Unit 4 - Part 1 Homework

W203 Statistics for Data Science

Student Name: Youzhi Wu

Section Number: 05

1. Best Game in the Casino

(a) How much do you get paid if the coin comes up heads 3 times?

Set X as a discrete R.V. such that:

$$X = \begin{cases} 0 & \text{if 0 head} \\ 2 & \text{if 1 head} \\ 4 & \text{if 2 heads} \\ y & \text{if 3 heads} \end{cases}$$

Because we know that $E(X) = 6$, we have the following:

$$\begin{aligned} E(X) &= \sum_{k \in O} k \cdot P(X = k) \\ &= 0 * 0.5^3 + 2 * 0.5^3 * P_{1,3} + 4 * 0.5^3 * P_{1,3} + y * 0.5^3 \\ &= 6 \\ y &= \frac{6 - 0 * 0.5^3 - 2 * 0.5^3 * P_{1,3} - 4 * 0.5^3 * P_{1,3}}{0.5^3} \\ &= 30 \end{aligned}$$

Therefore, we have that you will get paid $30 if the coin comes up heads 3 times.

(b) Write down a complete expression for the cumulative probability function for your win- nings from the game.

The cumulative probability function F(x) can be written as follows:

$$F(x) = \begin{cases} \frac{1}{8} & \text{if x = 0} \\ \frac{1}{2} & \text{if } x \leq 1 \\ \frac{7}{8} & \text{if } x \leq 2 \\ 1 & \text{if } x \leq 3 \end{cases}$$

2. Reciprocal Dice

Assuming the 6-sided die is a fair die, which means that each side has a 1/6 possibility to get rolled.

Option (a): Get 1/E(X) in dollars right away

$$\begin{aligned} \frac{1}{E(X)} &= \frac{1}{1/6 * 1 + 1/6 * 2 + 1/6 * 3 + 1/6 * 4 + 1/6 * 5 + 1/6 * 6} \\ &= \frac{2}{7} \\ &= 0.286 \end{aligned}$$

Option (b): Get 1/X in dollars

$$\begin{aligned} E(\frac{1}{X}) &= 1/6 * 1 + (1/6) * (1/2) + (1/6) * (1/3) + (1/6) * (1/4) + (1/6) * (1/5) + (1/6) * (1/6) \\ &= 0.408 \end{aligned}$$

In terms of expectation, Option (b) has a greater expectation, and thus is better.

3. The Baseline for Measuring Deviations

(a) Write down an expression for E(Y) and simplify it as much as you can

Because

$$Y = (X - t)^2$$

Therefore, we have

$$\begin{aligned} E(Y) &= E((X - t)^2) \\ &= E(X^2 - 2tX + t^2) \\ &= E(X^2) - E(2tX) + E(t^2) \\ &= E(X^2) - 2tE(X) + t^2 \end{aligned}$$

(b) Compute the value of t that minimizes E(Y)

$$\frac{dE(Y)}{dt} = 2t - 2E(X)$$

The value of t that minimizes E(Y) can be written as follows:

$$\begin{aligned} 2t - 2E(X) = 0 \\ t = E(X) \end{aligned}$$

Therefore, when t equals to the mean of X, it minimizes E(Y).

(c) What is the value of E(Y) for this choice of t?

$$\begin{aligned} E(Y)_{min} &= E(X^2) - 2[E(X)]^2 + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

Therefore, E(Y) reaches minimum value when it equals to the variance of X.

4. Optional Advanced Exercise: Heavy Tails

(a) Is E(M) finite?

$$
\begin{aligned}
E(M) &= \sum_{x \in 1,2,3,...} x \frac{c}{x^3} + \sum_{x \notin 1,2,3,...} x * 0 \\
&= \sum_{x \in 1,2,3,...} \frac{c}{x^2}
\end{aligned}
$$

The series $\sum_{x \in 1,2,3,...} \frac{1}{x^2}$ converges to $\frac{\pi^2}{6}$, therefore E(M) is finite.

(b) Is var(M) finite?

$$
\begin{aligned}
var(M) &= E(M^2) - [E(M)]^2 \\
&= \sum_{x \in 1,2,3,...} x^2 \frac{c}{x^3} - [\sum_{x \in 1,2,3,...} x \frac{c}{x^3}]^2 \\
&= \sum_{x \in 1,2,3,...} \frac{c}{x} - [\sum_{x \in 1,2,3,...} \frac{c}{x^2}]^2
\end{aligned}
$$

The series $\sum_{x \in 1,2,3,...} \frac{1}{x}$ does not converge, therefore var(M) is not finite.