SUMMARY OF STATISTICS FORMULAS

THE MEAN

Population: $\mu = \frac{\Sigma X}{N}$ Sample: $M = \frac{\Sigma X}{n}$

SUM OF SQUARES

Definitional: $SS = \Sigma(X - \mu)^2$

Computational: $SS = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$

VARIANCE

Population: $\sigma^2 = \frac{SS}{N}$ Sample: $s^2 = \frac{SS}{n-1}$

STANDARD DEVIATION

Population: $\sigma = \sqrt{\frac{SS}{N}}$ Sample: $s = \sqrt{\frac{SS}{n-1}}$

z-SCORE (FOR LOCATING AN X VALUE)

$$z = \frac{X - \mu}{\sigma}$$

BINOMIAL (NORMAL APPROXIMATION)

$z = \frac{X - pn}{\sqrt{npq}}$ or $z = \frac{X/n - p}{\sqrt{pq/n}}$

z-SCORE (FOR LOCATING A SAMPLE MEAN)

$z = \frac{M - \mu}{\sigma_M}$ where $\sigma_M = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}}$

t STATISTIC (SINGLE SAMPLE)

$t = \frac{M - \mu}{s_M}$ where $s_M = \sqrt{\frac{s^2}{n}}$

t STATISTIC (INDEPENDENT MEASURES)

$$t = \frac{(M_1 - M_2) - (\mu_1 - \mu_2)}{s_{(M_1 - M_2)}}$$

where $s_{(M_1 - M_2)} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$ and $s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$

t STATISTIC (RELATED SAMPLES)

$t = \frac{M_D - \mu_D}{s_{M_D}}$ where $s_{M_D} = \sqrt{\frac{s^2}{n}}$

ESTIMATION

t Statistic (Single Sample)

$$\mu = M \pm t s_M$$

t Statistic (Independent Measures)

$$\mu_1 - \mu_2 = M_1 - M_2 \pm t s_{(M_1 - M_2)}$$

t Statistic (Related Samples)

$$\mu_D = M_D \pm t s_{M_D}$$

INDEPENDENT-MEASURES ANOVA

$SS_{\text{total}} = \Sigma X^2 - \frac{G^2}{N}$ $df_{\text{total}} = N - 1$

$SS_{\text{between}} = \Sigma \frac{T^2}{n} - \frac{G^2}{N}$ $df_{\text{between}} = k - 1$

$SS_{\text{within}} = \Sigma SS_{\text{inside each treatment}}$ $df_{\text{within}} = N - k$

$F = \frac{MS_{\text{between}}}{MS_{\text{within}}}$ where each $MS = \frac{SS}{df}$

REPEATED-MEASURES ANOVA

$SS_{\text{between}} = \Sigma \frac{T^2}{n} - \frac{G^2}{N}$ $df_{\text{between}} = k - 1$

$SS_{\text{error}} = SS_{\text{within}} - SS_{\text{subjects}}$

$df_{\text{error}} = (N - k) - (n - 1)$

where $SS_{\text{within}} = \Sigma SS_{\text{inside each treatment}}$

and $SS_{\text{subjects}} = \Sigma \frac{P^2}{k} - \frac{G^2}{N}$

$F = \frac{MS_{\text{between}}}{MS_{\text{error}}}$ where each $MS = \frac{SS}{df}$

TWO-FACTOR ANOVA

$$SS_{\text{between treatments}} = \Sigma\frac{T^2}{n} - \frac{G^2}{N}$$

$$df_{\text{between treatments}} = \text{number of cells} - 1$$

$$SS_{\text{within treatments}} = \Sigma SS_{\text{each treatment}}$$

$$df_{\text{within treatments}} = \Sigma df_{\text{each treatment}}$$

$$SS_A = \Sigma\frac{T^2_{\text{ROW}}}{n_{\text{ROW}}} - \frac{G^2}{N}$$

$$df_A = (\text{number of levels of } A) - 1$$

$$SS_B = \Sigma\frac{T^2_{\text{COL}}}{n_{\text{COL}}} - \frac{G^2}{N}$$

$$df_B = (\text{number of levels of } B) - 1$$

$$SS_{A\times B} = SS_{\text{between treatments}} - SS_A - SS_B$$

$$df_{A\times B} = df_{\text{between treatments}} - df_A - df_B$$

$$F_A = \frac{MS_A}{MS_{\text{within}}} \qquad F_B = \frac{MS_B}{MS_{\text{within}}} \qquad F_{A\times B} = \frac{MS_{A\times B}}{MS_{\text{within}}}$$

where each $MS = \frac{SS}{df}$

PEARSON CORRELATION

$$r = \frac{SP}{\sqrt{SS_X SS_Y}}$$

where $SP = \Sigma(X - M_X)(Y - M_Y) = \Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{n}$

SPEARMAN CORRELATION

$$r_s = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)}$$

LINEAR REGRESSION

$\hat{Y} = bX + a$ where $b = \frac{SP}{SS_X}$ and $a = M_Y - bM_X$

$$SS_{\text{regression}} = r^2 SS_Y \quad df = 1 \quad MS_{\text{regression}} = \frac{SS}{df}$$

$$SS_{\text{residual}} = (1 - r^2)SS_Y \quad df = n - 2 \quad MS_{\text{residual}} = \frac{SS}{df}$$

MULTIPLE REGRESSION

$$\hat{Y} = b_1X_1 + b_2X_2 + a$$

$$R^2 = \frac{b_1SP_{X1Y} + b_2SP_{X2Y}}{SS_Y}$$

$$SS_{\text{regression}} = R^2 SS_Y \quad df = 2 \quad MS_{\text{regression}} = \frac{SS}{df}$$

$$SS_{\text{residual}} = (1 - R^2)SS_Y \quad df = n - 3 \quad MS_{\text{residual}} = \frac{SS}{df}$$

LINEAR AND MULTIPLE REGRESSION

Standard error of estimate $= \sqrt{MS_{\text{residual}}}$

$$F = \frac{MS_{\text{regression}}}{MS_{\text{residual}}}$$

CHI-SQUARE STATISTIC

$$\chi^2 = \Sigma\frac{(f_o - f_e)^2}{f_e}$$

MANN-WHITNEY *U*

$$U_A = n_An_B + \frac{n_A(n_A + 1)}{2} - \Sigma R_A$$

$$U_B = n_An_B + \frac{n_B(n_B + 1)}{2} - \Sigma R_B$$

KRUSKAL-WALLIS TEST

$$H = \frac{12}{N(N + 1)}\left(\Sigma\frac{T^2}{n}\right) - 3(N + 1)$$

FRIEDMAN TEST

$$\chi^2_r = \frac{12}{nk(k + 1)}\Sigma R^2 - 3n(k + 1)$$

MEASURES OF EFFECT SIZE

$$\text{Cohen's } d = \frac{\text{mean difference}}{\text{standard deviation}}$$

r^2 and η^2 (Percentage of Variance Accounted For)

$$r^2 = \frac{t^2}{t^2 + df}$$ (for Independent and Repeated t)

$$\eta^2 = \frac{SS_{\text{treatment}}}{SS_{\text{treatment}} + SS_{\text{error term}}}$$ (for Analysis of Variance)

$$\text{phi} = \sqrt{\frac{\chi^2}{n}}$$

$$\text{Cramér's } V = \sqrt{\frac{\chi^2}{n(df)}}$$

(for Chi-square Test for Independence)