CS109 Lecture 15: LLM Learning Guide

Adding Random Variables & the Central Limit Theorem

Summer 2026

Lecture 15: LLM Learning Guide

Before lecture: read the Lecture 15 slides (adding random variables and the Central Limit Theorem). 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: IID random variables

Learn: > "Define what it means for random variables $X_1, \ldots, X_n$ to be independent and identically distributed (i.i.d.): same PMF or PDF (hence same expectation and variance), and mutually independent. Give examples that are i.i.d. (repeated die rolls) and examples that are not (drawing cards without replacement; variables with different distributions)."

Test me: > "Describe several collections of random variables and make me decide for each whether they are i.i.d., and if not, whether independence or identical distribution fails. Check my justification."

Concept 2: Convolution — the distribution of a sum

Learn: > "Explain how to find the PMF of $X + Y$ for independent discrete random variables: $P(X + Y = n) = \sum_k P(X = k)P(Y = n - k)$. Explain each ingredient: the sum over the mutually exclusive ways to split n, and independence turning joint probabilities into products. Mention this is called convolution, show the integral version for continuous RVs, and work a small example with two dice or two small PMF tables."

Test me: > "Give me two small PMF tables for independent random variables and make me compute the full PMF of their sum. Check each term, and ask me where I used mutual exclusivity and where I used independence."

Concept 3: Sums with closed forms (Binomial, Poisson, Normal)

Learn: > "State and explain the three closed-form sum rules for independent random variables: $\text{Bin}(n_1, p) + \text{Bin}(n_2, p) = \text{Bin}(n_1 + n_2, p)$ (same p required — explain why), $\text{Poi}(\lambda_1) + \text{Poi}(\lambda_2) = \text{Poi}(\lambda_1 + \lambda_2)$, and $N(\mu_1, \sigma_1^2) + N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Contrast with the sum of two uniforms, which is *not* uniform (it's a triangle) — most distributions don't have closed-form sums."

Test me: > "Quiz me with several sums of independent random variables (binomials with same and different p, Poissons, normals, uniforms) and make me either name the distribution of the sum with parameters or say that there is no simple closed form. Check my reasoning on the trick cases."

Concept 4: Differences of normals and zero-sum games (ELO)

Learn: > "Explain how to compute the probability that one normal beats another: for independent $A \sim N(\mu_A, \sigma_A^2)$ and $B \sim N(\mu_B, \sigma_B^2)$, the difference $D = A - B$ is $N(\mu_A - \mu_B,\ \sigma_A^2 + \sigma_B^2)$ — emphasize that variances add even for a difference — and $P(A > B) = P(D > 0)$ via the standard normal CDF. Connect this to ELO ratings: each team's game performance is a normal around its rating, and the higher performance wins."

Test me: > "Give me two players with ELO-style ratings and performance variances and make me compute the probability the underdog wins. Check that I add (not subtract) the variances, and ask me what happens to the win probability as the variance grows."

Concept 5: The Central Limit Theorem

Learn: > "State the Central Limit Theorem: the sum (or average) of n i.i.d. random variables, each with mean μ and variance σ^2, is approximately $N(n\mu,\ n\sigma^2)$ (or $N(\mu, \sigma^2/n)$ for the average) as n gets large — regardless of the shape of the original distribution. Explain the requirements (i.i.d., finite variance, n reasonably large), why 'more mass in the middle' appears even for sums of 2 or 3 dice, and give real-world examples of why so many quantities look normal (Galton board, sums of many small effects)."

Test me: > "Give me a few scenarios of sums of random variables and make me decide whether the CLT applies (watch for: not identically distributed, not independent, tiny n), and for one valid scenario make me write down the approximating normal's parameters. Check my reasoning like a grader."

Concept 6: Using the CLT (with continuity correction)

Learn: > "Work through using the CLT to approximate a probability about a sum of discrete i.i.d. RVs, e.g. the sum of many dice: compute $n\mu$ and $n\sigma^2$, then use the normal CDF with a continuity correction (e.g. $P(S \le 25)$ becomes $P(S \le 25.5)$ in the normal). Explain when the continuity correction applies (integer-valued sums) and when it does not (sums of continuous RVs)."

Test me: > "Give me a problem about the sum of many dice rolls (or another integer-valued i.i.d. sum) with a threshold probability to compute. Make me set up the normal approximation, apply the continuity correction correctly, and express the answer with Φ. Then ask me why no continuity correction is needed for a sum of uniforms."

Wrap-up prompt (after all six concepts)

> "Give me one multi-part CS109-style problem in which I must (1) recognize which of several sums have closed forms and name them, (2) compute one convolution term by hand, (3) compute the probability that one normal exceeds another, and (4) use the CLT with continuity correction to

approximate a probability about a sum of many i.i.d. discrete random variables. 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."

Note: the CLT is why the normal distribution shows up everywhere, and it powers the next stretch of the course: sampling, bootstrapping, and hypothesis testing all lean on sums and averages of i.i.d. samples. Also: the midterm is coming up — this material is a great excuse to review Binomial, Poisson, Normal, and Uniform.

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