CS109 Lecture 10 Worksheet: Probabilistic Models

Joint Distributions, Marginals, and Bayes with Random Variables

Summer 2026

Lecture 10 Worksheet: Probabilistic Models

Topics: joint probability mass functions, marginal distributions, the law of total probability from a joint, independence of random variables, and Bayes' theorem with random variables (prior, likelihood, posterior, normalization), including the mix of discrete and continuous variables.

How to use this sheet. Work in small groups. A TA and the instructor will circulate to help. We will go over selected solutions on the board. You do not need to finish every part; aim for understanding over speed. You may leave answers as exact fractions unless a decimal is requested.

Problem 1 (Review of Lecture 9): Normal

IQ-like scores are $X \sim N(\mu = 100, \sigma^2 = 225)$.

(a) What is the standard deviation?

(b) Find $P(X > 130)$ using a z-score and Φ. (Check: $z = 2$, $\Phi(2) \approx 0.9772$.)

Problem 2: A Joint PMF

Let X be relationship status (S = single, R = in a relationship) and Y be class year (Fr, So, Jr). Their joint PMF $P(X, Y)$ is:

	Fr	So	Jr
S	0.18	0.12	0.10
R	0.12	0.18	0.30

(a) Verify this is a valid joint PMF.

(b) Find $P(X = \text{R}, Y = \text{Jr})$ and $P(X = \text{S}, Y = \text{Fr})$ directly from the table.

(c) Find $P(Y = \text{Jr})$ using the law of total probability (sum the relevant cells).

Problem 3: Marginal Distributions

Using the joint table from Problem 2:

(a) Find the marginal PMF of X: $P(X = \text{S})$ and $P(X = \text{R})$.

(b) Find the marginal PMF of Y: $P(Y = \text{Fr})$, $P(Y = \text{So})$, $P(Y = \text{Jr})$.

(c) Explain in one sentence why summing a row (or column) of the joint gives a marginal.

Problem 4: Conditioning and Independence

Using the same joint table:

(a) Find $P(X = \text{R} \mid Y = \text{Jr})$.

(b) Are X and Y independent? Check whether $P(X = \text{S}, Y = \text{Fr}) = P(X = \text{S})\, P(Y = \text{Fr})$, and state your conclusion.

Problem 5: Bayes with Random Variables (Baby Hearing Test)

A baby can hear ($Y = 1$) with prior probability 0.75. After a sound plays, the change in gaze (degrees) is modeled as: if the baby **can** hear, $X \sim N(15, 25)$; if it **cannot** hear, $X \sim N(8, 25)$ (both have $\sigma = 5$). You observe a gaze change of $X = 14$.

(a) Write Bayes' theorem for $P(Y = 1 \mid X = 14)$ in terms of the prior, the likelihood *densities* $f(14 \mid Y = 1)$ and $f(14 \mid Y = 0)$, and the normalization.

(b) Since both densities share the same constant $\frac{1}{\sigma\sqrt{2\pi}}$, it cancels. Using only the exponential parts $e^{-(14-\mu)^2/(2\cdot 25)}$, compute the posterior $P(Y = 1 \mid X = 14)$. (Check: about 0.86.)

Problem 6: The Size of a Joint (pset4: joint_size)

You are designing a probabilistic system with 20 random variables, each taking one of 5 values. You consider storing the full joint distribution as a table with one cell per unique assignment.

(a) How many cells would the table have?

(b) In one sentence, why does this motivate the models (independence assumptions, Bayes networks) we build next lecture?

Challenge Problem (optional): Tired Baby (pset4: tired_baby)

Your prior that the baby is tired is $P(\text{Tired}) = \frac{3}{4}$. If the baby is tired, the time until they rub their eyes is $\text{Exp}(\lambda = 3)$; if not tired, it is $\text{Exp}(\lambda = 1)$. The baby rubs their eyes after 2 minutes.

(a) Set up Bayes' theorem for $P(\text{Tired} \mid t = 2)$, using exponential **densities** $f(2) = \lambda e^{-\lambda \cdot 2}$ as the likelihoods.

(b) Compute the posterior. Is the baby more or less likely to be tired than the prior suggested, and why does that make sense given the exponential means?