STA 131A: Introduction to Probability Theory

Lecture 11: Review for Midterm 1

Dogyoon Song

Spring 2026, UC Davis

Announcements

Midterm 1 is in class on Fri, Apr 24

- You may bring *one **hand-written** letter-sized (8.5 $\times$ 11 inches), double-sided sheet of paper* with formulas and brief notes
- **Calculator**: Simple (non-graphing) calculators only
- **No textbooks** or other materials beyond the single cheat sheet
- **SDC accommodations:** Confirm scheduling with AES online
- Practice midterms are posted on the course webpage

Midterm 1 review

What you should know cold:

- Translate a word problem into a sample space, events, and random variables.
- Identify the right tool: conditioning, Bayes, independence, counting, PMFs, expectation, or variance.
- State the relevant formula with its assumptions before using it.
- Execute calculations cleanly and justify the key modeling step.

Today's review focuses on:

- Core definitions and formulas.
- Typical problem types from lecture, homework, and the practice midterm.
- Common pitfalls that cause otherwise-correct solutions to fail.

A problem-solving recipe

Before computing, ask yourself to answer:

1. **What is random?** Identify the experiment, outcomes, events, and random variables.
2. **What is given?** Probabilities, conditional probabilities, PMFs, independence, or equal likelihood?
3. **What is being asked?** Event probability, PMF, expectation, variance, or conditional quantity?
4. **Which tool fits?** Counting, conditioning, Bayes, independence, total expectation, or linearity of expectation?
5. **What assumption justifies the step?** Equal likelihood? Independence? Disjointness? Partition?

Good exam solutions are not just computations: They should clearly identify the event/model/tool being used.

Events and basic probability identities

Set theory:

- Events are sets: use unions, intersections, complements, and partitions carefully.
- Know these identities:

$$P(A^c) = 1 - P(A), \qquad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c$$

Probability axioms:

- Nonnegativity, additivity, normalization
- Probability is defined for events

Typical task: Rewrite a complicated event into simpler or disjoint pieces before computing probabilities; use complements or notice inclusion/inequality relations

Conditional probability toolkit

- Definition:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (P(B) > 0)$$

- Multiplicative rule:

$$P(A \cap B) = P(B)\, P(A \mid B) = P(A)\, P(B \mid A)$$

- Law of total probability: if $\{A_i\}$ is a partition, then

$$P(B) = \sum_i P(B \mid A_i)\, P(A_i)$$

- Bayes' rule: if $\{A_i\}$ is a partition and $P(B) > 0$, then

$$P(A_j \mid B) = \frac{P(B \mid A_j)\, P(A_j)}{\sum_i P(B \mid A_i)\, P(A_i)}$$

Typical task: Identify the observed event, condition on a useful partition, and weight conditional probabilities by the partition probabilities.

Independence of events

- Definition:
$$P(A \cap B) = P(A)\,P(B)$$
- If $P(B) > 0$, this is equivalent to
$$P(A \mid B) = P(A)$$
- Independence is different from disjointness: if $A \cap B = \emptyset$ and $P(A), P(B) > 0$, then A and B are *not* independent.
- For several events: pairwise independence does *not* imply mutual independence.
- Useful caution: if $A \subseteq B$, $0 < P(A)$, and $P(B) < 1$, then A and B cannot be independent.

Typical task: Test independence by comparing $P(A \cap B)$ with $P(A)P(B)$, or compare $P(A \mid B)$ with $P(A)$ when $P(B) > 0$.

Counting toolkit

- Product rule: break the process into stages and multiply the number of choices.
- Permutations (order matters):

$$\frac{n!}{(n-k)!}$$

- Combinations (order does not matter):

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- Multinomial coefficients / labeled partitions:

$$\binom{n}{n_1, \ldots, n_r} = \frac{n!}{n_1! \cdots n_r!}$$

Typical task: Decide whether order matters, whether repetitions are allowed, whether groups are labeled, and whether it is easier to count the complement.

Discrete random variables, PMFs, and named models

- A random variable is a function from outcomes to numbers:

$$X : \Omega \to \mathbb{R}.$$

- A PMF is

$$p_X(x) = P(X = x),$$

  with $p_X(x) \geq 0$ and $\sum_x p_X(x) = 1$.
- For a function $Y = g(X)$,

$$p_Y(y) = \sum_{x:\, g(x)=y} p_X(x).$$

- To compute a PMF, identify the event $\{X = x\}$ and add the probabilities of all outcomes therein.

Typical task: Derive the PMF of X or $g(X)$ from an experiment, and check that the PMF sums to 1.

Expectation and variance

- Expectation:

$$\mathbb{E}[X] = \sum_x x\, p_X(x), \qquad \mathbb{E}[g(X)] = \sum_x g(x)\, p_X(x)$$

- Variance:

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2$$

- Expectation and variance of an affine function:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b, \qquad \mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$$

  - Expectation is linear even without independence.
  - Pitfall: in general, $\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$.
- Indicator trick: if $X = \sum_i I_i$, then $\mathbb{E}[X] = \sum_i \mathbb{E}[I_i] = \sum_i P(I_i = 1)$.

Typical task: Compute $\mathbb{E}[X]$, $\mathbb{E}[g(X)]$, and $\mathrm{Var}(X)$; use indicators or transformations when direct computation is easier.

Some named models

- **Bernoulli(p)**: one success/failure trial

$$P(X=1)=p, \qquad P(X=0)=1-p, \qquad \mathbb{E}[X]=p, \quad \mathrm{Var}(X)=p(1-p)$$

- **Binomial(n, p)**: number of successes in n independent Bernoulli trials

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}, \qquad \mathbb{E}[X]=np, \quad \mathrm{Var}(X)=np(1-p)$$

- **Geometric(p)**: waiting time until the first success

$$P(X=k)=(1-p)^{k-1}p, \quad k=1,2,\ldots, \qquad \mathbb{E}[X]=\frac{1}{p}, \quad \mathrm{Var}(X)=\frac{1-p}{p^2}$$

- **Poisson(λ)**: count in a fixed window

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}, \qquad \mathbb{E}[X]=\mathrm{Var}(X)=\lambda$$

Typical task: Recognize the model from the story, write down its PMF.

Joint, marginal, and conditional PMFs

- Joint PMF:
$$p_{X,Y}(x, y) = P(X = x, Y = y)$$
- Marginals:
$$p_X(x) = \sum_y p_{X,Y}(x, y), \qquad p_Y(y) = \sum_x p_{X,Y}(x, y)$$
- Conditional PMFs:
$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x, y)}{p_Y(y)} \qquad \text{if } p_Y(y) > 0$$
- Reconstruction:
$$p_{X,Y}(x, y) = p_Y(y)\, p_{X|Y}(x \mid y)$$

Typical task: Navigate between joint, marginal, and conditional PMFs; compute probabilities such as $P(X < Y)$, $P(X + Y = z)$, or $P(X = x \mid Y = y)$.

Conditional expectation and total expectation

- Conditional expectation:

$$\mathbb{E}[X \mid Y = y] = \sum_{x} x\, p_{X|Y}(x \mid y)$$

- Law of total expectation:

$$\mathbb{E}[X] = \sum_{y} \mathbb{E}[X \mid Y = y]\, p_Y(y).$$

  More generally, if $\{A_i\}$ is a partition,

$$\mathbb{E}[X] = \sum_{i} P(A_i)\, \mathbb{E}[X \mid A_i].$$

Typical task: Compute a hard mean by conditioning on an easier variable, event, case, or state of the world.

Independence of random variables and consequences

- Definition:
$$p_{X,Y}(x,y) = p_X(x)p_Y(y) \qquad \text{for all } x, y$$
- If X and Y are independent, then
$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$
and more generally
$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\,\mathbb{E}[h(Y)]$$
- If X and Y are independent, then
$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$
- Important cautions:
  - linearity of expectation does *not* require independence;
  - $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ alone does *not* imply independence.

Typical task: Check independence by testing joint-PMF factorization for all relevant (x, y), then use independence to factor expectations or add variances.

Final checklist / common pitfalls

- Define the sample space, events, and random variables before computing.
- Identify the tool:
  - Conditioning / Bayes,
  - Independence,
  - Counting,
  - PMFs / expectations / variance.
- Distinguish carefully:
  - Disjoint vs. independent,
  - Pairwise vs. mutual independence,
  - Joint vs. marginal vs. conditional PMFs.
- Remember:
  - Order matters in permutations, not in combinations;
  - $\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$ in general;
  - Linearity of expectation does *not* require independence;
  - Independence guarantees $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.
- Justify why a formula applies: partition, disjointness, equal likelihood, or independence.

Wrap-up: How to prepare

Use the practice midterms actively

- First attempt them without notes and under time pressure.
- Mark which steps failed: identifying the tool, setting up the model, or doing the calculation.
- Then review and prepare your cheat sheet around those weak points.

In preparation of cheat sheet, you may want to

- Include definitions and assumptions, not only formulas.

During the exam

- Write down the relevant event or random variable before using a formula.
- Keep notation clear: distinguish events, values, PMFs, and expectations.
- If stuck, condition on a useful case or try counting the complement.

Advice: Be fluent in translating words to probabilistic objects and choosing the right tool.