# 1 | Introduction to Probability

Probability: the likelihood of something happening

- "Success"
- "Failure"

At the moment, those are the two categories.

## 1.1 | A Statistical Trial

Each trial is an event with an outcome. Say, your case for "success" is rolling a 6. All other events are considered as a "failure"

Therefore…

$$Prob(success) = \frac{\#\, times\ sucess}{\#\, trials} \tag{1}$$

Statistics is not a hard and fast mathematical thing.

### 1.1.1 | Definitions

In order to perform the trials, we need to define a few things.

$X$ - Random Variable: a statistic that we could measure

There is a basic bell curve that comes with most trials, given , and the $y$ is the count.

We don't want to sample the whole population, so we therefore want to sample a subset.

1. Mean The mean of our samples could be defined as follows:

$$\bar{x_{10}} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2}$$

Where little $x_{10}$ represents the random variable of a subset of the population, and hence $\bar{x_{10}}$ is the mean of the 10 samples. If we increase the sample size, it will become harder for the sample size to deviate.

Hence…

$$\bar{x_{10}} \approx \bar{X} \tag{3}$$

that the mean of 10 random samples is pretty close to the actual mean.

And, the distribution of the sampled means will become a bell curve slightly "fatter" than that in the original distribution. **As sample size increases, the skinnier the bell curve for the distribution of means increases.**

# 2 | Standard Deviation

Variance is calculated as follows:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \tag{4}$$

But, we don't know $\mu$, population mean, and $N$, total samples, we instead elect to use

$$\hat{\sigma^2} \approx \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \hat{x})^2}{n-1} \tag{5}$$

# 3 | Mean

## 3.1 | Sample Mean

$$\bar{x} = mean\ of\ x = \frac{\sum_{i=1}^{n} x_i}{n} \tag{6}$$

We know that, for the actual population, we have a bell curve that acts as a probability distribution of an element at any given height.

## 3.2 | Population Mean

$$\bar{X} = \mu \tag{7}$$

# 4 | Distribution of Means

If, for instance, we took the mean value of all possible random groups of $10$, we will deduct a similar bell curve for the value of $\bar{x_{10}}$, called the "Distribution of the Means"; the means of sample size $n = 10$. It would be a bell curve skinnier than the previous one.

## 4.1 | Properties

According to the `Central Limit Theorem`, we know that as $n$ increases, the bell curve will become sharper and sharper. The bigger the sample size, the more narrow the distribution will become — until it tends to a single value as $n = pop\_size$, and that will be a single line that is the mean.

It fact, the `Central Limit Theorem` states that even if the sampled distribution is not normal, the distribution of the means will be normal. Hence: even in a lopsided distribution will this normality of the distribution of the means hold

Note, that: $\sigma_{\bar{x_n}} = \frac{\sigma}{\sqrt{n}}$, with $\sigma_{\bar{x_n}}$ being the standard deviation of the distribution of the means, $n$ is the number of samples to calculate the means, and $\sigma$ is the standard deviation in the population. See Standard Deviation.

As $n$ increases, you reduce the distribution of the means to be more confident (lower standard deviation.) A good $n$ is therefore $n = 16$ or $n = 15$.

### 4.2 | General Formula of Normal Function

$$f(x) = A_e^{-Bx^2} \tag{8}$$

Where, $A$ is the amplitude at origin and $B$ is related to the standard deviation.

# 5 | Hypothesis Testing

If $A$ *were* true, what is the likelihood that $B$ is not true?

Say, we propose that…

> $A$: Ten percent of the World's Trees are Redwood Trees

We go out, randomly sample $50$ trees (but like actually randomly), and we find no Redwood Trees ($B$).

### 5.1 | Proportion Method

If $P_1 = 0.1$ is a Redwood Tree (any given tree is $10$% possible to be a redwood), the probability of finding $B$, $50$% *non* Redwoods, is $P_{50} = (1 - 0.1)^{50} = 0.005153775 = 0.51$%.

The standard acceptable values for a probability is, for disapprovals, $5$%, and approvals, $95$%. Given the probability of $B$ being true if $A$ is true is $<< 5$%, we disapprove $A$ as $B$ is measured to be true.

The means of the distribution of means is as follows:

$$\sigma_x = \frac{\sigma}{\sqrt{n}} \tag{9}$$

We could see, if you take 4 samples, you reduce the means by factor of $2$, 25, reduce by factor of $5$.

### 5.2 | Example: Comparing Launcher

One launcher has a mean launch distance $\mu_1$, and $\mu_2$. Essentially, we want to know if $\mu_1$ and $\mu_2$ are, when plotted against the sample distribution of distances, two different points in two overlapping fat curves forming one giant bell curve (not significantly different), or part of two different spiky curves (significantly different).

We need to figure, therefore, the standard deviation squared of the sample.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x - \bar{x})^2}{n} \tag{10}$$

This is the squared, mean distance between the mean of the sample population vs. each sample. Squared distance from the center.

**Why squared?** 1) This function has a well-behaved derivative vs. the absolute value distance. 2) You are biasing for the extremes as they are more far out.

# 6 | How to do the actual comparison

## 6.1 | Getting the Measured Variables

First, let's define the random variable we are interested in.

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{n=launchers} \bar{x_1} i}{n} \tag{11}$$

So our random variable is x-bar-bar: the mean of the launchers.

Therefore, the variants measured from the original variants are:

$$\sigma = \sqrt{{\sigma_1}^2 + {\sigma_2}^2 + {\sigma_3}^2 + \dots} \tag{12}$$

## 6.2 | Finding Statistical Significance

$$Y = \bar{\bar{x_1}} - \bar{\bar{x_0}} \tag{13}$$

Our null hypothesis is that $\mathbb{Y} = 0$, the two distributions have no difference. The distribution of $\mathbb{Y}$ is likely going to be normal, with some $\sigma_{\bar{\bar{x_1}} - \bar{\bar{x_0}}}$ as its width.

The significance of a given $y_0$ appearing in the distribution, therefore, is the integral of the normal curve between the normalized $y_0$ and infinity ("the probability of something being more "extreme").

For all normal curves, the area under $n$% of standard deviation is the same. Therefore, all you have to know is how many standard deviations away is $y_0$.