Probability and Mathematical Statistics in Data Science

Lecture 15: Section 5.5: Conditional Expectation
Section 5.6: Expectation by Conditioning

Conditional Expectation: An example

- Let X and Y be rvs with the distribution described below, and let $S = X + Y$:

x	1	2	3
P(X=x)	1/4	1/2	1/4

- Let's write down the joint distribution of X and S:

	$X=1$	$X=2$	$X=3$
$S=2$	0.0625	0	0
$S=3$	0.125	0.125	0
$S=4$	0.0625	0.25	0.0625
$S=5$	0	0.125	0.125
$S=6$	0	0	0.0625

Conditional Expectation: An example

As usual, we will display the joint distribution in a table.

	$X = 1$	$X = 2$	$X = 3$
$S = 2$	0.0625	0	0
$S = 3$	0.125	0.125	0
$S = 4$	0.0625	0.25	0.0625
$S = 5$	0	0.125	0.125
$S = 6$	0	0	0.0625

Recall that summing along the rows will give us the distribution of S:

s	2	3	4	5	6
$P(S = s)$	0.0625	0.25	0.375	0.25	0.0625

Conditional Expectation: An example

Recall that summing along the rows will give us the distribution of S:

s	2	3	4	5	6
$P(S=s)$	0.0625	0.25	0.375	0.25	0.0625

Now suppose someone runs the experiment and tells you that $S=3$. Given this information, what is the distribution of X?

Given $S=3$, you can ignore all the rows of the table except the one corresponding to $S=3$.

The calculation for the first cell is

$$P(X=1 \mid S=3) = \frac{P(X=1, S=3)}{P(S=3)} = \frac{0.125}{0.25} = 0.5$$

What is the P(X=1| S = 4)?

Conditional Expectation: An example

	$X=1$	$X=2$	$X=3$
Conditional distribution of X given $S=2$	1	0	0
Conditional distribution of X given $S=3$	0.5	0.5	0
Conditional distribution of X given $S=4$	0.1667	0.6667	0.1667
Conditional distribution of X given $S=5$	0	0.5	0.5
Conditional distribution of X given $S=6$	0	0	1

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*

	$X=1$	$X=2$	$X=3$	$E(X \mid S=s)$
Conditional distribution of X given $S=2$	1	0	0	1
Conditional distribution of X given $S=3$	0.5	0.5	0	1.5
Conditional distribution of X given $S=4$	0.1667	0.6667	0.1667	2
Conditional distribution of X given $S=5$	0	0.5	0.5	2.5
Conditional distribution of X given $S=6$	0	0	1	3

The values of $E(X \mid S=3)$ and $E(X \mid S=5)$ are clear by symmetry. The value of $E(X \mid S=4)$ can be calculated using the definition of expectation:

$$E(X \mid S=4) \;=\; 1(0.1667) + 2(0.6667) + 3(0.1667) \;=\; 2$$

Expectation by Conditioning

s	2	3	4	5	6
$E(X \mid S = s)$	1	1.5	2	2.5	3
$P(S = s)$	0.0625	0.25	0.375	0.25	0.0625

And now find the weighted average of the conditional expectations:

$$1(0.0625) + 1.5(0.25) + 2(0.375) + 2.5(0.25) + 3(0.0625) \;=\; 2 \;=\; E(X)$$

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?

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

Exercise 5.7.13

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

Let M – Number of Male children in a randomly chosen family
Let N – be the number of children

What is the distribution of M given N = n? M – Bin(n, 0.51)

=> E(M | N = n) = 0.51n