CS109 Lecture 10: LLM Learning Guide

Probabilistic Models

Summer 2026

Lecture 10: LLM Learning Guide

Before lecture: read the Lecture 10 slides (joint distributions, marginals, and Bayes with random variables). Then open your favorite LLM and work through the concepts below **in order**. Each concept has a *Learn* prompt and a *Test me* prompt. Attempt the "Test me" questions yourself before asking the LLM to grade you.

How to get the most out of this:

- Tell the LLM your level: "I'm a student in Stanford's CS109 (probability for computer scientists). Explain at that level."
- After any answer, ask "why?" at least once and ask it to show each step.
- When it states a formula, ask it to define every symbol before continuing.
- Attempt practice yourself first, then say: "Here is my reasoning: _____. Where exactly am I wrong, if anywhere?"

Concept 1: Joint probability mass functions

Learn: > "Explain a joint PMF $P(X = x, Y = y)$ for two discrete random variables: what it represents (the probability of the intersection of two events), why the entries must be nonnegative and sum to 1, and why a joint distribution is 'complete information' that can answer any question about X and Y. Use a small 2-by-3 table as an example."

Test me: > "Give me a small joint PMF table and make me verify it is valid and read off a couple of joint probabilities. Check my work."

Concept 2: Marginal distributions

Learn: > "Explain how to get a marginal PMF from a joint by summing over the other variable: $P(X = x) = \sum_y P(X = x, Y = y)$. Explain why this is the law of total probability and why it is called 'marginal' (sums written in the margins of the table). Work an example from a joint table."

Test me: > "Give me a joint table and make me compute the full marginal PMF of each variable, and confirm each marginal sums to 1. Check my sums."

Concept 3: Conditioning within a joint

Learn: > "Explain how to compute a conditional distribution from a joint: $P(X = x \mid Y = y) = \frac{P(X=x,Y=y)}{P(Y=y)}$, where the denominator is a marginal. Work an example and show that, for fixed y, the conditional probabilities over x sum to 1."

Test me: > "Give me a joint table and make me compute a conditional probability and then the whole conditional distribution of one variable given a value of the other. Check that my conditional sums to 1."

Concept 4: Independence of random variables

Learn: > "Explain what it means for two random variables to be independent: $P(X = x, Y = y) = P(X = x)P(Y = y)$ for all x, y. Show how to test independence from a joint table, and give one independent and one dependent example."

Test me: > "Give me a joint table and make me determine whether X and Y are independent by checking the factorization. Make sure I check enough cells and interpret the result correctly."

Concept 5: Bayes' theorem with random variables

Learn: > "Explain Bayes' theorem for random variables: posterior $\propto$ likelihood $\times$ prior, with the denominator as a normalization constant (the law of total probability). Define prior, likelihood, posterior, and normalization. Work a discrete example (e.g. updating belief in a hidden Bernoulli variable given an observation)."

Test me: > "Give me a scenario with a prior on a hidden variable and a likelihood for an observation, and make me compute the posterior with Bayes' theorem. Check that I normalized correctly."

Concept 6: Mixing discrete and continuous (using densities as likelihoods)

Learn: > "Explain how Bayes' theorem works when the hidden variable is discrete but the observation is continuous: use the PDF $f(x \mid \text{hypothesis})$ as the likelihood instead of a PMF. Emphasize that a shared normalizing constant in the densities cancels. Work an example like inferring a hidden Bernoulli given a normally or exponentially distributed observation."

Test me: > "Give me a discrete prior and continuous (normal or exponential) likelihoods, make me observe a value, and make me compute the posterior using the densities. Check that I used $f(x \mid \cdot)$ and normalized."

Wrap-up prompt (after all six concepts)

> "Give me one multi-part CS109-style problem with a joint distribution that requires me to (1) compute a marginal, (2) compute a conditional, (3) test independence, and (4) apply Bayes' theorem with a continuous observation (density as likelihood). Let me solve every part, then grade each part, tell me which concept it tested, and tell me the single concept I should review most before the next lecture."

Note: next lecture, "inference," is Bayes' theorem with random variables applied repeatedly, updating a whole belief distribution (a dictionary/table) as observations arrive.

Bring any sticking points to lecture; the TAs and instructor will help you work through them on the worksheet.