Confidence Interval Estimation

1. 100(1 – **α**)% Confidence Interval for the mean μ, Normal Population or Large Sample

σ known: $$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

σ unknown: $$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

Required sample size to estimate the mean, μ, with a maximum error *e* and with

confidence 1 - α $$n = \left(\frac{z_{\alpha/2}\ \sigma}{e} \right)^2$$

If σ is unknown, and we have a preliminary estimate s then

$$n = \left(\frac{z_{\alpha/2}\ s}{e} \right)^2$$

2. Small Sample 100(1 – **α**)% Confidence Interval for the mean μ of a Normal Population

σ known: $$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

σ unknown $$\bar{x} \pm t_{\alpha/2,f} \frac{s}{\sqrt{n}}, \quad \text{the number of degrees of freedom } f = n - 1$$

3. 100(1 – **α**)% Confidence Interval for the difference in the means $\mu_1 - \mu_2$ using two independent samples. Normal Populations or Large Samples

If σ_1 and σ_2 are known: $$\left(\bar{x}_1 - \bar{x}_2 \right) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

If σ_1 and σ_2 are unknown: $$\left(\bar{x}_1 - \bar{x}_2 \right) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

4. 100(1 – **α**)% Confidence Interval for the difference in the means $\mu_1 - \mu_2$ of Two Normal Populations with unknown equal variances, $\sigma_1^2 = \sigma_2^2$, using two independent small samples.

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\frac{\alpha}{2},\, f} \; s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

where $$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

and the number of degrees of freedom $f = n_1 + n_2 - 2$.

5. 100(1 – **α**)% Confidence Interval for the difference in the means $\mu_1 - \mu_2$ of Two Normal Populations with unknown and unequal variances, using two independent small samples.

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\frac{\alpha}{2},\, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Where the number of degrees of freedom

$$v = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

6. 100(1 – **α**)% Confidence Interval for the difference in the means of two related populations

$$\bar{d} \pm t_{\frac{\alpha}{2}, n-1} \frac{s_d}{\sqrt{n}}$$

Where $\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}$, $d_i = x_{1i} - x_{2i}$, and $s_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}$

7. <u>Large Sample</u> 100(1 – **α**)% Confidence Interval for π, a population proportion

$$p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

Required Sample Size to estimate a population proportion, π, with a maximum error e and with confidence 1 - α:

- if we have a preliminary estimate p

$$n = \frac{z_{\alpha/2}^2 \, p(1-p)}{e^2}$$

- if we do <u>not</u> have a preliminary estimate p

$$n_{\max} = \frac{z_{\alpha/2}^2}{4e^2}$$

8. $(1-\alpha)100\%$ C.I for the difference between two population proportions,

$\pi_1 - \pi_2$, based on <u>large samples</u>.

$$(p_1 - p_2) \pm z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Where $p_1 = \frac{x_1}{n_1}$, $p_2 = \frac{x_2}{n_2}$ are the sample proportions.

Assumptions:

1. $n_1\pi_1 \geq 5$, $n_1(1-\pi_1) \geq 5$
2. $n_2\pi_2 \geq 5$, $n_2(1-\pi_2) \geq 5$