Assignment 2

Probability Theory II
(EN.553.721, Spring 2026)

Assigned: February 16, 2026 Due: 11:59pm EST, March 2, 2026

Submit solutions in LaTeX. Write in complete sentences. Include and justify all steps of your arguments, but avoid writing excessive explanation that is not contributing to your solution.

See my website for policies about late submissions, collaboration, and use of AI assistants.

Problem 1 (Conditional expectation)**.** This problem will derive some further properties of the conditional expectation and parallels with the ordinary expectation. We always assume $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{F}$ are σ-algebras.

1. Let $X \in L^2$ be an $\mathcal{F}$-measurable random variable. Write

$$\mathrm{Var}[X \mid \mathcal{G}] := \mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{G}])^2 \mid \mathcal{G}].$$

   Note that $\mathrm{Var}[X] = \mathrm{Var}[X \mid \{\emptyset, \Omega\}]$ is just the usual variance. Show that

$$\mathrm{Var}[X] = \mathbb{E}[\mathrm{Var}[X \mid \mathcal{G}]] + \mathrm{Var}[\mathbb{E}[X \mid \mathcal{G}]].$$

   Make sure you are clear on why both terms on the right-hand side are well-defined scalars.

2. For any σ-algebras $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$, show that

$$\mathbb{E}[\mathrm{Var}[X \mid \mathcal{G}_1]] \geq \mathbb{E}[\mathrm{Var}[X \mid \mathcal{G}_2]].$$

3. Show that, if $X \in L^2$ is an $\mathcal{F}$-measurable random variable and $Y \in L^2$ is any $\mathcal{G}$-measurable random variable, then

$$\mathbb{E}(X - Y)^2 \geq \mathbb{E}(X - \mathbb{E}[X \mid \mathcal{G}])^2.$$

   Thus $Y^* = \mathbb{E}[X \mid \mathcal{G}]$ is a minimizer of the left-hand side among $\mathcal{G}$-measurable random variables Y, giving a precise sense in which the conditional expectation is an "optimal estimator" of a random variable under a measurability constraint.

   (**Hint:** Add and subtract $\mathbb{E}[X \mid \mathcal{G}]$ inside the parentheses on the left-hand side.)

4. Let $X \in L^1$ be an $\mathcal{F}$-measurable random variable with $X \geq 0$ almost surely and Y be a $\mathcal{G}$-measurable random variable with $Y > 0$ almost surely. Prove that

$$\mathbb{P}[X > Y \mid \mathcal{G}] \leq \frac{\mathbb{E}[X \mid \mathcal{G}]}{Y} \text{ almost surely.}$$

Here the left-hand side is defined to equal $\mathbb{E}[\mathbb{1}\{X > Y\} \mid \mathcal{G}]$.

Problem 2 (Martingales)**.** This problem clarifies some of the definitions and behaviors of martingales.

1. Let $(M_n)_{n\geq 0}$ be a martingale with $M_n \in L^2$ for each $n \geq 0$. Prove that M_n has *uncorrelated increments*: for any $0 \leq a < b \leq c < d$,

$$\mathbb{E}[(M_d - M_c)(M_b - M_a)] = 0.$$

2. Let $(S_n)_{n\geq 0}$ be the simple random walk on the integers $\mathbb{Z}$, i.e., $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i$ for $X_i \sim \mathrm{Unif}(\{\pm 1\})$ drawn i.i.d. Let $f : \mathbb{Z} \to \mathbb{R}$ satisfy, for all $k \in \mathbb{Z}$,

$$f(k) \geq \frac{1}{2}(f(k+1) + f(k-1)).$$

Show that $(f(S_n))_{n\geq 0}$ is a supermartingale, a submartingale if the inequality above is reversed, and a martingale if the inequality is an equality. Describe the f for which the inequality is an equality. Note that the directions of these inequalities match what you would expect from the terms "sub-" and "super-" here (supermartingales are produced by functions that are above their averages, and so forth).

3. Let $(Z_n)_{n\geq 0}$ be adapted to a filtration $(\mathcal{F}_n)_{n\geq 0}$ and have $Z_n \in L^1$. Show that there exist $(M_n)_{n\geq 0}$ and $(H_n)_{n\geq 0}$ adapted to the same filtration such that M_n is a martingale (with respect to that filtration), H_n is predictable (with respect to that filtration), and $Z_n = M_n + H_n$ almost surely for all $n \geq 0$. Further show that if Z_n is a submartingale, then we may take H_n to be almost surely non-decreasing (i.e., for any $i \geq 0$, $H_{i+1} - H_i \geq 0$ almost surely).

4. Construct a martingale $(M_n)_{n\geq 0}$ such that $M_n \to +\infty$ almost surely.

(**Hint:** Consider a sum of independent but not i.i.d. random variables.)

Problem 3 (Distributional distances)**.** This problem will show you a few different useful ways of measuring the distance between probability measures. In all cases, let $(\Omega, \mathcal{F})$ be a measurable space and μ, ν and μ_m, ν_n for $m, n \geq 1$ probability measures on it.

1. Let

$$d_1(\mu, \nu) := \sup_{A\in\mathcal{F}} |\mu(A) - \nu(A)|.$$

Show that d_1 is a metric on the set of probability measures on the measurable space above. Show also that if $d_1(\mu_n, \mu) \to 0$ as $n \to \infty$ then $\mu_n \to \mu$ weakly, but that the converse is not always true. You may look up and use a form of the Portmanteau theorem for general probability spaces.

2. Suppose that $\nu \ll \mu$, so that the Radon-Nikodym derivative $\frac{d\nu}{d\mu} \geq 0$ is defined. Show that
$$d_1(\mu, \nu) \leq \frac{1}{2} \int \left| \frac{d\nu}{d\mu} - 1 \right| d\mu.$$

3. Again supposing $\nu \ll \mu$, let
$$d_2(\mu, \nu) := \sqrt{\int \left(\sqrt{\frac{d\nu}{d\mu}} - 1 \right)^2 d\mu}.$$
   Show that, for all $\nu \ll \mu$,
$$\frac{1}{2} d_2(\mu, \nu)^2 \leq d_1(\mu, \nu) \leq d_2(\mu, \nu).$$

4. Show the following inequalities:
$$d_1(\mu_1 \otimes \mu_2, \nu_1 \otimes \nu_2) \leq d_1(\mu_1, \nu_1) + d_1(\mu_2, \nu_2),$$
$$d_2(\mu_1 \otimes \mu_2, \nu_1 \otimes \nu_2)^2 \leq d_2(\mu_1, \nu_1)^2 + d_2(\mu_2, \nu_2)^2.$$
   (**Hint:** Show the second inequality first.)

Problem 4 (Gaussian vectors)**.** A *Gaussian random vector* is, for $\mathbf{\Sigma} \in \mathbb{R}^{d \times d}$ a positive definite symmetric matrix and $\boldsymbol{\mu} \in \mathbb{R}^d$ a vector, the random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ whose density over $\mathbb{R}^d$ is given by
$$\rho(\boldsymbol{x}) = \frac{1}{\sqrt{\det(2\pi\mathbf{\Sigma})}} \exp\left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right).$$
There is also a sensible definition when $\mathbf{\Sigma}$ is only positive semidefinite and therefore not invertible, but we will omit that case here. You may assume all covariance matrices encountered below are invertible. You should already know and can use without proof the following properties of Gaussian random vectors:

- In the above setting, $\mathbb{E}\boldsymbol{X} = \boldsymbol{\mu}$ and $\mathrm{Cov}(\boldsymbol{X}) := \mathbb{E}(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^\top = \mathbf{\Sigma}$.
- If $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ and $\boldsymbol{X}' \sim \mathcal{N}(\boldsymbol{\mu}', \mathbf{\Sigma}')$ are independent Gaussian random vectors, then $\mathsf{Law}(\boldsymbol{X} + \boldsymbol{X}') = \mathcal{N}(\boldsymbol{\mu} + \boldsymbol{\mu}', \mathbf{\Sigma} + \mathbf{\Sigma}')$.
- If $\boldsymbol{X}$ is as above and $\boldsymbol{M} \in \mathbb{R}^{d' \times d}$, then $\mathsf{Law}(\boldsymbol{M}\boldsymbol{X}) = \mathcal{N}(\boldsymbol{M}\boldsymbol{\mu}, \boldsymbol{M}\mathbf{\Sigma}\boldsymbol{M}^\top)$.

Using these results, show the following. Suppose $(X_1, \dots, X_a, Y_1, \dots, Y_b) \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$.

1. Suppose further that $\mathrm{Cov}(X_i, Y_j) = 0$ for all $i \in [a], j \in [b]$. Show that $(X_1, \dots, X_a)$ and $(Y_1, \dots, Y_b)$ are independent.

2. Let $\boldsymbol{X} = (X_1, \ldots, X_a)$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_b)$. Suppose that, according to this same partition, $\boldsymbol{\Sigma}$ has a block structure

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B}^\top \\ \boldsymbol{B} & \boldsymbol{C} \end{bmatrix}.$$

   Describe, in terms only of $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, a matrix $\boldsymbol{R}$ such that $\boldsymbol{Z} := \boldsymbol{X} - \boldsymbol{R}\boldsymbol{Y}$ is independent of $\boldsymbol{Y}$. This $\boldsymbol{Z}$ is a Gaussian random vector; compute its mean and covariance. Deduce an expression for $\mathbb{E}[\boldsymbol{X} \mid \boldsymbol{Y}]$ in terms of only $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ (constant matrices), $\boldsymbol{\mu}$ (a constant vector), and $\boldsymbol{Y}$ (a random vector) that holds almost surely.

3. Consider the two-dimensional case $a = b = 1$, where $(X, Y) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} \in \mathbb{R}^2$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{2\times 2}$ symmetric and positive definite. Show that the conditional variance $\mathrm{Var}[X \mid Y] = \mathbb{E}[(X - \mathbb{E}[X \mid Y])^2 \mid Y]$ is a constant (almost surely), not depending on Y. Give a formula for this constant in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.