Midterm exam: solutions

Exercise 1. Quiz. (18 points) Answer each yes/no question below (1 pt) and provide a short justification (proof or counter-example) for your answer (2 pts).

a) Let $(\Omega, \mathcal{F})$ be a measurable space and let $g\colon \Omega \to \Omega'$ be some function onto the set Ω'. We construct a collection $\mathcal{F}'$ of subsets of Ω' in the following way:

$$\mathcal{F}' = \{F' \subseteq \Omega' : g(F) = F', \text{ for some } F \in \mathcal{F}\}.$$

where $g(F) = F'$ means that $F' \subset \Omega'$ is the image of the set F. Does $(\Omega', \mathcal{F}')$ always form a measurable space?

Answer: No. Here's a counter example. Consider $\Omega = \{1, 2, 3\}$, and the σ-field $\mathcal{F} = \{\{1\}, \{2, 3\}, \Omega, \phi\}$. Let $g(\omega)$ be a non-injective function from Ω onto $\Omega' = \{1, 2\}$ such that $g(1) = g(2) = 1$ and $g(3) = 2$.

Now, the set $\mathcal{F}' = \{\phi, \Omega', \{1\}\}$, which clearly is not a σ-field on Ω'. Thus, $(\Omega', \mathcal{F}')$ does not generally form a measurable space.

b) Let $\{A_1, A_2, A_3\}$ be a collection of pairwise independent events on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $B = A_1 \cup A_2$. Are the events B and A_3 also independent?

Answer: No. One counterexample is:

$$\begin{aligned} \Omega &= \{1, 2, 3, 4\} \\ \mathcal{F} &= \mathcal{P}(\Omega) \\ \mathbb{P}(\{i\}) &= 1/4 \text{ for } i \in \Omega \\ A_1 &= \{\{1, 2\}\} \\ A_2 &= \{\{1, 3\}\} \\ A_3 &= \{\{1, 4\}\} \end{aligned}$$

It is easy to verify that $\{A_1, A_2, A_3\}$ are pairwise independent. However, $B = A_1 \cup A_2$ and A_3 are not independent, since $\mathbb{P}(B \cap A_3) = 1/4 \neq (3/4)(1/2) = \mathbb{P}(B)\mathbb{P}(A_3)$.

c) Let X, Y be two square integrable random variables. For two constants $a, b \in \mathbb{R}$, is it true that $\text{Cov}(aX + b, Y) = a\text{Cov}(X, Y) + b$?

Answer: No, $\text{Cov}(aX + b, Y) = a\text{Cov}(X, Y)$.

$$\begin{aligned} \text{Cov}(aX + b, Y) &= \mathbb{E}((aX + b)Y) - \mathbb{E}(aX + b)\mathbb{E}(Y) \\ &= \mathbb{E}(aXY) + \mathbb{E}(bY) - (a\mathbb{E}(X) + b)\mathbb{E}(Y) \\ &= a\mathbb{E}(XY) + b\mathbb{E}(Y) - a\mathbb{E}(X)\mathbb{E}(Y) - b\mathbb{E}(Y) \\ &= a\mathbb{E}(XY) - a\mathbb{E}(X)\mathbb{E}(Y) = a\text{Cov}(X, Y). \end{aligned}$$

d) Let X be a standard Gaussian distribution i.e., $X \sim \mathcal{N}(0, 1)$. Let Y be a random variable defined as follows:

$$Y = \begin{cases} X, & |X| < a \\ -X, & |X| \geq a \end{cases}$$

where $a > 0$ is a constant. Then, is Y a standard Gaussian random variable?

Answer: Yes. Here, if we can show that the CDF of Y i.e., $F_Y(t)$ is same as the CDF of X i.e., $F_X(t)$ for every $t \in \mathbb{R}$, it will imply that Y is a standard Gaussian random variable as well.

Case - I : $t \leq -a$ or $t \geq a$: For $t \leq -a$, $F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(X \geq -t) = 1 - F_X(-t) = F_X(t)$. Note that the last equality holds due to the fact that a pdf of a standard Gaussian random variable is an even function. The case for $t \geq a$ follows analogously.

Case - II : $-a < t < a$: $F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(-a < Y \leq t) + \mathbb{P}(Y \leq -a) = \mathbb{P}(-a < X \leq t) + \mathbb{P}(X \leq -a) = F_X(t)$.

e) Let X, Y be the same as in part **d)**. Is $Z = (X, Y)$ a Gaussian random vector?

Answer: No. The vector $Z := (X, Y)$ is a Gaussian random vector if $\forall \alpha, \beta \in \mathbb{R}$, we have that $W := \alpha X + \beta Y$ is a Gaussian random variable. We see that:

$$W = \begin{cases} (\alpha + \beta)X & ; \text{ if } |X| < a \\ (\alpha - \beta)X & ; \text{ if } |X| \geq a \end{cases}$$

The CDF of this random variable W is discontinuous at a, and therefore cannot be Gaussian. Thus, Z is not a Gaussian random vector.

Another way of showing that Z is not a Gaussian random variable for $\forall \alpha, \beta$ is pick $\alpha, \beta = 1$. Then, $\mathbb{P}(W = 0) = \mathbb{P}(X + Y = 0) = \mathbb{P}(|X| \geq a) > 0$. It implies that W is not a Gaussian random variable since it is not ven continuous).

Answer:

f) Let $U \sim \text{Uniform}[0, 1]$ and define

$$X_n = n 1_{\left[0, \frac{1}{n^2}\right]}(U), \quad n = 1, 2, \ldots$$

Does X_n converge in quadratic mean to some random variable X?

Answer: No, X_n does not converge in quadratic mean to anything. First, note that X_n converges in probability to zero. Therefore, if it were to converge in quadratic mean to anything, it must be zero. Then

$$\mathbb{E}((X_n - 0)^2) = n^2 \frac{1}{n^2} = 1.$$

Thus, the limit as n goes to infinity is one and not zero which is required by the definition of quadratic convergence.

Exercise 2. (12 points) Let X, Y be two independent random variable with $X \sim \text{Binomial}(n, p)$ and $Y \sim \text{Binomial}(m, p)$. Recall that the pmf of a Binomial(n, p) random variable X is given by $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

a) Compute the probability mass function of $X + Y$ using convolution. *Hint: Recall Vandermonde's identity:* $\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k} \binom{n}{r-k}$

a) Compute the probability mass function of $X+Y$ using convolution. *Hint: Recall Vandermonde's identity:* $\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}$

Answer: For $k \in \{0, 1, \ldots, m+n\}$,

$$\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{j=0}^{n} \mathbb{P}(X=j, X+Y=k) \\
&= \sum_{j=0}^{n} \mathbb{P}(X=j)\mathbb{P}(Y=k-j) \\
&= \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{m-k+j} \\
&= p^k (1-p)^{m+n-k} \sum_{j=0}^{n} \binom{n}{j}\binom{m}{k-j} \\
&= \binom{n+m}{k} p^k (1-p)^{m+n-k}.
\end{aligned}$$

Thus, $X+Y \sim \text{Binomial}(n+m, p)$.

b) Compute the characteristic functions of X and Y. *Hint: Recall the Binomial identity* $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

Answer: Let's compute the characteristic function of X.

$$\begin{aligned}
\varphi_X(t) &= \mathbb{E}(e^{itX}) \\
&= \sum_{j=0}^{n} \binom{n}{j} e^{itj} p^j (1-p)^{n-j} \\
&= \sum_{j=0}^{n} \binom{n}{j} (pe^{it})^j (1-p)^{n-j} \\
&= (1-p+pe^{it})^n.
\end{aligned}$$

Similarly, $\varphi_Y(t) = (1-p+pe^{it})^m$.

c) Check your result for part **a)**. Compute the probability mass function of $X+Y$ using the characteristic functions in part **b)**.

Answer: Since X and Y are independent,

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = (1-p+pe^{it})^n (1-p+pe^{it})^m = (1-p+pe^{it})^{n+m}.$$

Thus, $X+Y \sim \text{Binomial}(n+m, p)$, verifying our result from part **a)**.

Exercise 3. (12 points) For this problem, we remind you of basic definitions related to vector spaces in the attached appendix.

a) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{V}$ be a space of all $\mathcal{F}$-measurable random variables on this space. Show that $\mathcal{V}$ forms a vector space over the reals.

In the interest of time you may skip axioms 1-2 and 5-8 (in the appendix) as these are trivial.

With regard to axioms 3 and 4 (in the appendix): Describe the zero-element of this vector space. If $X \in \mathcal{V}$ is an $\mathcal{F}$-measurable random variable, what is its additive inverse?

Answer: Axioms 1-2 and 5-8 follow from properties of reals, and how multiplication by a scalar interacts with random variables.

With regard to axioms 3 and 4: The zero element is the deterministic random variable that maps all of Ω to zero. The additive inverse of X is just $-X$.

b) Suppose Ω is finite. Describe a basis for $\mathcal{V}$.

Answer: Ω is finite implies that $\mathcal{F}$ is also finite. Recall that a finite σ-field is generated by its atoms and that this set of atoms partitions Ω. Let $\{A_1, \ldots, A_k\}$ be the atoms of $\mathcal{F}$. Then, one possible basis for $\mathcal{V}$ would be $\{1_{A_1}, \ldots, 1_{A_k}\}$, that is, the set of indicator random variables on the atoms of $\mathcal{F}$.

First, we can confirm that $\{1_{A_1}, \ldots, 1_{A_k}\}$ spans all of $\mathcal{V}$. For every $X \in \mathcal{V}$, $X(\omega)$ is constant on each A_i. This follows from the fact that X is $\mathcal{F}$-measurable, Thus, every such X can be written as a linear combination of elements in $\{1_{A_1}, \ldots, 1_{A_k}\}$.

Secondly, we can confirm that $\{1_{A_1}, \ldots, 1_{A_k}\}$ are linearly independent. This follows from the fact that $\{A_1, \ldots, A_k\}$ partitions Ω and thus 1_{A_i} cannot be written as a linear combination of $\{1_{A_1}, \ldots, 1_{A_{i-1}}, 1_{A_{i+1}}, \ldots 1_{A_k}\}$ for al $i \in \{1, \ldots k\}$.

c) Suppose Ω is finite. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-field and let $\mathcal{W}$ be the space of all $\mathcal{G}$-measurable random variables. Then, is $\mathcal{W}$ a linear subspace of $\mathcal{V}$? Why or why not?

Answer: By part (b), $\mathcal{W}$ is also a vector space. Likewise, $\mathcal{G}$ is generated by some (smaller) set of atoms $\{B_1, \ldots, B_k\}$ and $\{1_{B_1}, \ldots, 1_{B_k}\}$ forms a basis for $\mathcal{W}$. Since every element in $\mathcal{W}$ is $\mathcal{G}$-measurable, it can be written as a linear combination of $\{1_{B_1}, \ldots, 1_{B_k}\}$. On the other hand, random variables that are not $\mathcal{G}$-measurable must not be constant on at least one set B_i. Such random variables cannot be written as linear combinations of $\{1_{B_1}, \ldots, 1_{B_k}\}$.

Exercise 4. (16 points) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable random variable.

a) Assume the p^{th}-moment of X is finite i.e., $\mathbb{E}(|X|^p) < \infty$ for some $p \in \mathbb{N}$. Then, show that $\forall$ $p' \in [0, p]$, we have $\mathbb{E}(|X|^{p'}) < \infty$.

Answer: Note that $|X(\omega)|^{p'} \leq |X(\omega)|^p$ if $|X(\omega)| \geq 1$. However, if $|X(\omega)| < 1$, we have that $|X(\omega)|^{p'} < 1 + |X(\omega)|^p$. Therefore, for every $\omega \in \Omega$, we have that $|X(\omega)|^{p'} < 1 + |X(\omega)|^p$, which implies that $\mathbb{E}(|X|^{p'}) \leq 1 + \mathbb{E}(|X|^p)$. Hence, if $\mathbb{E}(|X|^p) < \infty$, then $\mathbb{E}(|X|^{p'}) < \infty$ as well.

b) Assume that X is integrable, e.i. $\mathbb{E}(|X|) < \infty$. Show that $\exp(\mathbb{E}(\log|X|)) \leq \mathbb{E}(|X|)$. Conclude from this that for any $y_1, \ldots, y_k \in \mathbb{R}$ and $\alpha_1, \ldots, \alpha_k \geq 0$ such that $\sum_{i=1}^{k} \alpha_k = 1$, we have the following inequality:

$$|y_1|^{\alpha_1}|y_2|^{\alpha_2} \ldots |y_k|^{\alpha_k} \leq \alpha_1|y_1| + \alpha_2|y_2| + \cdots + \alpha_k|y_k|.$$

Answer: Recall that $\log(\cdot)$ is a concave function. On using Jensen's inequality, we have the following result :$\mathbb{E}(\log(|X|)) \leq \log(\mathbb{E}(|X|))$, which directly implies the result to be proven.

For the second part, Let X is a discrete random variable which takes values $y_1, y_2, \cdots, y_k$ with probabilities $\alpha_1, \alpha_2, \cdots, \alpha_k$, respectively (note that $\sum_{i=1}^{k} \alpha_i = 1$). Now, on applying the above result on this random variable X, we have

$$\exp\left(\sum_{i=1}^{k} \alpha_i \log(|y_i|)\right) \leq \sum_{i=1}^{k} \alpha_i |y_i|$$
$$\exp\left(\log\left(\prod_{i=1}^{k} |y_i|^{\alpha_i}\right)\right) \leq \sum_{i=1}^{k} \alpha_i |y_i|$$

which gives us the desired result.

In class, we have seen the Cauchy-Schwarz inequality i.e., $\mathbb{E}(|XY|) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$. In the next part, we will extend it to higher moments of the random variables X and Y by proving a result known as Holder's inequality. That is, let $p, q \geq 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. We will show that

$$\mathbb{E}(|XY|) \leq (\mathbb{E}(|X|^p))^{1/p}(\mathbb{E}(|Y|^q))^{1/q}.$$

c) First, show that for any $x, y \in \mathbb{R}$ and $p, q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we have $|xy| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q}$. This result is well-known as Young's inequality for products.

Answer: Use the above result in part - b) for $k = 2$. Substitute $y_1 = x^p$ and $y_2 = y^q$ with $\alpha_1 = 1/p$ and $\alpha_2 = 1/q$.

d) Finally, let X, Y be $\mathcal{F}$-measurable random variables with $\mathbb{E}(|X|^{\max\{p,q\}}), \mathbb{E}(|Y|^{\max\{p,q\}}) < \infty$. Show that

$$\frac{\mathbb{E}(|XY|)}{(\mathbb{E}(|X|^p))^{1/p}(\mathbb{E}(|Y|^q))^{1/q}} \leq 1,$$

from which Holder's inequality follows immediately.

Answer: Here, we can use the Young's inequality for products. $\forall\ \omega \in \Omega$, let $x = \frac{X(\omega)}{(E(|X|^p))^{1/p}}$ and $y = \frac{Y(\omega)}{(E(|Y|^q))^{1/q}}$. On using the above result in part - c)., we have

$$\left|\frac{X(\omega)Y(\omega)}{(E(|X|^p))^{1/p}(E(|Y|^q))^{1/q}}\right| \leq \frac{|X(\omega)|^p}{pE(|X|^p)} + \frac{|Y(\omega)|^q}{qE(|Y|^q)}$$

Now, on taking expectation both sides w.r.t P_{XY}, we have

$$\frac{\mathbb{E}|X(\omega)Y(\omega)|}{(E(|X|^p))^{1/p}(E(|Y|^q))^{1/q}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

Thus, we have the desired result.