Homework 7
Probability: a Graduate Course

1. Conditional Jensen's inequality

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, $\mathcal{F} \subset \Sigma$ be a sigma-algebra, $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function, and X be a random variable satisfying $\mathbb{E}|X| < \infty$ and $\mathbb{E}|\varphi(X)| < \infty$. Prove that

$$\varphi\left(\mathbb{E}[X|\mathcal{F}]\right) \le \mathbb{E}\left[\varphi(X)|\mathcal{F}\right].$$

2. Conditional expectation is a contraction

Check that conditional expectation is a contraction in L^p. Specifically, let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, let $\mathcal{F} \subset \Sigma$ be a sigma-algebra, and let $p \in (1, \infty)$. For a random variable $X \in L^p = L^p(\Omega, \Sigma, \mathbb{P})$, denote $f(X) = \mathbb{E}[X|\mathcal{F}]$. Check that

$$\|f(X) - f(Y)\|_{L^p} \le \|X - Y\|_{L^p} \quad \text{for any } X, Y \in L^p.$$

3. Conditional Cauchy-Schwarz

Show that

$$(\mathbb{E}[XY|\mathcal{F}])^2 \le \mathbb{E}[X^2|\mathcal{F}] \cdot \mathbb{E}[Y^2|\mathcal{F}]$$

almost surely.

4. Conditioning and second moment

Let $Y = \mathbb{E}[X|\mathcal{F}]$. Show that if $\mathbb{E}[Y^2] = \mathbb{E}[X^2]$ then $X = Y$ a.s.

5. Properties of conditional expectation

Is each statement below true or false? Prove or give a counterexample.

(a) $\mathbb{E}\, X = 0$ implies $\mathbb{E}[X|\mathcal{F}] = 0$ a.s.
(b) $\mathbb{E}[X|\mathcal{F}] = 0$ a.s. implies $\mathbb{E}\, X = 0$.
(c) $\mathbb{E}[X|Y + Z] = \mathbb{E}[X|Y] + E[X|Z]$.

6. Geometry of conditional expectation

Fix a probability space $(\Omega, \Sigma, \mathbb{P})$. Denote by L^2 the space of all random variables with finite variance on $(\Omega, \Sigma, \mathbb{P})$, equipped with the inner product $\langle X, Y \rangle = \mathbb{E}\, XY$. Consider a random variable $Y \in L^2$. Let $H \subset L^2$ consist of the random variables that are independent of Y:

$$H := \left\{ X \in L^2 : \ X \perp\!\!\!\perp Y \right\}.$$

Let $L \subset L^2$ consist of all variables of the form $h(Y)$ where $h : \mathbb{R} \to \mathbb{R}$ is a measurable function:

$$L := \left\{ h(Y) \in L^2 : \; h : \mathbb{R} \to \mathbb{R} \text{ is measurable} \right\}.$$

Is each statement below true or false? Prove or give a counterexample.

(a) H is a closed linear subspace of L^2.
(b) L is a closed linear subspace of L^2.
(c) H is orthogonal to L (meaning, $\langle h, \ell \rangle = 0$ for any $h \in H$ and $\ell \in L$).
(d) H is the orthogonal complement of L.

7. Normal conditioning property

Extend the result from Lecture 28. Prove that if random variables $X, Y_1, \dots, Y_n$ are jointly normal, then

$$\mathbb{E}[X | Y_1, \dots, Y_n] = a_1 Y_1 + \cdots + a_n Y_n + b$$

for some numbers $a_1, \dots, a_n, b \in \mathbb{R}$.

8. Filtering

Extend the result from Lecture 28. Let $X \sim N(0,1)$ be an unknown signal. Suppose we observe n versions of the signal corrupted by noise:

$$Y_i = X + W_i, \quad i = 1, \dots, n,$$

where $W_i \sim N(0, \sigma_i^2)$ are all jointly independent of X and of each other. Find the best estimate $\widehat{X}$ of the signal X as a function of observations Y_i and the variances of the noise σ_i – one that minimizes the mean squared error.

Hint: For $n = 1$, we showed in class that $\widehat{X} = Y/(1 + \sigma^2)$. Here, a similar estimator should work with Y replaced by the arithmetic mean of Y_i, and σ replaced by the arithmetic mean σ_i.

9. Conditional variance

Study the law of total variance (see notes for Lecture 29). Let X and Y be random variables with finite variance. Is each statement below true or false? Prove or give a counterexample.

(a) $\text{Var}(X) \geq \text{Var}(X|Y)$ a.s.
(b) $\text{Var}(X) \geq \mathbb{E}\left[\text{Var}(X|Y) \right]$.

10. The law of total covariance

Extend the law of total variance (see lecture notes for Lecture 28) to random vectors: formulate and prove the corresponding law of total covariance.