

Unit 4, Lecture 2

*Numerical Methods and Statistics***Companion Reading**

Bulmer Chapter 5

1 Expected Values

The expected value of a sample x is

$$E[X] = \sum_Q P(X = x)x \quad (1)$$

An expected value is analogous to a mean, except you don't add data points. You add the elements in your sample space, weighted by their probability. Be careful not to be confused: the expected value is not the most likely outcome. The most likely outcome is the x which maximizes $P(x)$, written as:

$$\operatorname{argmax}_x P(x) \quad (2)$$

Think of an expected value as a measure of the center point of the sample space.

1.1 Die Roll Example

Consider a fair die example, where d is an element in our sample space. Our sample space is one through six and the probability of a sample occurring is $1/6$.

$$E[D] = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 \dots$$

$$E[D] = \frac{1}{6} \times 21 = 3.5$$

Notice that the expected value is NOT an element of the sample space.

1.2 Unfair Die Example

Now consider our unfair die, where

$$P(D = d) = \frac{d}{21}$$

so for example, the probability of rolling a 4 is $4/21$.

$$E[D] = \sum_{1, \dots, 6} \frac{d}{21} \times d = \frac{91}{21}$$

2 Expected Value of a Random Variable

To calculate the expected value of a random variable S , use:

$$E[S] = \sum_Q P(X = x)S(x) \quad (3)$$

Remember that the random variable S is defined by some $S(x)$ function, whose input is an element in the sample space and has an output of some real number.

2.1 Expected Value for Sum of 2 Dice

Consider the product sample space of two dice: $Q_2 = Q \otimes Q$, where Q is $1, \dots, 6$. Example of elements of this sample space are rolling a 2 for the first die and a 4 for the second die: $(2, 4)$. There are 36 elements in the sample space and we'll take the probability of each to be equivalent. This gives:

$$P(X = x) = P(X = (x_1, x_2)) = \frac{1}{36}$$

Define S as our random variable representing the sum of the two dice.

$$S(x) = S((x_1, x_2)) = x_1 + x_2$$

Now we can plug these into our expected value equation:

$$\begin{aligned} E[S] &= \sum_Q P(X = x)S(x) = \sum_{(x_1, x_2) \in Q_2} \frac{1}{36} S((x_1, x_2)) \\ &= \frac{1}{36} \times (1 + 1) + \frac{1}{36} \times (1 + 2) + \frac{1}{36} \times (1 + 3) + \dots + \frac{1}{36} \times (6 + 6) \end{aligned}$$

To calculate the answer, we need to sum all 36 possible dice rolls together. That turns out to be 252. So we have:

$$E[S] = \frac{252}{36} = 7$$

2.2 Expected Value for Sum of 2 Dice - Alternate

Another way to solve this problem is to DEFINE our sample space to be the possible sums. Then $Q = \{2, 3, \dots, 12\}$. Our corresponding probability is going to be the number of combinations to roll the sum divided by some normalizing constant. For example, the probability of rolling a 6 is $5 / Z$, where Z is some constant to make sure our probabilities sum to 1. The 5 comes from the number of ways to roll a 6: $(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)$. Now we to find Z :

$$\begin{aligned} Z &= \underbrace{1}_{\text{combinations to roll a 2}} + \underbrace{2}_{\text{combinations to roll a 3}}, \dots, \underbrace{1}_{\text{combinations to roll a 12}} \\ Z &= \underbrace{1}_2 + \underbrace{2}_3 + \underbrace{3}_4 + \underbrace{4}_5 + \underbrace{5}_6 + \underbrace{6}_7 + \underbrace{5}_8 + \underbrace{4}_9 + \underbrace{3}_{10} + \underbrace{2}_{11} + \underbrace{1}_{12} = 36 \end{aligned}$$

where the underbraces indicate which dice roll sum each term corresponds to. Doing this work also reveals an equation for the probability:

$$P(S = s) = \frac{6 - |s - 7|}{36}$$

Finally, we are ready to utilize the expected value equation:

$$\begin{aligned}
E[S] &= \sum_Q P(S=s)s = \sum_{2,\dots,12} \frac{6-|s-7|}{36} \times s \\
&= \frac{1 \times 2}{36} + \frac{2 \times 3}{36} + \frac{3 \times 4}{36} + \dots + \frac{1 \times 12}{36} \\
&= \frac{252}{36} = 7
\end{aligned}$$

Thus we've arrived at the same answer.

2.3 Continuous Expected Value

$$E[X] = \int_{\mathcal{X}} xp(x) dx \quad (4)$$

2.4 Continuous Example

$$p'(x) \propto x, \quad Q = [0, 5]$$

First, we must normalize it:

$$\begin{aligned}
\int_0^5 x dx &= \frac{25}{2} \\
p(x) &= \frac{2x}{25} \\
\int_0^5 x \times \frac{2x}{25} &= \frac{2 \times 125}{25 \times 3} = \frac{10}{3}
\end{aligned}$$

2.5 Conditional Expectation Value

$$E(X | Y = y) = \sum_{\mathcal{X}} P(X = x | Y = y) x \quad (5)$$

2.6 Die Roll Example

Our random variable X is the observation, and $Y = 0$ if the observation is odd and $Y = 1$ if the observation is even.

$$E(X | Y = 0) = \frac{1}{3} \times 1 + \frac{1}{3} \times 3 + \frac{1}{3} \times 5 = 3$$

Expected value gives the center of a random variable or probability distribution.

3 Variance

Variance gives the average deviation from that center. To make sure variation above and below the center contributes, it is the expected squared distance from the expected value:

$$\text{Var}(X) = E[(E[X] - X)^2] = E[X^2] - E[X]^2 \quad (6)$$

3.1 Die Roll Example

We already know $E[X] = 21/6 = 3.5$. To find $E[D^2]$, we can use our random variable expectation equation. Take F to be our random variable defined by $F(x) = x^2$:

$$E[D^2] = E[F] = \sum_Q P(x)x^2 =$$

$$= \frac{1}{6} \times 1^2 + \frac{1}{6} \times 2^2 + \frac{1}{6} \times 3^2 \dots$$

$$= \frac{91}{6}$$

$$\text{Var}(D) = \frac{91}{6} - \frac{21^2}{6^2} = \frac{35}{12} = 2.92$$