HW01: THEORETICAL QUESTIONS

Spring 2024, with Ethan P. Marzban

PSTAT 100: Data Science Concepts and Analysis

Welcome to the first PSTAT 100 Homework! This quarter, homework will consist of a mix of programming/coding questions, as well as some conceptual/theoretical questions.

> **!** Important
>
> - This PDF contains only the theoretical questions to Homework 01. Computing questions appear in a separate document.
> - Do **not** try to write your solutions to the theoretical questions on this PDF- please write your solutions on a separate sheet of paper.
> - When submitting, merge your PDFs (containing your computing and theoretical question answers), and upload this merged PDF to Gradescope.
>   - After submitting, make sure to **match pages**.

Probability Review

1) **Does Rounding Preserve Uniformity?** Let $U \sim \text{Unif}[0, 10]$ and let $Y := \text{round}(U)$, where $\text{round}(t)$ denotes the value of t rounded to the nearest integer. (For example, $\text{round}(2.5) = 3$ and $\text{round}(2.4) = 2$.)

   (a) Write down the support (state space) S_Y of Y. Is Y a discrete or a continuous random variable?

   Solution: We see that $S_Y = \{0, 1, 2, \cdots, 10\}$, meaning Y is a discrete random variable.

   (b) Define $A_3 := \{Y = 3\}$; i.e. A_3 is the event that Y equals 3. Find $\mathbb{P}(A_3)$.

   Solution: First, we would like to find the values of U that map to $\{Y = 3\}$. We can see that any number in the set $[2.5, 3.5)$ gets rounded to 3; hence

$$\begin{aligned} \mathbb{P}(A_3) &:= \mathbb{P}(Y = 3) \\ &= \mathbb{P}\left(U \in [2.5, 3.5)\right) = \frac{3.5 - 2.5}{10} = \boxed{\frac{1}{10}} \end{aligned}$$

(c) Define $A_0 := \{Y = 0\}$; i.e. A_0 is the event that Y equals 0. Find $\mathbb{P}(A_0)$.

Solution: We proceed similarly as in part (a), now noting that any number in the set $[0, 0.5)$ gets mapped to zero (U cannot attain negative values, so we ignore any numbers in the interval $[-0.5, 0)$). This gives:

$$\begin{aligned} \mathbb{P}(A_0) &:= \mathbb{P}(Y = 0) \\ &= \mathbb{P}\left(U \in [0, 0.5)\right) = \frac{0.5 - 0}{10} = \boxed{\frac{1}{20}} \end{aligned}$$

(d) Extend your logic from parts (a) and (b) to find an expression for $\mathbb{P}(Y = y)$, for appropriate values of y. (This amounts to finding the PMF/PDF of Y.)

Solution: With a bit of thought, we can see that $\mathbb{P}(Y = y) = 1/10$ for any $y \in \{1, 2, \cdots, 9\}$. We have already seen that $\mathbb{P}(Y = 0) = 1/20$; similar reasoning gives $\mathbb{P}(Y = 10) = 1/20$, meaning our final PMF for Y is:

$$\mathbb{P}(Y = y) = \begin{cases} 1/20 & \text{if } y = 0 \\ 1/10 & \text{if } y = 1, 2, \cdots, 9 \\ 1/20 & \text{if } y = 10 \\ 0 & \text{otherwise} \end{cases}$$

As a quick check, we can see that the PMF of Y does indeed sum to unity:

$$\frac{1}{20} + \underbrace{\frac{1}{10} + \cdots + \frac{1}{10}}_{9 \text{ terms}} + \frac{1}{20} = \frac{1}{20} + \frac{9}{10} + \frac{1}{20} = \frac{10}{10} = 1 \checkmark$$

(e) Is Y uniformly distributed (discrete or continuous)?

Solution: No; if it were (discrete) uniformly distributed on its support, we would have $\mathbb{P}(Y = y) = 1/11$ for any $y \in S_Y$, which, by part (d), is clearly not the case.

(f) Compute $\mathbb{E}[Y]$ and Var(Y). Some useful facts to remember:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$
$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

Solution:

$$\begin{aligned}
\mathbb{E}[Y] &:= \sum_y y\mathbb{P}(Y = y) \\
&= 0 \cdot \frac{1}{20} + \sum_{y=1}^{10} y \cdot \frac{1}{10} + 10 \cdot \frac{1}{20} \\
&= \frac{1}{\cancel{10}} \cdot \frac{\cancel{10} \cdot 11}{2} + \frac{1}{2} = \boxed{6} \\
\mathbb{E}[Y^2] &= 0^2 \cdot \frac{1}{20} + \sum_{y=1}^{10} y^2 \cdot \frac{1}{10} + 10^2 \cdot \frac{1}{20} \\
&= \frac{1}{\cancel{10}} \cdot \frac{\cancel{10} \cdot 11 \cdot 21}{6} + 5 = \frac{87}{2} \\
\mathrm{Var}(Y) &= \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = \frac{87}{2} - (6^2) = \boxed{\frac{15}{2} = 7.5}
\end{aligned}$$