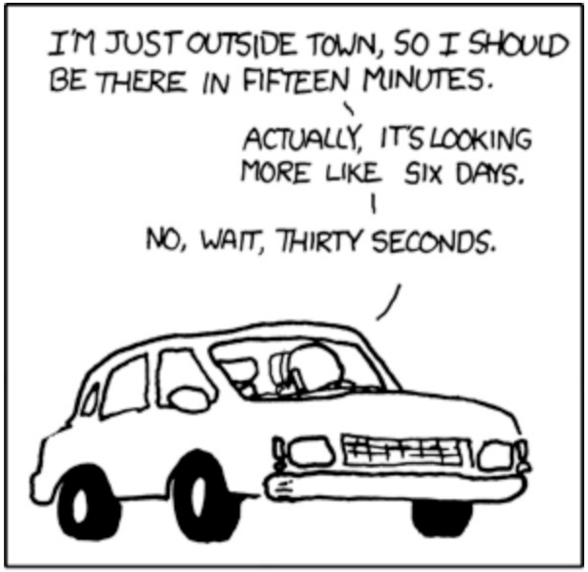
Stat 88: Probability & Statistics in Data Science



THE AUTHOR OF THE WINDOWS FILE COPY DIALOG VISITS SOME FRIENDS.

https://xkcd.com/612/

Lecture 12: 2/24/2022

Indicators, Unbiased estimators, Conditional expectation

Sections 5.3, 5.4, 5.5

Agenda

- Finish up joint distributions
- 5.3
  - Method of indicators
- 5.4
  - Unbiased Estimators
- 5.5
  - Conditional expectation

Joint distributions

$P(X=x, Y=y) = f(x,y)$
Joint prob. mass function for X, Y

- Draw two tickets without replacement from a box that has 5 tickets marked 1, 2, 2, 3, 3. Let X_1 and X_2 represent the values of the tickets drawn on the first and second draws respectively.
- Create a table of all possible outcomes for the pair (X_1, X_2) (which is also a random variable), and write down the probabilities using the multiplication rule.

Marginal dsn for X_1

	$X_2 = 1$	$X_2 = 2$	$X_2 = 3$	$P(X_1 = x)$
$X_1 = 1$	0	$\frac{2}{20}$	$\frac{2}{20}$	$P(X_1=1)=\frac{4}{20}$
$X_1 = 2$	$2/20$	$2/20$	$4/20$	$P(X_1=2)=\frac{8}{20}$
$X_1 = 3$	$2/20$	$4/20$	$2/20$	$P(X_1=3)=\frac{8}{20}$
Marginal dsn for X_2	$P(X_2=1)=\frac{4}{20}$	$P(X_2=2)=\frac{8}{20}$	$P(X_2=3)=\frac{8}{20}$	

$$E(X_1) = \sum_x x \cdot f(x) = 1 \cdot \frac{4}{20} + 2 \cdot \frac{8}{20} + 3 \cdot \frac{8}{20} = \frac{4+16+24}{20} = 2.2$$

Marginal distributions

$$E(X_2) = 1 \cdot \frac{4}{20} + 2 \cdot \frac{8}{20} + 3 \cdot \frac{8}{20} = \frac{4+16+24}{20} = \frac{44}{20} = 2.2$$

- What is $P(X_1 = 1)$? Write down the pmf for X_1 and X_2 — done last time

  $4/20$

- Are they independent? No $P(X_1 = 1 \overset{\text{"and"}}{,} X_2 = 1) = 0 \neq P(X_1 = 1) \cdot P(X_2 = 1)$

- Use the table to compute $P(X_1 + X_2 = 5)$ $= P(X_1 = 2, X_2 = 3) + P(X_1 = 3, X_2 = 2)$

  $= \frac{4}{20} + \frac{2}{20} = \frac{6}{20} = 3/10$

- Use the table to compute $E\big(g(X_1, X_2)\big)$, where $g(X_1, X_2) = |X_1 - X_2|$ (the expected distance between the two draws)

$$E(g(X_1, X_2)) = \sum_{(x,y)} g(x,y) f(x,y)$$

$$= 0 \cdot 0 + 1 \cdot 2/20 + 2 \cdot 2/20 + 1 \cdot \frac{2}{20} + 0 \cdot \frac{2}{20} + 1 \cdot \frac{4}{20} + 2 \cdot \frac{4}{20} + 1 \cdot \frac{2}{20} + 0 \cdot \frac{2}{20}$$

$g(X_1, X_2) = |X_1 - X_2|$

X_1 \ X_2	1	2	3
1	0, 2, 0	1, 3, 2/20	2, 4, 2/20
2	1, 3, 2/20	0, 4, 2/20	1, 5, 4/20
3	2, 4, 4/20	1, 5, 2/20	0, 6, 2/20

Exercise

- If $S = X_1 + X\_2$, find $E(S)$.

Show that $E(X_1 + X_2) = E(X_1) + E(X_2)$

$S = X_1 + X_2$

Let X, Y be r.v.

$$f(x,y) = P(X=x, Y=y)$$

Marginal dsn of X is denoted by $P(X=x) = f_X(x)$

Marginal dsn for Y is $f_Y(y) = P(Y=y)$

$$f_X(x) = P(X=x) = \sum_y P(X=x, Y=y) = \sum_y f(x,y)$$

$$f_Y(y) = P(Y=y) = \sum_x P(X=x, Y=y) = \sum_x f(x,y)$$

We say that X & Y are independent if

$$f(x,y) = f_X(x) \cdot f_Y(y)$$

for every possible pair (x,y) $\Big\}$ (☆)

And if (☆) is true then X & Y are independent.

Example A joint p.m.f for 2 r.v. M, S is given below.

① Find $E(M)$

② Are M & S indep.

$S \backslash M$	$M=2$	$M=3$	$f_S(s)$
$S=2$	0	$1/3$	$1/3$
$S=3$	$1/3$	$1/3$	$2/3$
$P(M=m)$	$1/3$	$2/3$	

① $E(M) = 2 \cdot f_M(2) + 3 \cdot f_M(3) = 2 \cdot \frac{1}{3} + 3 \cdot \frac{2}{3} = 8/3$

② $P(M=2, S=2) = 0 \neq P(M=2)P(S=2)$

Indicators A r.v. I_A is called the indicator function for an event A if

$$I_A = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false} \end{cases}$$

$$E(I_A) = 1 \cdot P(A) + 0 \cdot P(A^c)$$
$$= P(A)$$

Method of indicators

- Additivity of Expectation: This is a very useful property – no matter what the joint distribution of X and Y may be, we have:
$$E(X+Y) = E(X) + E(Y)$$
- Whether X and Y are dependent or independent, this additivity property holds, making it enormously useful.
- We also have linearity: $E(aX + bY) = aE(X) + bE(Y)$

- Recall that we talked about "classifying and counting" – that is, we divide up the outcomes into those that we are interested in (successes), and everything else (failures), and then count the number of successes.
- We can represent these outcomes as 0 and 1, where 1 marks a success and 0 and failure, so if we model the trials as **draws from a box**, we can count the number of success by counting up the number of times we drew a 1.

- We can represent each draw as a Bernoulli trial, where $p = P(S)$

I_A , $\mathbb{1}_A \rightarrow$ 1 b/c indicator is 1 if A is true

Using indicators and additivity

- For example, say we roll a die 10 times, and success is rolling a 1. ⚀
- Then $p = 1/6$, and we can define a Bernoulli rv as $X = \begin{cases} 0, & w.p. 5/6 \leftarrow \\ 1, & w.p. 1/6 \leftarrow \end{cases}$
- We can also define an event A: let A be the event of rolling a 1 and define a rv $\boldsymbol{I_A}$ aka $\mathbf{1}_{\boldsymbol{A}}$ that takes the value 1 if A occurs and 0 otherwise.
- This is a Bernoulli rv, what is its expectation (in terms of A)?

$P(A)$

- Now let $X \sim Bin(10, \frac{1}{6})$, so X counts the number of successes in 10 rolls. Let's find $E(X)$ using additivity and indicators: $s \Leftrightarrow$ ⚀

Let A_1 be the event of rolling ⚀ on 1st roll
Let A_2 " " " ⚀ " 2nd roll
Let A_k " " " ⚀ " kth roll

$1 \leq k \leq 10$

$X = I_{A_1} + I_{A_2} + \cdots + I_{A_{10}}$

$\mathbb{E}(X) = \sum_{k=1}^{10} \mathbb{E}(I_{A_k}) = 10 \cdot \frac{1}{6}$

$\mathbb{E}(I_{A_1} + I_{A_2} + \cdots + I_{A_{10}})$

$= \mathbb{E}(I_{A_1}) + \mathbb{E}(I_{A_2}) + \cdots + \mathbb{E}(I_{A_{10}}$

$$\mathbb{E}(X) = \sum_x x \cdot f(x)$$

$X \sim Bin(n, p)$

Using indicators to compute expected value

$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$

- Binomial $\mathbb{E}(X) = \sum_{x=0}^{n} \underbrace{\left[\binom{n}{x} p^x (1-p)^{n-x}\right]}_{f(x)} x$

Define A_1 to the event of a S in 1^{st} trial $\to I_{A_1}$, $\mathbb{E}(I_{A_1}) = 1 \cdot p + 0(1-p) = p$

A_2 " " " S in 2^{nd} trial $\to I_{A_2}$

$X = I_{A_1} + I_{A_2} + I_{A_3} + \cdots + I_{A_n}$

$\mathbb{E}(X) = \mathbb{E}(I_{A_1} + I_{A_2} + \cdots + I_{A_n}) = \sum_{k=1}^{n} \mathbb{E}(I_{A_k}) = \sum_{k=1}^{n} p = n \cdot p$

- Hypergeometric: Did we use the independence of the trials for the binomial? If not, we can use the same method to compute the expected value of a hypergeometric rv:

$X \sim HG(N, G, n)$

Again let $X = I_{A_1} + I_{A_2} + \cdots + I_{A_n}$

$I_{A_k} = \begin{cases} 1 & \text{if } S \text{ on } k^{th} \text{ draw} \\ 0 & \text{o/w} \end{cases} = G/N$

Show $\mathbb{E}(X) = n \cdot G/N$

Using indicators

Exercise 5.7.6: A die is rolled 12 times. Find the expectation of:

a) the number of times the face with five spots appears

b) the number of times an odd number of spots appears

c) the number of faces that don't appear

d) the number of faces that do appear

(a) Let X = # of times a ⚄ is rolled in 12 rolls

Want $\mathbb{E}(X)$. $X \sim Bin(12, 1/6)$ $\mathbb{E}(X) = 2 = 12 \cdot \frac{1}{6}$

Exercise: Find $\mathbb{E}(X)$ using indicator functions

Let A_k be the event of rolling ⚄ on k^{th} roll.

(c) A die is rolled 12 times. Want the expected $\underbrace{\text{\# of faces that don't appear}}$.

Let X = # of faces that don't appear.

Let A_k be the event that k^{th} face does not appears, $k = 1, 2, \dots 6$

$$I_{A_k} = \begin{cases} 1, & \text{if } A_k \text{ is true} \\ 0, & \text{o/w} \end{cases}$$

A_1 is true. $\boxed{\cdot}$ does not appear

$P(A_1) = \left(\frac{5}{6}\right)^{12}$, $P(A_k) = \left(\frac{5}{6}\right)^{12}$

$$X = I_{A_1} + I_{A_2} + \dots + I_{A_6}$$

$$E(X) = E(I_{A_1}) + E(I_{A_2}) + \dots + E(I_{A_6})$$

$$= 6 \cdot \left(\frac{5}{6}\right)^{12}$$

# of rolls

# of faces

(d)

$$I_{A_1} = \begin{cases} 1, & \text{if } \boxed{\cdot} \text{ appears at least once} \\ 0 & \text{o/w} \end{cases}$$

Example

- Let X be the number of spades in 7 cards dealt **with replacement** from a well shuffled deck of 52 cards containing 13 spades. Find E(X).

1. Write down what X is X = # of spades in 7 cards.
2. Define an indicator for the kth trial: I_k $I_k = \begin{cases} 1 & \text{if } k^{th} \text{ draw is } \spadesuit \\ 0 & \text{o/w} \end{cases}$ $P(I_k = 1) = \frac{1}{4} = \frac{13}{52}$
3. Find $p = P(I_k = 1)$ $= \frac{1}{4}$
4. Write X as a sum of indicators $X = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7$
5. Now compute $E(X)$ using additivity

$$\mathbb{E}(X) = 7 \cdot \frac{1}{4} = 7/4$$

- Do the same thing if we deal 7 cards **without replacement.**

Same as above, by symmetry.

Missing classes

- We can use indicators to compute the chance that something *doesn't* occur.

- For example, say we have a box with balls that are red, white, or blue, with 35% being red, 30% being white, and 35% blue. If we draw n times with replacement from this box, what is the expected number of colors that *don't* appear in the sample?

Examples

1. An instructor is trying to set up office hours during RRR week. On one day there are 8 available slots: 10-11, 11-noon, noon-1, 1-2, 2-3, 3-4, 4-5, and 5-6. There are 6 GSIs, each of whom picks one slot. Suppose the GSIs pick the slots at random, independently of each other. Find the expected number of slots that no GSI picks.

Examples

- A building has 10 floors above the basement. If 12 people get into an elevator at the basement, and each chooses a floor at random to get out, independently of the others, at how many floors do you expect the elevator to make a stop to let out one or more of these 12 people?