Population: a set of numbers from which a sample will be drawn. (think: type of random variable)

Random Sample: A collection of independent & identically distributed random vars (measurements from population)

Statistic: Function of random vars $X_1, \ldots, X_n$ $\quad X_1, \ldots, X_n$

ex: sample mean $\quad \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$

$$S^2 = \frac{\sum (X_i - \overline{X})^2}{n-1} \quad \text{sample variance.}$$

Basic Question

How well does $\overline{X}$ reflect μ? — describes the population.

$$E[\overline{X}] = E\left(\sum X_i/n\right) = \frac{1}{n} \sum E[X_i]$$
$$= \frac{1}{n} \sum \mu = \frac{1}{n} n \mu = \mu.$$

$$Var(\overline{X}) = Var\left[\sum X_i/n\right]$$
$$= \frac{1}{n^2} Var\left(\sum X_i\right) = \frac{1}{n^2} Cov\left(\sum X_i, \sum X_i\right)$$

$Cov(X_i, X_j)_{i \neq j} = 0$

$Cov(X_i, X_i)$

$$\frac{1}{n^2} \sum Var(X_i) = \frac{1}{n^2} n \cdot \sigma^2 = \frac{1}{n} \sigma^2.$$

Suppose population w/ variance $\sigma^2 = 64$.

Sample size of $n = 32$

$$\sigma_{\bar{X}}^2 = \frac{1}{n}\sigma^2 = \frac{64}{32} = 2$$

$$\sigma_{\bar{X}} = \sqrt{2}$$

$$P(|\bar{X} - \mu| < 2) \; ?$$

<u>Chebyshev:</u> $P(|X - \mu_X| < k\sigma_X) \geq 1 - \frac{1}{k^2}$

$$P(|\bar{X} - \mu| < 2) \geq 1 - \frac{1}{(\sqrt{2})^2} = 1 - \frac{1}{2} = \frac{1}{2}$$

$$\underset{|\mu - \bar{X}|}{2} = k\sigma_{\bar{X}} = k\sqrt{2}$$

$$k = \sqrt{2}$$

$$P(\bar{X} - 2 < \mu < \bar{X} + 2) \geq \frac{1}{2}$$

we are at least 50% certain that μ is between $\bar{X} - 2$ & $\bar{X} + 2$

What if we know that the pop was normally distributed?

$\bar{X}$ has mean μ & Variance 2

$\bar{X}$ is normally distributed

$$\rightsquigarrow \frac{\bar{X} - \mu}{\sqrt{2}} = Z \leftarrow \text{std normal.}$$

$$P(|\overline{X}-\mu|<2) = P\left(\left|\frac{\overline{X}-\mu}{\sqrt{2}}\right| < \frac{2}{\sqrt{2}}\right)$$

$$= P(|Z| \leq \sqrt{2})$$

$$= P(-\sqrt{2} \leq Z \leq \sqrt{2})$$

$$\quad -1.4 \qquad\qquad 1.4$$

$$\approx .84$$

$$P(\overline{X}-2 < \mu < \overline{X}+2) \approx .84$$