Solutions to Midterm Exam

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

a) Let $\mathcal{F}$ be a σ-field on Ω and $\mathcal{G}, \mathcal{H}$ be two sub-σ-fields of $\mathcal{F}$.

a1) Does it always hold that $\mathcal{G} \cap \mathcal{H}$ is a σ-field ?

Note: $\mathcal{G} \cap \mathcal{H}$ is by definition the list of subsets of Ω belonging to both $\mathcal{G}$ and $\mathcal{H}$.

Answer: Yes. $\emptyset, \Omega \in \mathcal{G} \cap \mathcal{H}$; if $A \in \mathcal{G} \cap \mathcal{H}$, then $A^c \in \mathcal{G} \cap \mathcal{H}$ also; and if $(A_n, n \geq 1)$ is a sequence of events such that $A_n \in \mathcal{G} \cap \mathcal{H}$ for every $n \geq 1$, then $\cup_{n\geq 1} A_n \in \mathcal{G}$ and $\cup_{n \geq 1} A_n \in \mathcal{H}$, so $\cup_{n\geq 1} A_n \in \mathcal{G} \cap \mathcal{H}$.

a2) Does it always hold that $\mathcal{G} \cup \mathcal{H}$ is a σ-field ?

Note: $\mathcal{G} \cup \mathcal{H}$ is by definition the list of subsets of Ω belonging to either $\mathcal{G}$ or $\mathcal{H}$.

Answer: No. For example, if $A \in \mathcal{G}$ and $B \in \mathcal{H}$, then it is not always the case that $A \cup B \in \mathcal{G}$ or $A \cup B \in \mathcal{H}$. Example: $\Omega = \{1, 2, 3, 4\}$, $\mathcal{G} = \sigma(\{1, 2\})$, $\mathcal{H} = \sigma(\{1, 3\})$; then $A = \{1, 2\} \in \mathcal{G}$, $B = \{1, 3\} \in \mathcal{G}$, but $A \cup B = \{1, 2, 3\}$ does not belong to $\mathcal{G}$, nor to $\mathcal{H}$.

b) Let X, Y, Z be three random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

b1) Does it always hold that $\sigma(X, Y, Z) = \sigma(X + Y, Y + Z, Z + X)$?

Answer: Yes, as the transformation $(x, y, z) \mapsto (x + y, y + z, z + x)$ is a one-to-one transformation (as the computation $\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = 2 \neq 0$ shows). The knowledge of (X, Y, Z) is therefore equivalent to that of $(X + Y, Y + Z, Z + X)$.

b2) If $\sigma(X + Y + Z) = \sigma(X, Y, Z)$, does that necessarily imply that X, Y, Z are independent ?

Answer: No. Here is a simple counter-example: $X = Y = Z$, and X is a non-constant random variable.

c) Let X be a continuous random variable and Y be a discrete random variable which is independent of X and also such that $Y(\omega) \neq 0$ for all $\omega \in \Omega$. Let finally $Z = X \cdot Y$.

c1) Is it always the case that Z is a continuous random variable ?

Answer: Yes. Let C be the countable set of (non-zero) values of Y and let us compute

$$
\begin{aligned}
F_Z(t) &= \sum_{y \in C} \mathbb{P}(\{X \cdot Y \leq t\} | \{Y = y\})\, \mathbb{P}(\{Y = y\}) \\
&= \sum_{y \in C, y > 0} \mathbb{P}(\{X \leq t/y\})\, \mathbb{P}(\{Y = y\}) + \sum_{y \in C, y < 0} \mathbb{P}(\{X \geq t/y\})\, \mathbb{P}(\{Y = y\})
\end{aligned}
$$

so

$$F_Z(t) = \sum_{y \in C, y>0} F_X(t/y)\, \mathbb{P}(\{Y = y\}) + \sum_{y \in C, y<0} (1 - F_X(t/y))\, \mathbb{P}(\{Y = y\})$$

which is differentiable as F_X is (and y never takes the value 0).

c2) Assume now that $X \sim \mathcal{N}(0,1)$ and $\mathbb{P}(\{Y = +1\}) = \mathbb{P}(\{Y = +2\}) = \frac{1}{2}$. Is Z a Gaussian random variable in this case ?

Answer: No. Using the above formula, we obtain

$$F_Z(t) = \frac{1}{2}\, F_X(t) + \frac{1}{2}\, F_X(t/2)$$

whose derivative is

$$p_Z(t) = \frac{1}{2\sqrt{2\pi}} \left(\exp(-t^2/2) + \frac{1}{2} \exp(-t^2/8) \right)$$

which is not a Gaussian (it is a centered random variable, but there is no value of $\sigma > 0$ such that $p_Z(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-t^2/2\sigma^2)$),

Exercise 2. (15 points)
a) Let $\lambda > 0$ and $X \sim \mathcal{E}(\lambda)$, i.e., $p_X(x) = \lambda\, e^{-\lambda x}$ for $x \geq 0$. What is the distribution of $Y = \mu X$, where $\mu > 0$?

Answer: $Y \sim \mathcal{E}(\lambda/\mu)$. The simplest way to see this is to note that since Y is a scaled version of X, it also has an exponential distribution, and that $\mathbb{E}(Y) = \mu\, \mathbb{E}(X) = \mu/\lambda$.

b) Let X be a discrete random variable taking values in $\mathbb{N}$ and such that $\mathbb{P}(\{X \geq k\}) = \frac{2}{k(k+1)}$ for every $k \geq 1$. Compute $\mathbb{E}(X)$.

Answer: (Side note: observe that $\mathbb{P}(\{X = 0\}) = 0$, as $\mathbb{P}(\{X \geq 1\}) = 1$). Let us then compute

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k \geq 1} \mathbb{P}(\{X \geq k\}) = \sum_{k \geq 1} \frac{2}{k(k+1)} = \sum_{k \geq 1} \left(\frac{2}{k} - \frac{2}{k+1} \right) \\
&= \lim_{n \to \infty} \sum_{k=1}^{n} \left(\frac{2}{k} - \frac{2}{k+1} \right) = \lim_{n \to \infty} 2 - \frac{2}{n+1} = 2
\end{aligned}$$

Another option is to compute:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k \geq 1} k\, \mathbb{P}(\{X = k\}) = \sum_{k \geq 1} k\, \left(\mathbb{P}(\{X \geq k\}) - \mathbb{P}(\{X \geq k+1\}) \right) \\
&= \sum_{k \geq 1} k \left(\frac{2}{k(k+1)} - \frac{2}{(k+1)\,(k+2)} \right) = \sum_{k \geq 1} \frac{2k}{k+1}\, \frac{2}{k(k+2)} \\
&= 2 \sum_{k \geq 1} \frac{2}{(k+1)\,(k+2)} = 2 \sum_{\ell \geq 2} \frac{2}{\ell(\ell+1)} = 2\,(\mathbb{E}(X) - 1)
\end{aligned}$$

so $\mathbb{E}(X) = 2$.

c) Let X be a $\mathcal{U}([0,1])$ random variable, Y be independent of X and such that $\mathbb{P}(\{Y=+1\}) = \mathbb{P}(\{Y=-1\}) = \frac{1}{2}$ and $Z = X \cdot Y$. Compute $\phi_Z(t) \in \mathbb{R}$ for $t \in \mathbb{R}$.

Answer: Let us compute

$$\begin{aligned}\phi_Z(t) &= \mathbb{E}(\exp(itXY)) = \frac{1}{2}\,\mathbb{E}(\exp(itX)) + \frac{1}{2}\mathbb{E}(\exp(-itX)) \\ &= \frac{1}{2}\,\frac{e^{it}-1}{it} + \frac{1}{2}\,\frac{e^{-it}-1}{-it} = \frac{e^{it}-e^{-it}}{2it} = \frac{\sin(t)}{t}\end{aligned}$$

Another way to obtain this result is to observe that $Z \sim \mathcal{U}([-1,+1])$ and to compute directly the characteristic function.

d) Let X_1, X_2 be two square-integrable random variables such that $\mathrm{Var}(X_1+X_2) = \mathrm{Var}(X_1-X_2)$. Compute $\mathrm{Cov}(X_1, X_2)$.

Answer: Let us expand the variance on both sides:

$$\begin{aligned}\mathrm{Var}(X_1+X_2) &= \mathrm{Var}(X_1) + 2\mathrm{Cov}(X_1,X_2) + \mathrm{Var}(X_2) \\ \mathrm{Var}(X_1-X_2) &= \mathrm{Var}(X_1) - 2\mathrm{Cov}(X_1,X_2) + \mathrm{Var}(X_2)\end{aligned}$$

If the 2 variances are equal, then $2\mathrm{Cov}(X_1,X_2) = -2\mathrm{Cov}(X_1,X_2)$, meaning that $\mathrm{Cov}(X_1,X_2) = 0$.

e) Does there exist a non-negative random variable X such that $\mathbb{E}(X) = 10$ and $\mathbb{E}(2^X) = 1000$? Justify your answer.

Answer: No. By Jensen's inequality, it should hold that $2^{\mathbb{E}(X)} \leq \mathbb{E}(2^X)$, but $1024 \not\leq 1000$.

Exercise 3. (13 points)
Let $n \geq 1$ and $X_1, \ldots, X_n$ be i.i.d. random variables with common cdf $F(t) = \exp(-\exp(-t))$ for $t \in \mathbb{R}$.

a) Verify that F is indeed a cdf.

Answer: $F'(t) = \exp(-\exp(-t))\,\exp(-t) > 0$ for all $t \in \mathbb{R}$, so F is increasing; $\lim_{t\to+\infty} F(t) = \exp(-0) = 1$, $\lim_{t\to-\infty} F(t) = \exp(-\infty) = 0$; and F is continuous.

b) Compute *both* the cdf and the pdf of

$$Y_n = \max\{X_1, \ldots, X_n\} - \ln(n)$$

Answer: Let us compute

$$\begin{aligned}F_{Y_n}(t) &= \mathbb{P}(\{\max\{X_1,\ldots,X_n\} \leq t + \ln(n)\}) = \mathbb{P}(\{X_1 \leq t+\ln(n)\})^n \\ &= \exp(-\exp(-t-\ln(n)))^n = \exp(-n\exp(-t)/n) = F(t)\end{aligned}$$

So for every $n \geq 1$, Y_n has the same cdf as X_1, and its pdf is given by

$$p_{Y_n}(t) = \exp(-\exp(-t))\,\exp(-t)$$

c) Compute *both* the cdf and the pdf of

$$Z_n = \min\{\exp(-X_1), \ldots, \exp(-X_n)\}$$

Answer: Two possibilities here: either compute directly

$$\mathbb{P}(\{Z_n \geq t\}) = \mathbb{P}(\{\exp(-X_1) \geq t\})^n = \mathbb{P}(\{X_1 \leq -\ln(t)\})^n = \exp(-n \exp(\ln(t))) = \exp(-nt)$$

so that $F_{Z_n}(t) = 1 - \exp(-nt)$ and $p_{Z_N}(t) = n \exp(-nt)$ [in other words, $Z_n \sim \mathcal{E}(n)$].

Or observe that

$$Z_n = \exp(-\max\{X_1, \ldots, X_n\}) = \exp(-Y_n - \ln(n)) = \frac{\exp(-Y_n)}{n}$$

which by part b) has the same distribution as $\frac{\exp(-X_1)}{n}$. Then

$$\mathbb{P}(\{Z_n \leq t\}) = \mathbb{P}(\{\exp(-X_1) \leq nt\}) = \mathbb{P}(\{X_1 \geq -\ln(nt)\}) = 1 - \exp(-nt)$$

reaching the same conclusion.

Exercise 4. (14 points)

Let $\alpha > 0$ and $(X_n,\, n \geq 1)$ be a sequence of independent random variables such that

$$\mathbb{P}(\{X_n = +n^\alpha\}) = \mathbb{P}(\{X_n = -n^\alpha\}) = \frac{1}{2n} \quad \text{and} \quad \mathbb{P}(\{X_n = 0\}) = 1 - \frac{1}{n} \quad \text{for } n \geq 1$$

Let also $S_n = X_1 + \ldots + X_n$ for $n \geq 1$.

a) Compute $\mathbb{E}(S_n)$ and $\mathrm{Var}(S_n)$, then estimate both quantities as a function of n using the approximation (valid for a generic value of $\gamma \in \mathbb{R}$):

$$\sum_{j=1}^n j^\gamma \simeq \int_1^n dx\, x^\gamma$$

$\Big($as an example, such an approximation allows to estimate $\displaystyle\sum_{j=1}^n j \simeq \int_1^n dx\, x \simeq \frac{n^2}{2}$.$\Big)$

Answer: $\mathbb{E}(S_n) = \sum_{j=1}^n \mathbb{E}(X_j) = 0$ and as the X_j are independent,

$$\mathrm{Var}(S_n) = \sum_{j=1}^n \mathrm{Var}(X_j) = \sum_{j=1}^n \mathbb{E}(X_j^2) = \sum_{j=1}^n j^{2\alpha-1}$$

Using the hint, the variance behaves as n gets large as

$$\mathrm{Var}(S_n) \simeq \int_1^n dx\, x^{2\alpha-1} \simeq \frac{n^{2\alpha}}{2\alpha}$$

b) For what values of $\beta > 0$ can you show that $\frac{S_n}{n^\beta} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$?

Answer: By Chebyshev's inequality, for every $\varepsilon > 0$,

$$\mathbb{P}\left(\left\{\left|\frac{S_n}{n^\beta}\right| \geq \varepsilon\right\}\right) = \mathbb{P}(\{|S_n| \geq n^\beta \varepsilon\}) \leq \frac{\mathbb{E}(S_n^2)}{n^{2\beta}\,\varepsilon^2} = \frac{\mathrm{Var}(S_n)}{n^{2\beta}\,\varepsilon^2} \simeq \frac{n^{2(\alpha-\beta)}}{2\alpha\,\varepsilon^2}$$

which tends to 0 as $n \to \infty$ as soon as $\beta > \alpha$.

c) For what values of $\beta > 0$ can you show that $\frac{S_n}{n^\beta} \underset{n\to\infty}{\to} 0$ almost surely ?

Answer: In order to show almost sure convergence via the Borel-Cantelli lemma, we need to show that

$$\sum_{n\geq 1} \mathbb{P}\left(\left\{\left|\frac{S_n}{n^\beta}\right| \geq \varepsilon\right\}\right) < +\infty$$

As

$$\sum_{n\geq 1} \mathbb{P}\left(\left\{\left|\frac{S_n}{n^\beta}\right| \geq \varepsilon\right\}\right) \lesssim \sum_{n\geq 1} \frac{n^{2(\alpha-\beta)}}{2\alpha\,\varepsilon^2}$$

the sum is finite if $2(\alpha - \beta) < -1$, i.e., $\beta > \alpha + \frac{1}{2}$.