

# Homework 06: Expected Value and Variance

---

**Name:** CJ Kennedy

---

This assignment is due on Canvas by **6:00PM on Friday October 7**. Your solutions to theoretical questions should be done in Markdown directly below the associated question. Your solutions to computational questions should include any specified Python code and results as well as written commentary on your conclusions. Remember that you are encouraged to discuss the problems with your classmates, but **you must write all code and solutions on your own**.

## NOTES:

- Any relevant data sets should be available in the Homework 01 assignment write-up on Canvas. To make life easier on the grader if they need to run your code, do not change the relative path names here. Instead, move the files around on your computer.
  - If you're not familiar with typesetting math directly into Markdown then by all means, do your work on paper first and then typeset it later. Remember that there is a [reference guide](#) linked on Canvas on writing math in Markdown. **All** of your written commentary, justifications and mathematical work should be in Markdown.
  - Because you can technically evaluate notebook cells in a non-linear order, it's a good idea to do **Kernel → Restart & Run All** as a check before submitting your solutions. That way if we need to run your code you will know that it will work as expected.
  - It is **bad form** to make your reader interpret numerical output from your code. If a question asks you to compute some value from the data you should show your code output **AND write a summary of the results** in Markdown directly below your code.
  - This probably goes without saying, but... For any question that asks you to calculate something, you **must show all work and justify your answers to receive credit**. Sparse or nonexistent work will receive sparse or nonexistent credit.
- 

Import Pandas, NumPy, and matplotlib.pyplot.

```
In [131... import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

## Problem 1

---

Consider a continuous random variable  $X$  with a PDF given by:

$$f(x) = \begin{cases} \frac{1}{10}(1 - \frac{1}{20}x) & \text{if } 0 \leq x \leq 20 \\ 0 & \text{elsewhere} \end{cases}$$

Furthermore, consider  $Y$  to be a random variable with values  $Y = 3X + 5$ .

## Part A

**(2 points)** Show that  $f(x)$  is a valid PDF.

**solution:** Put your solution to Part A here:

We must show:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= 1 \\ \Rightarrow \int_0^{20} f(x)dx &= \int_0^{20} \frac{1}{10}(1 - \frac{1}{20}x)dx \\ &= \frac{x}{10} - \frac{x^2}{400} \Big|_0^{20} \\ &= \frac{20}{10} - \frac{400}{400} = 2 - 1 = 1 \end{aligned}$$

Furthermore, it is evident that  $f(x) \geq 0$  for all of  $x$  from 0 to 20:

$$f(0) = 0.1$$

$$f(20) = 0$$

## Part B

**(2 points)** Find  $E[X]$ .

**solution:** Put your solution to Part B here:

We are going to use a substitution with  $u$  and  $du$  instead of reusing  $x$ :

$$\begin{aligned} E[X] &= \int_0^{20} uf(u)du \\ &= \int_0^{20} u \frac{1}{10}(1 - \frac{1}{20}u)du \\ &= \left( \frac{u^2}{20} - \frac{u^3}{600} \right) \Big|_0^{20} \\ &= \frac{20^2}{20} - \frac{20^3}{600} = \frac{20}{3} \\ \therefore E[X] &= \frac{20}{3} \end{aligned}$$

## Part C

**(2 points)** Find  $Var(X)$ .

**solution:** Put your solution to part C here:

$$\text{Var}(X) = E[X^2] - (E[X])^2$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx$$

$$\Rightarrow \int_0^{20} u^2 f(u) du$$

$$= \int_0^{20} \frac{u^2}{10} \left(1 - \frac{u}{20}\right) du$$

$$= \left( \frac{u^3}{30} - \frac{u^4}{800} \right) \Big|_0^{20}$$

$$= \frac{20^3}{30} - \frac{20^4}{800} = \frac{200}{3}$$

$$\text{Var}(X) = \frac{200}{3} - \frac{400}{9}$$

$$\therefore \text{Var}(X) = \frac{200}{9}$$

## Part D

**(2 points)** Find  $E[Y]$ .

**solution:** Put your solution to part D here:

Recall  $Y = 3X + 5$ . Then,

$$E[Y] = E[3X + 5]$$

We can distribute:

$$E[Y] = 3E[X] + 5$$

From, part B:

$$E[Y] = 3 * \frac{20}{3} + 5$$

$$\therefore E[Y] = 25$$

## Part E

**(2 points)** Find  $\text{Var}(Y)$ .

**solution:** Put your solution to part E here:

Note the relation:

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Then,

$$\text{Var}(Y) = \text{Var}(3X + 5) = 3^2 \text{Var}(X) = 9 \text{Var}(X)$$

$$\therefore \text{Var}(Y) = 200$$

## Problem 2

---



Consider the American roulette wheel as pictured above.

### Part A

Suppose you decide to bet on red23 over and over and over again UNTIL you win, then you'll stop playing.

Winning means the ball lands on the red23 slot - and it has an equal chance of landing in any of the slots.

So, you might play only once (if you win in one game.) Or, perhaps you'll play twice, or three times, etc. before you win.

Let  $X$  be the random variable, "Number of times you play till you win."

**(2 points)** How many times do you expect to play before you win?

**solution:** Put your answer to Part A here:

There are 38 total possible outcomes on the roulette wheel. The probability of winning,  $p$ , is  $\frac{1}{38}$ .

Thus, we would expect to play the game 38 times until winning.

### Part B

**(3 points)** What is the probability that you don't win until your 4th attempt?

**solution:** Put your answer to Part B here:

$$P(X = 4) = (1 - p)^{k-1}p$$

$$P(X = 4) = \frac{109}{4487}$$

### Part C

Recall from Calculus that a geometric series with ratio  $r$  diverges if  $|r| > 1$ , but if  $0 < |r| < 1$  then the series converges:

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}.$$

Therefore, for  $0 < p < 1$ , we have  $\sum_{k=0}^{\infty} p(1-p)^k = p \cdot \frac{1}{1-(1-p)} = 1$ .

**(5 points)** Explain (show) why  $E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}$ .

**solution:** Put your answer to Part C here:

Denote the variable  $q$  as  $1-p$ :

$$E[X] = \sum_{k=1}^{\infty} kp(q)^{k-1}$$

We know  $\frac{d}{dq}(q)^k = k(q)^{k-1}$ .

$$E[X] = p \frac{d}{dq} (\sum_{k=1}^{\infty} (q)^k)$$

We are given the convergence  $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$  which allows us to write:

$$\sum_{n=0}^{\infty} ar^n = \frac{ar}{1-r}$$

So,  $a = 1$  and  $r = q$

$$E[X] = p \frac{d}{dq} \left( \frac{q}{1-q} \right) = p \frac{q'(1-q) - (1-q)'q}{(1-q)^2}$$

$$= p \frac{(1-q) + q}{(1-q)^2}$$

We can substitute  $p$ :

$$E[X] = p \frac{1}{p^2} = \frac{1}{p}$$

## Part D

**(3 points)** What is  $E[X]$  ?

**solution:** Put your answer to Part D here:

Applying Part C's answer:

$$E[X] = \frac{1}{p}$$

$$\therefore E[X] = 38$$

## Part E

You are interested in knowing how much money you should expect to win **each time you play**. Afterall, you have decided to play over and over again till you win.

Now, let  $X$  be the random variable, "Amount of money you win."

You are still betting on 23red, and each bet costs \$1.

If you lose, you lose your dollar.

If you win, you get your dollar back **and** you get an additional \$35 for winning.

**(3 points)** What is  $E[X]$  ?

**solution:** Put your answer to Part E here:

We know  $E[X] = \sum_i a_i \cdot P(X = a_i)$ .

So, we are able to write:

$$E[X] = p \cdot 35 - q \cdot 1$$

$$= \frac{35}{38} - \frac{37}{38}$$

$$= -\frac{2}{38}$$

Thus, the expected winnings each time played is  $-\frac{2}{38}$  dollars.

## Problem 3

---

**Using the definition** of expected value and variance,

Discrete:  $E[X] = \sum_i a_i \cdot P(X = a_i)$ .

Continuous:  $E[X] = \int_{-\infty}^{\infty} x f(x) dx$

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Find the following:

### Part A

**(3 points)** Suppose  $X \sim U[\alpha, \beta]$ , find  $E[X]$ .  $X$  takes on all real values between  $\alpha$  and  $\beta$ .

**solution:** Put your answer to Part A here:

The uniform distribution can be written:

$$f(x) = \begin{cases} \frac{1}{\alpha-\beta} & \text{if } \alpha \leq x \leq \beta \\ 0 & \text{elsewhere} \end{cases}$$

We are given:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

So,

$$\begin{aligned} E[X] &= \int_{\alpha}^{\beta} x \left( \frac{1}{\alpha-\beta} \right) dx \\ &= \frac{x^2}{2} \frac{1}{\alpha-\beta} \Big|_{\alpha}^{\beta} = \frac{(\beta^2 - \alpha^2)}{2} \frac{1}{\alpha-\beta} \\ &= \frac{(\beta-\alpha)(\beta+\alpha)}{2} \frac{1}{\alpha-\beta} = \frac{(\alpha-\beta) \cdot -(\beta+\alpha)}{-2(\alpha-\beta)} \\ &= \frac{-(\beta+\alpha)}{-2} = \frac{\beta+\alpha}{2} \\ \therefore E[X] &= \frac{\beta+\alpha}{2} \end{aligned}$$

## Part B

**(3 points)** Suppose  $X \sim U[\alpha, \beta]$ , find  $Var[X]$ .

**solution:** Put your answer to Part B here:

We are given  $Var[X] = E[X^2] - (E[X])^2$ .

So we need to find  $E[X^2]$ :

$$\begin{aligned} E[X^2] &= \int_{\alpha}^{\beta} x^2 \left( \frac{1}{\alpha-\beta} \right) dx \\ &= \frac{x^3}{3} \frac{1}{\alpha-\beta} \Big|_{\alpha}^{\beta} = \frac{(\beta^3 - \alpha^3)}{3} \frac{1}{\alpha-\beta} \\ &= \frac{(\beta-\alpha)(\beta^2 + \beta\alpha + \alpha^2)}{3} \frac{1}{\alpha-\beta} \\ &= \frac{(\alpha-\beta) \cdot -(\beta^2 + \beta\alpha + \alpha^2)}{-3(\alpha-\beta)} \\ &= \frac{-(\beta^2 + \beta\alpha + \alpha^2)}{-3} \\ E[X^2] &= \frac{(\beta^2 + \beta\alpha + \alpha^2)}{3} \end{aligned}$$

$$\begin{aligned} Var[X] &= E[X^2] - (E[X])^2 = \frac{(\beta^2 + \beta\alpha + \alpha^2)}{3} - \left( \frac{\beta+\alpha}{2} \right)^2 \\ &= \frac{(\beta^2 + \beta\alpha + \alpha^2)}{3} - \frac{(\beta^2 + 2\beta\alpha + \alpha^2)}{4} \\ &= \frac{4(\beta^2 + \beta\alpha + \alpha^2)}{12} - \frac{3(\beta^2 + 2\beta\alpha + \alpha^2)}{12} \end{aligned}$$

$$= \frac{\beta^2 - 2\alpha\beta + \alpha^2}{12}$$

$$\therefore \text{Var}[X] = \frac{\beta^2 - 2\alpha\beta + \alpha^2}{12}$$

## Part C

**(3 points)** Suppose  $X \sim \text{Ber}(p)$ , find  $E[X]$ .

**solution:** Put your answer to Part C here:

The bernoulli distribution can be written:

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{if } x = (0, 1) \\ 0 & \text{elsewhere} \end{cases}$$

We use the discrete version:  $E[X] = \sum_i a_i \cdot P(X = a_i)$ .

$$E[X] = \sum_i a_i \cdot (p^{a_i} \cdot (1-p)^{1-a_i})$$

We only need to evaluate the summation for 0 and 1:

$$E[X] = 0(p^0(1-p)^1) + 1(p(1-p)^0) = p$$

$$\therefore E[X] = p$$

## Part D

**(3 points)** Suppose  $X \sim \text{Ber}(p)$ , find  $\text{Var}[X]$ .

**solution:** Put your answer to Part D here:

We know we need to find the second moment:  $E[X^2] = \sum_i a_i^2 \cdot P(X = a_i)$ .

$$\begin{aligned} E[X^2] &= \sum_i a_i^2 \cdot (p^{a_i} \cdot (1-p)^{1-a_i}) \\ &= 0^2(p^0(1-p)^1) + 1^2(p(1-p)^0) = p \end{aligned}$$

Then, the variance:

$$\text{Var}[X] = E[X^2] - (E[X])^2 = p - p^2$$

$$\therefore \text{Var}[X] = p - p^2$$

## Problem 4

---

Consider a card game played with a standard deck of 52 cards.

The cards are shuffled, a card is chosen, recorded, and returned to the deck.



This is done three times and the record of three choices is observed.

The game costs \$1 to play.

If all three cards have the same number, then you get your dollar back plus \$3.

If you only have two cards with the same number, then you get your dollar back plus \$2.

If all three cards are of the same suit, then you get your dollar back plus \$2.

Examples:

Ace of spades, Ace of clubs, 5 of diamonds: dollar back plus \$2.

Ace of spaces, 2 of spades, 7 of spades: dollar back plus \$2.

5 of diamonds, 8 of diamonds, 8 of diamonds: dollar back plus \$2, plus \$2.

4 of hearts, 4 of hearts, 4 hearts: dollar back plus \$3, plus \$2.

6 of diamonds, 6 of hearts, 6 of clubs: dollar back plus \$3.

3 of hearts, 5 of diamonds, Queen of clubs: Lose your dollar.

**(7 points)** Write a function or functions that will create a random draw of three cards as described above. Simulate this game (at least 1000 times) and determine from the simulation the expected winnings per dollar of this game.

```
In [164... # Put your code for Problem 4 here:
def pick_card(): # choose 1-13 card, choose 1-4 possible suit
    numbers = np.array([str(ii) for ii in range(1,13+1)])
    number = np.random.choice(numbers)
    suits = np.array(["Hearts", "Diamonds", "Spades", "Clubs"])
    suit = np.random.choice(suits)
    #print(number+" of "+suit)
    return number,suit;

def earnings(money):
    money = money - 1; # subtract cost
    num1,suit1 = pick_card() # pick 3 cards
    num2,suit2 = pick_card()
    num3,suit3 = pick_card()
    cond1 = 0; cond2 = 0; cond3 = 0; # preallocate return conditions
    if(num1==num2==num3): # all card numbers are same
        cond1 = (3+1)
    if(num1==num2 or num1==num3 or num2==num3): # 2 card numbers are same
        cond2 = (2+1)
    if(suit1==suit2==suit3): # all suits are same
        cond3 = (2+1)
    money = money+(cond1+cond2+cond3) # return money
    return money

money = 1 # preallocate 1$
n = 10000;
```

```
for i in range(0,n): # 10000 sims
    money = earnings(money);

result = money/n # calculate money/game
result = "{:.3f}".format(result)
print("The expected winnings per game is $", result, ".")
```

The expected winnings per game is \$ -0.130 .

## Rubric Check

---

**(5 points)** Make sure your answers are thorough but not redundant. Explain your answers, don't just put a number. Make sure you have matched your questions on Gradescope. Make sure your PDF is correct and your LaTeX is correct. etc. etc. BE NEAT.

In [ ]: