Conditional Expectation

Robert Y. Lewis

CS 0220 2024

April 15, 2024

Overview

Expectation definition

Recall:

Definition: If R is a random variable defined on a sample space $\mathcal{S}$, then the expectation of R is

$$\mathbb{E}[R] ::= \sum_{\omega \in \mathcal{S}} R(\omega) \Pr[\omega].$$

Shortcut formula for expectation

Instead of summing over all outcomes, sometimes easier to group outcomes by the blocks defined by the random variable.

$$\mathbb{E}[R] = \sum_{x \in \text{range}(R)} x \cdot \Pr[R = x].$$

Proof:

$$\begin{aligned}
\mathbb{E}[R] &= \textstyle\sum_{\omega \in \mathcal{S}} R(\omega) \Pr[\omega] && \text{defn expt} \\
&= \textstyle\sum_{x \in \text{range}(R)} \sum_{\omega \in [R=x]} R(\omega) \Pr[\omega] && \text{range is partition} \\
&= \textstyle\sum_{x \in \text{range}(R)} \sum_{\omega \in [R=x]} x \Pr[\omega] && \text{defn } R, \omega \\
&= \textstyle\sum_{x \in \text{range}(R)} x \sum_{\omega \in [R=x]} \Pr[\omega] && \text{factor out } x \\
&= \textstyle\sum_{x \in \text{range}(R)} x \Pr[R = x]. && \text{defn event prob}
\end{aligned}$$

Definition of conditional expectation

Definition: The conditional expectation of a random variable R given event A is:

$$\mathbb{E}[R|A] := \sum_{r \in \text{range}(R)} r \cdot \Pr[R = r|A].$$

Example: R is the deck of cards example from before: value of card, 1 for ace, 10 for face.

$$\begin{aligned}
&\mathbb{E}[R|R < 10] \\
&= \textstyle\sum_{i=1}^{9} i \Pr[R = i|R < 10] \\
&= \textstyle\sum_{i=1}^{9} i \frac{\Pr[R=i \wedge R<10]}{\Pr[R<10]} \\
&= \textstyle\sum_{i=1}^{9} i \ \frac{1/13}{9/13} \\
&= 1/9 \ \textstyle\sum_{i=1}^{9} i \\
&= 45/9 = 5.
\end{aligned}$$

The height of elephants

How tall are elephants on average? Don't know. Web search didn't help. Conditional expectations can!

$$
\begin{aligned}
\mathbb{E}[H] &= \mathbb{E}[H|\text{Af}] \cdot \Pr[\text{Af}] + \mathbb{E}[H|\text{As}] \cdot \Pr[\text{As}] && \text{split} \\
&= 10 \cdot 0.9 + \mathbb{E}[H|\text{As}] \cdot 0.1 && \text{Google} \\
&= 10 \cdot 0.9 + (\mathbb{E}[H|\text{As} \wedge m] \Pr[m|\text{As}] \\
&\qquad + \mathbb{E}[H|\text{As} \wedge f] \cdot \Pr[f|\text{As}]) \cdot 0.1 && \text{split} \\
&= 10 \cdot 0.9 + (9 \cdot 1/2 + 7 \cdot 1/2) \cdot 0.1 && \text{Google} \\
&= 9.8 && \text{calculator}
\end{aligned}
$$

Memoryless crash time

A program crashes at the end of each hour of use with probability p. Let C be the random variable for the time until a crash occurs. What is $\mathbb{E}[C]$?

Let A be the event that the program crashes after the first hour and $\overline{A}$ be the complementary event. Mean time to failure is

$$\mathbb{E}[C] = \mathbb{E}[C|A] \cdot \Pr[A] + \mathbb{E}[C|\overline{A}] \cdot \Pr[\overline{A}]$$

$\mathbb{E}[C|A] = 1$ because A is the event that the program crashes after one hour.

$\mathbb{E}[C|\overline{A}] = 1 + \mathbb{E}[C]$ because the program runs for an hour, then we're back in the original situation.

$$\begin{aligned}
\mathbb{E}[C] &= 1 \cdot p + (1 + \mathbb{E}[C])(1 - p) \\
\mathbb{E}[C] &= p + 1 + \mathbb{E}[C] - p - p\mathbb{E}[C] \\
p\mathbb{E}[C] &= 1 \\
\mathbb{E}[C] &= 1/p
\end{aligned}$$

Mean time to failure

General principle:

If a system independently fails at each time step with probability p, then the expected number of steps up to the first failure is $1/p$.

More failures

I'm driving down a street with lots of intersections with traffic lights. Each light is (independently) red 50% of the time and green 50% of the time. How many intersections do I expect to drive through before I stop?

failure	red light
hour	intersection
p	$1/2$
$1/p$	2

- Coin flips until heads? 2
- Cards until face card? 13/3 = 4.3 ish
- Insertions into a mostly-empty hash table before collision?

Median

Not the same as mean or expected value. Intuitively, it's the "middle" of the set of values.

1, 3, **4**, 10, 1000

Useful sometimes for disregarding outlier values of a random variable. But expectation has better mathematical properties, so we'll focus on that.

Median

Definition (from an earlier version of the textbook): The *median* of a random variable R is a value $x \in \text{range}(R)$ such that

$$\Pr[R \leq x] \geq \frac{1}{2}$$

and

$$\Pr[R \geq x] \geq \frac{1}{2}.$$

Why is this wrong?

Fixing the book’s definition

Can you think of a better way to define this in the context of probability spaces? When does the concept make sense?