STAT111: Mathematical Statisics **Spring 2025**

Week 9 Bayesian Inference: Binary Decision Problem

Instructor: Prof. Phil Everson Presenter: Zhengfei Li (Alex)

I would like to acknowledge that this handout is created with reference to Introduction to Probability by Blitzstein, J.K. I would also like to acknowledge the instructions from Prof. Everson.

Problem 1a

"The two envelopes paradox" is a classic example of a binary decision problem (Blitzstein 9.1.6). Imagine two wealthy-looking visitors recruit you and a friend to give them a tour of the college. As a reward, the visitors present you with two envelopes and tell you that one envelope contains twice as much money as the other. You and your friend flip a coin to assign the envelopes and you see that yours contains \$100. You realize your friend has either \$50 or \$200, and because you flipped a coin to decide, you figure the expected value of your friend's amount is $(1/2)(50) + (1/2)(200) = \$125$. So you're thinking you'd like to switch envelopes. Meanwhile, your friend hasn't looked yet, but she figures if she has X dollars, then you have $X/2$ or $2X$, for an expected value of $(1/2)X/2 + (1/2)(2X) = 1.25X > X$. So she'd like to switch envelopes too. Explain the flaw in this reasoning, using θ to represent the larger amount and X to represent the amount in your envelope (and Y the amount in your friend's envelope).

Solution

This this part, I would like to present the solution from Blitzstein's textbook.

According to the problem setup, we can put events together to two possible events with probability

$$P(Y = 2X) = P(Y = X/2) = 1/2$$

We may start from Law of Total Expectation,

$$\begin{aligned}
E[Y] &= E[Y|Y = 2X]P(Y = 2X) + E[Y|Y = X/2]P(Y = X/2) \\
&= \frac{1}{2}E[Y|Y = 2X] + \frac{1}{2}E[Y|Y = X/2] \\
&= \frac{1}{2}\left[E[2X|Y > X] + E[X/2|Y < X]\right] \\
&= \frac{1}{2}\left[2E[X|Y > X] + \frac{1}{2}E[X|Y < X]\right]
\end{aligned}$$

The reasoning suggested in the prompt assumes that $E[X|Y > X] = E[X|Y < X] = E[X]$, therefore

$$E[Y] = \frac{1}{2}\left[2E[X] + \frac{1}{2}E[X]\right] = \frac{5}{4}E[X]$$

The assumption means that knowing whether X is larger or smaller than Y does not provide any information about the value of X.

Intuitively, this might not be realistic. There are constraints on the actual value one may tip including their finite wealth and the representation limit of our currency system.

We will take a statistical point of view into this problem with reference to the prior distribution of the larger amount between the two envelops (θ).

Problem 1b

Give a Bayesian solution to the 2-envelopes paradox.

Solution

The situation is that we observe one of the envelopes $X = x$, and we are wondering whether the other envelope contains more money or less. Therefore, we would work out the probability of event $Y > X$ (a.k.a. $X = \theta/2$) with Bayes' Theorem.

$$\begin{aligned}
P(X = \theta/2 | X = x) &= \frac{f(X = x | X = \theta/2) \cdot P(X = \theta/2)}{f(X = x)} \\
&= \frac{\frac{1}{2} f(X = x | X = \theta/2)}{f(X = x | \theta = 2x) f(\theta = 2x) + f(X = x | \theta = x) f(\theta = x)} \\
&= \frac{\frac{1}{2} f(\theta = 2x)}{\frac{1}{2} f(\theta = 2x) + \frac{1}{2} f(\theta = x)} \\
&= \frac{f(\theta = 2x)}{f(\theta = 2x) + f(\theta = x)}
\end{aligned}$$

The paradox argues that

$$\begin{aligned}
P(X = \theta/2 | X = x) = \frac{f(\theta = 2x)}{f(\theta = 2x) + f(\theta = x)} &= \frac{1}{2} \\
2f(\theta = 2x) &= f(\theta = 2x) + f(\theta = x) \\
f(\theta = 2x) &= f(\theta = x) \quad \forall x \in S(X)
\end{aligned}$$

In addition, the bayesian approach also suggests that the decision of whether we should switch the envelope depends on our prior belief of the distribution of θ.

How would the prior change if this were a gameshow prize rather than a tip?

Solution

Based on the amount of money in my wallet, my tip would have $\theta \sim \text{Unif}(0, 170)$. Therefore, if the tour guide receive an envelope with \$100 from me, they would have guessed that the \$100 is more likely to be the larger amount, and we would not switch the envelope.

If the prize is a gameshow prize, however, the participant would expect a typical prize to be $\theta \sim \text{Unif}(0, 10^6)$ for TV show effect. In this case, we would think that the \$100 is more likely to be the smaller amount, and we would switch the envelope.

Identify the prior specification that results in the paradox of both friends wanting to switch, then switch back, and so on.

Solution

$X \sim \text{Unif}(0, \infty)$

Briefly, due to non-negativity of a tip, $X \in \mathbb{R}^+ \cup \{0\}$. Because $X = 100 \neq 0$ is a non-trivial event, we know that $\{0\}$ is a strict subset of the support of X.

Take arbitrary $x \in \mathbb{R}^+$, we know that $f(2^k \cdot x) = f(x), k \in \mathbb{N}$. Therefore, we know that as $k \to \infty$ the support of f spans towards infinity.

Take arbitrary $x \in \mathbb{R}^+$, we know that $f(2^{-k} \cdot x) = f(x), k \in \mathbb{N}$. Let's setup a constant sequence $\{f(x), f(2^{-1}x), \cdots, f(2^{-k}x), \cdots\}, k \in \mathbb{N}$, and by continuity of f, we suggest that $f(2^{-k}x) \to f(0)$. As we pick x to be arbitrary, we know that $f(x)$ is a constant function: $X \sim \text{Unif}(0, \infty)$.

However, $X \sim \text{Unif}(0, \infty)$ is not a valid distribution, so this is the loophole of the "infinite switching" paradox.

Consider Exponential and Uniform as possible prior distributions for θ, and work out the posterior probabilities for θ given the observed value $X = x$. Discuss the difference between deciding based on the higher probability or based on the posterior mean of your friend's envelope value.

Solution

Let's first consider $\theta \sim \text{Unif}(a, b)$, and assume $b/2 > a$ for a general case.

$$
\begin{aligned}
P(\theta = 2X|X = x) = P(X = \theta/2|X = x) &= \frac{f(\theta = 2x)}{f(\theta = 2x) + f(\theta = x)} \\
&= \begin{cases} 1 & \text{if } x \in (a/2, a) : \text{ switch by prob} \\ 1/2 & \text{if } x \in (a, b/2) : \text{ indifferent by prob} \\ 0 & \text{if } x \in (b/2, b) : \text{ not switch by prob} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&E(Y|X = x) \\
&= P(Y = 2x|X = x)E(Y|Y = 2x, X = x) + P(Y = x/2|X = x)E(Y = x/2|Y = x/2, X = x) \\
&= P(Y = 2x|X = x)2x + (1 - P(Y = 2x|X = x))\frac{x}{2} \\
&= \begin{cases} 2x & \text{if } x \in (a/2, a) : \text{ switch by exp} \\ 2x \cdot \frac{1}{2} + \frac{x}{2} \cdot \frac{1}{2} = \frac{5}{4}x & \text{if } x \in (a, b/2) : \text{ switch by exp} \\ \frac{x}{2} & \text{if } x \in (b/2, b) : \text{ not switch by exp} \end{cases}
\end{aligned}
$$

Considering $\theta \sim \text{Expo}(\lambda)$.

$$
\begin{aligned}
P(\theta = 2X|X = x) = P(X = \theta/2|X = x) &= \frac{f(\theta = 2x)}{f(\theta = 2x) + f(\theta = x)} \\
&= \frac{\lambda e^{-\lambda(2x)}}{\lambda e^{-\lambda(2x)} + \lambda e^{-\lambda x}} \\
&= \frac{\lambda e^{-\lambda x} \cdot e^{-\lambda x}}{(e^{-\lambda x} + 1)\lambda e^{-\lambda x}} \\
&= \frac{1}{e^{\lambda x} + 1}
\end{aligned}
$$

$P(\theta = 2X|X = x)$ achieves 1/2 at $x = 0$ and $\frac{1}{e^{\lambda x}+1}$ is a decreasing function. This means that, assuming we observed $X > 0$, it is always more possible that we are holding the smaller envelope. Therefore, we would want to switch the envelope by probability.

$$
\begin{aligned}
&E(Y|X = x) \\
&= P(Y = 2x|X = x)E(Y|Y = 2x, X = x) + P(Y = x/2|X = x)E(Y = x/2|Y = x/2, X = x) \\
&= P(Y = 2x|X = x)2x + (1 - P(Y = 2x|X = x))\frac{x}{2} \\
&= \frac{1}{e^{\lambda x} + 1} \cdot 2x + \left(1 - \frac{1}{e^{\lambda x} + 1}\right) \cdot \frac{x}{2} \\
&= \frac{1}{e^{\lambda x} + 1} \cdot 2x + \left(\frac{e^{\lambda x}}{e^{\lambda x} + 1}\right) \cdot \frac{x}{2} \\
&= \left[\frac{2 + \frac{1}{2}e^{\lambda x}}{e^{\lambda x} + 1}\right] x \\
&= \left[\frac{1}{2} + \frac{3}{2(e^{\lambda x} + 1)}\right] x
\end{aligned}
$$

Let $\left[\frac{1}{2} + \frac{3}{2(e^{\lambda x}+1)}\right] = 1$

$$
\begin{aligned}
\frac{1}{2} + \frac{3}{2(e^{\lambda x} + 1)} &= 1 \\
\frac{3}{2(e^{\lambda x} + 1)} &= \frac{1}{2} \\
e^{\lambda x} &= 2 \\
\lambda x &= \ln(2) \\
x &= \frac{\ln(2)}{\lambda}
\end{aligned}
$$

By expectation, the conclusion above means that if we receive $x < \frac{\ln(2)}{\lambda}$, we want to switch. If we receive $x > \frac{\ln(2)}{\lambda}$, we do not want to switch. If we receive $x = \frac{\ln(2)}{\lambda}$, we are indifferent.

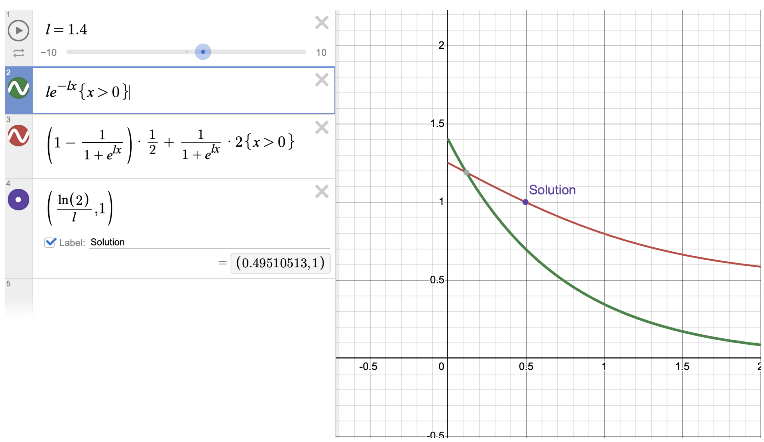

Figure 1: Red: $\frac{E(Y|X=x)}{x}$, Green: PMF of Expo(l), Purple: solution of $\frac{E(Y|X=x)}{x} = 1$