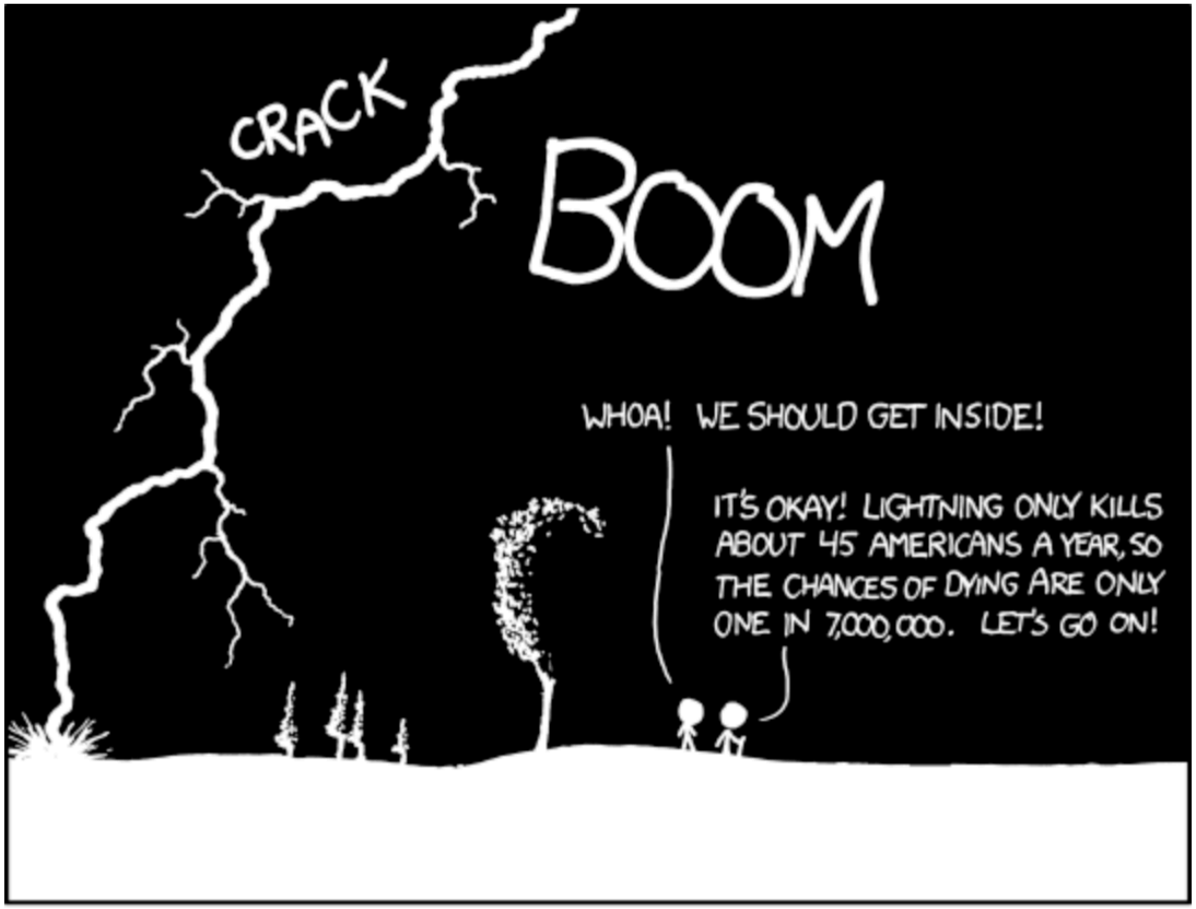
Stat 88: Probability & Math. Statistics in Data Science



THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

https://xkcd.com/795/

Lecture 17: 3/1/2021

Expectation by conditioning

Sections 5.6

Expectation by Conditioning

- 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: $E(\,X \mid S = s\,) = \sum xP(\,X = x \mid S = s\,)$, with the sum over all values of X.

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

- 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*

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 $g(S)$?

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?

Catching misprints

- The number of misprints is a rv N ~ 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?

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?

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

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 $x = E(X)$
- 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:*

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

Expected waiting time until k sixes have been rolled

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

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

Example

A fair die is rolled repeatedly.

a) Find the expected waiting time (number of rolls) till a total of 5 sixes appear.

b) Find the expected waiting time until two *different* faces are rolled.

Tossing coins

A fair coin is tossed 3 times. Let *X* be the number of heads in the first 2 tosses and *Y* the number of heads in the last two tosses.

a) Find $E(Y \mid X = 2)$.

b) Find $E(Y \mid X = 1)$

c) Find $E(Y \mid X = 0)$

d) Find $E(Y)$

All who wander might be lost…

A lost tourist arrives at a point with 2 roads. Road A brings him back to the same point after 1 hour of walking. Road B leads to the city in 2 hours. Assuming the tourist randomly chooses a road at all times, what is the expected time until the tourist arrives to the city?

Expectation computation.

- Tamara chooses an integer X uniformly at random from 1 to 425. She then chooses an integer Y uniformly at random from $1, \ldots, X$. Find $E(Y)$.

Tamara, again

- Tamara chooses an integer *N* uniformly at random from *Pois*(μ). She then picks *N* cards from a deck with replacement. Find the expected number of ace cards.