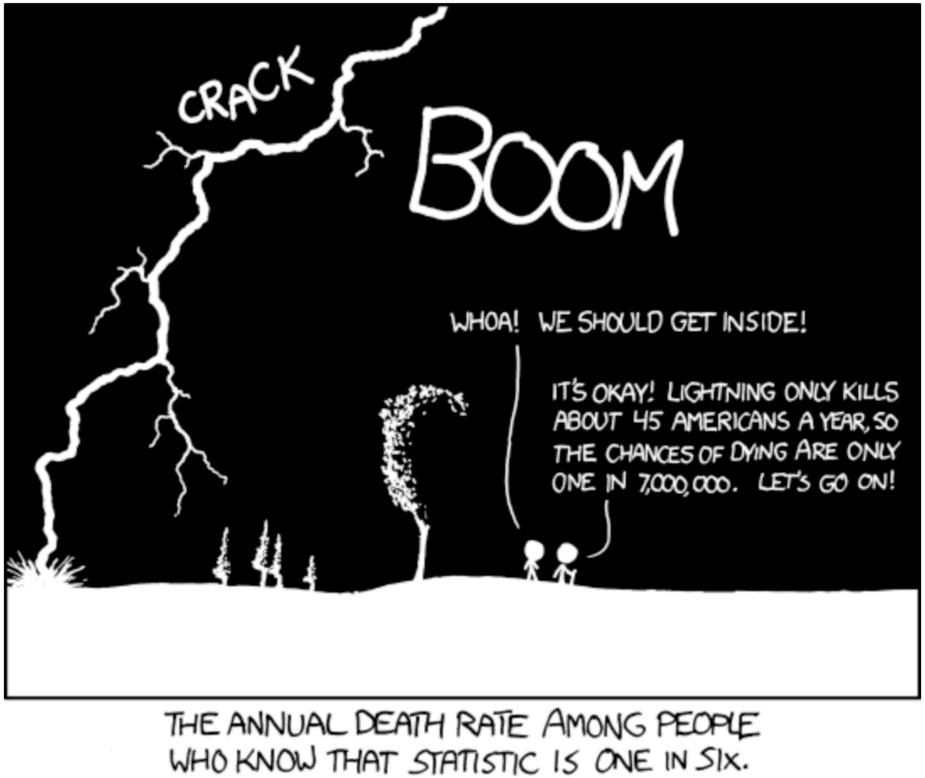
Stat 88: Probability & Math. Statistics in Data Science



https://xkcd.com/795/

Lecture 17: 3/1/2021

Expectation by conditioning

Sections 5.6

Lec 17 : First recall from last time

S \ X	1	2	3	$f_S(s)$ Marg. dsn of
2	$\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{16}$	0	0	1/16
3	$\frac{1}{4}\cdot\frac{1}{2}=\frac{1}{8}$	1/8	0	$1/4 = 4/16$
4	$\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{16}$	$\frac{1}{4}\left(=\frac{1}{2}\cdot\frac{1}{2}\right)$	$\frac{1}{16}$ $P(X=3, Y=1)$	6/16
5	0	$X=2, Y=3$ $\frac{1}{2}\cdot\frac{1}{4}=1/8$	1/8	4/16
6	0	0	1/16 $X=Y=3$	1/16
$f_X(x)$	4/16	8/16	4/16	

joint dsn of X and S

$$\boxed{E(X) = 1\cdot\frac{4}{16} + 2\cdot\frac{8}{16} + 3\cdot\frac{4}{16} = 2}$$

Cond'l. dsn of X|S

Cond'l exp. of X|S

conditional expectation

Given	$P(X=1)$	$P(X=2)$	$P(X=3)$	$E(X\mid S=s)$
$S=2$	1	0	0	1
$S=3$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{3}{2}=1.5$
$S=4$	$=P(X=1,S=4)/P(S=4)$ 1/6	4/6	1/6	$1\cdot\frac{1}{6}+2\cdot\frac{4}{6}+3\cdot\frac{1}{6}=$ $12/6=2$
$S=5$	0	$\frac{P(X=2,S=5)}{P(S=5)}$ $\frac{1}{2}$	$\frac{1}{2}$	$5/2 = 2\cdot\frac{1}{2}+\frac{3}{2}$
$S=6$	0	0	1	$3\cdot 1=3$

fixing s

$g(s) = E(X \mid S=s)$ is a function of s, not x

$E(X \mid S)$ is a r.v.

Writing down the dsn of g(S)

s	2	3	4	5	6
$P(S=s)$	1/16	4/16	6/16	4/16	1/16
$E(X\mid S=s)$	1	1.5	2	2.5	3

$$X = \begin{cases} 1 & \text{w/pr. } 1/4 \\ 2 & \text{w/pr. } 1/2 \\ 3 & \text{w/pr } 1/4 \end{cases}$$

$F_X = F_Y$, X, Y iid.

$S = X + Y$

$g(S) = E(X \mid S)$

r.v. ; what varies is S.

$$E(g(S)) = 1\cdot\frac{1}{16} + (1.5)\left(\frac{4}{16}\right) + (2)\left(\frac{6}{16}\right) + (2.5)\left(\frac{4}{16}\right) + 3\left(\frac{1}{16}\right)$$

$$= \frac{1+6+12+10+3}{16} = \frac{32}{16} = 2.$$

$$E(g(S)) = E(E(X \mid S)) = E(X)$$

Expectation by Conditioning

cond'l ≡ conditional

- In the example we just worked out, once we fix a value *s* for S, then we have a distribution for X, and can compute its expectation using that distribution that depends on *s*: $E(\,X \mid S = s\,) = \sum_x \underbrace{xP(\,X = x \mid S = s\,)}_{\text{cond'l dsn of } X \mid S=s}$, with the sum over all values of X.

- Note that $E(\,X \mid S = s\,)$ is a ***function of*** s. We can think of $\boxed{E(\,X \mid S\,)}$ as a rv.

  $E(X|S) \leftarrow$ r.v.

- This means that if we want to compute $E(X)$, we can just take a weighted average of these conditional expecations $E(\,X \mid S = s\,)$:

$$\boldsymbol{E(X) = \sum_{s} E(\,X \mid S = s\,)\,P(S = s)}$$

- This is the *law of iterated expectation*

averages out the X

$$E(E(X|S)) = E(X)$$

averages out the S

Law of iterated expectation

- Note that $E(X \mid S = s)$ is a function of s. That is, if we change the value of s we get a different value. (It is not a function of x, though.)

- Therefore we can define the function $g(s) = E(X \mid S = s)$, and the random variable $g(S) = E(X \mid S)$.

- In general, recall that $E\big(g(S)\big) = \sum_s g(s)f(s) = \sum_s g(s)P(S = s)$.
- How can we use this to find the expected value of the rv $g(S)$?

$$E(g(S)) = \underbrace{\sum_s E(X \mid S=s)\,P(S=s)} = E(X)$$

To prove this explicitly:

write out $E(X \mid S=s) = \sum_x x \cdot P(X=x \mid S=s)$

Now you have a double sum,

write out $P(X=x \mid S=s) = \dfrac{P(X=x, S=s)}{P(S=s)}$, $P(S=s) \neq 0$

$g(S) = \mathbb{E}(X|S$, $\mathbb{E}(g(S)) = \mathbb{E}(X)$

Examples from the text: Time to reach campus

- 2 routes to campus, student prefers route A (expected time =15 minutes) and uses it 90% of the time. 10% of the time, forced to take route B which has an expected time of 20 minutes. What is the expected duration of her trip on a randomly selected day?

$X =$ duration of the trip $\begin{cases} 15m \\ 20 \text{ mins} \end{cases}$

$\mathbb{E}(X) = 15(0.9) + 20(0.1) = 15.5$ minutes.

$S =$ Route taken

Conditioning X on

$$\mathbb{E}(X) = \mathbb{E}\left(\underbrace{\mathbb{E}(X|S=s)}\, P(S=s) \right)$$

$$= 15 \cdot P(S=A) + 20\, P(S=B)$$

$$= 15(0.9) + 20(0.1)$$

Catching misprints

$E(X \mid S=s)$ is NOT a r.v., a fixed #

$E(X \mid S)$ is a r.v.

- The number of misprints is a rv $N \sim$ Pois(5) dsn. Each misprint is caught before printing with chance 0.95 independently of all other misprints. What is the expected number of misprints that are caught before printing?

P(catching a misprint before printing) = 0.95

Let M = # of misprints caught before printing

Sp. we know document has 8 misprints.

$M \sim Bin(8, 0.95)$, $E(M \mid N=8) = (8)(0.95)$

What about if $N = n$? $M \sim Bin(n, 0.95)$, $E(M \mid N=n) = 0.95n$

$$E(M) = E(\underbrace{\overbrace{E(M \mid N)}^{\text{r.v. in } N.}}_{\text{(function of N) dependent on values of N}})$$

$$= \sum_{n=0}^{\infty} E(M \mid N=n) \cdot P(N=n) = \sum_{n=0}^{\infty} 0.95n\, P(N=n)$$

$$= 0.95 \sum_{n=0}^{\infty} n \cdot P(N=n) = (0.95)(5)$$

$\mathbb{E}(\mathbb{E}(M|N=n)) = \sum_n \left(\sum_m m \cdot P(M=m|N=n)\right) P(N=n)$

$g(N) = \mathbb{E}(M|N)$

$\mathbb{E}(g(N)) = \sum_n g(n) P(N=n)$

Exercise 5.7.13

- A survey organization in a town is studying families with children. Among the families that have children, the distribution of the number of children is shown below. Suppose each child has chance 0.51 of being male, independently of all other children. What is the expected number of male children in a family picked at random from those that have children?

$\mathbb{E}(N) = (1)(0.2) + 2(0.4) + 3(0.2) + 4(0.15) + 5(0.05)$

n	1	2	3	4	5
prop. with n children	0.2	0.4	0.2	0.15	0.05

M = # of male children in a randomly picked family

N = # of children

What is the dsn of $M|N=n$, $M \sim Bin(n, 0.51)$

$\mathbb{E}(M|N=n) = 0.51n$

$\mathbb{E}(M) = \mathbb{E}(\mathbb{E}(M|N)) = \sum_{n=1}^{5} \mathbb{E}(M|N=n) P(N=n)$

$= \sum_{n=1} 0.51n \cdot P(N=n) = (0.51)(\underbrace{1(0.2) + 2(0.4) + 3(0.2) + \cdots}_{\text{from table}})$

≈ 1.25

$= 0.51 \sum_{n=1}^{\infty} n P(N=n) = 0.51 \cdot E(N)$

Expectation of a Geometric waiting time

- $X \sim Geom(p)$: X is the number of trials until the first success
- $P(X = k) = (1-p)^{k-1}\, p, \quad k = 1, 2, 3, \ldots$
- Let $x = E(X)$ going to solve for x
- Recall that $P(X > 1) = P(first\ trial\ is\ F) = 1 - p$
- We can split the possible situations into when the first trial is a success and the first trial is a failure, and condition on this and compute the *conditional expectation:*

$X = T_1$:

Y = result of first trial. either first trial S or F

$$E(X) = E(X \mid X = 1)P(X = 1) + E(X \mid X > 1)P(X > 1)$$

$= 1 \cdot p \qquad (x+1) \quad (1-p)$

just like starting over but with 1 step more

Exp # of trials $= x+1$

$x = p + (x+1)(1-p)$

$\cancel{x} = \cancel{p} + \cancel{x} + 1 - px - \cancel{p}$ $\Rightarrow \boxed{x = 1/p = E(X)}$

Expected waiting time until k sixes have been rolled

- Let T_k be the waiting time until the kth six is rolled. (fair die $P(\boxed{6}) = \frac{1}{6}$)
- For example, consider $k = 1$ (geometric) , expected waiting time=6 rolls $= \frac{1}{\frac{1}{6}}$
- What about if $k = 2$? $k = 3$? $k = 10$?

k=1

start - - - - - - $\overset{\uparrow}{\text{1st } \boxed{6}}$ | $\underbrace{\overset{\downarrow \text{ starting over}}{\text{— — — — — — —}}}$ } k=2.

$\mathbb{E}(T_1) = 6$

$$T_2 = T_1 + T_1 \rightarrow \mathbb{E}(T_2) = \mathbb{E}(T_1) \cdot 2$$
$$= 12$$

$$\mathbb{E}(T_{10}) = 60 = 10 \cdot \mathbb{E}(T_1)$$

$\underbrace{\text{wait for 1st 5}}_{T_1}$ | $\underbrace{\text{wait for 2nd}}_{T_1}$ | wait for 3rd

- General formula for expected waiting time until kth success where P(S) = p

$= \underline{k \cdot 1/p}$??