Formula Sheet

Value of Test Statistic	Confidence Interval estimation
$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ known	$\bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}},$
$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n - 1,$ σ unknown	$\bar{x} - t_{\alpha/2}\dfrac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\dfrac{s}{\sqrt{n}},$
$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$; σ_1 and σ_2 known	$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}},$
$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$; $v = n_1 + n_2 - 2,$ $\sigma_1 = \sigma_2$ but unknown, $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}},$
$t' = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$; $v = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}},$ $\sigma_1 \neq \sigma_2$ and unknown	$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}},$
$t = \dfrac{\bar{d} - d_0}{s_d/\sqrt{n}}$; $v = n - 1$	$\bar{d} - t_{\alpha/2}\dfrac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\alpha/2}\dfrac{s_d}{\sqrt{n}},$

Simple linear regression and correlation co-efficients and test statistic for correlation co-efficient

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{\left[n\Sigma x^2 - (\Sigma x)^2\right]\left[n\Sigma y^2 - (\Sigma y)^2\right]}}$$

$$b_0 = \frac{\sum_{i=1}^{n} y_i - b_1\sum_{i=1}^{n} x_i}{n} = \bar{y} - b_1\bar{x}.$$

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, \qquad v = n - 2$$

ANOVA Formulae

Sum of squares	Defining formula	Computing formula
Total, SST	$\Sigma(x_i - \bar{x})^2$	$\Sigma x_i^2 - (\Sigma x_i)^2/n$
Treatment, SSTR	$\Sigma n_j(\bar{x}_j - \bar{x})^2$	$\Sigma(T_j^2/n_j) - (\Sigma x_i)^2/n$
Error, SSE	$\Sigma(n_j - 1)s_j^2$	$SST - SSTR$

Correlation co-efficient and co-variance formulae for joint PMF/PDF

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y. \qquad \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}.$$

Inference on regression co-efficients ($(v = n - 2)$)

$$b_1 - t_{\alpha/2}\frac{s}{\sqrt{S_{xx}}} < \beta_1 < b_1 + t_{\alpha/2}\frac{s}{\sqrt{S_{xx}}},$$

$$b_0 - t_{\alpha/2}\frac{s}{\sqrt{nS_{xx}}}\sqrt{\sum_{i=1}^{n} x_i^2} < \beta_0 < b_0 + t_{\alpha/2}\frac{s}{\sqrt{nS_{xx}}}\sqrt{\sum_{i=1}^{n} x_i^2},$$

T-Statistic for β_0

$$T = \frac{B_0 - \beta_0}{S\sqrt{\sum_{i=1}^{n} x_i^2/(nS_{xx})}}$$

T-Statistic for β_1

$$T = \frac{(B_1 - \beta_1)/(\sigma/\sqrt{S_{xx}})}{S/\sigma} = \frac{B_1 - \beta_1}{S/\sqrt{S_{xx}}}$$

Total probability and Baye's Rule formulae

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i). \qquad P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$$

Gradient Descent Algorithm

$$b_0^{new} = b_0^{old} - \alpha\frac{\partial L}{\partial b_0}$$

$$b_1^{new} = b_1^{old} - \alpha\frac{\partial L}{\partial b_1}$$

$$\frac{\partial L}{\partial b_0} = \frac{1}{n}\sum_{1}^{n}(b_0 + b_1 x - y) : \qquad \frac{\partial L}{\partial b_1} = \frac{1}{n}\sum_{1}^{n}(b_0 + b_1 x - y)x$$