Sample Mean $\bar{X} = \frac{\sum x}{n}$

Sample Variance $s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

Grouped Data:

x_i^* = mid point of interval *i*

f_i = Frequency of interval *i*

Sample Mean $\frac{\sum x_i^* f_i}{\sum f_i}$

Sample Variance $\frac{\sum x_i^{*2} f_i - \left(\sum x_i^* f_i\right)^2 / n}{n-1}$

Random Variables:

Binomial

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,\ldots,n$$

Hypergeometric

$$P(X = x) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}, \quad x = 0,1,\ldots,\max(n,a)$$

Geometric $P(X = x) = (1-p)^{x-1} p, \quad x = 1,2,\ldots$

Poisson $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,\ldots$

Uniform $f(x) = \begin{cases} \frac{1}{\beta-\alpha}, \alpha < x < \beta \\ 0 \quad elsewhere \end{cases}$

Exponential $f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, x > 0 \\ 0 \quad elsewhere \end{cases}$

Confidence Interval Estimation

I. One Sample Problem:

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

sample size: $n = \left(\frac{z_{\alpha/2}\ \sigma}{E}\right)^2$

σ unknown,

small sample: $\bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

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

Large Sample Confidence Interval Estimation of *p*, a population proportion

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

Sample Size: $n = \frac{z_{\alpha/2}^2\, p(1-p)}{E^2}$

Maximum Sample Size: $\frac{z_{\alpha/2}^2}{4E^2}$

II. Two Sample Problem:

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 unknown, Large samples

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

σ_1 and σ_2 unknown, $\sigma_1^2 = \sigma_2^2$ small samples

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

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}, \quad f = n_1 + n_2 - 2.$$

Hypothesis Testing

I. One Sample Problem:

σ^2 known, Test Statistic $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

σ^2 unknown, small sample,

Test Statistic $t = \dfrac{\bar{x} - \mu_0}{s / \sqrt{n}}$

σ^2 unknown, large sample

Test Statistic $z = \dfrac{\bar{x} - \mu_0}{s / \sqrt{n}}$

A population proportion, large sample Test

$$z = \frac{p - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

II. Two Sample Problem

σ_1^2 and σ_2^2 known

Test Statistic $z = \dfrac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

σ_1^2 and σ_2^2 unknown, large samples

Test Statistic $z = \dfrac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

σ_1^2 and σ_2^2 unknown, $\sigma_1^2 = \sigma_2^2$ small samples

Test Statistic $z = \dfrac{\bar{x}_1 - \bar{x}_2 - d_0}{s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

Simple linear Regression

Estimated regression model:

$\hat{y} = a + bx$, where: $b = \dfrac{S_{xy}}{S_{xx}}, \quad a = \bar{y} - b\bar{x}$

$$S_{xx} = \sum\left(x_i - \bar{x}\right)^2 = \sum x_i^2 - n\left(\bar{x}\right)^2$$

$$S_{yy} = \sum\left(y_i - \bar{y}\right)^2 = \sum y_i^2 - n\left(\bar{y}\right)^2$$

$$S_{xy} = \sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = \sum x_i\, y_i - n\,\bar{x}\,\bar{y}$$

$$SSE = \sum\left(y_i - y_i\right)^2 = S_{yy} - b\,S_{xy}$$

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

$$s_e^2 \equiv MSE = \frac{SSE}{n-2}$$

Inference about the regression coefficients

C.I for β: $\quad b \pm t_{\alpha/2, n-2} \dfrac{s_e}{\sqrt{S_{xx}}}$

C.I. for α: $\quad a \pm t_{\alpha/2, n-2} \dfrac{s_e\sqrt{\sum x^2}}{\sqrt{n\,S_{xx}}}$

Testing β: $\quad \mathrm{t} = \dfrac{b - \beta_0}{s_e \big/ \sqrt{S_{xx}}}$

Testing α: $\quad \mathrm{t} = \dfrac{a - \alpha_0}{s_e\sqrt{\dfrac{\sum x^2}{n\,S_{xx}}}}$

C. I for the mean response at x_0 : $\mu_{Y|x_0}$

$$\hat{y}(x_0) \pm t_{\alpha/2, n-2}\, s_e \sqrt{\frac{1}{n} + \frac{\left(x_0 - \bar{x}\right)^2}{S_{xx}}}$$

P.I for a future response at x_0 is:

$$\hat{y}(x_0) \pm t_{\alpha/2, n-2}\, s_e \sqrt{1 + \frac{1}{n} + \frac{\left(x_0 - \bar{x}\right)^2}{S_{xx}}}$$

Coefficient of determination

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

Correlation coefficient

$$r = b\sqrt{\frac{S_{xx}}{S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx}\; S_{yy}}}, \quad r^2 = R^2$$

Testing the correlation coefficient ρ = 0

Test Statistic $Z = \sqrt{n-3}\,\ln\left(\dfrac{1+r}{1-r}\right)$