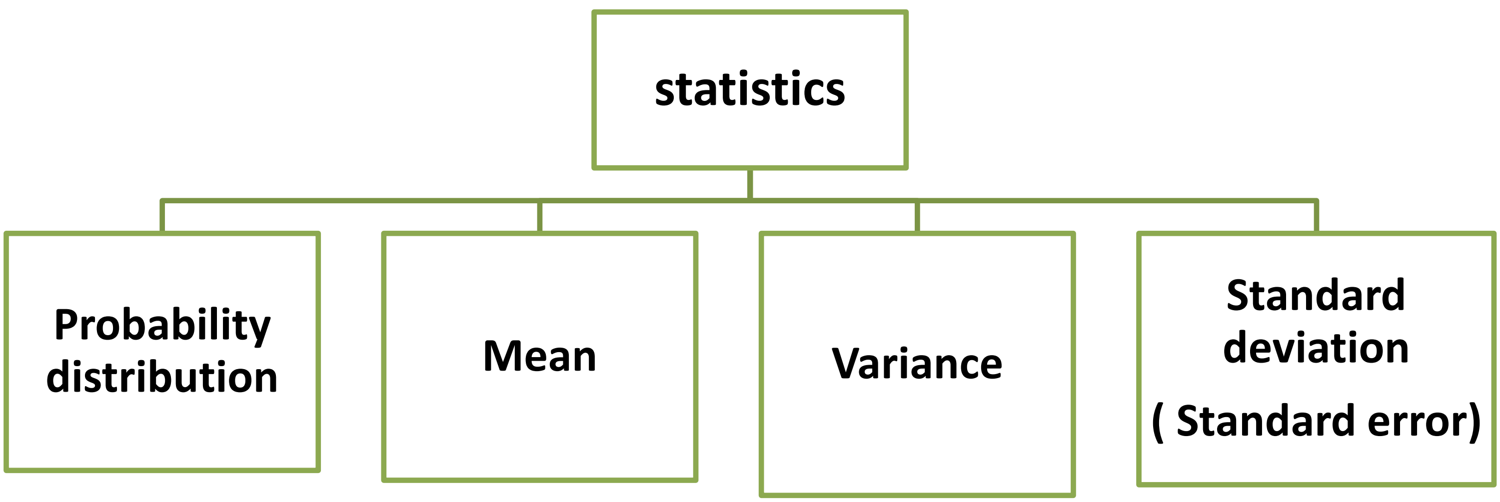
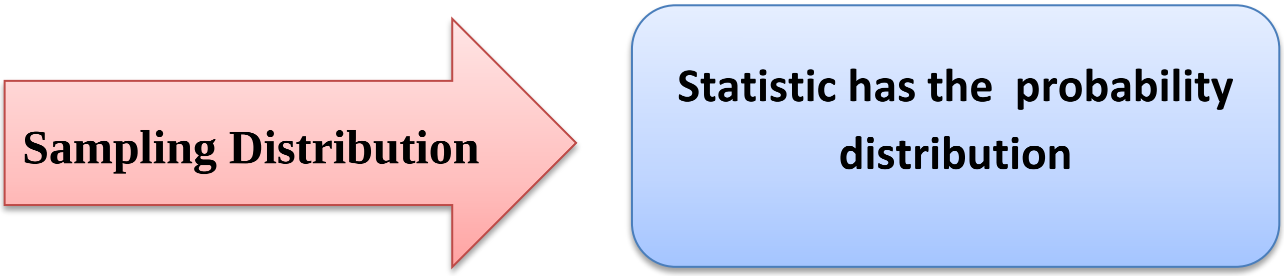
CHAPTER 5: Probabilistic Features of the Distributions of Certain Sample Statistics

Statistics: is measure computed form the random sample.

$\bar{x}$ S^2 S $\hat{p}$ } Example of statistics

* In this Chapter we will discuss the probability distributions of some statistics.





Sampling Distribution of the Sample Mean $\overline{X}$

- **Normal + σ^2 Known**
- **Non-normal + σ^2 Known + n ≥ 30 (large)**

Mean ($\overline{X}$) $= \mu$

Variance ($\overline{X}$) $= \frac{\sigma^2}{n}$

Standard deviation ($\overline{X}$) $= \frac{\sigma}{\sqrt{n}}$

$$\overline{X} \sim Normal\left(\mu, \frac{\sigma^2}{n}\right)$$

$$Z = \frac{\overline{X} - mean}{S.D} = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

- **Normal + σ^2 Unknown + n < 30 (small)**

Mean ($\overline{X}$) $= \mu$

Variance ($\overline{X}$) $= \frac{s^2}{n}$

Standard deviation ($\overline{X}$) $= \frac{s}{\sqrt{n}}$

$\overline{X} \sim t_{(n-1)}$, **df=** n-1

$$T = \frac{\overline{X} - mean}{S.D} = \frac{\overline{X} - \mu}{s / \sqrt{n}}$$

Sampling Distribution of the difference between two Sample Mean

$\bar{X}_1 - \bar{X}_2$

- **Normal + σ_1^2, σ_2^2 Known**
- **Non-normal + σ_1^2, σ_2^2 Known + $n_1, n_2 \geq 30$ (large)**

$$\text{Mean}\ (\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$$

$$\text{Variance}\ (\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$\text{Standard deviation}\ (\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\bar{X}_1 - \bar{X}_2 \sim Normal\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - mean}{S.D} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Sampling Distribution of the Sample Proportion $\hat{p}$

- $n \geq 30$ (large) + $np > 5$ + $nq > 5$, (q =1-p)

Mean $(\hat{p}) = P$

Variance $(\hat{p}) = \frac{pq}{n}$

Standard deviation $(\bar{X}) = \sqrt{\frac{pq}{n}}$

$$\hat{p} \sim Normal\left(p, \frac{pq}{n}\right)$$

$$Z = \frac{\hat{p} - mean}{S.D} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Sampling Distribution of the difference between two Sample Proportions

$\hat{p}_1 - \hat{p}_2$

- $n_1, n_2 \geq 30$ (large) + $n_1 p_1 > 5$ + $n_2 p_2 > 5$ + $n_1 q_1 > 5$ + $n_2 q_2 > 5$

Mean $(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$

Variance $(\hat{p}_1 - \hat{p}_2) = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$

Standard deviation $(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

$$\hat{p}_1 - \hat{p}_2 \sim Normal\left(p_1 - p_2, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)$$

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - mean}{S.D} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$