CS109: Probability for Computer Scientists

Lecture 14 Worksheet — Beta

Feb 6, 2026

---

1 Unknown Probability of Heads

Let X be the unknown probability that a coin lands heads. You flip the coin 10 times and observe 9 heads and 1 tail. Let H be the number of heads and T be the number of tails. This question is asking us to solve for:

$$f(X = x \mid H = 9, T = 1).$$

Using Bayes' theorem we can write this as:

$$f(X = x \mid H = 9, T = 1) = K \cdot P(H = 9, T = 1 \mid X = x)\, f(X = x).$$

Where K is the normalization constant we will solve for later.

a) Write an expression for the likelihood $P(H = 9, T = 1 \mid X = x)$.
   Solution: $P(H = 9, T = 1 \mid X = x) = \binom{10}{9} x^9 (1 - x)^1$.

b) What do you think we should use for our prior $f(X = x)$?
   Solution: Use a uniform prior on $[0, 1]$: $X \sim \text{Uni}(0, 1)$, so $f(X = x) = 1$ for $0 \leq x \leq 1$.

c) Write an expression for the posterior density of X. You can use K to represent a normalization constant.
   Solution: $f(X = x \mid H = 9, T = 1) = K\, x^9 (1 - x)$ for $0 \leq x \leq 1$.

d) Now instead of observing 9 heads and 1 tails, assume we flip a coin n times and observe h tails. Write and simplify an expression for our posterior belief $f(X = x|H = h)$.
   Solution: If we observe h tails and $(n - h)$ heads with a uniform prior, $f(X = x \mid H = h) = K\, x^{n-h} (1 - x)^h$, i.e. $X \mid H = h \sim \text{Beta}(n - h + 1, h + 1)$.

e) Use that equation to write an updated belief after observing 7 successes and 1 failure.

Solution: $X \mid \text{data} \sim \text{Beta}(8,2)$, with density $f(x) = Kx^7(1-x)^1$.

2 Beta Prior Update

Now suppose your prior is $X \sim \text{Beta}(a_0, b_0)$. Then you flip your coin and observe h heads and t tails. Derive the posterior distribution of X and give its parameters in terms of a_0, b_0, h, t.

Solution: $X \mid \text{data} \sim \text{Beta}(a_0 + h,\ b_0 + t)$.

3 Beta(1, 1)

What is the pdf of X where $X \sim \text{Beta}(a = 1, b = 1)$? **Solution:** $f(x) = 1$ for $0 \le x \le 1$ (and 0 otherwise). So it is $\text{Uni}(0,1)$.

4 A Beta Example

A medicine is believed to work about 80% of the time before testing. Then you try the medicine on 20 patients. It "works" for 14 and "doesn't work" for 6. What is your updated belief that the drug works?

(a) To start, first write down how you would represent your prior belief that the medicine works 80% of the time?

Solution: One simple choice is $X \sim \text{Beta}(8,2)$ (mean $= 0.8$).

(b) Now, write an expression for the posterior belief.

Solution: Posterior: $X \mid \text{data} \sim \text{Beta}(8 + 14,\ 2 + 6) = \text{Beta}(22, 8)$.

4. Laplace Smoothing and Small Sample Effects

Suppose a recommendation system is comparing two videos. Let p_A, p_B be the probabilities that a user likes Video A and Video B.

- Video A: 10,000 likes, 50 dislikes
- Video B: 10 likes, 0 dislikes

Assume independent Beta priors with Laplace smoothing (Beta$(1, 1)$ prior for each video).

a) Write the posterior distribution for p_A and the posterior distribution for p_B.

Solution: $p_A \mid \text{data} \sim \text{Beta}(10001, 51)$ and $p_B \mid \text{data} \sim \text{Beta}(11, 1)$.

b) Compute the posterior mean of each probability. Based on posterior means alone, which video appears better?

Solution: $E[p_A] = \frac{10001}{10001+51} = \frac{10001}{10052} \approx 0.9949$, $E[p_B] = \frac{11}{12} \approx 0.9167$. By posterior mean, Video A looks better.

5. Decision-Making with Beta Beliefs

You are choosing between Drug A and Drug B. Let X_A, X_B be success probabilities.

- Prior for each drug: $\text{Beta}(1, 1)$
- You tested Drug B five times and observed 2 successes and 3 failures.

a) Write the posterior for X_B.

Solution: $X_B \mid \text{data} \sim \text{Beta}(1 + 2,\ 1 + 3) = \text{Beta}(3, 4)$.

b) Compute $E[X_B]$.

Solution: $E[X_B] = \frac{3}{3+4} = \frac{3}{7} \approx 0.429$.

c) Describe the Thompson Sampling algorithm.

Solution: For each drug, keep a Beta belief. Sample one value from each Beta. Choose the drug with the larger sample. Update that drug's Beta using the observed success/failure.