

# Statistics for Data Science

## Unit 4 Part 1 Homework: Discrete Random Variables

May 25, 2019

### 1. Best Game in the Casino<sup>†</sup>

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.

#### GIVENS

$X$  is a binomial random variable based on  $n$  trials with success probability  $p$ .

Let  $x$  be the number of H's among the  $n = 3$  trials.

$$b(x; n, p) = \begin{cases} \binom{n}{k} p^x (1-p)^{n-x}, & x \in \{0, 1, 2, 3, \dots\} \\ 0, & \text{otherwise.} \end{cases}$$

$$\text{Probability of 0 Heads: } P(X = 0) = b(0; 3, \frac{1}{2}) = \binom{3}{0} (\frac{1}{2})^3 = \frac{1}{8}$$

$$\text{Probability of 1 Heads: } P(X = 1) = b(1; 3, \frac{1}{2}) = \binom{3}{1} (\frac{1}{2})^3 = \frac{3}{8}$$

$$\text{Probability of 2 Heads: } P(X = 2) = b(2; 3, \frac{1}{2}) = \binom{3}{2} (\frac{1}{2})^3 = \frac{3}{8}$$

$$\text{Probability of 3 Heads: } P(X = 3) = b(3; 3, \frac{1}{2}) = \binom{3}{3} (\frac{1}{2})^3 = \frac{1}{8}$$

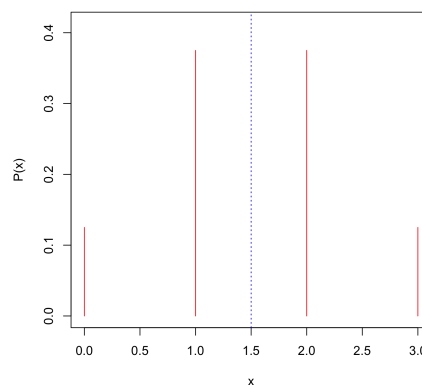


Figure 1: Plot of binomial distribution of  $n=3$ ,  $p=\frac{1}{2}$

- (a) How much do you get paid if the coin comes up heads 3 times?

Given fair coins, the expectation of  $X$ ,  $E(X) = np = 3 \cdot \frac{1}{2} = \frac{3}{2}$  (also shown in graph above)

The expected winnings from the game is \$6, such that  $g(E(X)) = 6$

We also have:

$$G(X = 0) = 0$$

$$G(X = 1) = 2$$

$$G(X = 2) = 4$$

$$G(X = 3) = A$$

Using the expectation that  $E(g(X)) = g(E(X))$  we can solve for  $A$ :

$$E(g(X)) = \sum_x g(x) \cdot p(x) = 6$$

$$(0 \cdot \frac{1}{8}) + (2 \cdot \frac{3}{8}) + (4 \cdot \frac{3}{8}) + (A \cdot \frac{1}{8}) = 6$$

$$(\frac{6}{8}) + (\frac{12}{8}) + (\frac{A}{8}) = 6$$

$$6 + 12 + A = 48$$

$$A = 30$$

- (b) Write down a complete expression for the cumulative probability function for your winnings from the game.

$$g(x) = \begin{cases} 0 & ; \quad x < 1 \\ \frac{16}{3} & ; \quad 1 \leq x < 3 \\ 208 & ; \quad x \leq 3 \end{cases}, \quad p = \frac{1}{2}$$

$$\text{Winnings} = g(P(X \leq x)) = \sum_{y=0}^x g(y) \binom{3}{y} p^y (1-p)^{3-y}$$

$$G(P(X \leq 0)) = (\frac{1}{8} \cdot 0) = 0$$

$$G(P(X \leq 1)) = (\frac{1}{8} \cdot 0) + (\frac{3}{8} \cdot \frac{16}{3}) = 2$$

$$G(P(X \leq 2)) = (\frac{1}{8} \cdot 0) + (\frac{3}{8} \cdot \frac{16}{3}) + (\frac{3}{8} \cdot \frac{16}{3}) = 4$$

$$G(P(X \leq 3)) = (\frac{1}{8} \cdot 0) + (\frac{3}{8} \cdot \frac{16}{3}) + (\frac{3}{8} \cdot \frac{16}{3}) + (\frac{1}{8} \cdot 208) = 30$$

## 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.
- (b) You wait until the die is rolled, then get  $1/X$  in dollars.
- (c) Which option is better for you, in expectation?

### 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.

- (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  $B$ ,  $E(A + B) = E(A) + E(B)$ .
- (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)
- (c) What is the value of  $E(Y)$  for this choice of  $t$ ? (Hint: this should also be a very familiar value)

### 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:

$$p_M(x) = \begin{cases} c/x^3, & x \in \{1, 2, 3, \dots\} \\ 0, & \text{otherwise.} \end{cases}$$

where  $c$  is a constant (you can calculate its value if you like, but it's not important).

- (a) Is  $E(M)$  finite?  
Yes it is finite because the sum of  $c/x^3$  from 1 to  $\infty$  will converge. It converges because this is a p-series ( $1/n^p$ ) that converges when  $p > 1$  and diverges when  $0 < p < 1$ .
- (b) Is  $V(M)$  finite?  
Also finite for the same reason.

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 problematic 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.