Problem set 4: More summarizing distributions

Due October 23, 2023, at 9pm

(Your name here)

NOTE: Start with the file `ps4_2023_more_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.

The entire problem set refers to the following joint distribution of two random variables X and Y:

$$f(x,y) = \begin{cases} 1/2 & x=0, y=0 \\ 1/6 & x=1, y=0 \\ 1/3 & x=1, y=1 \\ 0 & \text{otherwise} \end{cases}$$

Part 1: Covariance and correlation

(1a) Compute the covariance of X and Y. Show your work.

(1b) Compute the correlation of X and Y. Show your work. (There is no need to repeat calculations from (1a).)

(1c) Write a function that takes as arguments `x` (a vector of possible values of a random variable X), `y` (a vector of possible values of a random variable Y), and `fxy` (a vector of frequencies $f(x,y)$ for each combination of x and y) and returns $\text{Cov}[X,Y]$. Use it to confirm the covariance calculation you did above.

```
# your code here
```

(1d) Write a function with the same arguments that computes correlation between two random variables X and Y. Use it to confirm the correlation calculation you did above.

```
# your code here
```

Part 2: Conditional expectations, LIE, and BLP

(2a) What is $E[Y \mid X=0]$? What is $E[Y \mid X=1]$?

(2b) Show that $E[Y] = E[E[Y \mid X]]$, i.e. that the law of iterated expectations holds in this case.

(2c) What is the best linear predictor (BLP) of Y given X? Express your answer in terms of an intercept α and a slope β.

(2d) How close does the BLP of Y given X come to the conditional expectation function (CEF) of Y given X?

Part 3: Law of total variance

(3a) What is $V[Y \mid X=0]$? What is $V[Y \mid X=1]$?

(3b) Compute the two components of the Law of Total Variance (i.e. Eve's law) and confirm that they add up to $V[Y]$.