Computational Formulas: Descriptive Statistics, Correlation, Regression and *t* Tests

Mean	$\bar{X} = \frac{\sum X}{n} \quad \mu = \frac{\sum X}{N}$
Median	$Median_{odd\ number\ of\ scores} = \left[\frac{n+1}{2}\right]^{th} score$ $Median_{even\ number\ of\ scores} = \frac{\left[\frac{n+2}{2}\right]^{th} score + \left[\frac{n}{2}\right]^{th} score}{2}$
Variance	$S^2 = \frac{\sum X^2 - \frac{\sum (X)^2}{n}}{n-1}$
Standard deviation	$S = \sqrt{S^2}$
z score	$z = \frac{X - \bar{X}}{S}$
Covariance	$\text{cov}_{XY} = \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{n-1}$ $\text{cov}_{XY} = \frac{\sum XY - \frac{\sum X \cdot \sum Y}{n}}{n-1}$
Correlation	$r = \frac{\text{cov}_{XY}}{S_X \cdot S_Y}$ $r = \frac{(n \cdot \sum XY) - (\sum X \cdot \sum Y)}{\sqrt{[(n \cdot \sum X^2) - (\sum X)^2] \cdot [(n \cdot \sum Y^2) - (\sum Y)^2]}}$ $df = n - 2$
Coefficient of determination	r^2

Linear regression	$\hat{Y} = bX + a$ $b = \frac{\text{cov}_{XY}}{S_X^2} = r_{XY} = \frac{S_Y}{S_X}$ or $b = \frac{\text{cov}_{XY}}{S_Y^2} = r_{XY} = \frac{S_X}{S_Y}$ $a = \bar{Y} - (b \cdot \bar{X})$ or $a = \bar{X} - (b \cdot \bar{Y})$
Standard error of the estimate	$S_{XY} = S_X \cdot \sqrt{1-r^2}$ or $S_{YX} = S_Y \cdot \sqrt{1-r^2}$
Standard error of the mean	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$
z test	$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$
Estimated standard error of the mean	$Est\sigma_{\bar{X}} = \frac{S}{\sqrt{n}}$
t Tests	$t = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$ $df = n - 1$ $t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sigma_{diff}}$ $est\sigma_{diff} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ $df = (n_1 - 1) + (n_2 - 1)$ $est\sigma_{diff} = \sqrt{(estimated\sigma_{\bar{X}_1})^2 + (estimated\sigma_{\bar{X}_2})^2 - (2 \cdot \text{cov})}$ $\sigma_{diff} = \sqrt{(estimated\sigma_{\bar{X}_1})^2 + (estimated\sigma_{\bar{X}_2})^2 - (2 \cdot r \cdot estimated\sigma_{\bar{X}_1} \cdot estimated\sigma_{\bar{X}_2})}$ $estimated\sigma_{diff} = \sqrt{\frac{\frac{\sum D^2}{n} - \bar{D}^2}{n-1}}$ $t = \frac{\bar{D}}{estimated\sigma_{diff}}$ $df =$ number of pairs - 1

Computational Formulas: One-Way Analysis of Variance

$$MS_{wg} = \frac{\sum(X_1 - \bar{X}_1)^2 + \sum(X_2 - \bar{X}_2)^2 + \sum(X_3 - \bar{X}_3)^2 + \cdots + \sum(X_k - \bar{X}_k)^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \cdots + (n_k - 1) +}$$

$$MS_{bg} = \frac{n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + \cdots + n_k(\bar{X}_k - \bar{\bar{X}})^2}{k - 1}$$

$$\bar{\bar{X}} = \frac{\sum \bar{X}}{k} = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \cdots + \bar{X}_k}{k}$$

$$\bar{\bar{X}} = \frac{\sum\left(\sum X\right)}{N_{total}} = \frac{\sum X_1 + \sum X_2 + \sum X_3 + \cdots + \sum X_k}{N_{total}}$$

$$F = \frac{MS_{bg}}{MS_{wg}} \qquad df_{bg} = k - 1 \qquad df_{wg} = (n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \cdots + (n_k - 1)$$

$$MS_{bg} = \frac{SS_{bg}}{df_{bg}} \qquad MS_{wg} = \frac{SS_{wg}}{df_{wg}}$$

$$SS_{bg} = \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \cdots + \frac{(\sum X_k)^2}{n_k}\right] - \left[\frac{(\sum X_1 + \sum X_1 + \cdots + \sum X_1)^2}{N_{total}}\right]$$

$$SS_{wg} = \left[\sum X_1^2 + \sum X_2^2 + \cdots + \sum X_k^2\right] - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \cdots + \frac{(\sum X_k)^2}{n_k}\right]$$

$$SS_{total} = \left[\sum X_1^2 + \sum X_2^2 + \cdots + \sum X_k^2\right] - \left[\frac{\left(\sum X_1 + \sum X_1 + \cdots + \sum X_1\right)^2}{N_{total}}\right]$$

$$SS_{total} = SS_{bg} + SS_{wg} \qquad HSD = q \cdot \sqrt{\frac{MS_{wg}}{n}}$$

Computational Formulas: Two-Way Analysis of Variance

$$SS_{total} = \sum\sum X^2 - \frac{\left(\sum\sum X\right)^2}{N_{total}}$$

$$SS_{wg} = \sum\sum X^2 - \frac{\sum\left(\sum X_{cell}\right)^2}{n_{cell}} \qquad SS_r = \frac{\sum\left(\sum X_{row}\right)^2}{n_{row}} - \frac{\left(\sum\sum X\right)^2}{N_{total}}$$

$$SS_c = \frac{\sum\left(\sum X_{col}\right)^2}{n_{col}} - \frac{\left(\sum\sum X\right)^2}{N_{total}} \qquad SS_{rxc} = SS_{Total} - (SS_{wg} + SS_r + SS_c)$$

$df_r =$ number of rows – 1 $\qquad df_c =$ number of columns –1

$df_{rxc} = df_r \cdot df_c =$ (number of rows-1) · (number of columns –1)

$$df_{wg} = (n_{cell_1} - 1) + (n_{cell_2} - 1) + \cdots + (n_{cell_k} - 1)$$

$df_{wg} = N_{Total}$ – number of cells $\qquad df_{Total} = N_{Total} - 1$

$$MS_r = \frac{SS_r}{df_r} \qquad MS_c = \frac{SS_c}{df_c} \qquad MS_{rxc} = \frac{SS_{rxc}}{df_{rxc}} \qquad MS_{wg} = \frac{SS_{wg}}{df_{wg}}$$

$$F_r = \frac{MS_r}{MS_{wg}} \qquad F_c = \frac{MS_c}{MS_{wg}} \qquad F_{rxc} = \frac{MS_{rxc}}{MS_{wg}}$$

Computational Formulas: Nonparametric Statistics

Chi-square	$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$ $f_e = \frac{(row\ total)\cdot(column\ total)}{grand\ \ total}$ *df* = (number of rows –1) · (number of columns – 1)
Mann Whitney U	$U_1 = (n_1 \cdot n_2) + \frac{n_1(n_1+1)}{2} - \sum R_1$ $U_2 = (n_2 \cdot n_1) + \frac{n_2(n_2+1)}{2} - \sum R_2$ $U_1 + U_2 = n_1 \cdot n_2$
Kruskal Wallis	$H = \left[\frac{12}{N_{total}\cdot(N_{total}+1)}\right]\cdot\left[\frac{(\sum R_1)^2}{n_1} + \frac{(\sum R_2)^2}{n_2} + \cdots + \frac{(\sum R_k)^2}{n_k}\right] - [3\cdot(N_{total}+1)]$