

# Problem set 3: Summarizing distributions

Due October 16, 2023, at 9pm

(Your name here)

NOTE: Start with the file `ps3_2023_summarizing_distributions.qmd` (available from the github repository at <https://github.com/UChicago-pol-methods/IntroQSS-F23/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.

$x$	$f(x)$
-1	.3
2	.2
3	.5

(1a) Compute  $E[X]$ . Show your work.

**Answer:**

$$E[X] =$$

(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 ( $E[(X - c)^2]$ ) 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 R code here
```

(1e) Create a vector `cs` that contains numbers in the sequence 1.0, 1.1, 1.2, . . . , 3. 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 `mse`.

**Answer:**

```
# your R code here
```

(1f) Using the `plot()` command, make a scatterplot showing the MSE of  $X$  (vertical axis) at each value of  $c \in \{1, 1.1, 1.2, \dots, 3\}$  (horizontal axis). Use `abline()` to add a vertical line at  $E[X]$ .

**Answer:**

```
# your R code here
```

## Question 2: Variance

Consider the random variable  $X$  characterized by the following PMF:

$x$	$f(x)$
0	.6
1	.1
2	.3

(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 R code here
```

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

For random variable  $X$  and  $c \in \mathbb{R}$ ,  $V[X + c] = V[X]$ ,

$$\begin{aligned} V[X + c] &= E[(X + c - E[X + c])^2] && \text{(Step 1)} \\ &= E[(X + c - E[X] - c)^2] && \text{(Step 2)} \\ &= E[(X - E[X])^2] && \text{(Step 3)} \\ &= V[X] && \text{(Step 4)} \end{aligned}$$

**Answer:**

In Step 1, . . .

In Step 2, . . .

In Step 3, . . .

In Step 4, . . .

(2d) Use your variance 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:**

```
# your R code here
```

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

**Answer:**

```
# your R code here
```