



Expectation

A random variable is fully represented by its probability mass function (PMF), which represents each of the values the random variable can take on, and the corresponding probabilities. A PMF can be a lot of information. Sometimes it is useful to summarize the random variable! The most common, and arguably the most useful, summary of a random variable is its "**Expectation**".

Definition: Expectation

The expectation of a random variable X , written $E[X]$ is the average of all the values the random variable can take on, each weighted by the probability that the random variable will take on that value.

$$E[X] = \sum_x x \cdot P(X = x)$$

Expectation goes by many other names: Mean, Weighted Average, Center of Mass, 1st Moment. All of which are calculated using the same formula.

Recall that $P(X = x)$, also written as $P(x)$, is the probability mass function of the random variable X . Here is code that calculates the expectation of the sum of two dice, based off the probability mass function:

```
def expectation_sum_two_dice():
    exp_sum_two_dice = 0
    # sum of dice can take on the values 2 through 12
    for x in range(2, 13):
        pr_x = pmf_sum_two_dice(x) # pmf gives Pr(x)
        exp_sum_two_dice += x * pr_x
    return exp_sum_two_dice
```

If we worked it out manually we would get that if X is the sum of two dice, $E[X] = 7$:

$$E[X] = \sum_x x \cdot P(X = x) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \dots + 12 \cdot \frac{1}{36} = 7$$

7 is the "average" number you expect to get if you took the sum of two dice near infinite times. In this case it also happens to be the same as the mode, the most likely value of the sum of two dice, but this is not always the case!

Properties of expectation

Property: Linearity of Expectation

$$E[aX + b] = a E[X] + b$$

Property: Expectation of the Sum of Random Variables

$$E[X + Y] = E[X] + E[Y]$$

Property: Law of Unconscious Statistician

$$E[g(X)] = \sum_x g(x) P(X = x)$$

One can also calculate the expected value of a function $g(X)$ of a random variable X when one knows the probability distribution of X but one does not explicitly know the distribution of $g(X)$. This theorem has the humorous name of "the Law of the Unconscious Statistician" (LOTUS), because it is so useful that you should be able to employ it unconsciously.

Property: Expectation of a Constant

$$E[a] = a$$

Sometimes in proofs, you will end up with the expectation of a constant (rather than a random variable). For example what does the $E[5]$ mean? Since 5 is not a random variable, it does not change, and will always be 5, $E[5] = 5$.