W203 Statistics for Data Science

Aditya Mengani

aditya.mengani@berkeley.edu

Monday, 4pm

Summer 2020

Unit 2 Homework: Discrete Random Variables

May 10, 2020

1. 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. Your expected winnings from the game are \$6.

1. (a) How much do you get paid if the coin comes up heads 3 times? (3 points)
2. (b) Write down a complete expression for the cumulative probability function for your win- nings from the game. (3 points)

olution:-

1 (a). Given a fair coin is flipped 3 times, the total no. of permutations can be written as $2 \times 2 \times 2 = 8$ (Since a fair coin has an equal probability of getting an head or tail)

According to pascal's triangle, a probability of 3 heads can be written as $P\{H, H, H\} = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

Similarly we can derive for other possibilities as below:

$P\{T, T, T\} = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

$P\{1 \text{ head}, 2 \text{ tails}\} = 3\left[\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right] = \frac{3}{8}$; outcomes being $\{HTT, TTH, THT\}$

$P\{2 \text{ head}, 1 \text{ tail}\} = 3\left[\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right] = \frac{3}{8}$; outcomes being $\{HHT, HTH, T, HH\}$

$P\{H, H, H\} = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ ——— (a)

Given from problem, $\{T, T, T\}$ you get 0 \$

$\{1 \text{ head}, 2 \text{ tails}\}$ you get 2\$

$\{2 \text{ head}, 1 \text{ tail}\}$ you get 4 \$

$\{3 \text{ heads}\}$ find ?

We can denote a random variable X, for the amount of dollars won

~~number of outcomes of heads in an event, the possible values of $x \in X$ can be $\{0,1,2,3\}$~~

Given x, we have already found $P(x)$ earlier, from (a)

We can combine all of them into a table,

	0	1	2	3
$P(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
Amount (X)	0\$	2\$	4\$	y

— (c)

we already know from problem statement, ~~for getting heads~~ is 6\$ $E(X)$, where X is no. of winnings

$E(X) = 6$

from the formula of $E(X)$, we know that,

$$\boxed{E(X) = \mu_x = \sum_{x \in D} x \cdot p(x) \quad \text{where } D \text{ is set of possible values and } x \in X}$$ — (b)

Let y be the amount in \$, recived for the event of flipping a fair coin and getting head 3 times

substituting (c) into (b) we get,

$$\frac{1}{8} \times 0 + \frac{3}{8} \times 2 + \frac{3}{8} \times 4 + \frac{1}{8} \times y = 6$$

$$\Rightarrow \frac{6+12}{8} + \frac{y}{8} = 6 \Rightarrow \frac{y}{8} = 6 - \frac{9}{4} = \frac{15}{4} \Rightarrow y = \frac{15}{4} \times 8 = 30\$$$

(thirty dollars)

Thus, if the expected winnings from a game is 6\$, then you will have to be paid 30\$, if coin comes up heads 3 times (b). Expression for C.P.F. for winnings from the game:

The Cumulative distribution function (cdf) $F(x)$ of a discrete random variable X with pmf $p(x)$ is defined for every number x by

$$F(x) = P(X \le x) = \sum_{y:\, y \le x} p(y)$$

for any number x, $F(x)$ is the probability that the observed value of X will be atmost x

Let Y = amount of dollars won

y	0	2	4	30
$P(y)$	1/8	3/8	3/8	1/8

Lets determine $F(y)$ for all possible values of Y:

$F(0) = P(Y \le 0) = P(Y=0) = P(0) = 1/8$

$F(2) = P(Y \le 2) = P(Y = 0 \text{ or } 2) = P(0) + P(2) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$

$F(4) = P(Y \le 4) = P(Y = 0 \text{ or } 2 \text{ or } 4) = P(0) + P(2) + P(4)$
$= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$

$F(30) = P(Y \le 30) = P(Y = 0 \text{ or } 4 \text{ or } 30) = P(0) + P(4) + P(30) + P(2)$
$= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$

Thus the cumulative probability function for winnings from game, is defined by F(y)

$$F(y) = \begin{cases} 0 & y < 0 \\ 1/8 & 0 \le y < 2 \\ 1/2 & 2 \le y < 4 \\ 7/8 & 4 \le y < 30 \\ 1 & 30 \le y < \infty \end{cases}$$

2. Reciprocal Dice
Let X be a random variable representing the outcome of rolling a 6-sided die. Before the die

is rolled, you are given two options:

(a) You get 1/E(X) in dollars right away. (3 points)
(b) You wait until the die is rolled, then get 1/X in dollars. (3 points)

(c) Which option is better for you, in expectation? (1 point)

2. (a). Let X be the random variable representing outcome of rolling six sided die:
Assuming the die has a fair probability:
It can be written as:

x	1	2	3	4	5	6
$P(x)$	1/6	1/6	1/6	1/6	1/6	1/6

$$E(X) = 1\times\frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 6\times\frac{1}{6}$$

$$= \frac{1+2+3+4+5+6}{6} = \frac{1}{6}\times\frac{6(7)}{2} = 3.5$$

from option (a), we get $\frac{1}{E(X)}$ in dollars right away

$$= \frac{1}{3.5} = \frac{2}{7} = 0.28\,\$$$

2. (b). wait until die is rolled, get $\frac{1}{X}$ in dollars.

Let Y be a random variable denoting

$$Y = h(X) = \frac{1}{X}$$

The pmf of X and derived pmf of Y are as follows:

x	1	2	3	4	5	6
$P(x)$	1/6	1/6	1/6	1/6	1/6	1/6

$\Rightarrow$

y	1	1/2	1/3	1/4	1/5	1/6
$P(y)$	1/6	1/6	1/6	1/6	1/6	1/6

$$E(Y) = E[(h(X)] = \sum_{D} y \cdot P(y)$$

$$= 1 \times \frac{1}{6} + \frac{1}{2} \times \frac{1}{6} + \frac{1}{3} \times \frac{1}{6} + \frac{1}{4} \times \frac{1}{6} + \frac{1}{5} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6}$$

$$= \frac{1}{6}\left[1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}\right]$$

$$= \frac{1}{6}\left[\frac{60+30+20+15+\cancel{12}+10}{60}\right] = \frac{1}{6} \times \frac{147}{60} = 0.40\,\$$$

2 (c). In expectation terms, option (b) which shows $E(Y) = 0.40\$$ is better compared to option (a) where $E(X) = 0.28\$$

hence ~~is~~ waiting until the die is rolled and then get $E(\frac{1}{X}) = 0.40\$$ is better

3. The Baseline for Measuring Deviations

Given any random variable X and a real number t, we can define another random variable Y = $(X - t)^2$. In other words, for any random variable X, we can choose a real number, t, as a baseline and calculate the squared deviation of X away from t.

You might wonder why we often square deviations (instead of taking an absolute value, or cubing them, etc.). This exercise will shed some light on why this is a natural choice.

1. (a) Write down an expression for E(Y) and simplify it as much as you can. Even though we haven't proved this yet, you can use the fact that for any two random variables, A and B, E(A + B) = E(A) + E(B). (3 points)
2. (b) Taking a partial derivative with respect to t, compute the value of t that minimizes E(Y). (Hint: Your answer should be a very familiar value) (3 points)
3. (c) What is the value of E(Y) for this choice of t? (3 points) (Hint: this should also be a very familiar value)

Given rv $Y = (X-t)^2$

3.(a). Find an expression for $E(Y)$

assume $Y = h(X) = (X-t)^2$

we know, $E(Y) = E[h(x)] = \cancel{\sum_{D} y \cdot p(y)}$

$$= \sum_{D} (X-t)^2 \cdot p(x)$$

$$\boxed{E(Y) = \sum_{D} (x^2 + t^2 - 2xt) \cdot p(x)}$$

where 'D' has a set of all possible values in X and $p(x)$ is pmf

$$= \underbrace{\sum_{D} x^2 \cdot p(x)}_{E(X^2)} + t^2 \underbrace{\sum_{D} p(x)}_{1} - 2t \underbrace{\sum_{D} x \cdot p(x)}_{E(X)}$$

(this can be written as expectation of X^2)

1 (Since sum of all probabilities = 1)

Thus the equation simplifies as;

$$\boxed{E(Y) = E(X^2) + t^2 - 2t \cdot E(X)}$$

3(b). By taking a partial derivate to t, the above expression can be written as,

$$\frac{\partial E(Y)}{\partial t} = \underbrace{\frac{\partial}{\partial t} \cdot E(X^2)}_{0} + \frac{\partial}{\partial t} \cdot t^2 - 2 \cdot \frac{\partial}{\partial t} \cdot t \cdot E(X)$$

we knew that $\frac{\partial E(X^2)}{\partial t} = 0$, as we are doing partial derivate with respect to t, and that case $E(X^2)$ would be a constant in

~~similarly~~

$$= 2t - 2 \cdot E(X) \cdot \frac{\partial}{\partial t} t$$

$$\boxed{E'(Y) = \frac{\partial E(Y)}{\partial t} = 2t - 2E(X)}$$

~~Assuming this is a linear expression of the form y = mx+c~~

we can find minimum value of y, at a point where derivative = 0

$$\Rightarrow 2t - 2E(X) = 0$$

$$\Rightarrow \boxed{t = E(X)}$$

3) c. Substituting $t = E(X)$ from solution (b) in to expression (a)

$$E(Y) = E(x^2) + t^2 - 2t \cdot E(X)$$

$$= E(x^2) + [E(x)]^2 - 2 \cdot E(x) \cdot E(x)$$

$$= E(x^2) + [E(x)]^2 - 2[E(x)]^2$$

$$E(Y) = E(x^2) - [E(x)]^2$$

this is ~~[illegible]~~ also the variance of X represented as $V(X)$

Thus $\boxed{E(Y) = V(X) = E(x^2) - [E(x)]^2}$

⇒ for a constant 't' and random variable X, expectation of random variable Y, which is represented as $Y = (X-t)^2$, is equal to

$$\sigma^2 = \text{variance of } X = E(x^2) - (E(x))^2$$

4. Optional Advanced Exercise: Heavy Tails

One reason to study the mathematical foundation of statistics is to recognize situations where common intuition can break down. An unusual class of distributions are those we call heavy- tailed. The exact definition varies, but we'll say that a heavy-tailed distribution is one for which not all moments are finite. Consider a random variable M with the following pmf:

? $c/x^3, x \in \{1, 2, 3, ...\}$ $p_M(x) =$ 0, otherwise.

where c is a constant (you can calculate its value if you like, but it's not important). (a) Is E(M) finite? (Bonus +3 points)

(b) Is V (M) finite? (Bonus +3 points)

Heavy-tailed distributions may seem odd, but they're not as rare as you might suspect. Researchers argue that the distribution of wealth is heavy-tailed; so is the distribution of computer file sizes, insurance payouts, and area burned by forest fires. These random variables are prob- lematic in that a lot of common statistical techniques don't work on them. For this class, we'll assume that all of our variables don't have heavy-tails.

Note: Maximum score on any homework is 100%

(a). Is E(M) finite

$$P(M) = \begin{cases} c/x^3, & x \in \{1, 2, 3 \ldots\} \\ 0, & \text{otherwise} \end{cases}$$

From the pmf we can choose c so that $\sum_{x=1}^{\infty} (c/x^3) = 1$

The expected value of M is

$$\mu = E(M) = \sum_{x=1}^{\infty} x \cdot P(M) = \sum_{x=1}^{\infty} x \cdot \left(\frac{c}{x^3}\right) = c \sum_{x=1}^{\infty} \frac{1}{x^2}$$

$$= c\left[1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} \cdots\right]$$

$$E(M) = c\left[1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} \cdots\right]$$

~~Exp~~ the sum of reciprocal of squares of integers converges to a finite value

$$\boxed{\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}} \quad \text{———} \quad (e)$$

Thus $\boxed{E(M) = c \cdot \frac{\pi^2}{6}}$ which is a finite value — (e)

(b) Is $V(M)$ finite?

$$\sigma^2 = V(M) = \sum_{x=1}^{\infty} (M-\mu)^2 \cdot p(M)$$

$$= E(M^2) - [E(M)]^2 \quad \text{———} \quad (f)$$

substituting (e) in (f)

$$\sigma^2 = V(M) = \underbrace{E(M^2)}_{\text{Need to calculate}} - \underbrace{[(\pi^2 \cdot c)/6]^2}_{\text{constant value}} \quad \text{———} \quad (g)$$

$$E(M^2) = \sum_{x=1}^{\infty} x^2 \cdot p(M) = \sum_{x=1}^{\infty} x^2 \cdot \left(\frac{c}{x^3}\right) = c \sum_{x=1}^{\infty} x^2 \times \frac{1}{x^3}$$

$$= c \sum_{x=1}^{\infty} \left[\frac{1}{x}\right]$$

$$= c\left[1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \cdots\cdots\right]$$

Thus harmonic ~~...~~ series is resulted above

$$E(M^2) = c\left[1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\cdots\right] \quad \text{—— (h)}$$

The summation of reciprocal of integers is diverging ~~...~~ ~~...~~ in nature and leads to infinity

substituting (h) to (g)

$$E(M^2) = c\underbrace{\left[1+\frac{1}{2}+\frac{1}{3}+\cdots\right]}_{\text{Infinite}} - \underbrace{\left[(\pi^2 \times c)/6\right]^2}_{\text{constant}}$$

$$= \infty$$

Thus $E(M^2)$ is infinite in nature,

where as $E(M)$ is finite