Sampling distribution

$$\mu_{\bar{X}} = \mu \quad , \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2,$$

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Student t-distribution

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}, \quad df = n-1,$$

Chi-Square Distribution:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}, \quad df = n-1$$

F-distribution

$$F = \frac{s_1^2}{s_2^2}, \quad df_1 = n_1 - 1, df_2 = n_2 - 1$$

$$F_{\alpha, n_1, n_2} = \frac{1}{F_{1-\alpha, n_2, n_1}}$$

Confidence Interval Estimation

1) single Sample :C.I for μ

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

b. when n small : $\bar{x} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$

c. Large sample: $\bar{x} \pm z_{\alpha/2} \frac{S}{\sqrt{n}}$

d. Error $= e = \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}}$, $e = \frac{t_{\alpha/2}\,S}{\sqrt{n}}$ or

$$e = \frac{z_{\alpha/2}\,S}{\sqrt{n}}$$

e. Required sample size= n = $\left(\frac{z_{\alpha/2}\,\sigma}{\sqrt{n}}\right)^2$

2) single Sample : C.I for P

a. $\hat{P} \pm z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

b. Error $= z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

c. Required sample size= n=

$$\frac{Z^2_{\alpha/2}\hat{P}(1-\hat{P})}{e^2}$$

3) One sample: C.I for σ

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$$

4) Two Samples :C.I for $\mu_1 - \mu_2$

1) If σ_1 and σ_2 are known

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

2) Large samples

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

3) Small samples:

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

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

$$df = n_1 + n_2 - 2$$

4) Paired samples: C.I for μ_D

$$\bar{d} \pm t_{\alpha/2}\frac{S_d}{\sqrt{n}}, \quad df = n-1$$

5) Two Samples :C.I for $P_1 - P_2$

$$\left(\hat{P}_1 - \hat{P}_2\right) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_1\hat{q}_1}{n_1} + \frac{\hat{P}_2\hat{q}_2}{n_2}}, \text{ where}$$

$$\hat{P}_i = \frac{X_i}{n_i}, \quad \hat{q}_i = 1 - \hat{P}_i$$

Testing Hypotheses

1) single Sample: Testing about μ

a. when σ known: $z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$

b. when n small : $t = \dfrac{\overline{x} - \mu_0}{S / \sqrt{n}}, df = n - 1$

c. Large sample: $Z = \dfrac{\overline{x} - \mu_0}{S / \sqrt{n}}$

2) single Sample :testing about P

Large sample test:

$$z = \frac{X - np_0}{\sqrt{np_0 q_0}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

3) Two Samples :Testing about $\mu_1 - \mu_2$

a. If σ_1 and σ_2 are known

$$z = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

b. Large Samples:

$$z = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

c. Small samples:

$$t = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - d_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

d. Paired samples: Testing about μ_D

$$t = \frac{\overline{d} - d_0}{s_d / \sqrt{n}}$$

4) Two Samples :Testing about : $P_1 - P_2$

Large Samples under $H_0 : P_1 - P_2 = 0$

$$Z = \frac{\widehat{P}_1 - \widehat{P}_2 - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$, where

$$\hat{p} = \frac{n_1 \widehat{P}_1 + n_2 \widehat{P}_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

Simple linear Regression

1) Estimated regression model:

$\hat{y} = a + bx$, where:

$$b = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sum (x - \overline{x})^2} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2}$$

And $a = \overline{y} - b\,\overline{x}$

2) Sum of squares:

a. $S_{XX} = \sum \left(x - \overline{x}\right)^2 = \sum x^2 - n\left(\overline{x}\right)^2$

b. $S_{YY} = \sum \left(y - \overline{y}\right)^2 = \sum y^2 - n\left(\overline{y}\right)^2$

c. $S_{XY} = \sum \left(x - \overline{x}\right)\left(y - \overline{y}\right) = \sum xy - n\left(\overline{x}\right)\left(\overline{y}\right)$

d. $SSE = \sum \left(y_i - \widehat{y}_i\right)^2 = S_{YY} - b\,S_{XY}$

e. $SST = \sum \left(y_i - \overline{y}\right)^2$

3) $S^2 = \dfrac{SSE}{n - 2}$

4) Inference about the regression coefficients

a. C.I for β : $b \pm t_{\alpha/2} \dfrac{S}{\sqrt{S_{XX}}}, df = n - 2$

b. C.I. for α : $a \pm \dfrac{t_{\alpha/2} S \sqrt{\sum x^2}}{\sqrt{n S_{XX}}}$

c. Testing about β**:** $t = \dfrac{b - \beta_0}{S \big/ \sqrt{S_{XX}}}$

d. Testing about α**:** $t = \dfrac{a - \alpha_0}{S\sqrt{\dfrac{\sum x^2}{n S_{XX}}}}$

5) C. I for the mean response $\mu_{Y|X_0}$

$$\hat{y}_0 \pm t_{\alpha/2} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}}$$

6) P.I for a single response y_0 **is:**

$$\hat{y}_0 \pm t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}}$$

7) Coefficient of determination

$$R^2 = 1 - \frac{SSE}{SST}$$

8) Correlation coefficient

$$r = b\sqrt{\frac{S_{XX}}{S_{YY}}} = \frac{S_{XY}}{\sqrt{S_{XX}\ S_{YY}}},\ r^2 = R^2$$

9) Testing about the population correlation coefficient:

$$t = \frac{b}{S\big/\sqrt{S_{XX}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\ df = n - 2$$