Probability
Problem Set 6

Name: Jonah Aden   **Date:** December 8, 2025

   Problem 1. (Monty Hall: Sober and dizzy)

The Monty Hall problem: Monty hosts a game show. There are three doors: one hides a car and two hide goats. The contestant picks a door, which is not opened. Monty then opens another door which has a goat behind it. Finally, contestant must decide whether to stay with her original choice or switch to the other unopened door. The problem asks which is the better strategy: staying or switching?

To be precise, let's label the door that contestant picks by A, and the other two doors by B and C. Hypothesis H_A is that the car is behind door A, and similarly for hypotheses H_B and H_C.

(a) In the usual formulation, Monty is sober and knows the locations of the car and goats. So if the contestant picks a door with a goat, Monty always opens the other door with a goat. And if the contestant picks the door with a car, Monty opens one of the other two doors at random. Suppose that sober Monty Hall opens door B, revealing a goat. So the data is: "Monty showed a goat behind B". Our hypotheses are "the car is behind door A", etc. Make a Bayes table with prior, likelihood and posterior. Use the posterior probabilities to determine the best strategy.

Hypothesis	Prior	Likelihood	Prior × Lik	Posterior
H_A	1/3	1/2	1/6	1/3
H_B	1/3	0	0	0
H_C	1/3	1	1/3	2/3

Since $P(H_C|\text{data}) = 2/3 > 1/3 = P(H_A|\text{data})$, the best strategy is to switch.

(b) Now suppose that Monty is dizzy, i.e. he has completely forgotten where the car is and is only aware enough to randomly open one of the two doors not chosen by the contestant. It's entirely possible he might accidentally reveal the car, ruining the show.

Dizzy Monty Hall opens door B, revealing a goat. Make a Bayes table with prior, likelihood and posterior. Use the posterior probabilities to determine the best strategy. (Hint: the data is the same but the likelihood function is not.)

Hypothesis	Prior	Likelihood	Prior × Lik	Posterior
H_A	1/3	1/2	1/6	1/2
H_B	1/3	0	0	0
H_C	1/3	1/2	1/6	1/2

Since $P(H_A|\text{data}) = P(H_C|\text{data}) = 1/2$, the contestant is indifferent between staying and switching.

(c) Based on Monty's pre-show behavior, we think that Monty is sober with probability 0.7 and dizzy with probability 0.3. Repeat the analysis from parts (a) and (b) in this situation.

Combined likelihoods: $P(\text{data}|H_A) = 1/2$, $P(\text{data}|H_B) = 0$, $P(\text{data}|H_C) = 0.7(1) + 0.3(1/2) = 17/20$.

Hypothesis	Prior	Likelihood	Prior $\times$ Lik	Posterior
H_A	1/3	1/2	1/6	10/27
H_B	1/3	0	0	0
H_C	1/3	17/20	17/60	17/27

Since $P(H_C|\text{data}) = 17/27 > 10/27 = P(H_A|\text{data})$, the best strategy is to switch.

Problem 2. Prediction

We are going to explore the dice problem from class further. I have five dice (4, 6, 8, 12, or 20 sides) and pick one at random (uniform probability). I then roll this die n times and tell you that, miraculously, every roll resulted in the value 7.

(a) First, consider just the first roll. Find the prior predictive probability that the first roll will be a 7 and the posterior (after the first roll) predictive probability that the second roll will be a 7. Also find the posterior (after the first roll) probabilities for the chosen die.

Prior predictive probability:

$$P(x_1 = 7) = \frac{1}{5}\left(0 + 0 + \frac{1}{8} + \frac{1}{12} + \frac{1}{20}\right) = \frac{31}{600}$$

Posterior probabilities after observing $x_1 = 7$:

$$P(\text{8-sided}|x_1 = 7) = \frac{15}{31}, \quad P(\text{12-sided}|x_1 = 7) = \frac{10}{31}, \quad P(\text{20-sided}|x_1 = 7) = \frac{6}{31}$$

Posterior predictive probability:

$$P(x_2 = 7|x_1 = 7) = \frac{1}{8} \cdot \frac{15}{31} + \frac{1}{12} \cdot \frac{10}{31} + \frac{1}{20} \cdot \frac{6}{31} = \frac{361}{3720}$$

(b) Find the posterior probability $P(H \mid \text{data})$ for each die given the data of all n rolls (your answers should involve n). What is the limit of each of these probabilities as n grows to infinity? Explain why this makes sense.

Let $S_n = (1/8)^n + (1/12)^n + (1/20)^n$. The posterior probabilities are:

$$P(\text{8-sided}|\text{data}) = \frac{(1/8)^n}{S_n}, \quad P(\text{12-sided}|\text{data}) = \frac{(1/12)^n}{S_n}, \quad P(\text{20-sided}|\text{data}) = \frac{(1/20)^n}{S_n}$$

As $n \to \infty$: $P(\text{8-sided}|\text{data}) \to 1$ and the others go to 0. This makes sense because the 8-sided die has the highest probability of rolling 7, so repeated 7s provide the strongest evidence for that die.

(c) Given that my first 10 rolls resulted in 7 (i.e., $n = 10$), rank the possible values for my next roll from most likely to least likely. Note any ties in rank and explain your reasoning carefully. You need not do any computations to solve this problem.

Most likely to least likely:

1. Values 1, 2, 3, 4, 5, 6, 7, 8 (all tied)

2. Values 9, 10, 11, 12 (all tied)

3. Values 13, 14, 15, 16, 17, 18, 19, 20 (all tied)

After 10 sevens, the 8-sided die has the largest posterior probability. Values 1–8 can come from all three candidate dice. Values 9–12 can only come from the 12 or 20-sided dice. Values 13–20 can only come from the 20-sided die.

(d) Let x_i be the result of the ith roll. Find the posterior predictive pmf for the $(n+1)$st roll given the data. That is, find $P(x_{n+1} \mid x_1 = 7, \ldots, x_n = 7)$ for $x_{n+1} = 1, \ldots, 20$. (Hint: use part (b) and the law of total probability. Many values of the pmf coincide, so you do not need to do 20 separate computations. You should check that your answer is consistent with your ranking in part (c) for $n = 10$.)

Using the law of total probability with $S_n = (1/8)^n + (1/12)^n + (1/20)^n$:

$$P(x_{n+1} = k \mid \text{data}) = \begin{cases} \dfrac{(1/8)^{n+1} + (1/12)^{n+1} + (1/20)^{n+1}}{S_n} & k = 1, \ldots, 8 \\[2mm] \dfrac{(1/12)^{n+1} + (1/20)^{n+1}}{S_n} & k = 9, \ldots, 12 \\[2mm] \dfrac{(1/20)^{n+1}}{S_n} & k = 13, \ldots, 20 \end{cases}$$

(e) What function does the pmf in part (d) converge to as n grows to infinity? Explain why this makes sense.

As $n \to \infty$, the pmf converges to:

$$P(x_{n+1} = k) \to \begin{cases} 1/8 & k = 1, \ldots, 8 \\ 0 & k = 9, \ldots, 20 \end{cases}$$

This is the uniform distribution on $\{1, 2, \ldots, 8\}$. It makes sense because in the limit we become certain the 8-sided die was chosen, so the next roll is just a fair roll of that die.

Problem 3. Suppose that a random sample of size n is to be taken from a distribution for which the mean is μ and the standard deviation is 3. Use the central limit theorem to determine approximately the smallest value of n for which the following relation will be satisfied:
$$\Pr(|\bar{X}_n - \mu| < 0.3) \geq 0.95.$$

By CLT, $\bar{X}_n \sim N(\mu, 9/n)$ approximately. Standardizing:

$$P(|\bar{X}_n - \mu| < 0.3) = P\left(|Z| < \frac{0.3\sqrt{n}}{3}\right) = 2\Phi(0.1\sqrt{n}) - 1 \geq 0.95$$

So $\Phi(0.1\sqrt{n}) \geq 0.975$, which means $0.1\sqrt{n} \geq 1.96$, giving $\sqrt{n} \geq 19.6$ and $n \geq 384.16$.

The smallest such integer is $n = 385$.

Problem 4. Suppose that 16 digits are chosen at random with replacement from the set $\{0, \ldots, 9\}$. What is the probability that their average will lie between 4 and 6?

For a single digit X uniform on $\{0, 1, \ldots, 9\}$: $E[X] = 4.5$ and $\text{Var}(X) = 8.25 = 33/4$.
By CLT, $\bar{X}_{16} \sim N(4.5, 33/64)$ approximately. The standard deviation is $\sqrt{33}/8$. Standardizing:

$$P(4 < \bar{X}_{16} < 6) = \Phi\left(\frac{6 - 4.5}{\sqrt{33}/8}\right) - \Phi\left(\frac{4 - 4.5}{\sqrt{33}/8}\right) = \Phi\left(\frac{12}{\sqrt{33}}\right) - \Phi\left(\frac{-4}{\sqrt{33}}\right) \approx 0.7385$$

Problem 5. Suppose that the proportion θ of defective items in a large manufactured lot is unknown, and the prior distribution of θ is the uniform distribution on the interval $[0, 1]$. When eight items are selected at random from the lot, it is found that exactly three of them are defective. Determine the posterior distribution of θ. Compute the expected mean and variance of θ.

Prior: $\theta \sim \text{Uniform}[0, 1]$, i.e., $\xi(\theta) = 1$.
Likelihood: $P(3 \text{ defective out of } 8|\theta) \propto \theta^3(1 - \theta)^5$.
Posterior: $f(\theta|\text{data}) \propto 1 \cdot \theta^3(1 - \theta)^5 = \theta^3(1 - \theta)^5$, which is the kernel of Beta$(4, 6)$.
So $\theta|\text{data} \sim \text{Beta}(4, 6)$.
Posterior mean: $E[\theta|\text{data}] = \dfrac{4}{4 + 6} = \dfrac{2}{5}$
Posterior variance: $\text{Var}(\theta|\text{data}) = \dfrac{4 \cdot 6}{(4 + 6)^2(4 + 6 + 1)} = \dfrac{24}{1100} = \dfrac{6}{275}$

Problem 6. Consider again the problem described in Problem 5, but suppose now that the prior p.d.f. of θ is as follows:

$$\xi(\theta) = \begin{cases} 2(1 - \theta), & 0 < \theta < 1, \\ 0, & \text{otherwise.} \end{cases}$$

As in previous Exercise, suppose that in a random sample of eight items exactly three are found to be defective. Determine the posterior distribution of θ.

Prior: $\xi(\theta) = 2(1 - \theta)$, which is Beta$(1, 2)$.
Likelihood: $P(3 \text{ defective out of } 8|\theta) \propto \theta^3(1 - \theta)^5$.
Posterior: $f(\theta|\text{data}) \propto (1 - \theta) \cdot \theta^3(1 - \theta)^5 = \theta^3(1 - \theta)^6$, which is the kernel of Beta$(4, 7)$.
So $\theta|\text{data} \sim \text{Beta}(4, 7)$.