Exam Introduction to Probability Theory

Friday November 7 2025 — Duration: 2 hours —Documents, computers and calculators are **NOT** allowed

Questions

Exercise 1. (5 points)

Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbb{E}[X_i] = \mu$, and suppose $\mathbb{E}[|X_i|^{2+\delta}] < \infty$ for some $\delta > 0$. Prove that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu$ in probability.

Exercise 2. (15 points)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(A_n)_{n\geq 1}$ be a sequence of events.

(a) Define

$$\limsup_{n\to\infty} A_n \quad \text{and} \quad \liminf_{n\to\infty} A_n,$$

and give an intuitive interpretation of these sets.

(b) Show that

$$\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n.$$

(c) Prove that if $(A_n)_{n\geq 1}$ is a sequence of independent events

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n^c) < \infty \text{ implies } \mathbb{P}(\limsup_{n\to\infty} A_n) = 1$$

Prove that the same result holds even without the assumption of independence of the A_n

(d) Let $A_n = \{|X_n| > n\}$, where (X_n) are i.i.d. random variables with finite variance. Show that $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$.

Exercise 3. 10 points

Let $X_n = \mathbf{1}_{\{U\leq 1/n\}}$, where $U \sim \text{Uniform}(0,1)$. Determine whether $X_n \to 0$ (the zero random variable):

(i) almost surely,

(ii) in probability,

(iii) in L^1.

Exercise 3, 20 points

Let (X, Y) be a pair of continuous random variables with joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} 2, & 0 < y < x < 1, \\ 0, & \text{otherwise}. \end{cases}$$

(a) Compute the marginal densities $f_X(x)$ and $f_Y(y)$.

(b) Determine the conditional density $f_{Y|X}(y \mid x)$.

(c) Compute the conditional expectation $\mathbb{E}[Y \mid X]$.

(d) Compute $\mathbb{E}[\mathbb{E}[Y \mid X]]$ and compare it with $\mathbb{E}[Y]$.

(e) Compute $\mathrm{Var}(Y)$ and $\mathrm{Var}(\mathbb{E}[Y \mid X])$, and verify the law of total variance:

$$\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(\mathbb{E}[Y \mid X]).$$

Exercise 4. (20 points)

Let $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], \mathcal{B}([0,1]), \text{Lebesgue measure})$.

We define a sequence of random variables $(X_n)_{n\geq 1}$ in *stages* as follows.

- Stage 1: define one function

$$X_1(\omega) = \mathbf{1}_{[0,1]}(\omega).$$

- Stage 2: define two functions

$$X_2(\omega) = \mathbf{1}_{[0,\frac{1}{2}]}(\omega), \qquad X_3(\omega) = \mathbf{1}_{[\frac{1}{2},1]}(\omega).$$

- Stage 3: define three functions

$$X_4(\omega) = \mathbf{1}_{[0,\frac{1}{3}]}(\omega), \quad X_5(\omega) = \mathbf{1}_{[\frac{1}{3},\frac{2}{3}]}(\omega), \quad X_6(\omega) = \mathbf{1}_{[\frac{2}{3},1]}(\omega).$$

- Stage 4: define four functions – split $[0,1]$ into four equal subintervals and let X_7, X_8, X_9, X_{10} be their indicator functions.

- And so on: at Stage k, split $[0,1]$ into k equal subintervals, and define k new random variables, each one being the indicator of one of those subintervals.

The full sequence (X_n) is obtained by concatenating Stage 1, then Stage 2, then Stage 3, etc.

(a) Show that for every X_n that appears in Stage k, we have

$$\mathbb{E}[|X_n|] = \int_0^1 X_n(\omega)\, d\omega = \frac{1}{k}.$$

Deduce that $X_n \to 0$ in L^1.

(b) Prove that X_n does not converge almost surely (Hint: use the fact that for any $\omega \in [0,1]$ and for each Stage k, there is *at least one* random variable from that stage, call it $X_{n(k)}$, such that $X_{n(k)}(\omega) = 1$.)

Exercise 5. (15 points)

Let $S_n = X_1 + \cdots + X_n$ where X_i takes values $+1$ with probability p and -1 with probability $1-p$, independently.

(a) Compute $\mathbb{E}[S_n]$ and $\mathrm{Var}(S_n)$.

(b) Provide an approximation for $\mathbb{P}(S_n > 0)$ for large n (the final answer shall be expressed in terms of the standard normal cumulative distribution function)

Exercise 6. (15 points)

Consider a Galton–Watson branching process with offspring distribution Z being a Poisson random variable of parameter λ, $\lambda \in \mathbb{R}^+$.

- Write a closed equation for the extinction probability, p
- Suppose $\lambda = 1 + \varepsilon$ with $\varepsilon > 0$ small. Obtain the first-order expansion of p in terms of ε

Solutions

1. Weak Law under Moment Assumptions

Let $Y_i = X_i - \mu$. Then $\mathbb{E}[Y_i] = 0$ and by assumption $\mathbb{E}[|Y_i|^{2+\delta}] < \infty$, which in particular implies $\mathrm{Var}(Y_i) = \sigma^2 < \infty$. For any $\varepsilon > 0$,

$$\mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) = \mathbb{P}\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} Y_i\Big| > \varepsilon\Big) \leq \frac{\mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n} Y_i)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0.$$

Thus $\bar{X}_n \to \mu$ in probability.

2 Lim sup and inf

(a) **Definition and interpretation.**

$$\limsup_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n\geq k} A_n,$$
$$\liminf_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} \bigcap_{n\geq k} A_n.$$

Intuitively:

- $\limsup A_n$ = set of outcomes that occur in infinitely many A_n's.
- $\liminf A_n$ = set of outcomes that occur in all but finitely many A_n's (i.e. eventually always true).

(b) Since $\bigcap_{n\ge k} A_n \subseteq \bigcup_{n\ge k} A_n$ for each k, taking unions and intersections yields

$$\liminf_n A_n = \bigcup_k \bigcap_{n\ge k} A_n \subseteq \bigcap_k \bigcup_{n\ge k} A_n = \limsup_n A_n.$$

(c) **Borel–Cantelli application.** If $\sum_n \mathbb{P}(A_n^c) < \infty$, then by the first Borel–Cantelli lemma, only finitely many A_n^c occur. Hence all but finitely many A_n occur, i.e. $\omega \in \liminf A_n \subseteq \limsup A_n$ and $\mathbb{P}(\limsup A_n) = 1$.

Conversely, if the A_n are independent and $\sum_n \mathbb{P}(A_n^c) = \infty$, then by the *second* Borel–Cantelli lemma, $\mathbb{P}(\limsup A_n^c) = 1$. Thus infinitely many A_n^c occur, so not all A_n occur infinitely often, giving $\mathbb{P}(\limsup A_n) < 1$.

(d) Let $A_n = \{|X_n| > n\}$ where X_n are i.i.d. with $\mathbb{E}[X_1^2] < \infty$. Then by Markov's inequality,

$$\mathbb{P}(A_n) = \mathbb{P}(|X_1| > n) \le \frac{\mathbb{E}[X_1^2]}{n^2}.$$

Hence $\sum_n \mathbb{P}(A_n) < \infty$. By the first Borel–Cantelli lemma, $\mathbb{P}(\limsup A_n) = 0$. That is, only finitely many X_n exceed n in magnitude almost surely.

Exercise 3

Here $U \sim \text{Uniform}(0,1)$ is a single random variable.

(i) *Almost surely*: For any fixed outcome ω with $U(\omega) = u \in (0,1]$, for all sufficiently large n we have $u > 1/n$, hence $X_n(\omega) = 0$ eventually. Thus $X_n(\omega) \to 0$. (If $u = 0$, which has probability 0, then $X_n(\omega) = 1$ for all n.) Therefore $X_n \to 0$ almost surely.

(ii) *In probability*: Since $\mathbb{P}(X_n = 1) = \mathbb{P}(U \le 1/n) = 1/n \to 0$, we have $\mathbb{P}(|X_n - 0| > \varepsilon) \to 0$ for any $\varepsilon \in (0,1]$. Hence $X_n \to 0$ in probability.

(iii) *In L^1*: We have

$$\mathbb{E}[|X_n - 0|] = \mathbb{E}[X_n] = \mathbb{P}(U \le 1/n) = \frac{1}{n} \to 0,$$

so $X_n \to 0$ in L^1.

Exercise 4

(a) L^1 **convergence.**

In Stage k, the interval $[0,1]$ is divided into k disjoint subintervals of equal length $1/k$. Each X_n in that stage is the indicator of one of these subintervals. Therefore

$$\mathbb{E}[|X_n|] = \int_0^1 X_n(\omega)\,d\omega = (\text{length of that subinterval}) = \frac{1}{k}.$$

As we go through stages $k = 1, 2, 3, \ldots$, $k \to \infty$ and so $1/k \to 0$. Hence $\mathbb{E}[|X_n|] \to 0$ as $n \to \infty$, which means

$$\|X_n - 0\|_{L^1} = \mathbb{E}[|X_n|] \longrightarrow 0.$$

Therefore $X_n \to 0$ in L^1.

(b) **Failure of pointwise convergence.**

Fix $\omega \in [0,1]$. At Stage k, we split $[0,1]$ into k equal subintervals

$$\left[0, \tfrac{1}{k}\right], \left[\tfrac{1}{k}, \tfrac{2}{k}\right], \ldots, \left[\tfrac{k-1}{k}, 1\right].$$

These subintervals cover $[0,1]$ and are pairwise disjoint except for endpoints. Thus ω belongs to *exactly one* of them (up to a set of measure zero coming from the endpoints). Call that subinterval the j-th one. Then there is a corresponding random variable X_n in Stage k whose indicator is precisely that j-th subinterval. For that n, we have $X_n(\omega) = 1$.

Since this happens *for every* Stage $k = 1, 2, 3, \ldots$, we can find infinitely many indices n with $X_n(\omega) = 1$. On the other hand, for most other n, $X_n(\omega) = 0$. Hence the numeric sequence $X_n(\omega)$ keeps jumping between 0 and 1 infinitely many times and does *not* converge.

The only possible exceptions are ω at subinterval boundaries (like $\omega = 1/2, 1/3, 2/3, \ldots$), which form a countable set and therefore have Lebesgue measure 0. Thus, for almost every ω,

$$\lim_{n\to\infty} X_n(\omega) \text{ does not exist.}$$

In particular, $X_n(\omega)$ does not converge to 0. Therefore $X_n \not\to 0$ almost surely.

Exercise 5.

Let $X_i \in \{+1, -1\}$ with $\mathbb{P}(X_i = +1) = p$ and $\mathbb{P}(X_i = -1) = 1 - p$.

(a) $\mathbb{E}[X_i] = (+1)p + (-1)(1-p) = 2p - 1$. Hence

$$\mathbb{E}[S_n] = n(2p-1).$$

Also $\mathrm{Var}(X_i) = 1 - (2p-1)^2 = 4p(1-p)$, so

$$\mathrm{Var}(S_n) = 4np(1-p).$$

(b) By the CLT,

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \Rightarrow N(0,1).$$

Thus for large n,

$$\begin{aligned}\mathbb{P}(S_n > 0) &= \mathbb{P}\left(\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}} > \frac{-\mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}}\right) \\ &\approx 1 - \Phi\left(\frac{-n(2p-1)}{\sqrt{4np(1-p)}}\right) = \Phi\left(\frac{(2p-1)\sqrt{n}}{2\sqrt{p(1-p)}}\right).\end{aligned}$$

Exercise 6. Branching Process Near Criticality

Let $Z \sim \text{Poisson}(\lambda)$ so $G(s) = \exp(\lambda(s-1))$. Let q be the extinction probability.

(a) By Galton–Watson theory, q is the smallest fixed point of G: $q = G(q)$. Thus

$$q = \exp(\lambda(q-1)).$$

The trivial solution $q = 1$ always exists. For $\lambda > 1$ there is another solution $q < 1$.

(b) Write $\lambda = 1 + \varepsilon$ with $\varepsilon > 0$ small and set $q = 1 - \delta$ with δ small. Then

$$1 - \delta = \exp((1+\varepsilon)((1-\delta) - 1)) = \exp(-(1+\varepsilon)\delta).$$

Taking log and expanding $\log(1-\delta) = -\delta - \delta^2/2 + o(\delta^2)$ gives

$$-\delta - \frac{\delta^2}{2} + o(\delta^2) = -(1+\varepsilon)\delta.$$

Cancel $-\delta$ and keep leading terms:

$$-\frac{\delta^2}{2} \approx -\varepsilon\delta \quad \Longrightarrow \quad \delta \approx 2\varepsilon.$$

Hence

$$1 - q = \delta \sim 2(\lambda - 1) \qquad \text{as } \lambda \downarrow 1^+.$$

(c) Interpretation: For $\lambda \leq 1$, $q = 1$ (extinction a.s.). For $\lambda > 1$, survival has positive probability $1 - q \approx 2(\lambda - 1)$ when λ is just above 1. Thus $\lambda = 1$ is the *critical* threshold separating certain extinction from possible indefinite growth (supercritical phase).

conditional exp

We are given (X, Y) with joint density

$$f_{X,Y}(x,y) = \begin{cases} 2, & 0 < y < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) **Marginal densities f_X and f_Y.**

For f_X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_0^x 2\,dy = 2x, \quad 0 < x < 1,$$

and $f_X(x) = 0$ otherwise.

For f_Y:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_y^1 2\,dx = 2(1-y), \quad 0 < y < 1,$$

and $f_Y(y) = 0$ otherwise.

(b) **Conditional density $f_{Y|X}(y \mid x)$.**

By definition,

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

For $0 < y < x < 1$, we have $f_{X,Y}(x, y) = 2$ and $f_X(x) = 2x$, hence

$$f_{Y|X}(y \mid x) = \frac{2}{2x} = \frac{1}{x}, \quad 0 < y < x < 1.$$

For other values of y, $f_{Y|X}(y \mid x) = 0$.

In particular, conditional on $X = x$, the variable Y is uniform on $(0, x)$.

(c) **Compute $\mathbb{E}[Y \mid X]$.**

For a fixed $X = x$, since $Y \mid X = x \sim \text{Unif}(0, x)$, its conditional expectation is

$$\mathbb{E}[Y \mid X = x] = \int_0^x y \cdot \frac{1}{x}\, dy = \frac{1}{x} \cdot \frac{x^2}{2} = \frac{x}{2}.$$

Therefore,

$$\mathbb{E}[Y \mid X] = \frac{X}{2}.$$

(d) **Compare $\mathbb{E}[\mathbb{E}[Y \mid X]]$ and $\mathbb{E}[Y]$.**

First,

$$\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}\left[\frac{X}{2}\right] = \frac{1}{2}\mathbb{E}[X].$$

Using $f_X(x) = 2x$ on $(0, 1)$,

$$\mathbb{E}[X] = \int_0^1 x \cdot 2x\, dx = 2\int_0^1 x^2\, dx = 2 \cdot \frac{1}{3} = \frac{2}{3}.$$

Hence

$$\mathbb{E}[\mathbb{E}[Y \mid X]] = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

Now compute $\mathbb{E}[Y]$ directly from f_Y:

$$\mathbb{E}[Y] = \int_0^1 y \cdot 2(1 - y)\, dy = 2\int_0^1 (y - y^2)\, dy = 2\left(\frac{1}{2} - \frac{1}{3}\right) = 2 \cdot \frac{1}{6} = \frac{1}{3}.$$

Therefore,

$$\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[Y] = \frac{1}{3},$$

illustrating the tower property $\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[Y]$.

(e) *(Optional)* **Law of total variance.**

We verify

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

- *Step 1:* Compute $\text{Var}(Y)$.

  First find $\mathbb{E}[Y^2]$:

  $$\mathbb{E}[Y^2] = \int_0^1 y^2 \cdot 2(1 - y)\, dy = 2\int_0^1 (y^2 - y^3)\, dy = 2\left(\frac{1}{3} - \frac{1}{4}\right) = 2 \cdot \frac{1}{12} = \frac{1}{6}.$$

  We already have $\mathbb{E}[Y] = \frac{1}{3}$, so

  $$\text{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.$$

- *Step 2:* Compute $\mathrm{Var}(Y \mid X = x)$ and take expectation.
  Conditional on $X = x$, we have $Y \mid X = x \sim \mathrm{Unif}(0, x)$, so

$$\mathrm{Var}(Y \mid X = x) = \frac{x^2}{12}.$$

  Therefore

$$\mathbb{E}[\mathrm{Var}(Y \mid X)] = \mathbb{E}\left[\frac{X^2}{12}\right] = \frac{1}{12}\mathbb{E}[X^2].$$

  Compute $\mathbb{E}[X^2]$ using $f_X(x) = 2x$:

$$\mathbb{E}[X^2] = \int_0^1 x^2 \cdot 2x \, dx = 2\int_0^1 x^3 \, dx = 2 \cdot \frac{1}{4} = \frac{1}{2}.$$

  Thus

$$\mathbb{E}[\mathrm{Var}(Y \mid X)] = \frac{1}{12} \cdot \frac{1}{2} = \frac{1}{24}.$$

- *Step 3:* Compute $\mathrm{Var}(\mathbb{E}[Y \mid X])$.
  We found $\mathbb{E}[Y \mid X] = X/2$. Hence

$$\mathrm{Var}(\mathbb{E}[Y \mid X]) = \mathrm{Var}\left(\frac{X}{2}\right) = \frac{1}{4}\,\mathrm{Var}(X).$$

  Compute $\mathrm{Var}(X)$.
  We already have $\mathbb{E}[X] = \frac{2}{3}$. Also

$$\mathbb{E}[X^2] = \frac{1}{2}.$$

  Therefore

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9}{18} - \frac{8}{18} = \frac{1}{18}.$$

  Thus

$$\mathrm{Var}(\mathbb{E}[Y \mid X]) = \frac{1}{4} \cdot \frac{1}{18} = \frac{1}{72}.$$

- *Step 4:* Check the identity.

$$\mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(\mathbb{E}[Y \mid X]) = \frac{1}{24} + \frac{1}{72} = \frac{3}{72} + \frac{1}{72} = \frac{4}{72} = \frac{1}{18} = \mathrm{Var}(Y).$$

  The law of total variance is verified.