PSYCO 413 Formula Sheet

If you are viewing this in your browser you can use `ctrl shift +` to zoom in on windows and `command +` on a mac.

Note:

- N = Total sample size
- n = Level/Group sample size

Spread

Sample Standard Deviation and Variance

$S = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}} \qquad S^2 = \frac{\sum(X-\bar{X})^2}{n-1}$

Z-Score:

$Z = \frac{X - \bar{X}}{S}$

One-Sample T-test / Paired T-test

Classic

Standard Error:

$S_{\bar{X}} = \frac{S}{\sqrt{N}}$

Degrees of Freedom:

$df = N - 1$

Test Statistic:

$T = \frac{\bar{X} - \mu_0}{s_{\bar{X}}}$

Trimmed

Standard Error:

$s_{\bar{X}_t} = \frac{S_w}{(1-2G)\sqrt{N}}$ G = Proportion trimmed (on each side)

Degrees of Freedom:

$df_t = h - 1 \qquad h = N - 2 \cdot \lfloor G \cdot N \rfloor$

Test Statistic:

$T_t = \frac{(1-2G)(\bar{X}_t - \mu_0)}{\frac{S_w}{\sqrt{N}}}$

Two-Sample Independent T-Test

Classic

Standard Error (pooled):

$$S_{\bar{X}_2-\bar{X}_1} = \sqrt{\frac{S_p^2}{n_2} + \frac{S_p^2}{n_1}} \qquad S_p^2 = \frac{(n_2-1)\, S_2^2 + (n_1-1)\, S_2^2}{n_2 + n_1 - 2}$$

Degrees of Freedom:

$$df = n_2 + n_1 - 2$$

Test-Statistic:

$$T = \frac{(\bar{X}_2-\bar{X}_1)-(\mu_2-\mu_1)}{S_{\bar{X}_2-\bar{X}_1}}$$

Effect Size: d

$$d = \frac{\bar{X}_2-\bar{X}_1}{S_p}$$

Effect Size: g

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

Welch:

Standard Error:

$$S_{\bar{X}_2-\bar{X}_1} = \sqrt{\frac{S_2^2}{n_2} + \frac{S_1^2}{n_1}}$$

Test Statistic:

$$T = \frac{(\bar{X}_2-\bar{X}_1)-(\mu_2-\mu_1)}{S_{\bar{X}_2-\bar{X}_1}}$$

Degrees of Freedom:

$$df = \frac{(q_1+q_2)^2}{\frac{q_1^2}{n_1-1}+\frac{q_2^2}{n_2-1}} \qquad q_j = \frac{S_j^2}{n_j}$$

Trimmed Means:

Squared-Standard Error (for the jth group):

$$d_j = \frac{(n_j-1)S_{wj}^2}{h_j(h_j-1)} \qquad h_j = n_j - 2 \cdot \lfloor G \cdot n_j \rfloor$$

Degrees of Freedom:

$$df = \frac{(d_2+d_1)^2}{\frac{d_2^2}{h_2-1}+\frac{d_1^2}{h_1-1}}$$

Test-Statistic:

$$T_y = \frac{(\bar{X}_{t2} - \bar{X}_{t1}) - (\mu_2 - \mu_1)}{\sqrt{d_2 + d_1}}$$

Confidence Interval:

$$(\bar{X}_{t2} - \bar{X}_{t1}) \pm \sqrt{d_2 + d_1} \cdot T_{crit}$$

Simple Linear Regression

(i.e., Linear Regression with one predictor)

Pearson's Correlation Coefficient R:

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

Standard Error of R:

$$S_R = \sqrt{\frac{1-R^2}{N-2}}$$

Degrees of Freedom of R:

$$df_R = N - 2$$

Slope:

$$b_1 = R\frac{S_y}{S_x}$$

Intercept:

$$b_0 = \bar{Y} - b_1 \cdot \bar{X}$$

Slope and Intercept Standard Error:

$$S_{b_1} = \sqrt{\frac{S^2_{resid}}{SS_X}} \qquad S_{b_0} = \sqrt{\frac{S^2_{resid} \cdot \sum X_i^2}{N \cdot SS_X}}$$

$$S^2_{resid} = \frac{\sum(Y - \hat{Y})^2}{N-2} \qquad SS_X = \sum(X_i - \bar{X})^2$$

Degrees of Freedom:

$df = N - P - 1$ $\qquad$ P = The number of predictors

Test-Statistics:

$$T_{b_1} = \frac{b_1 - \beta_1}{S_{b_1}} \qquad T_{b_0} = \frac{b_0 - \beta_0}{S_{b_0}}$$

Multiple Regression and ANOVA

Multiple R-Squared:

$$R^2 = \frac{SS_M}{SS_T}$$

Adjusted R-Squared:

$$R^2_{adj} = 1 - \frac{(1 - R^2)(N - 1)}{N - P - 1}$$

Total Sum of Squares Formulas (SS_T):

$$\begin{aligned} SS_T &= SS_M + SS_R \\ &= \Sigma(Y_i - \bar{Y}_{grand})^2 \\ &= S^2_{grand} \cdot (N - 1) \end{aligned}$$

i = individual observations

$df_T = N - 1$

Model Sum of Squares Formulas (SS_M):

$$\begin{aligned} SS_M &= SS_T - SS_R \\ &= \Sigma(\hat{Y}_i - \bar{Y}_{grand})^2 \\ &= \Sigma\, n_j \cdot (\bar{Y}_j - \bar{Y}_{grand})^2 \end{aligned}$$

j = number of levels/groups across predictors

$df_M = j - 1$

Main Effect and Interaction Sum of Squares (SS_A, SS_A, $SS_{A\times B}$):

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

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

$$SS_{A\times B} = SS_M - SS_A - SS_B$$

j = levels/groups within the predictor

$df_A = j - 1 \qquad df_B = j - 1 \qquad df_{A\times B} = df_A \cdot df_B$

Residual Sum of Squares Formulas (SS_R or SS_E):

$$\begin{aligned} SS_R &= SS_T - SS_M \\ &= \Sigma(Y_i - \hat{Y}_i)^2 \\ &= \Sigma(Y_{ij} - \bar{Y}_j)^2 \\ &= \Sigma S_j^2 \cdot (n_j - 1) \end{aligned}$$

$df_R = df_T - df_M$

Mean-Squares:

$$MS_x = \frac{SS_x}{df_x}$$

F-stat to Compare Regression Slopes:

$$F = \left(\frac{N-P-1}{P}\right)\left(\frac{R^2}{1-R^2}\right) \qquad df_1 = P \qquad df_2 = N - P - 1$$

F-stat to Compare Two Hierarchical Regression Models:

$$F_{diff} = \left(\frac{N-P_{mod2}-1}{P_{diff}}\right)\left(\frac{R^2_{diff}}{1-R^2_{mod2}}\right) \qquad df_1 = P_{diff} \qquad df_2 = N - P_{mod2} - 1$$

F-stat to test Main Effects and Interactions in an ANOVA

$$F = \frac{MS_x}{MS_R} \qquad df_1 = SS_x \qquad df_2 = SS_R$$

Effect Size

Two Independent Means:

Cohen's d:

$$d = \frac{\bar{x}_2 - \bar{x}_1}{S_p}$$

$$S_p = \sqrt{\frac{(n_1-1)\cdot S_1^2 + (n_2-1)\cdot S_2^2}{n_1+n_2-1}}$$

Hedge's g:

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

Two Dependent Means:

Cohen's d:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{S_{rm}}$$

$$S_{rm} = \frac{S_D}{\sqrt{2(1-R)}}$$

S_D = Standard Deviation of the difference scores.

R = Pearson's Correlation Coefficient

Hedge's g:

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

ANOVA Main Effects and Interaction

Note: A one way ANOVA is just a Two-way ANOVA with a single Main effect and no interaction.

$\omega_x^2 = \frac{SS_x - df_x \cdot MS_R}{SS_T + MS_R}$

x = predictor or interaction

ANOVA Planned Comparisons in terms of R:

$R = \frac{T^2}{T^2 + df}$

$df = N - P - 1$

T = The test-statistic for the contrast's b value.

P = the number of predictors (i.e., dummy variables/contrasts)

ANOVA Pairwise Comparisons in terms of D:

see Cohen's d for a Two-Sample Independent T-test.

Chi-Square Test

$\chi^2 = \sum \frac{(O-E)^2}{E}$

- O = observed frequency
- E = Expected frequency
- Sum across cells.

Degrees of Freedom for One Variable:

$df = k - 1$

k = Number of categories.

Expected Frequencies for Two Variables:

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

- R = row total
- C = Column total

Degrees of Freedom for Two Variables:

$df = (R - 1)(C - 1)$

- R = Number of rows in the table.
- C = Number of columns in the table.

Miscellaneous

MAD-Median Rule:

$$\frac{|X - \text{median}(x)|}{\text{MADN}} > 2.24$$

Biserial Correlation:

$$R_b = \frac{R_{pb} \cdot \sqrt{p \cdot q}}{u}$$

- R_{pb} = Point-Biseral Correlation.
- p = Proportion of cases falling into the largest category.
- q = Proportion of cases falling into the smallest category.
- u = the ordinate (i.e., y-axis value) of a standard normal distribution at p.
- Standard Error: $S_{R_b} = \frac{\sqrt{pq}}{u\sqrt{N}}$
- Test Statistic: $Z_{R_b} = \frac{R_b - \rho_b}{S_{R_b}}$