Midterm exam #2: solutions

Please pay attention to the presentation of your answers and always provide justification. Correct answer alone will not get you full points.

Exercise 1. Quiz. (18 points) Answer each question below (1 pt) and provide a short justification (proof or counter-example) for your answer (2 pts).

a) Let X and Y be discrete random variables with probability mass functions

$$p_X(0) = p_X(1) = \frac{1}{2} \quad \text{and} \quad p_Y(0) = p_Y(1) = p_Y(2) = \frac{1}{3}.$$

What is $d_{TV}(X, Y)$? Find the maximal coupling of (X, Y); that is, find a coupling that achieves the coupling inequality.

Solution:

$$d_{TV}(X,Y) = \frac{1}{2}\sum_{k=0}^{2} |p_X(k) - p_Y(k)| = \frac{1}{2}\left(\frac{1}{2} - \frac{1}{3} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3}\right) = \frac{1}{3}$$

The maximal coupling (X', Y') will be the one that achieves $\mathbb{P}(\{X' = Y'\}) = 1 - d_{TV}(X, Y) = \frac{2}{3}$. By inspection, this is the coupling with pmf

$$p_{X'Y'}(k,m) = \begin{cases} \frac{1}{3}, & k = m \\ \frac{1}{6}, & m = 2 \end{cases}$$

b) Let X_1 and X_2 be independent random variables. Let Y_1 and Y_2 be independent random variables such that $X_i \succeq Y_i$ for $i = 1, 2$. That is, X_i stochastically dominates Y_i. Then, is it true that

$$X_1 + X_2 \succeq Y_1 + Y_2?$$

Solution: Yes, this is true. We know that $X_i \succeq Y_i$ implies that there exists a coupling $(\hat{X}_i, \hat{Y}_i)$ with $\hat{X}_i \geq \hat{Y}_i$ a.s. Pick such couplings $(\hat{X}_1, \hat{Y}_1)$ and $(\hat{X}_2, \hat{Y}_2)$ to be independent of each other. Then,

$$\hat{X}_1 + \hat{X}_2 \geq \hat{Y}_1 + \hat{Y}_2 \quad \text{a.s}$$

But, we know that $\hat{X}_1 + \hat{X}_2$ has the same distribution as $X_1 + X_2$. Likewise, $\hat{Y}_1 + \hat{Y}_2$ has the same distribution as $Y_1 + Y_2$. It follows that $X_1 + X_2 \succeq Y_1 + Y_2$.

c) Let X and Y be independent Uniform(0,1) random variables. Use convolution to find the pdf of $X + Y$.

Solution: The pdf of a Uniform$(0, 1)$ random variable is

$$f_X(x) = \mathbf{1}_{[0,1]}(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Since X and Y are independent, the pdf of $X+Y$ is the convolution

$$f_{X+Y}(t) = (f_X * f_Y)(t) = \int_{-\infty}^{\infty} f_X(s)\, f_Y(t-s)\, ds.$$

When $0 \le t \le 1$ and $0 \le s \le t \le 1$, then $0 \le t-s \le t \le 1$ and the integral is positive:

$$\begin{aligned} f_{X+Y}(t) &= \int_0^t f_X(s)\, f_Y(t-s)\, ds \\ &= \int_0^t 1\, ds \\ &= t. \end{aligned}$$

When $1 \le t \le 2$ and $t-1 \le s \le 1$, then $0 \le t-1 \le t-s \le 1$ and the integral is positive:

$$\begin{aligned} f_{X+Y}(t) &= \int_{t-1}^1 f_X(s)\, f_Y(t-s)\, ds \\ &= \int_{t-1}^t 1\, ds \\ &= 2-t. \end{aligned}$$

When $t<0$ or $t>2$, it is easy to see that $f_{X+Y}=0$. From these, we conclude that:

$$f_{X+Y}(t) = \begin{cases} t, & 0 \le t \le 1, \\ 2-t, & 1 \le t \le 2, \\ 0, & \textit{otherwise}. \end{cases}$$

d) Let X and Y be two random variables with mean 0, variance 1 and covariance ρ. Show that $\mathbb{E}[\max(X^2, Y^2)] \le 1 + \sqrt{1-\rho^2}$. *Hint : use that* $\max(a,b) = (a+b+|a-b|)/2$ *for any* $a, b \in \mathbb{R}$.

Solution:

$$\begin{aligned} E(\max\{X^2, Y^2\}) &= \frac{1}{2}E(X^2+Y^2) + \frac{1}{2}E|(X-Y)(X+Y)| \\ &\le 1 + \frac{1}{2}\sqrt{E((X-Y)^2)E((X+Y)^2)} \\ &= 1 + \frac{1}{2}\sqrt{(2-2\rho)(2+2\rho)} = 1 + \sqrt{1-\rho^2}, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality.

e) Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two Gaussian random vectors. Is it true that $X+Y = (X_1+Y_1, X_2+Y_2)$ is a Gaussian random vector?

Solution: The statement is not true in general. The sum of two Gaussian vectors is a Gaussian vector if they are jointly Gaussian (for instance, if they are independent), but this is not guaranteed by the premises.

Counterexample: Let $X_1 \sim \mathcal{N}(0,1)$ be a standard normal random variable. Let S be an independent random variable, with $\mathbb{P}(S=1) = \mathbb{P}(S=-1) = 1/2$.

Define the vectors X and Y as follows:

$$X = (X_1, 0) \quad \text{and} \quad Y = (0, SX_1)$$

First, we show that both X and Y are Gaussian vectors.

- For vector $X = (X_1, 0)$, any linear combination of its components is $aX_1 + b \cdot 0 = aX_1$, which is a normal random variable (possibly degenerate if $a = 0$). Thus, X is a Gaussian vector.
- For vector $Y = (0, SX_1)$, we first check the distribution of the component $Y_2 = SX_1$. For any $y \in \mathbb{R}$, the CDF is $\mathbb{P}(SX_1 \leq y) = \frac{1}{2}\mathbb{P}(X_1 \leq y) + \frac{1}{2}\mathbb{P}(-X_1 \leq y) = \frac{1}{2}\Phi(y) + \frac{1}{2}\mathbb{P}(X_1 \geq -y) = \frac{1}{2}\Phi(y) + \frac{1}{2}\Phi(y) = \Phi(y)$, where Φ is the standard normal CDF. So $Y_2 \sim \mathcal{N}(0, 1)$. Any linear combination of the components of Y is $a \cdot 0 + b(SX_1) = bSX_1$, which is a normal random variable. Thus, Y is a Gaussian vector.

Now, consider the sum vector $Z = X + Y$:

$$Z = (X_1, SX_1)$$

To check if Z is a Gaussian vector, we examine a linear combination of its components, for example, their sum $Z_1 + Z_2 = X_1 + SX_1 = X_1(1 + S)$. The random variable $(1 + S)$ takes the value 2 with probability $1/2$ (when $S = 1$) and the value 0 with probability $1/2$ (when $S = -1$). Therefore, the distribution of $X_1(1 + S)$ is a mixture: it is the random variable $2X_1$ with probability $1/2$ and the constant 0 with probability $1/2$. This distribution is not normal (it has a point mass at 0). Since not all linear combinations of the components of Z are normally distributed, $Z = X + Y$ is not a Gaussian vector.

f) Suppose that X and Y are i.i.d. random variables with mean 0 and variance 1. Can it be the case that $\frac{X+Y}{\sqrt{2}}$ has the same distribution as X?

Solution: Yes. Let X and Y be Gaussian random variables. Since they are independent, they form are Gaussian vector. Hence, $X + Y \sim \mathcal{N}(0, \sqrt{2})$ and $\frac{X+Y}{\sqrt{2}} \sim \mathcal{N}(0, 1)$. $\frac{X+Y}{\sqrt{2}}$ has the same distribution as X.

Exercise 2. (16 points)

Part 1. Preliminaries Consider the sequences of random variables $\{U_n\}_{n \geq 1}$ and $\{W_n\}_{n \geq 1}$, and random variables U and W, all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

a) Show that if $W_n \underset{n\to\infty}{\overset{d}{\to}} c$, where c is a constant, then $W_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} c$.

Solution: Since $W_n \underset{n\to\infty}{\overset{d}{\to}} c$, we have $\mathbb{P}(X_n \leq x) \to 0$ if $x < c$ and $\mathbb{P}(X_n \leq x) \to 1$ if $x > c$. This implies, for any $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}(|X_n - c| > \epsilon) &= \mathbb{P}(X_n > c + \epsilon) + \mathbb{P}(X_n < c - \epsilon) \\
&\leq 1 - \mathbb{P}(X_n \leq c + \epsilon) + \mathbb{P}(X_n \leq c - \epsilon) \underset{n\to\infty}{\to} 0.
\end{aligned}$$

Thus, $W_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} c$.

Slutsky's theorem: Suppose that

$$U_n \underset{n\to\infty}{\overset{d}{\to}} U \quad \text{and} \quad W_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} c$$

where c is a constant. Then, Slutsky's theorem says that

$$\text{i) } U_n W_n \underset{n\to\infty}{\overset{d}{\to}} cU \quad \text{and} \quad \text{ii) } U_n + W_n \underset{n\to\infty}{\overset{d}{\to}} c + U.$$

b) Argue that Slutsky's theorem cannot be true under a more general convergence condition $W_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} W$. It is enough to provide a counterexample for either i) or ii).

Solution: One possible counter example is as follows. Let W be a non-zero, symmetrically distributed random variable. Let $W_n = W$ and $U_n = -W$. Then $W_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} W$ and $U_n \underset{n\to\infty}{\overset{d}{\to}} W$. However, $W_n + U_n = 0$ for all n, and this does not converge in distribution to $2W$.

c) Show that if $U_n \underset{n\to\infty}{\overset{d}{\to}} U$ then $f(U_n) \underset{n\to\infty}{\overset{d}{\to}} f(U)$ for any continuous function f. *Hint: If $g(x)$ is a continuous and bounded function, and $f(x)$ is a continuous function, then $g(f(x))$ is a continuous and bounded function.*

Solution: We know from class notes that $U_n \underset{n\to\infty}{\overset{d}{\to}} U$ if and only if $\mathbb{E}(g(U_n)) \underset{n\to\infty}{\to} \mathbb{E}(g(U))$ for all continuous and bounded U. This implies that $\mathbb{E}(g(f(U_n))) \underset{n\to\infty}{\to} \mathbb{E}(g(f(U)))$ for all continuous and bounded g. This, in turn, implies that $f(U_n) \underset{n\to\infty}{\overset{d}{\to}} f(U)$.

Part 2. (Self-Centralized Central Limit Theorem) Let $\{X_k\}_{k\geq 1}$ be i.i.d. random variables with mean zero and finite variance $\sigma^2 = \mathbb{E}[X_1^2]$. Consider the sums

$$S_n = \sum_{k=1}^n X_k, \quad V_n = \left(\sum_{k=1}^n X_k^2\right), \quad \text{and} \quad T_n = \frac{S_n}{V_n^{1/2}}.$$

d) Determine the limiting distribution of $\frac{S_n}{\sqrt{n\sigma^2}}$.

Solution: By the Central Limit Theorem, since the X_k are i.i.d. with mean 0 and finite variance $\sigma^2 = \mathbb{E}[X_1^2]$,

$$\frac{S_n}{\sqrt{n\,\sigma^2}} = \frac{\sum_{k=1}^n X_k}{\sqrt{n\,\sigma^2}} \overset{d}{\to} Z, \qquad Z \sim \mathcal{N}(0,1).$$

Thus the limiting distribution of $S_n/\sqrt{n\sigma^2}$ is standard normal.

e) Does $\frac{V_n}{n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} V$ for some random variable V? If yes, what is V?

Solution: By the Weak Law of Large Numbers, applied to the nonnegative random variables X_k^2 which have mean σ^2 (i.e. $\mathbb{E}(|X_k|) < \infty$),

$$\frac{V_n}{n} = \frac{1}{n}\sum_{k=1}^n X_k^2 \overset{\mathbb{P}}{\to} \sigma^2.$$

f) Find the limiting distribution of T_n. *Hint: You may use any of the results in Part 1 of this exercise.*

Solution: From part (e), we know that

$$\frac{V_n}{n} \xrightarrow{\mathbb{P}} \sigma^2 \implies \frac{V_n}{n} \xrightarrow{d} \sigma^2.$$

We use the result in part (c) with the continuous function $f(U_n) = \sigma\, U_n^{-1/2}$ to see that

$$f\left(\frac{V_n}{n}\right) = \frac{\sigma\sqrt{n}}{V_n^{1/2}} \xrightarrow{d} f(\sigma^2) = \frac{\sigma}{\sqrt{\sigma^2}} = 1.$$

From part (a), this implies that

$$\frac{\sigma\sqrt{n}}{V_n^{1/2}} \xrightarrow{\mathbb{P}} 1.$$

We write T_n as a multiplication of two quantities:

$$T_n = \frac{S_n}{V_n^{1/2}} = \frac{S_n}{\sigma\sqrt{n}} \cdot \frac{\sigma\sqrt{n}}{V_n^{1/2}} =: U_n \cdot W_n.$$

From part (d) $\frac{S_n}{\sigma\sqrt{n}} = U_n \xrightarrow{d} Z \sim \mathcal{N}(0,1)$. We found that $\frac{\sigma\sqrt{n}}{V_n^{1/2}} = W_n \xrightarrow{\mathbb{P}} 1$. Applying the product form of Slutsky's theorem given above, we obtain

$$T_n = U_n W_n \xrightarrow{d} 1 \cdot Z = Z.$$

Hence $T_n \xrightarrow{d} \mathcal{N}(0,1)$, which completes the proof.

Exercise 3. (16 points) Consider independent random variables $\{X_n\}_{n\geq 1}$ defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with

$$\mathbb{P}(X_n = e^n) = e^{-n} \quad \text{and} \quad \mathbb{P}(X_n = 0) = 1 - e^{-n}.$$

a) Compute $\mathbb{E}\,(X_n)$ and $\mathrm{Var}(X_n)$ for $n \geq 1$.

Solution: The random variable X_n takes values in $\{0, e^n\}$. Its expectation is: $\mathbb{E}[X_n] = 0 \cdot \mathbb{P}(X_n = 0) + e^n \cdot \mathbb{P}(X_n = e^n) = e^n \cdot e^{-n} = 1$. For the variance, we first compute the second moment:

$$\mathbb{E}[X_n^2] = 0^2 \cdot \mathbb{P}(X_n = 0) + (e^n)^2 \cdot \mathbb{P}(X_n = e^n) = e^{2n} \cdot e^{-n} = e^n.$$

Thus, the variance is:

$$\mathrm{Var}(X_n) = \mathbb{E}[X_n^2] - (\mathbb{E}[X_n])^2 = e^n - 1^2 = e^n - 1.$$

b) Show that $X_n \underset{n\to\infty}{\xrightarrow{\mathbb{P}}} 0$.

Solution: To show convergence in probability to 0, we check the condition $\lim_{n\to\infty} \mathbb{P}(|X_n| > \epsilon) = 0$ for any $\epsilon > 0$. Since $X_n \geq 0$, we have $|X_n| = X_n$. For any given $\epsilon > 0$, there exists an integer N such that for all $n \geq N$, $e^n > \epsilon$. For such n, the event $\{X_n > \epsilon\}$ is equivalent to $\{X_n = e^n\}$. Therefore, for $n \geq N$:

$$\mathbb{P}(|X_n| > \epsilon) = \mathbb{P}(X_n = e^n) = e^{-n}.$$

As $n \to \infty$, $e^{-n} \to 0$, which proves that $X_n \to 0$ in probability.

c) Show that $X_n \underset{n\to\infty}{\to} 0$ a.s. *Hint: consider* $A_n = \{X_n \neq 0\}$.

Solution: Set $A_n = \{X_n \neq 0\} = \{X_n = e^n\}$. Then $\mathbb{P}(A_n) = e^{-n}$ and

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} e^{-n} = \frac{1/e}{1 - 1/e} = \frac{1}{e-1} < \infty.$$

By the Borel–Cantelli lemma,

$$\mathbb{P}\big(\{\omega : \omega \in A_n \text{ infinitely often}\}\big) = 0.$$

Hence with probability one, ω belongs to only finitely many A_n, i.e. there exists $N(\omega)$ such that for all $n > N(\omega)$, $X_n(\omega) = 0$. Therefore $X_n \to 0$ almost surely.

Define partial sums $Y_n = \sum_{k=1}^{n} X_k$ for $n \geq 1$.

d) Does $Y_n \underset{n\to\infty}{\to} Y$ a.s., for some random variable Y?

Solution: From (c), we know that for almost every $\omega \in \Omega$, the sequence of real numbers $(X_n(\omega))_{n\geq 1}$ is eventually zero. That is, there exists an $N(\omega)$ such that $X_k(\omega) = 0$ for all $k > N(\omega)$. For such an ω, the sequence of partial sums $Y_n(\omega) = \sum_{k=1}^{n} X_k(\omega)$ becomes constant for $n \geq N(\omega)$:

$$\forall n \geq N(\omega), \quad Y_n(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega) = Y_{N(\omega)}(\omega).$$

An eventually constant sequence is convergent. Therefore, for almost every ω, the sequence $(Y_n(\omega))$ converges. This means the sequence of random variables (Y_n) converges almost surely to a limit, which we denote by $Y = \sum_{k=1}^{\infty} X_k$.

e) Does $Y_n \underset{n\to\infty}{\overset{L^1}{\to}} Y$, for some random variable Y?

Solution: If (Y_n) converged in L^1, then the sequence of expectations $(\mathbb{E}[Y_n])$ would have to converge to a finite limit. By linearity of expectation,

$$\mathbb{E}[Y_n] = \mathbb{E}\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n} \mathbb{E}[X_k] = \sum_{k=1}^{n} 1 = n.$$

Since $\lim_{n\to\infty} \mathbb{E}[Y_n] = \lim_{n\to\infty} n = +\infty$, the sequence of expectations diverges. Therefore, (Y_n) does not converge in L^1.