end{cases}$$

3. (**Mixed r.v., d.f. and density**) Let X be a r.v. with absolutely continuous d.f. $F_X(x)$ and Lebesgue density $f(x)$. Assume $f(x) > 0$ for all $x \in \mathbb{R}$. Let $Y = \max\{a, X+b\}$ for $a, b \in \mathbb{R}$.

   (a) Derive the d.f. of Y. **Solution**: recall that $\max\{a, X+b\} = a\mathbf{1}(X \leq a-b) + (X+b)\mathbf{1}(X > a-b)$. Then,

$$\begin{aligned}
\mathbb{P}(Y \leq y) &= \mathbb{P}(\max\{a, X+b\} \leq y) \\
&= \mathbb{P}(\max\{a, X+b\} \leq y, X \leq a-b) + \mathbb{P}(\max\{a, X+b\} \leq y, X > a-b) \\
&= \mathbb{P}(a \leq y, X \leq a-b) + \mathbb{P}(X \leq y-b, X > a-b)
\end{aligned}$$

which yields:

$$F_Y(y) = \mathbb{P}(Y \le y) = \begin{cases} 0 & \text{if } y < a \\ F_X(a-b) & \text{if } y = a \\ F_X(y-b) & \text{if } y > a \end{cases}$$

(b) Derive the density of Y w.r.t. the carrying measure $\mu = \text{Leb} + \delta_a$. **Solution**: note that the d.f. is discontinuous at point a (Y is a mixed r.v.). The density will be:

$$f_Y(y) = \frac{\mathrm{d}\mathbb{P}_Y}{\mathrm{d}\mu}(y) = \begin{cases} 0 & \text{if } y < a \\ F_X(a-b) & \text{if } y = a \\ f_X(y-b) & \text{if } y > a \end{cases}$$

4. (**Weak Convergence**) Let $X_1, \dots, X_n$ be i.i.d., $X_{(n)} = \max_{1 \le i \le n} X_i$ and $X_\infty \sim \text{Gumbel}(0,1)$. Show that:

(a) If $X_i \sim \text{Exp}(1)$ then $X_{(n)} - \log n \to_d X_\infty$.
Solution:

$$\begin{aligned} \mathbb{P}(X_{(n)} - \log(n) \le x) &= \mathbb{P}(X_{(n)} \le x + \log(n)) \\ &= F(x + \log(n))^n \\ &= \left(1 - e^{-x-\log(n)}\right)^n \\ &= \left(1 - \frac{e^{-x}}{n}\right)^n \\ &\to e^{-e^{-x}}. \end{aligned}$$

(b) If $X_n \sim \text{Logistic}(0,1)$ then $X_{(n)} - \log(n) \to_d X_\infty$.
Solution:

$$\begin{aligned} \mathbb{P}(X_{(n)} - \log(n) \le x) &= F(x + \log(n))^n \\ &= \left(\frac{1}{1 + e^{-x-\log(n)}}\right)^n \\ &= \left[\left(1 + \frac{e^{-x}}{n}\right)^n\right]^{-1} \\ &\to \left[e^{e^{-x}}\right]^{-1} = e^{-e^{-x}}. \end{aligned}$$

5. (**Modes of Convergence**) Let $\{X_n : n \ge 1\}$ be a sequence of random variables.

(a) ($\to_d$ implies $O_p(1)$) Let $X_n \to_d X_\infty$. Show that X_n is bounded in probability. **Solution**: let $F_n(x)$ and $F(x)$ denote the cdf of X_n and X_∞ at x, respectively. We want to show that for any $\varepsilon > 0 \; \exists \, M_\varepsilon : \mathbb{P}(|X_n| > M_\varepsilon) \le \varepsilon$. Fix an arbitrary $\varepsilon > 0$. Since $F(\infty) = 1$ and $F(-\infty) = 0$, we can find a large enough M_ε^* such that $F(-M_\varepsilon^*) \le \varepsilon/4$ and $F(M_\varepsilon^*) \ge 1 - \varepsilon/4$. Since a cdf has at most a countable number of discontinuity points, we can choose this M_ε^* such that F is continuous at M_ε^*. On the other hand, by convergence in distribution we have that $F_n(x) \to F(x)$ for all x at which F is continuous. In particular, for M_ε^*, there is an n_ε such that, for $n \ge n_\varepsilon$,

$$|F_n(-M_\varepsilon^*) - F(-M_\varepsilon^*)| \le \varepsilon/4 \quad \text{and} \quad |F_n(M_\varepsilon^*) - F(M_\varepsilon^*)| \le \varepsilon/4$$

In particular, note that this implies that for $n \geq n_\varepsilon$, $F_n(-M^*_\varepsilon) \leq F(-M^*_\varepsilon) + \varepsilon/4$ and $1 - F_n(M^*_\varepsilon) \leq 1 - F(M^*_\varepsilon) + \varepsilon/4$. Now, using that $\mathbb{P}(|X_n| > M_\varepsilon) = 1 - F_n(M^*_\varepsilon) + F_n(-M^*_\varepsilon)$ and applying the previous inequalities, we get:

$$\begin{aligned}
\mathbb{P}(|X_n| > M^*_\varepsilon) &\leq 1 - F(M^*_\varepsilon) + \varepsilon/4 + F(-M^*_\varepsilon) + \varepsilon/4 \\
&\leq 1 - F(M^*_\varepsilon) + F(-M^*_\varepsilon) + \varepsilon/2 \\
&\leq \varepsilon/4 + \varepsilon/4 + \varepsilon/2 = \varepsilon
\end{aligned}$$

for all $n \geq n_\varepsilon$. This means that X_n is bounded in probability for $n \geq n_\varepsilon$. We are left with the first $n_\varepsilon - 1$ terms in the sequence. However, since $\mathbb{P}(|X_n| > M) \to 0$ as $M \to \infty$, we can always find a large enough M^{**}_ε such that $\mathbb{P}(|X_n| > M^{**}_\varepsilon) \leq \varepsilon$ for $n = 1, 2, \cdots, n_\varepsilon - 1$. By simply choosing $M_\varepsilon = \max\{M^*_\varepsilon, M^{**}_\varepsilon\}$ we can bound the sequence for all n, which proves that X_n is bounded in probability.

(b) ($\to_{L_p}$ implies $\to_p$) Let $\sup_{n\geq 1} \mathbb{E}[|X_n|^p] < \infty$ for $p > 0$. Show that $X_n \to_{L_p} X_\infty \implies X_n \to_p X_\infty$. **Solution**: by Chebyshev's inequality, for all n and for all $\varepsilon > 0$ we have:

$$\mathbb{P}(|X_n - X_\infty| > \varepsilon) \leq \frac{\mathbb{E}\left[|X_n - X_\infty|^p\right]}{\varepsilon^p}$$

Now take the limit with $n \to \infty$:

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X_\infty| > \varepsilon) \leq \lim_{n\to\infty} \frac{\mathbb{E}\left[|X_n - X_\infty|^p\right]}{\varepsilon^p} = 0$$

by L_p convergence.

(c) ($\to_p$ without moments) Give an example of a sequence X_n such that (i) $X_n \to_p X_\infty$ with $\mathbb{E}[X_n]$ undefined for all $n = 1, 2, \ldots$, and (ii) $\mathbb{E}[X_\infty] < \infty$. **Solution**: let $X_n = X_\infty + \frac{Z}{n}$ where, for example, $X_\infty \sim \mathcal{N}(0, 1)$ and $Z \sim \text{Cauchy}(0, 1)$. Then the expectation is undefined for any n but $X_n \xrightarrow{p} X_\infty$ because $\mathbb{P}(|X_n - X_\infty| > \varepsilon) = \mathbb{P}(|Z| > n\varepsilon) \to 0$ as $n \to \infty$ and $\mathbb{E}[X_\infty] = 0$.

6. (**Non-iid LLNs**) Let $\{X_n : n \geq 1\}$ be a sequence of random variables. Show the following.

(a) (Heteroskedasticity) If $\{X_n : n \geq 1\}$ are independent with

$$\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] \to \mu \qquad \text{and} \qquad \bar{\sigma}^2_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{V}[X_i] = o(n),$$

then $\bar{X}_n \to_p \mu$. **Solution**: note that

$$\mathbb{E}[\bar{X}] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] = \bar{\mu}_n \quad \text{and } \mathbb{V}[\bar{X}] = \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{V}[X_i] = o(1)$$

then

$$\begin{aligned}
\mathbb{E}\left[\left(\bar{X} - \mu\right)^2\right] &= \mathbb{E}\left[\left(\bar{X} - \bar{\mu}_n + \bar{\mu}_n - \mu\right)^2\right] \\
&= \mathbb{E}\left[\left(\bar{X} - \bar{\mu}_n\right)^2 + \left(\bar{\mu}_n - \mu\right)^2 + 2\left(\bar{X} - \bar{\mu}_n\right)\left(\bar{\mu}_n - \mu\right)\right] \\
&= \mathbb{E}\left[\left(\bar{X} - \bar{\mu}_n\right)^2\right] + \left(\bar{\mu}_n - \mu\right)^2 + 2\left(\bar{\mu}_n - \mu\right)\mathbb{E}\left[\bar{X} - \bar{\mu}_n\right] \\
&= \mathbb{V}[\bar{X}] + \left(\bar{\mu}_n - \mu\right)^2 \to 0 \text{ as } n \to \infty
\end{aligned}$$

Then, by Markov inequality, for any $\varepsilon > 0$,

$$\mathbb{P}\left[|\bar{X}_n - \mu| > \varepsilon\right] \leq \frac{1}{\varepsilon^2}\mathbb{E}\left[\left(\bar{X} - \mu\right)^2\right] \to 0$$

Thus $\bar{X}_n \to_p \mu$.

(b) (Serial correlation) If $\{X_n : n \geq 1\}$ are identically distributed with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$ for all i, and $\text{Cov}[X_i, X_j] = \rho(|i-j|)$ for a function $\rho(x) \to 0$ as $x \to \infty$, then $\bar{X}_n \to_p \mu$. **Solution**: note that $\mathbb{E}[\bar{X}] = \mu$. Then

$$\begin{aligned}
\mathbb{E}\left[\left(\bar{X}-\mu\right)^2\right] &= \mathbb{V}[\bar{X}] \\
&= \frac{1}{n^2}\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{C}ov\left(X_i, X_j\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathbb{V}\left[X_i\right] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathbb{C}ov\left(X_i, X_j\right)\right) \\
&= \frac{\sigma^2}{n} + \frac{2}{n^2}\sum_{i>j}\rho(i-j)
\end{aligned}$$

Let's break the double sum into two sets, $S_1 := \{(i,j) : |i-j| < K\}$ and $S_2 := \{(i,j) : |i-j| \geq K\}$ for some K (which will be specified shortly):

$$\sum_{i>j}\rho(i-j) = \sum_{S_1}\rho(i-j) + \sum_{S_2}\rho(i-j)$$

Fix some $\varepsilon > 0$. Because $\rho(x) \to 0$ as $x \to \infty$, we can pick $K = K(\varepsilon)$ such that $\rho(x) < \varepsilon$ for $x \geq K$. Going back to our definition of the set S_2, we can see that our choice of K implies:

$$\sum_{S_2}\rho(i-j) < \sum_{S_2}\varepsilon < \sum_{i,j}\varepsilon = n^2\varepsilon$$

On the other hand, by Cauchy-Schwarz inequality:

$$\rho(|i-j|) = \mathbb{C}ov\left[X_i X_j\right] \leq \left(\mathbb{V}\left[X_i\right]\mathbb{V}\left[X_j\right]\right)^{1/2} = \sigma^2 < \infty$$

for all i, j. This will allow us to bound the first term in the double sum:

$$\sum_{S_1}\rho(i-j) < \sum_{S_1}\sigma^2 < nK\sigma^2$$

since, by construction of S_1, there are less than nK terms in the sum (to see this, fix, say, $i = 1$. In this case j will go from 2 to K, which means there are K pairs of the form $(1, j)$ in S_1. We can repeat this n times by varying $i = 1, 2, ..., n$ to see that there are nK pairs). This implies:

$$\sum_{i>j}\rho(i-j) = \sum_{S_1}\rho(i-j) + \sum_{S_2}\rho(i-j) < nK\sigma^2 + n^2\varepsilon$$

and therefore:

$$\frac{1}{n^2}\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) < \frac{\sigma^2}{n} + \frac{2(nK\sigma^2 + n^2\varepsilon)}{n^2} = \frac{(2K+1)n\sigma^2 + n^2\varepsilon}{n^2} = \frac{(2+1)K\sigma^2}{n} + \varepsilon$$

So

$$\lim_{n\to\infty}\mathbb{E}\left[\left(\bar{X}-\mu\right)^2\right] = \lim_{n\to\infty}\mathbb{V}[\bar{X}] < \varepsilon \quad \forall\varepsilon > 0$$

Since this is true for any arbitrary ε, we get $\lim_{n\to\infty}\mathbb{E}\left[\left(\bar{X}-\mu\right)^2\right] = 0$ and thus by the Markov inequality $\bar{X}_n \to_p \mu$.