Problem set 3: Summarizing distributions

Due October 21, 2024, at 10am

(Your name here)

NOTE: Start with the file `ps3_2024_summarizing_distributions.qmd` *(available from the github repository at https://github.com/UChicago-pol-methods/IntroQSS-F24/tree/main/assignments). Modify that file to include your answers. Make sure you can "render" the file (e.g. in RStudio by clicking on the* `Render` *button). Submit both the qmd file and the PDF via Canvas.*

Question 1: Expected value

Consider the random variable X characterized by the following PMF.

x	$f(x)$
0	.3
1	.5
2	.15
3	.05

(1a) Compute $\mathrm{E}[X]$. Show your work.

Answer:

(1b) Write an `R` function to compute the expectation of any discrete random variable. The arguments to your function should include the values the random variable can take on (`x`) and the probability it takes on each value (`probs`). Use your function to confirm your answer from question (1a).

Answer:

```
# your code here
```

(1c) Compute the MSE ($\mathrm{E}\left[(X-c)^2\right]$) for $c = 0$ and $c = 1$. Show your work.

Answer:

(1d) Write a function to compute the MSE for any discrete random variable at a value `c`. The arguments to your function should include the values the random variable can take on (`x`), the probability it takes on each value (`probs`), and the value `c` being considered. Use your function to confirm your answers from (1c).

Answer:

```
# your code here
```

(1e) Create a vector `cs` that contains numbers in the sequence `0.5, 0.55, 0.6, 0.65, 0.7, . . ., 2`. Using a for-loop and your function from (1d), compute the MSE for the random variable X whose PMF was given above at each value of `cs` and store the result in a vector called `mses`.

Answer:

```
# your code here
```

(1f) Using the `plot()` command, make a scatterplot showing the MSE of X (vertical axis) at each value of $c \in \{0.5, 0.55, 0.6, \ldots, 2\}$ (horizontal axis). Use `abline()` to add a vertical line at $E[X]$.

Answer:

Question 2: Variance

Consider the random variable X characterized by the following PMF:

x	$f(x)$
1	.2
2	.7
3	.1

(2a) Confirm that the variance of X is the same whether we compute it by the formula in Definition 2.1.12 or the Alternative Formula in Theorem 2.1.13. (We want the two variance computations, not the proof.)

Answer:

(2b) Write an `R` function to compute the variance of any discrete random variable. The arguments to your function should include the values the random variable can take on (`x`) and the probability it takes on each value (`probs`). Use your function to confirm your answer from question (2a).

Answer:

```
# your code here
```

(2c) Explain what definition/property/mathematical operation is being used in each step of the following proof.

For random variable X and $a \in \mathbb{R}$, $\mathrm{V}[aX] = a^2\mathrm{V}[X]$,

$$\begin{aligned}
\mathrm{V}[aX] &= \mathrm{E}\left[(aX - \mathrm{E}[aX])^2\right] && \text{(Step 1)}\\
&= \mathrm{E}\left[(aX - a\mathrm{E}[X])^2\right] && \text{(Step 2)}\\
&= \mathrm{E}\left[a^2(X - \mathrm{E}[X])^2\right] && \text{(Step 3)}\\
&= a^2\mathrm{E}\left[(X - \mathrm{E}[X])^2\right] && \text{(Step 4)}\\
&= a^2\mathrm{V}[X] && \text{(Step 5)}
\end{aligned}$$

Answer:

In Step 1, . . .

In Step 2, . . .

In Step 3, . . .

In Step 4, . . .

In Step 5, . . .

(2d) Use your variance function to confirm that $\mathrm{V}[aX] = a^2\mathrm{V}[X] \forall a \in \mathbb{R}$. (You can just show this for one value of a.)

Answer

(2e) Use your function to confirm that $V[X + c] = V[X] \,\forall\, c \in \mathbb{R}$. (You can just show this for one value of c.)

Answer