- Std. Variables
- Probability & Normal Distribution
- Literature: Part 2

Standardized Variables

- Useful when making comparisons between samples
- Useful when trying to determine probability of randomly obtaining a value from a population or sample
- Standardized Values: Z & T

Calculating Z

$$Z_i = \frac{X_i - \bar{X}}{S}$$

Statistics As Probability

- Statistics is based on the assumption that an observed value or sample of values is (or is not) **statistically different** from the population and that this difference is **not** a function of 'chance'

Simple Probability

Getting 5 Heads in a row. The probability of getting 1 head on on flip is ½ or .50

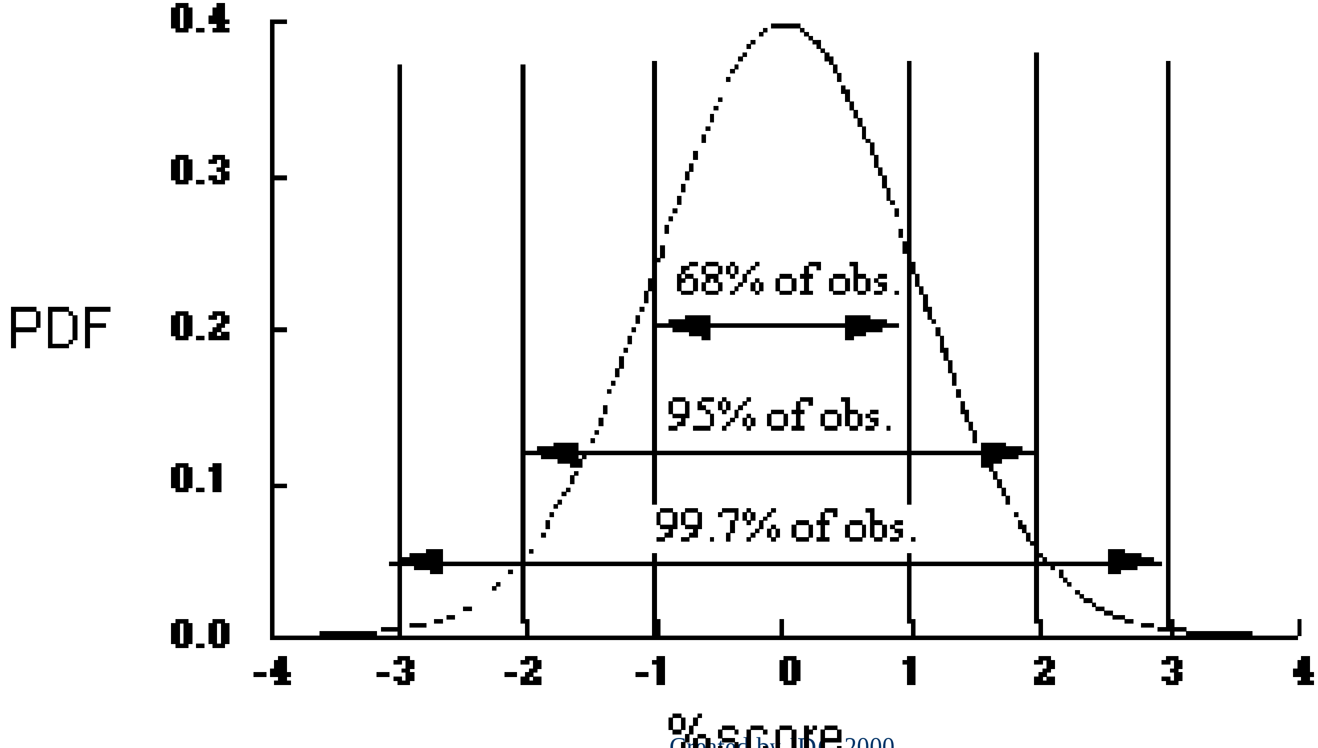
	0	1	2	3	4	5
p(x)	.031	.156	.312	.312	.156	.031

Probability and Normal Distribution

- A normal distribution of all observations--the Bell Curve--makes certain assumptions about central tendency and dispersion
- Based on these assumptions, the probability of a single observation being near the mean (or far away from it) can be calculated
- To apply the principles of probability to a normal distribution, several assumptions are made (and/or conditions met) with respect to the sample distribution: **central limit theorem** & **law of large numbers**

Normal Distribution

Population	Sample
$\mu = 0$	$\bar{X} = 0$
$\sigma = 1$	$s = 1$



Central Limit Theorem

- If repeated (an infinite number of times) samples of size ‘n’ are drawn from a normal population with μ and σ^2, then the sampling distribution of $\overline{X}$ will be normal with μ & σ^2/n

Law of Large Numbers

- If repeated (an infinite number of times) samples of size ‘n’ are drawn from any probability distribution the sample distribution of $\bar{X}$ will be normal with μ & $\sigma 2/\nu$, **<u>IF</u>** the sample size is large enough (e.g., Magic Number 32)

Normal?

- If all conditions are satisfied, a variety of tests can be performed on single or multiple samples to see if they are statistically different (beyond 3 deviations from the mean) from the rest of the population

- Depending on how often you're willing to be wrong, you can accept that an observation is different if 90%, 95%, 99% of all sample means are most likely different
  - This is the basic concept behind 'significance level' (a.k.a. α)

- Depending on how often you're willing to be wrong determines your research's susceptibility to error
  - Type I (proportional to α) & Type II (difficult to assess risk)