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:

```python
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, 12 + 1):
        pr_x = pmf_sum_two_dice(x) # pmf gives Pr sum is x
        exp_sum_two_dice += x * pr_x
    return exp_sum_two_dice

def pmf_sum_two_dice(x):
    # Return the probability that two dice sum to x
    count = 0
    # Loop through all possible combinations of two dice
    for dice1 in range(1, 6 + 1):
        for dice2 in range(1, 6 + 1):
            if dice1 + dice2 == x:
                count += 1
    return count / 36   # There are 36 possible outcomes (6x6)
```

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

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

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

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

Where a and b are constants and not random variables.

Property: Expectation of the Sum of Random Variables

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

This is true regardless of the relationship between X and Y. They can be dependent, and they can have different distributions. This also applies with more than two random variables.

$$E\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} E[X_i]$$

Property: Law of Unconcious Statistician

$$E[g(X)] = \sum_{x} g(x)\, \mathrm{P}(X = x)$$

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 unconciously.

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 $\mathrm{E}[5]$ mean? Since 5 is not a random variable, it does not change, and will always be 5, $\mathrm{E}[5] = 5$.

Expectation of Sums Proof

One of the most useful properties of expectation is that the sum of expectation of random variables can be calculated by summing the expectations of each random variable on its own. Later in [Adding Random Variables](#) we will learn that computing the full probability mass function (or probability density function) when adding random variables is quite hard, especially when the random variables are dependent. However the expectation of sums is always decomposable:

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

This very useful result is somewhat suprising. I believe that the best way to understand this result is via a proof. This proof, however, will use some concepts from the next section in the course reader, Probabilistic Models. Notice how the proof never needs to assume that X and Y are independent, or have the same distribution. In this proof X and Y are discrete, but if you made them continuous you would just replace the sum with an integral:

$E[X + Y]$

$\displaystyle = \sum_{x,y} (x + y)\, \mathrm{P}(X = x, Y = y)$ $\qquad$ Expected value of a sum

$\displaystyle = \sum_{x,y} \Big[x\, \mathrm{P}(X = x, Y = y) + y\, \mathrm{P}(X = x, Y = y) \Big]$ $\qquad$ Distributive property of sums

$\displaystyle = \sum_{x,y} x\, \mathrm{P}(X = x, Y = y) + \sum_{x,y} y\, \mathrm{P}(X = x, Y = y)$ $\qquad$ Commutative property of sums

$\displaystyle = \sum_{x} \sum_{y} x\, \mathrm{P}(X = x, Y = y) + \sum_{y} \sum_{x} y\, \mathrm{P}(X = x, Y = y)$ $\qquad$ Expanding sums

$\displaystyle = \sum_{x} x \sum_{y} \mathrm{P}(X = x, Y = y) + \sum_{y} y \sum_{x} \mathrm{P}(X = x, Y = y)$ $\qquad$ Distributive property of sums

$\displaystyle = \sum_{x} x\, \mathrm{P}(X = x) + \sum_{y} y\, \mathrm{P}(Y = y)$ $\qquad$ Marginalization

$= E[X] + E[Y]$ $\qquad$ Definition of Expectation

Example of Law of Unconcious Statistician

The property of expectation:

$$E[g(X)] = \sum_x g(x) \, \mathrm{P}(X = x)$$

is both useful and hard to understand just by reading the equation. It allows us to calculate the expectation of a function of any function applied to a random variable! One function that will turn out to be very useful when computing [Variance](#) is $E[X^2]$. According to the Law of Unconcious Statistician:

$$E[X^2] = \sum_x x^2 \, \mathrm{P}(X = x)$$

In this case g is the squaring function. To calculate $E[X^2]$ we calculate expectation in a way similar to $E[X]$ with the exception that we square the value of x before multiplying by the probability mass function. Here is code that calculates $E[X^2]$ for the sum of two dice:

```python
def expectation_sum_two_dice_squared():
    exp_sum_two_dice_squared = 0
    # sum of dice can take on the values 2 through 12
    for x in range(2, 12 + 1):
        pr_x = pmf_sum_two_dice(x) # pmf gives Pr(x)
        exp_sum_two_dice_squared += x**2 * pr_x
    return exp_sum_two_dice_squared
```