M5-01: Expected Value and Standard Error

Part of the "Polling, Confidence Intervals, and the Normal Distribution" Learning Badge

Video Walkthrough: *https://discovery.cs.illinois.edu/m5-01/*

Mean and Variance of Discrete Random Variables

Previously, we summarized a list of numbers by computing their average and SD. Now we'll do the analogous summaries for random variables (#'s generated by a chance process).

The mean of a random variable is also known as the **expected value** (EV). The expected value of a discrete random variable X is shown by:

$$\mu_X = E(X) = X_1 P_1 + \ldots + X_n P_n$$

....where: X_i: Value of event **#i**

P_i: Probability of event **#i**

The standard deviation of a random variable as the **standard error** (SE) and measures the spread. The SE of a discrete random variable X is shown by:

$$\sigma_X = \sqrt{(X_1 - \mu_X)^2 P_1 + \ldots + (X_n - \mu_X)^2 P_n}$$

...where X_i: Value of event **#i**

μ_X: Expected value of X (see left)

P_i: Probability of event **#i**

Puzzle #1: Let's say X is a random variable that looks at the number of workouts that I will do in a week. Here's the distribution. Find the expected value (EV) of X and the standard error (SE) of X:

X	P(X)
0	0.1
1	0.15
2	0.4
3	0.25
4	0.1

We can also make a histogram of all possible outcomes of a chance process and their probabilities. **This type of histogram is known as a probability histogram.** Probability histograms for discrete random variables are also known as probability mass functions (**pmf**). Probability histograms for continuous random variables are also known as probability density functions (**pdf**). Draw the pmf below:

Probability Histogram drawn by hand:

Probability Histogram from Python