Unit 3 Homeowrk: Summarizing Distributions

w203: Statistics for Data Science, Spring 2022, Section 6 - S. Williams

This homework, like all homework, is due at the time of your live session.

Applied Practice

(6 points total) Best Game in the Casino

You flip a fair coin 3 times, and get a different amount of money depending on how many heads you get.

- For 0 heads, you get $0.
- For 1 head, you get $2.
- For 2 heads, you get $4.

Another piece of information is also flashing at you as you stand before the lights:

- Overall, across all possible outcomes your expected winnings from the game are $6.

1. (3 points) How much do you get paid if the coin comes up heads 3 times?

P(h) = (0,0,0)(0,0,1)(0,1,0)(1,0,0)(1,1,0)(1,0,1)(0,1,1)(1,1,1)

P(X=$)	P(h)	Calculated P(h)	P(h)*P(X=$)
P(x=$0)	P(0 heads)	0.125	0
P(x=$2)	P(1 heads)	0.375	0.75
P(x=$4)	P(2 heads)	0.375	1.5
P(x=$d)	P(3 heads)	0.125	0.125*d

$$E[X] = 6$$
$$E[6] = 0 + 0.75 + 1.5 + (0.125 * d)$$
$$E[6] = 2.25 + (0.125 * d)$$
$$3.75 = 0.125 * d$$
$$d = \frac{3.75}{0.125} = 30$$

2. (3 points) Write down a complete expression for the cumulative probability function for your winnings from the game.

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \le x < 2 \\ \frac{1}{2}, & 2 \le x < 4 \\ \frac{7}{8}, & 4 \le x < 30 \\ 1, & x > 30 \end{cases}$$

(3 points total) The Warranty is Worth It

Suppose the life span of a particular (shoddy) server is a continuous random variable, T, with a uniform probability distribution between 0 and 1 year. The server comes with a contract that guarantees you money if the server lasts less than 1 year. In particular, if the server lasts t years, the manufacturer will pay you $g(t) = \$100(1-t)^{1/2}$. Let $X = g(T)$ be the random variable representing the payout from the contract.

1. (3 points) Compute the expected payout from the contract, $E(X) = E(g(T))$. Given the nature of the function, you might have to use integration by parts, or help from an integral solver.

$$f_T(t) = \begin{cases} 1, & 0 \le t \le 1 \\ 0, & otherwise \end{cases}$$

$$E[x] = \int_0^1 g(t)f(t)dt = \int_0^1 100(1-t)^{\frac{1}{2}}dt = \frac{200}{3} = 66.67$$

Great Time to Watch Async

Suppose your waiting time in minutes for the Caltrain in the morning is uniformly distributed on $[0,5]$, whereas waiting time in the evening is uniformly distributed on $[0,10]$. Each waiting time is independent of all other waiting times.

$E[X] = \frac{1}{2}(a+b)$ Or $E[X] = \frac{a+b}{2}$ where a=minimum of the uniform distribution and b=maximum of the uniform distribution.

a. (3 points) If you take the Caltrain each morning and each evening for 5 days in a row, what is your total expected waiting time?

$$E[X] = \frac{0+5}{2} = 2.5$$

$$E[Y] = \frac{0+10}{2} = 5$$

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

$$E[X+Y] = (5 \times 2.5) + (5 \times 5)$$

$$E[X+Y] = 12.5 + 25 = 37.5$$

b. (3 points) What is the variance of your total waiting time?
For a uniform random variable $V[X] = \frac{1}{12}(b-a)^2$ Or $V[X] = \frac{(b-a)^2}{12}$ where a=minimum of the uniform distribution and b=maximum of the uniform distribution.

$$V[X] = \frac{(5-0)^2}{12} = \frac{25}{12} = 2.083$$

$$V[Y] = \frac{(10-0)^2}{12} = \frac{100}{12} = 8.33$$

$$V[5X+5Y] = (5 \times 2.083) + (5 \times 8.33) = 10.415 + 41.65 = 52.07$$

c. (3 points) What is the expected value of the difference between the total evening waiting time and the total morning waiting time over all 5 days?

$$E[5Y-5X] = 5E[Y] - 5E[X] = (5 \times 5) - (5 \times 2.5) = 25 - 12.5 = 12.5$$

d. (3 points) What is the variance of the difference between the total evening waiting time and the total morning waiting time over all 5 days?

$$V[5Y-5X] = 5V[Y] - 5V[X] = (5 \times 8.33) - (5 \times 2.083) = 41.65 - 10.415 = 31.235$$