STA 130A: Mathematical Statistics: Brief Course

Lecture 19: Review for Midterm 2

Dogyoon Song

Fall 2026, UC Davis

Announcements

Midterm 2: Mon, Nov 9, 11:00–11:50 AM (in class)

- Arrive early: The exam ends at 11:50 AM sharp.
- One hand-written, letter-sized (8.5×11 inches), double-sided sheet of formulas/notes.
- A simple non-graphing scientific calculator is allowed; no other reference materials or electronic devices.
- **SDC accommodations:** Confirm scheduling with AES online.

Preparation and support

- Primary coverage: Lectures 12–19.
- Earlier probability tools remain prerequisites, although not to be actively asked.
- Check Canvas and the course webpage for practice materials.

After midterm 2

- No class on Wed, Nov 11 (Veterans day)
- The lecture on Fri, Nov 13 will also be held on Zoom; see Canvas for the link

Review for Midterm 2

What you should know cold:

- Translate the question into an event, distribution, moment, or bound.
- Choose the right tool and state what justifies using it.
- Set up the calculation with the correct support, weights, and parameters.
- Interpret the result and check that it answers the question.

Today's review focuses on:

- Joint densities, conditioning, and one-variable transformations.
- Covariance, correlation, and conditional/total variance.
- MGFs and independent sums; Markov and Chebyshev bounds.
- Typical tasks and common pitfalls—not new theory.

Tentative problem map

Tentative problem map and study priorities:

- **Problem 1: Concept checks**
  Apply core concepts through short calculations and brief justifications.
- **Problem 2: Continuous distributions**
  Work with marginal and conditional PDFs and transformations; identify supports.
- **Problem 3: Independent sums and correlation**
  Use MGFs for independent sums and analyze dependence between related variables.
- **Problem 4: Conditioning and probability bounds**
  Combine conditioning, covariance, total variance, and tail bounds in context.
- **Problem 5: Optional bonuses**
  Apply familiar ideas in less routine settings; complete the main problems first.

Note: This is a study guide, not a fixed blueprint; details and organization may change.

A problem-solving recipe and time management

Before computing, ask yourself:

1. **What is given?** A joint PDF, conditional model, MGF, or moments?
2. **What is asked?** A probability, distribution, conditional quantity, or bound?
3. **Which tool fits?** Integrate, condition, transform, differentiate an MGF, or apply an inequality?
4. **What justifies the step?** Support, normalization, independence, monotonicity, or nonnegativity?
5. **Does the answer make sense?** Check support, signs, units, normalization, and the direction of the bound.

Use the 50 minutes deliberately:

- Scan the exam first and allocate time in proportion to point values.
- If stuck, write the setup, move on, and return; reserve the last few minutes for checks.

Good solutions show the event, the method, and the key justification.

Review: Joint PDFs, support, and marginals

A joint PDF assigns probability to regions:

$$f_{X,Y}(x,y) \geq 0, \qquad \iint f_{X,Y}(x,y)\,dx\,dy = 1,$$

$$P((X,Y) \in A) = \iint_A f_{X,Y}(x,y)\,dx\,dy.$$

Marginals: Fix the coordinate retained; integrate out the other.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx.$$

- Intersect the event with the joint support before choosing limits.
- A constant joint density need not give constant marginal densities.
- Continuous marginals alone do not guarantee a joint PDF.

Typical task: Normalize a joint PDF, sketch its support, compute an event probability, and obtain its marginals.

Review: Joint cumulative distribution functions

Definition: Both coordinates lie below their respective cutoffs.

$$F_{X,Y}(x, y) = P(X \leq x,\ Y \leq y).$$

From a joint PDF:

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s, t)\, dt\, ds.$$

Recovering a density: Where the CDF is twice continuously differentiable,

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x, y).$$

- Use the intersection of the support with the lower-left cutoff region.
- Include the cases outside the support; the CDF need not be zero there.

Typical task: Derive a joint CDF piecewise from its support, or recover a PDF by differentiation.

Review: Expectations with a joint PDF

Average the value of interest over the joint distribution:

$$\mathbb{E}[g(X,Y)] = \iint g(x,y) f_{X,Y}(x,y)\, dx\, dy.$$

Common choices: $g(x,y) = x,\ y,\ xy,\ x^2,\ y^2$.

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

Linearity and affine rules:

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

- Linearity does not require independence.
- For an expectation, integrate *value* $\times$ *density* over the support; no derived PDF is needed.

Typical task: Compute $\mathbb{E}[XY]$, $\mathbb{E}[X - Y]$, or $\mathrm{Var}(X)$ directly from a joint model.

Review: Conditional PDFs and reconstruction

Conditioning on an event: For $A = \{X \in C\}$ with $P(A) > 0$,

$$f_{X|A}(x) = \frac{f_X(x)\mathbf{1}_{\{x \in C\}}}{P(A)}.$$

Conditioning on a value: Normalize the corresponding joint-density slice.

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad f_Y(y) > 0.$$

Reconstruction and total probability:

$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x \mid y), \qquad f_X(x) = \sum_i P(A_i) f_{X|A_i}(x),$$

where the A_i's form a partition with positive probabilities.

Check: State the conditional support and verify that the conditional density integrates to 1; do not divide by $P(Y = y) = 0$.

Typical task: Restrict and normalize a distribution, or build a joint/mixture model from conditional densities.

Review: Conditional and total expectation

Conditional mean: Average using the conditional density.

$$m(y) = \mathbb{E}[X \mid Y = y] = \int x f_{X|Y}(x \mid y)\, dx.$$

For fixed y, $m(y)$ is a number; $\mathbb{E}[X \mid Y] = m(Y)$ is a random variable.

Total expectation: Average the conditional means using the correct weights.

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = \int_{\{y: f_Y(y) > 0\}} m(y) f_Y(y)\, dy.$$

For discrete Y, replace the integral by $\sum_y m(y) p_Y(y)$; for a partition,

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

Caution: Conditional means are not averaged equally unless cases are equally likely.

Typical task: Find conditional means, then recover the unconditional mean by averaging.

Review: Independence of random variables

Joint-density factorization:

$$X, Y \text{ independent} \quad \Longleftrightarrow \quad f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \text{almost everywhere.}$$

Equivalently, $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ for all x, y, and conditioning leaves the marginal distribution unchanged.

Consequences:

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)], \qquad \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

- Check the support as well as the algebraic form of the density.
- The discrete version uses $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ for every pair.
- Independence implies zero covariance, but the converse can fail.

Typical task: Test factorization of a joint distribution, then justify any product-of-expectations or variance-addition step.

Review: Derived distributions—CDF first

For a known distribution of X and $Z = g(X)$:

1. **Support:** Identify the possible values of Z.
2. **CDF:** Rewrite the event using the original variable:

$$F_Z(z) = P(g(X) \leq z).$$

3. **PDF:** If Z has a PDF, differentiate where valid and set it to zero outside the transformed support.

Keep every relevant input: For $z \geq 0$,

$$\{X^2 \leq z\} = \{-\sqrt{z} \leq X \leq \sqrt{z}\}.$$

The CDF method works even when g is not one-to-one.

Typical task: Translate a transformed event correctly, give the full CDF, and obtain its density.

Review: Monotone and affine transformations

Smooth monotone shortcut: On an interval supporting X, suppose g is continuously differentiable and strictly monotone, with $g' \neq 0$ on the interior. For $h = g^{-1}$,

$$f_Z(z) = f_X(h(z))\,|h'(z)|$$

on the transformed interior, and zero outside the transformed support.

Affine special case: For $Z = aX + b$, $a \neq 0$,

$$f_Z(z) = \frac{1}{|a|} f_X\left(\frac{z-b}{a}\right).$$

- Map the support before substituting into the formula.
- Differentiate the *inverse* map and take the absolute value.
- Use CDF-first if the transformation is not one-to-one on the support.

Typical task: Find a transformed PDF and its support, including decreasing transformations.

Review: Covariance and variances of sums

Covariance:

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Useful rules:

$$\begin{aligned} \mathrm{Cov}(X, X) &= \mathrm{Var}(X), \\ \mathrm{Cov}(X, aY + b) &= a\mathrm{Cov}(X, Y), \\ \mathrm{Cov}(X, Y + Z) &= \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z). \end{aligned}$$

Variance of a linear combination:

$$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y).$$

- Independence gives zero covariance; uncorrelated variables need not be independent.
- Square the variance coefficients; keep the signs in the covariance term.

Typical task: Compute covariance from a joint model and find the variance of a sum, difference, or weighted combination.

Review: Correlation

Unit-free covariance: When both variances are positive,

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}, \qquad -1 \le \rho(X, Y) \le 1.$$

- The sign indicates the direction of linear association.
- $\rho = 0$ does not imply independence.
- $|\rho| = 1$ corresponds to an exact affine relationship with probability one.

Shifts and nonzero rescalings: For $a, c \neq 0$,

$$\rho(aX + b, cY + d) = \frac{ac}{|a||c|}\,\rho(X, Y).$$

Shifts and positive scales leave correlation unchanged; one sign reversal flips it.

Typical task: Calculate a correlation, interpret it, or determine how an affine change affects it.

Review: Conditional variance

Center at the conditional mean:

$$\begin{aligned}\mathrm{Var}(X \mid Y = y) &= \mathbb{E}\big[(X - \mathbb{E}[X \mid Y = y])^2 \mid Y = y\big] \\ &= \mathbb{E}[X^2 \mid Y = y] - \big(\mathbb{E}[X \mid Y = y]\big)^2.\end{aligned}$$

- For fixed y, this is a number: the spread in that conditional distribution.
- $\mathrm{Var}(X \mid Y)$ is a random variable whose value depends on Y.
- $\mathbb{E}[\mathrm{Var}(X \mid Y)]$ averages those spreads across conditioning cases.

Caution: Do not center conditional squared deviations at the unconditional mean $\mathbb{E}[X]$.

Typical task: Compute conditional first and second moments, obtain the conditional variance, and distinguish it from its average.

Review: Law of total variance

Two sources of variability:

$$\mathrm{Var}(X) = \underbrace{\mathbb{E}[\mathrm{Var}(X \mid Y)]}_{\text{average within-group variance}} + \underbrace{\mathrm{Var}(\mathbb{E}[X \mid Y])}_{\text{variance of group means}}.$$

Compute separately: First obtain $m(Y) = \mathbb{E}[X \mid Y]$ and $v(Y) = \mathrm{Var}(X \mid Y)$, then calculate $\mathbb{E}[v(Y)]$ and $\mathrm{Var}(m(Y))$.

- Weight groups by their probabilities, not automatically equally.
- Conditioning cannot increase variance *on average*, but a particular conditional variance can be larger.

Prediction interpretation: For $R = X - m(Y)$,

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

Typical task: Recover total variance from a conditional model and interpret the within-group and between-mean contributions.

Review: MGFs and moments

Definition: Fix t, then average e^{tX}.

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_x e^{tx} p_X(x) \quad \text{or} \quad \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx.$$

Moment extraction: When the MGF is finite on an open interval around 0,

$$M_X(0) = 1, \qquad \text{and} \qquad M_X^{(k)}(0) = \mathbb{E}[X^k],$$

Therefore,

$$\mathbb{E}[X] = M_X'(0), \qquad \mathrm{Var}(X) = M_X''(0) - \big(M_X'(0)\big)^2.$$

Caution: $M_X''(0)$ is the second moment, not the variance.

Typical task: Compute an MGF from a PMF/PDF, or differentiate a given MGF to obtain a mean and variance.

Review: Identifying distributions from MGFs

Uniqueness: Equal finite MGFs on an open interval around 0 imply the same distribution.

Familiar patterns:

Distribution	$M_X(t)$	Domain
$\text{Bernoulli}(p)$	$1 - p + pe^t$	$t \in \mathbb{R}$
$\text{Poisson}(\lambda)$	$\exp\{\lambda(e^t - 1)\}$	$t \in \mathbb{R}$
$\text{Exponential}(\lambda)$	$\lambda/(\lambda - t)$	$t < \lambda$
$N(\mu, \sigma^2)$	$\exp\{\mu t + \sigma^2 t^2/2\}$	$t \in \mathbb{R}$

Check: $M_X(t) = e^{2t+3t^2}$ identifies $N(2, 6)$, not $N(2, 3)$.

- Matching only a mean and variance does not identify a general distribution.

Typical task: Simplify an MGF, recognize its form, and state the distribution with the correct parameters.

Review: MGFs of transformations and independent sums

Affine transformation:

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

Independent sums: For mutually independent $X_1, \ldots, X_n$,

$$M_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t).$$

Recognize the resulting distribution:

- Independent Bernoulli(p) variables: sum is $\mathrm{Binomial}(n, p)$.
- Independent Poisson variables: add the Poisson parameters.
- Independent normal variables: the sum is normal, with means and variances added.

Workflow: Transform/multiply MGFs $\longrightarrow$ simplify $\longrightarrow$ match a known MGF.

Typical task: Find the distribution of an affine transform or independent sum, then use that distribution to answer a probability question.

Review: Markov's inequality

Nonnegative variable, positive threshold: For $X \geq 0$ and $a > 0$,

$$P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

Reason: On $\{X \geq a\}$, the value of X is at least a; elsewhere it is nonnegative.

$$X \geq a\mathbf{1}_{\{X \geq a\}} \quad \Longrightarrow \quad \mathbb{E}[X] \geq aP(X \geq a).$$

- The bound uses only nonnegativity and the mean; no distributional family is assumed.
- It is an upper bound, not an exact probability.
- A bound above 1 is uninformative. Compare candidate bounds for the same event.

Typical task: Check nonnegativity, match the requested threshold, and report a valid upper bound or complementary lower bound.

Review: Chebyshev's inequality

Deviations from the mean: For $\mu = \mathbb{E}[X]$, $\sigma^2 = \mathrm{Var}(X)$, and $c > 0$,

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}.$$

This is Markov's inequality applied to $(X - \mu)^2$.

Standardized form and complements:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \qquad (k > 0,\ \sigma > 0),$$

$$P(|X - \mu| < c) \geq 1 - \frac{\sigma^2}{c^2}.$$

- No normality or nonnegativity assumption on X is needed.
- For an upper tail, $\{X \geq \mu + c\} \subseteq \{|X - \mu| \geq c\}$.

Typical task: Rewrite a deviation event, apply the variance bound, and keep the inequality and endpoint directions correct.

Review: Finite-sample bounds for averages

For independent, identically distributed $X_1, \ldots, X_n$, with mean μ and variance σ^2, let $\bar{X}_n = (X_1 + \cdots + X_n)/n$.

First find the moments of the average:

$$\mathbb{E}[\bar{X}_n] = \mu, \qquad \mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}.$$

Then apply Chebyshev: For $\epsilon > 0$,

$$P(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}.$$

To guarantee this probability is at most $\delta \in (0, 1)$, it suffices that

$$n \geq \frac{\sigma^2}{\delta\epsilon^2}.$$

Caution: Round up; this sample size is sufficient, not necessarily minimal for the distribution.

Typical task: Bound error probability for fixed n, or choose n for a tolerance and guarantee.

Final checklist / common pitfalls

- **Support and normalization:** Integrate only over the allowed region; conditional and transformed PDFs must integrate to 1.
- **Different tasks:** Restrict the region for a probability; integrate out a coordinate for a marginal; normalize a slice for a conditional PDF.
- **Transformations:** Include all relevant inputs, the transformed support, and the absolute inverse derivative when using the monotone formula.
- **Dependence:** Means add without independence; variances need the covariance correction; product MGFs require independence.
- **Conditioning:** Average conditional variances *and* include the variance of conditional means; weight cases correctly.
- **MGFs:** $M_X''(0)$ gives $\mathbb{E}[X^2]$, not variance; identify the full distribution only after matching the MGF.
- **Bounds:** Check the event, threshold, and direction; report an inequality, not an equality or a normal approximation.

Wrap-up: How to prepare

Use the practice materials to find your weak steps:

- Attempt problems under exam conditions, using only the permitted aids.
- Diagnose errors: choosing the tool, setting up the calculation, or carrying it out.
- Rework missed problems without looking at the solution.

Prepare a usable note sheet:

- Include the conditions and workflow alongside each formula.
- Make sure you can use the tools you have written, not merely recognize them.

Unifying workflow: Identify the object and support $\longrightarrow$ choose and justify the tool $\longrightarrow$ calculate $\longrightarrow$ check and interpret.

Suggested reading: Lectures 12–18, assigned homework, and practice materials; [BT08, Secs. 3.4–3.5, 4.1–4.4, 5.1 (covered material)].

References


Dimitri Bertsekas and John N Tsitsiklis.
Introduction to probability, volume 1.
Athena Scientific, 2nd edition, 2008.