General

Median:

$$\tilde{x} = \begin{cases} x_{\left(\frac{N+1}{2}\right)} & \text{if } N \text{ is odd,} \\ \\ \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2} & \text{if } N \text{ is even.} \end{cases}$$

IQR:

$$Q_3 - Q_1$$

Sample Mean:

$$\bar{x} = \frac{\sum x}{N}$$

Sample Standard Deviation:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{N - 1}}$$

Z-Score:

$$z = \frac{x - \bar{x}}{s}$$

One Sample T-Test

Degrees of Freedom:

$$df = N - 1$$

Standard Error:

$$s_{\bar{x}} = \frac{s}{\sqrt{N}}$$

Confidence Interval:

$$\bar{x} \pm t_{\mathrm{crit}} \cdot s_{\bar{x}}$$

Test Statistic:

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}$$

Cohen's d:

$$d = \frac{\bar{x} - \mu_0}{s}$$

Hedge's g:

$$g = d \cdot \left(1 - \frac{3}{4(N) - 9}\right)$$

Paired T-test

Degrees of Freedom:

$$df = N_D - 1$$

Standard Error:

$$s_{\bar{D}} = \frac{s_D}{\sqrt{N}}$$

Confidence Interval:

$$\bar{D} \pm t_{\mathrm{crit}} \cdot s_{\bar{D}}$$

Test Statistic:

$$t = \frac{\bar{D} - \mu_D}{s_{\bar{D}}}$$

Cohen's d:

$$d = \frac{\bar{D} - \mu_D}{s_{rm}}$$

$$s_{rm} = \frac{s_D}{\sqrt{2(1-R)}}$$

Hedge's g:

$$g = d \cdot \left(1 - \frac{3}{4(N-1)-1}\right)$$

Independent T-test

Degrees of Freedom:

$$df = N - 2$$

Standard Error:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_{n_1}^2}{n_1} + \frac{s_{n_2}^2}{n_2}}$$

Pooled Variance:

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

Pooled Standard Error:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

Confidence Interval:

$$\bar{x}_1 - \bar{x}_2 \pm t_{\text{crit}} \cdot s_{\bar{x}_1 - \bar{x}_2}$$

Test Statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}_1 - \bar{x}_2}}$$

Cohen's d:

$$d = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2}}$$

Hedge's g:

$$g = d \cdot \left(1 - \frac{3}{4(n_1 + n_2) - 9}\right)$$

Regression

Pearson's Correlation Coefficient:

$$R = \frac{\sum (Z_x \cdot Z_y)}{N - 1}$$

Simple Linear Model:

$$\hat{y} = b_0 + b_1 x$$

Slope:

$$b_1 = R \frac{s_y}{s_x}$$

Intercept:

$$b_0 = \bar{y} - b_1 \cdot \bar{x}$$

Sum of Squared Residuals:

$$SSR = \sum (y - \hat{y})^2$$

Total Sum of Squares:

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

Standard Error of the Residuals:

$$s_{resid} = \sqrt{\frac{\sum (y - \hat{y})^2}{df}}$$

Degrees of freedom:

$$df = N - P - 1$$

Standard Error of the Intercept:

$$s_{b_0} = \sqrt{\frac{s_{\text{resid}}^2 \cdot \sum x_i^2}{N \cdot SS_x}}$$

Standard Error of the Slope:

$$s_{b_1} = \sqrt{\frac{s_{\text{resid}}^2}{SS_x}}$$

Sum of Squared Deviations (of x):

$$SS_x = \sum (x_i - \bar{x})^2$$

Regression Tests

Confidence Interval on Intercept:

$$b_0 \pm t_{\rm crit} \cdot s_{b_0}$$

Test Statistic on Intercept:

$$t = \frac{b_0 - \beta_0}{s_{b_0}}$$

Confidence Interval on Slope:

$$b_1 \pm t_{\text{crit}} \cdot s_{b_1}$$

Test Statistic on Slope:

$$t = \frac{b_1 - \beta_1}{s_{b_1}}$$

Robust Outlier Detection

MADN:

$$MADN = \frac{MAD}{0.6745}$$

MADN Rule:

$$\frac{|x - \text{med}|}{\text{MADN}} > 2.24$$

Chi-Squared Analysis

Goodness of Fit Test:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_j - E_j)^2}{E_j}$$

Degrees of Freedom:

$$df = k - 1$$

Test of Independence:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Cell Expected Value:

$$E_{ij} = \frac{R_i \cdot C_j}{N}$$

Degrees of Freedom:

$$df = (r-1)(c-1)$$

One-Way Independent ANOVA

Structural Model for ANOVA:

$$\hat{y}_{ij} = \mu + (\mu_j - \mu) + (y_{ij} - \mu_j)$$

= $\mu + \tau_j + \epsilon_{ij}$

Grand Mean:

$$\bar{Y}_{grand} = \frac{\sum Y_i}{N}$$

Total Sum of Squares:

$$SST = \sum_{i=1}^{N} (Y_i - \bar{Y}_{grand})^2$$

$$df_T = N - 1$$

Model Sum of Squares:

$$SSM = \sum_{n=1}^{j} n_j \cdot (\bar{Y}_j - \bar{Y}_{grand})^2$$

$$df_M = j - 1$$

Residual Sum of Squares:

$$SSR = \sum_{i=1}^{n} (Y_{ij} - \bar{Y}_j)^2$$

$$= SST - SSM$$

$$df_R = df_T - df_M$$

Mean Squares:

$$MS_M = \frac{SSM}{df_M}$$

$$MS_R = \frac{SSR}{df_R}$$

F-Ratio:

$$F = \frac{MS_M}{MS_R}$$

Probability

General Addition Rule:

If A and B are any two events, disjoint or not, then the probability that at least one of them will occur is

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

where P(A and B) is the probability that both events occur.

Multiplication Rule for Independent Processes:

If A and B represent events from two different and independent processes, then the probability that both A and B occur can be calculated as the product of their separate probabilities:

$$P(A \text{ and } B) = P(A) \times P(B)$$

Conditional Probability:

The conditional probability of outcome A given condition B is computed as the following:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

General Multiplication Rule:

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A|B) \times P(B)$$