Final exam: 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. (20 points) Answer each question below making sure to provide a short justification (proof or counter-example) for your answer.

a) Let $\Omega = [0, 1]$ and $\mathcal{F} = \{A\colon A \text{ or } A^c \text{ is countable}\}$. Is $\mathcal{F}$ a σ-field over Ω?

Solution:

We need to check that the three properties of the σ-field are satisfied.

(i) $\emptyset, \Omega \in \mathcal{F}$ since the empty set is countable, and Ω is its complement.

(ii) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$ follows directly from how we defined $\mathcal{F}$.

(iii) For the last property we can consider two cases. First, the sequence $(A_n, n \geq 1)$ contains countable sets only. Then, a countable union of countable sets is countable, so this property is satisfied. For the second case, suppose at least one of the sets A_i is a complement of a countable set. Then, by DeMorgan's law we have that

$$(\cup_{n=1}^{\infty} A_n)^c = \cap_{n=1}^{\infty} A_n^c.$$

Since A_i^c is countable, the intersection $\cap_{n=1}^{\infty} A_n^c$ is countable, and $\cup_{n=1}^{\infty} A_n$ is a complement of a countable set.

b) Let $\Omega = \{1, 2, \ldots, p\}$ where p is prime, $\mathcal{F}$ is the complete σ-field over Ω, and $\mathbb{P}(A) = \frac{|A|}{p}$ for all $A \in \mathcal{F}$. Show that if A and B are independent events, then at least one of A and B is either $\emptyset$ or Ω.

Solution:

If A and B are independent then we have $\frac{|A|}{p} \cdot \frac{|B|}{p} = \frac{|A \cap B|}{p}$, and it follows that

$$\frac{|A||B|}{p} = |A \cap B|.$$

Since the RHS is an integer, p must be a prime deviser of either $|A|$ or $|B|$. But, this is only possible if either $|A|$ or $|B|$ are 0 or p. This, in turn, means that at least one of A and B is either $\emptyset$ or Ω.

c) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let N be a random variable taking values in $\mathbb{N}^+$ and $(X_n)_{n \in \mathbb{N}^+}$ be a sequence of random variables. We define X_N by

$$\forall \omega \in \Omega, \quad X_N(\omega) = X_{N(\omega)}(\omega).$$

Show that X_N is a random variable.

Solution: The goal is to show that the function $X_N : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable.

Let $B \in \mathcal{B}(\mathbb{R})$. We consider

$$\begin{aligned} A = (X_N)^{-1}(B) &= \{X_N \in B\} = \{\omega \in \Omega : X_{N(\omega)}(\omega) \in B\} \\ &= \bigcup_{n \in \mathbb{N}^+} \{\omega \in \Omega : N(\omega) = n \text{ and } X_n(\omega) \in B\} \\ &= \bigcup_{n \in \mathbb{N}^+} (\{N = n\} \cap \{X_n \in B\}). \end{aligned}$$

Since N and X_n are random variables, the sets $\{N = n\}$ and $\{X_n \in B\}$ are in $\mathcal{F}$. It follows that $A \in \mathcal{F}$ and therefore that X_N is measurable.

d) Let X be a real square-integrable random variable defined on $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. Show that, if we define $\text{Var}(X \mid \mathcal{G}) = E[(X - E[X \mid \mathcal{G}])^2 \mid \mathcal{G}]$ then we have

$$\text{Var}(X) = E(\text{Var}(X \mid \mathcal{G})) + \text{Var}(E[X \mid \mathcal{G}]).$$

Solution: We write

$$X - E[X] = X - E[X \mid \mathcal{G}] + E[X \mid \mathcal{G}] - E[X].$$

The random variable $E[X \mid \mathcal{G}] - E[X]$ is $\mathcal{G}$-measurable. However, by the definition of conditional expectation in $L^2(\Omega, \mathcal{F}, P)$, $X - E[X \mid \mathcal{G}]$ is orthogonal to any r.v. that is $\mathcal{G}$-measurable.

We deduce

$$\begin{aligned} E\left[(X - E[X])^2\right] &= E\left[(X - E[X \mid \mathcal{G}])^2\right] + E\left[(E[X \mid \mathcal{G}] - E[X])^2\right] \\ &= E[\text{Var}(X \mid \mathcal{G})] + \text{Var}(E[X \mid \mathcal{G}]). \end{aligned}$$

Alternatively, we can calculate directly the cross-term without mentioning orthogonality. Let $Y = E[X \mid \mathcal{G}] - E[X]$. Note that Y is $\mathcal{G}$-measurable. Using the Law of Iterated Expectations (Tower Property):

$$\begin{aligned} E[(X - E[X \mid \mathcal{G}])Y] &= E\big[E[(X - E[X \mid \mathcal{G}])Y \mid \mathcal{G}]\big] \\ &= E\big[Y \cdot E[X - E[X \mid \mathcal{G}] \mid \mathcal{G}]\big] \quad \text{(since } Y \text{ is } \mathcal{G}\text{-measurable)} \\ &= E\big[Y \cdot (E[X \mid \mathcal{G}] - E[X \mid \mathcal{G}])\big] \\ &= E\big[Y \cdot 0\big] = 0. \end{aligned}$$

Exercise 2. (20 points) Monte-Carlo integration

Let f be a measurable function on $[0,1]$ with $\int_0^1 |f(x)|dx < \infty$. Let $U_1, U_2, \dots$ be independent and uniformly distributed on $[0,1]$, and let

$$I_n = \frac{1}{n}\left(f(U_1) + \cdots + f(U_n)\right).$$

a) Show that $I_n \underset{n \to \infty}{\overset{\mathbb{P}}{\to}} I$, where $I = \int_0^1 f(x)dx$.

Solution: Let $X_i = f(U_i)$ for $i \geq 1$. Then, $X_1, X_2, \dots$ is a sequence of iid random variables. Since $\mathbb{E}(|X_i|) = \int_0^1 |f(x)|dx < \infty$, from the law of large numbers,

$$I_n = \frac{1}{n}\sum_{i=1}^n X_i \overset{\mathbb{P}}{\to} \mathbb{E}[X_1] = \mathbb{E}[f(U_i)] = \int_0^1 f(x)dx = I.$$

b) Suppose that $\int_0^1 |f(x)|^2 dx < \infty$. Use the Chebyshev's inequality to estimate $\mathbb{P}(|I_n - I| \geq t)$ for $t > 0$.

Solution: First, we calculate the following:

$$\begin{aligned}
\mathbb{E}[I_n^2] &= \frac{1}{n^2}\left(n\mathbb{E}[X_i^2] + n(n-1)\mathbb{E}[X_i]^2\right) \\
&= \frac{1}{n}\left(\mathbb{E}[X_i^2] + (n-1)\mathbb{E}[X_i]^2\right) \\
&= \frac{1}{n}\left(\int_0^1 f(x)^2 dx + (n-1)I^2\right) \\
Var(I_n) &= \mathbb{E}[(I_n - \mathbb{E}[I_n])^2] = \mathbb{E}[(I_n - I)^2] = \mathbb{E}[I_n^2] - I^2 \\
&= \frac{1}{n}\left(\int_0^1 f(x)^2 dx + (n-1)I^2\right) - I^2 \\
&= \frac{1}{n}\left(\int_0^1 f(x)^2 dx - I^2\right).
\end{aligned}$$

Since $\int_0^1 |f(x)|^2 dx < \infty$, we can use Chebyshev's inequality with the function $\psi(t) = t^2$:

$$\begin{aligned}
\mathbb{P}(|I_n - I| \geq t) &\geq \frac{\mathbb{E}[\psi(|I_n - I|)]}{\psi(t)} \\
&\geq \frac{\mathbb{E}[(|I_n - I|)^2]}{t^2} \geq \frac{\sigma}{nt^2},
\end{aligned}$$

where $\sigma = \left(\int_0^1 f(x)^2 dx - I^2\right)$.

c) Suppose $\sup_{x\in[0,1]} |f(x)| < \infty$. Use Hoeffding's inequality to estimate $\mathbb{P}(|I_n - I| \geq t)$ for $t > 0$.

Solution: Since $\sup_{x\in[0,1]} |f(x)| < \infty$, $|I_n - I|$ is bounded. Assuming $a \leq |I_n - I| \leq b$,

$$\mathbb{P}(|I_n - I| \geq t) \leq 2exp\left(-\frac{nt^2}{2(b-a)^2}\right)$$

d) Briefly compare the results in parts (b) and (c). For example, which result is stronger? Which is more general?

Solution: The result in part (c), decays exponentially, so it is stronger. However, the assumption for part (c) is stricter, thus the bound in part (b) is more general.

Exercise 3. (30 points) Let $X_1, X_2, \ldots$ be independent random variables with

$$X_n = \begin{cases} 1 & \text{with probability } (2n)^{-1} \\ 0 & \text{with probability } 1 - n^{-1} \\ -1 & \text{with probability } (2n)^{-1} \end{cases}$$

Let $M_0 = 0$ and for $n \geq 1$

$$M_n = \begin{cases} X_n & \text{if } M_{n-1} = 0 \\ nM_{n-1}|X_n| & \text{if } M_{n-1} \neq 0 \end{cases}$$

a) Show that $(M_n, n \in \mathbb{N})$ is a martingale with respect to $\mathcal{F}_n = \sigma(M_0, M_1, \ldots, M_n)$.

Solution:

We need to check that the three properties of the martingale hold.

i) First, lets check that M_n is integrable for all n. $|M_0| = 0$ and $|M_1| = 1$. For $n \geq 2$

$$\begin{aligned}\mathbb{E}(|M_n|) &= \mathbb{E}(|X_n| 1\{M_{n-1} = 0 + n|M_{n-1}||X_n|\}) = \mathbb{E}(|X_n|1\{M_{n-1} = 0) + n\mathbb{E}(|M_{n-1}||X_n|\}) \\ &= \mathbb{E}(|X_n|)\mathbb{E}(1\{M_{n-1} = 0) + n\mathbb{E}(|M_{n-1}|)\mathbb{E}(|X_n|\}) = \frac{1}{n} \cdot \frac{n-2}{n-1} + \mathbb{E}(|M_{n-1}|) < \infty\end{aligned}$$

as needed.

ii) $(M_n, n \in \mathbb{N})$ is adapted to the filtration $(\mathcal{F}_n, n \in \mathbb{N})$ by definition

iii) Finally, for the martingale property:

$$\begin{aligned}\mathbb{E}(M_{n+1}|\mathcal{F}_n) &= \mathbb{E}(X_{n+1}1\{M_n = 0\} + (n+1)M_n|X_{n+1}||\mathcal{F}_n) \\ &= \mathbb{E}(X_{n+1}1\{M_n = 0\}|\mathcal{F}_n) + \mathbb{E}((n+1)M_n|X_{n+1}||\mathcal{F}_n) \\ &= \mathbb{E}(X_{n+1})\mathbb{E}(1\{M_n = 0\}|\mathcal{F}_n) + (n+1)\mathbb{E}(M_n|\mathcal{F}_n)\mathbb{E}(|X_{n+1}||\mathcal{F}_n) \\ &= 0 \cdot \mathbb{E}(1\{M_n = 0\}|\mathcal{F}_n) + (n+1)M_n\frac{1}{n+1} = M_n.\end{aligned}$$

b) Show that $(M_n, n \in \mathbb{N})$ does not converge almost surely.

Solution:

Intuitively, a typical realization of $(M_n, n \in \mathbb{N})$ looks as follows. We have sequences of zeros, followed by an occasional 1 or -1, followed by occasional exponential growth that gets zeroed out again. Note that the terms on the order of n or higher actually stop appearing after a while. But, for a typical ω the martingale will take values 1 or -1 infinitely often and so the sequence $(M_n(\omega), n \in \mathbb{N})$ does not converge to zero.

To show this formally, we proceed as follows. First, we show that $(M_n, n \in \mathbb{N})$ converges in probability to zero. Observe that $\mathbb{P}(\{M_n = 0\}) = \mathbb{P}(\{X_n = 0\})$. For any $\epsilon > 0$ we have

$$\mathbb{P}(|M_n| > \epsilon) = \mathbb{P}(|X_n| > \epsilon) \leq \frac{1}{n} \to 0$$

and so convergence in probability is satisfied.

We can disprove almost sure convergence by appealing to Lemma 9.2 in the notes. Fix, $0 < \epsilon < 1$. Then,

$$\begin{aligned}\mathbb{P}\left(\omega \in \Omega \colon |M_n(\omega)| > \epsilon \text{ i.o.}\right) &= \mathbb{P}\left(\omega \in \Omega \colon |X_n(\omega)| > \epsilon \text{ i.o.}\right) \\ &= 1 - \mathbb{P}\left(\omega \in \Omega \colon \exists N \text{ s.t. } X_n(\omega) = 0 \forall n > N\right) = 1.\end{aligned}$$

The last equality follows since for any $N > 0$,

$$\mathbb{P}\left(\omega \in \Omega \colon X_n(\omega) = 0 \forall n > N\right) = \prod_{n=N}^{\infty} \frac{n-1}{n} = 0.$$

c) Why does the martingale convergence theorem not apply?

The MCT.V2 fails because $\sup_{n\in\mathbb{N}} \mathbb{E}(|M_n|) = \infty$. This can be seen from part a) where we derived the recursion for $n > 1$,

$$\begin{aligned}\mathbb{E}(|M_n|) &= \frac{1}{n} \cdot \frac{n-2}{n-1} + \mathbb{E}(|M_{n-1}|) = \frac{1}{n} \cdot (1 - \frac{1}{n-1}) + \mathbb{E}(|M_{n-1}|) \\ &= \frac{1}{n} \cdot (1 - \frac{1}{n-1}) + \frac{1}{n-1} \cdot (1 - \frac{1}{n-2}) + \mathbb{E}(|M_{n-2}|) = \dots \\ &= \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k(k-1)}\end{aligned}$$

which clearly diverges as n goes to infinity.

The MCT.V1 fails because $\sup_{n\in\mathbb{N}} \mathbb{E}(|M_n|) = \infty$ implies $\sup_{n\in\mathbb{N}} \mathbb{E}(M_n^2) = \infty$.

Solution:

d) Define a stopping time $T = \inf\{n \in \mathbb{N} \colon M_n \leq -1 \text{ or } M_n \geq 1\}$. Show that $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.

Hint: You should use a stopped martingale $M_{T\wedge n}$.

Solution:

This problem is more trivial than intended since $T = 1$ always. Thus, $\mathbb{E}(M_T) = \mathbb{E}(M_1) = \mathbb{E}(M_0)$.

To make it more interesting, we should have defined T as follows:

$$T = \inf\{n \in \mathbb{N} \colon n \geq 2 \text{ and } (M_n \leq -1 \text{ or } M_n \geq 1)\}.$$

The argument follows along the same lines as OSTV3. Namely, we define a stopped martingale $M_{T\wedge n}$. This new martingale is bounded between -1 and 1 and so MCTV1 applies. This new martingale is closed at infinity and admits a limit $M_{T\wedge\infty}$. Thus,

$$\mathbb{E}(M_T) = M_{T\wedge\infty} = M_{T\wedge 0} = M_{T_0}.$$

Exercise 4. (30 points) Skorokhod's representation theorem

This exercise guides you through a proof of a special case of Skorokhod's representation theorem. The theorem states that if a sequence of random variables converges in distribution, we can construct a new sequence on a common probability space that has the same distributions and converges almost surely. We will prove this for the case where the limiting random variable has a continuous distribution function.

Let $(X_n, n \geq 1)$ be a sequence of real-valued random variables converging in distribution to a random variable X. Let F_n be the cumulative distribution function (CDF) of X_n, and F be the CDF of X. Assume that F is continuous on $\mathbb{R}$.

For a CDF G, define its **generalized inverse** or **quantile function** $G^{\leftarrow} : (0,1) \to \mathbb{R}$ by

$$G^{\leftarrow}(u) = \inf\{x \in \mathbb{R} : G(x) \geq u\}.$$

We recall the facts established in Exercise 1(h) of the first midterm: for any CDF G, any $u \in (0,1)$ and any $x \in \mathbb{R}$,

$$G^{\leftarrow}(u) \leq x \iff u \leq G(x).$$

Moreover, if $U \sim \text{Uniform}(0,1)$, then the random variable $Y = G^{\leftarrow}(U)$ has G as its CDF.

Convergence of Quantile Functions

Let $x_0 \in \mathbb{R}$ be a point of continuity for F. That is, for any $\epsilon > 0$, that there exists $\delta > 0$ such that if $|x - x_0| < \delta$, then $|F(x) - F(x_0)| < \epsilon$.

a) Fix $u \in (0, 1)$ and set $x_0 = F^{\leftarrow}(u) = \inf\{x : F(x) \geq u\}$. Prove that for every $\varepsilon > 0$ there exists N such that for all $n \geq N$,

$$\left|F_n^{\leftarrow}(u) - F^{\leftarrow}(u)\right| < \varepsilon.$$

Hint: Using continuity of F at x_0, choose $\alpha < x_0 < \beta$ with $F(\alpha) < u < F(\beta)$. Show that for large n, $F_n(\alpha) < u < F_n(\beta)$ and apply the equivalence $G^{\leftarrow}(u) \leq x \iff u \leq G(x)$ twice to trap $F_n^{\leftarrow}(u)$ between α and β.

Solution: First note that what you are asked to prove has a counter example. Namely, if F is flat on some interval, then the convergence is not guaranteed and we would need to handle this case separately. From now on, assume F is strictly increasing. Let $u \in (0, 1)$ be fixed. Let $x = F^{\leftarrow}(u)$. Since F is continuous, for any $\epsilon > 0$, we have $F(x - \epsilon) < u < F(x + \epsilon)$. The condition $F_n \to F$ in distribution means $F_n(y) \to F(y)$ for all points y where F is continuous. Since we assumed F is continuous everywhere, this convergence holds for all $y \in \mathbb{R}$. So, for our chosen $\epsilon > 0$, we can find an integer N such that for all $n \geq N$:

$$F_n(x - \epsilon) < u < F_n(x + \epsilon)$$

This is because $F_n(x - \epsilon) \to F(x - \epsilon) < u$ and $F_n(x + \epsilon) \to F(x + \epsilon) > u$. Now we use the property from 1(a) on this pair of inequalities:

- $F_n(x - \epsilon) < u$ implies that $x - \epsilon < F_n^{\leftarrow}(u)$.
- $u < F_n(x + \epsilon)$ implies that $F_n^{\leftarrow}(u) \leq x + \epsilon$.

Combining these gives, for $n \geq N$:

$$x - \epsilon < F_n^{\leftarrow}(u) \leq x + \epsilon$$

which means $|F_n^{\leftarrow}(u) - x| \leq \epsilon$. Since $x = F^{\leftarrow}(u)$, this is exactly

$$|F_n^{\leftarrow}(u) - F^{\leftarrow}(u)| \leq \epsilon.$$

This holds for any $\epsilon > 0$ (by choosing N large enough), so we have shown that $\lim_{n\to\infty} F_n^{\leftarrow}(u) = F^{\leftarrow}(u)$ for any $u \in (0, 1)$.

Almost Sure Convergence

b) Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P}) = ((0, 1), \mathcal{B}(0, 1), \lambda)$, where λ is the Lebesgue measure. Propose a way to construct the sequence $(Y_n, n \geq 1)$ and the limiting variable Y on this space that satisfy the claim of the theorem. Do Y_n and Y have the correct marginal distributions in your construction?

Solution: Let U be the random variable defined by $U(\omega) = \omega$. This is a random variable with a uniform distribution on $(0, 1)$. Define $Y_n = F_n^{\leftarrow}(U)$ and $Y = F^{\leftarrow}(U)$. From part (a), the random variable $Y_n = F_n^{\leftarrow}(U)$ has the CDF F_n. This is the same distribution as the original random variable X_n. Similarly, the random variable $Y = F^{\leftarrow}(U)$ has the CDF F, which is the distribution of X.

c) Using the result from part (a), show that $Y_n \to Y$ almost surely. Conclude the proof of the theorem.

Solution: By part (a), we showed that for any fixed $u \in (0,1)$, the sequence of real numbers $F_n^{\leftarrow}(u)$ converges to $F^{\leftarrow}(u)$. In our probability space, this means that for any $\omega \in (0,1)$, the sequence of real numbers $Y_n(\omega) = F_n^{\leftarrow}(\omega)$ converges to $Y(\omega) = F^{\leftarrow}(\omega)$. A sequence of random variables that converges for every outcome ω in the sample space is said to converge almost surely (in this case, it converges everywhere, which is stronger). Thus, $Y_n \to Y$ almost surely.

We have successfully constructed a sequence of random variables (Y_n) and a random variable Y on a single probability space $((0,1), \mathcal{B}(0,1), \lambda)$ such that:

1. For each n, Y_n has the same distribution as X_n.
2. Y has the same distribution as X.
3. $Y_n \to Y$ almost surely as $n \to \infty$.

This completes the proof of Skorokhod's representation theorem for the case where the limiting distribution function F is continuous.

Application

d) Let $(X_n, n \geq 1)$ with $X_n \sim \mathcal{N}(0, 1 + 1/n)$ and let $X \sim \mathcal{N}(0,1)$. First verify directly (e.g. via point-wise convergence of characteristic functions) that $X_n \xrightarrow{d} X$.

Solution: Their characteristic functions are

$$\varphi_{X_n}(t) = \exp\Big(-\tfrac{1+1/n}{2}t^2\Big) \xrightarrow[n\to\infty]{} \exp\Big(-\tfrac{t^2}{2}\Big) = \varphi_X(t), \quad t \in \mathbb{R},$$

so $X_n \xrightarrow{d} X$.

e) Give an explicit Skorokhod representation $(Y_n, n \geq 1)$ and Y in this Gaussian case. Show that $Y_n \stackrel{d}{=} X_n$, $Y \stackrel{d}{=} X$, and $Y_n \to Y$ almost surely.

Solution: On $((0,1), \mathcal{B}(0,1), \lambda)$ let $U(\omega) = \omega$. Denote the standard normal CDF with Φ. We set $Z = \Phi^{-1}(U)$; then $Z \sim \mathcal{N}(0,1)$. Define

$$Y_n = \sqrt{1 + \tfrac{1}{n}}\, Z = \sqrt{1 + \tfrac{1}{n}}\, \Phi^{-1}(U), \qquad Y = Z = \Phi^{-1}(U).$$

Since scaling a standard normal by $\sqrt{1 + 1/n}$ produces $\mathcal{N}(0, 1+1/n)$, we have $Y_n \stackrel{d}{=} X_n$ and $Y \stackrel{d}{=} X$. For each ω, $Z(\omega)$ is a fixed real number and $\sqrt{1 + 1/n} \to 1$, hence $Y_n(\omega) \to Y(\omega)$; thus $Y_n \to Y$ almost surely. This gives an explicit Skorokhod representation in the Gaussian case: the original weak convergence is realized by a.s. convergence of (Y_n).