Formula Sheet: Probability and Stats

1. t-tests: (a) $\frac{\bar{x}-\mu}{s/\sqrt{n}}$ (b) t = $\frac{\bar{X}_1-\bar{X}_2-\mu_1-\mu_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}}$
2. Coefficient of correlation = r = $\frac{n\sum YX-\sum Y\sum X}{\sqrt{n\sum X^2-(\sum X)^2}\sqrt{n\sum Y^2-(\sum Y)^2}}$
3. $SS_T = \sum\sum y_{ij}^2 - \frac{y_{..}^2}{N}$ *and* $SS_{treatments} = \sum\frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}$
4. Test of linearity for regression coefficient: $t = \frac{b_1}{s/\sqrt{S_{xx}}}$

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 \quad S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2, \quad S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}).$$

$$s^2 = \frac{S_{yy} - b_1 S_{xy}}{n-2}$$

Determining Regression Equation

$$a = \frac{(\Sigma\, y)(\Sigma\, x^2) - (\Sigma\, x)(\Sigma\, xy)}{n(\Sigma\, x^2) - (\Sigma\, x)^2}$$
$$b = \frac{n(\Sigma\, xy) - (\Sigma\, x)(\Sigma\, y)}{n(\Sigma\, x^2) - (\Sigma\, x)^2}$$

Hypothesis testing for Correlation :

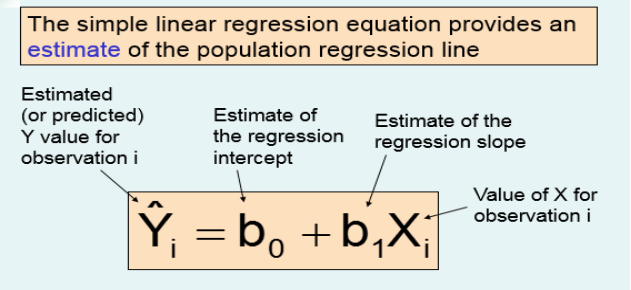
$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

with degrees of freedom equal to $n - 2$.

Simple Linear Regression Model: $Y = \beta_0 + \beta_1 x + \epsilon.$

ANOVA Table

Source of Variation	Sum of Squares	d.f.	Mean Square	F_o
Between Treatment (groups)	$k\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})^2$	k − 1	$MS_{Treatments}$	$Fo = MS_{Treatments} / MS_E$
Error (within treatment)	$SS_E = SS_T - SS_{treatment}$	N − k	MS_E	
Total	$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2$	N − 1		



ANOVA : $SS_T = \sum\sum y_{ij}^2 - \frac{y_{..}^2}{N}$ *and* $SS_{treatments} = \sum \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = n\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2$$

$$SSTR = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2 + n_4(\bar{x}_4 - \bar{x})^2$$

$$SS_{\text{Treatments}} = \sum_{i=1}^{a}\frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N}$$

Multiple regression Model:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Normal Equation:

$$\sum Y = na + b_1\sum X_1 + b_2\sum X_2$$
$$\sum X_1 Y = a\sum X_1 + b_1\sum X_1^2 + b_2\sum X_1 X_2$$
$$\sum X_2 Y = a\sum X_2 + b_1\sum X_1 X_2 + b_2\sum X_2^2$$

t-test: (a) $\frac{\bar{x}-\mu}{s/\sqrt{n}}$ **,** $t = \frac{\bar{X}_1 - \bar{X}_2 - \mu_1 - \mu_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $t = \frac{\bar{X}_1 - \bar{X}_2 - \mu_1 - \mu_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$

Coefficient of correlation= r = $\frac{n\sum YX - \sum Y\sum X}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$

Multiple correlation coefficient: R = $\sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}.r_{yx_2}.r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$

Standard error = $\sqrt{\frac{\sum(y-y')^2}{n-2}}$ **,** **pooled Variance :** $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1-1)+(n_2-1)}$

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Related Populations The Paired Difference Test : $D_i = X_{1i} - X_{2i}$,

$$\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n} \qquad t_{STAT} = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1}}$$

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown and not assumed equal:

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$ **Or Take : V = n-2 , with the least data value (Rule of thumb) see problem in book or ppts.**

Note : Students this formula sheet is prepared from PPts however some formulas are missing while you practice problem Remember them.

Only selected formulas will be provided in Final Exam Papers , so make sure you have remember some Formulas.