Chapter 5: The Sampling Distributions of Sample Statistics:

Case 1: The Sampling Distribution of the Sample Mean $\bar{X}$

The Sampling Distribution of the Sample Mean $\bar{X}$

σ^2 is known & Normal or Non-Normal Distribution

$$\text{Mean}(\bar{X}) = \mu \;;\; \text{Variance}(\bar{X}) = \frac{\sigma^2}{n}$$

$$\text{Standard error}(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right);\; Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

σ^2 is unknown & Normal Distribution

$$\text{Mean}(\bar{X}) = \mu \;;\; \text{Variance}(\bar{X}) = \frac{S^2}{n}$$

$$\text{Standard error}(\bar{X}) = \frac{S}{\sqrt{n}}$$

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)\; d.f = n - 1$$

Case 2: The Sampling Distribution of the Difference between two Sample Means $\bar{X}_1 - \bar{X}_2$:

When $n \geq 30$ & σ_1^2 and σ_2^2 are Known & Normal or Non-Normal distribution.

Then

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

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

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

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

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

<u>Case 3:</u> The Sampling Distribution of the Sample Proportion $\hat{\boldsymbol{p}}$:

When $n \geq 30, np > 5, nq > 5$ and $\hat{p} = \frac{X}{n}$.

Then

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

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

$$\text{Standard error}(\hat{p}) = \sqrt{\frac{pq}{n}}$$

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

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$$

<u>Case 4:</u> The Sampling Distribution of Difference between two Sample Proportions $\hat{\boldsymbol{p}}_1 - \hat{\boldsymbol{p}}_2$:

When $n_1 \geq 30, n_2 \geq 30, n_1 p_1 > 5, n_1 q_1 > 5, n_2 p_2 > 5, n_2 q_2 > 5$ and $\hat{p}_1 = \frac{X_1}{n_1}$, $\hat{p}_2 = \frac{X_2}{n_2}$;

$\hat{p}_1 - \hat{p}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2}$.

Then

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

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

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

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

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

Chapter 6:

1. The Point Estimates for the Population Parameters:

	Population Parameter	Point Estimator
Mean	μ	$\bar{X}$
Variance	σ^2	S^2
Standard Deviation	σ	S
Proportion	p	$\hat{p}$
The Difference between two Means	$\mu_1 - \mu_2$	$\bar{X}_1 - \bar{X}_2$
The Difference between two Proportions	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$

2. The Confidence Intervals for the Population Parameters:

Case 1: The Confidence Interval for the Population Mean μ:

The $(1-\alpha)100\%$ Confidence Interval for the Population Mean μ

σ^2 is known + Normal or Non-Normal Distribution

$$\bar{X} \pm Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

σ^2 is unknown + Normal Distribution

$$\bar{X} \pm t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

Case 2: The Confidence Interval for the Difference between two Population Means $\mu_1 - \mu_2$:

The $(1-\alpha)100\%$ Confidence Interval for the Difference between two Population Means $\mu_1 - \mu_2$

σ_1^2 and σ_2^2 are known

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$\sigma_1^2 = \sigma_2^2 = \sigma$ are unknown but equal

$$(\bar{X}_1 - \bar{X}_2) \pm t_{1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}, \; d.f = n_1 + n_2 - 2$$

<u>Case 3:</u> The Confidence Interval for the Population Proportion p:

When $n \geq 30, np > 5, nq > 5$ and $\hat{p} = \frac{X}{n}$.

Then the $(1-\alpha)100\%$ confidence interval for p is

$$\hat{p} \pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

<u>Case 4:</u> The Confidence Interval for the Difference between two Population Proportions $p_1 - p_2$:

When $n_1 \geq 30, n_2 \geq 30, n_1p_1 > 5, n_1q_1 > 5, n_2p_2 > 5, n_2q_2 > 5$ and $\hat{p}_1 = \frac{X_1}{n_1}$, $\hat{p}_2 = \frac{X_2}{n_2}$,

$\hat{p}_1 - \hat{p}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2}$.

Then the $(1-\alpha)100\%$ confidence interval for $p_1 - p_2$ is

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

The General Formulas:

$$Z = \frac{value - mean}{standard\ error}$$

$$estimator \pm (reliability\ cofficient \times standard\ error)$$

Where

Standard error = standard deviation

Estimator = Point estimate

Reliability coefficient = table value = $Z_{1-\frac{\alpha}{2}}$ or $t_{1-\frac{\alpha}{2}}$
