12.5 Confidence Intervals for the Difference Between Two Populations

Measure	Formula for Two Means	Formula for Two Proportions
Point Estimate	$\bar{x}_1 - \bar{x}_2$	$\bar{p}_1 - \bar{p}_2$
Standard Error	$se(\bar{X}_1 - \bar{X}_2) = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$	$se(\bar{P}_1 - \bar{P}_2) = \sqrt{\frac{\bar{p}_1(1 - \bar{p}_1)}{n_1} + \frac{\bar{p}_2(1 - \bar{p}_2)}{n_2}}$
Critical Value	$t^*_{\frac{\alpha}{2}, df}$	$z^*_{\frac{\alpha}{2}}$
Degrees of Freedom	$df = (n_1 - 1) + (n_2 - 1)$	N/A
Margin of Error (ME)	$ME = t^*_{\frac{\alpha}{2}, df} \times se(\bar{X}_1 - \bar{X}_2)$	$ME = z^*_{\frac{\alpha}{2}} \times se(\bar{P}_1 - \bar{P}_2)$

Measure	Excel Formula
Standard Normal: Critical Values	=NORM.S.DIST(z, 1) =NORM.S.INV$\left(\frac{\alpha}{2}\right)$
T Distribution: Critical Values	=T.INV.2T(α, df) =T.INV($\frac{\alpha}{2}$, df)

12.6 Testing the Difference Between Two Populations

Measure	Formula for Two Means	Formula for Two Proportions
Critical Value	$t^*_{\frac{\alpha}{2},\,df}$ for two-tailed hypothesis tests $t^*_{\alpha,\,df}$ for one-tailed hypothesis tests	$z^*_{\frac{\alpha}{2}}$ for two-tailed hypothesis tests z^*_{α} for one-tailed hypothesis tests
Test Statistic	$t_{df} = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$	$z = \dfrac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)_0}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$
Degrees of Freedom	$df = (n_1 - 1) + (n_2 - 1)$	N/A

Measure	Excel Formula
Standard Normal: P-value and Critical Values	=NORM.S.DIST(z, 1) =NORM.S.INV$\left(\frac{\alpha}{2}\right)$ or =NORM.S.INV(α)
T Distribution: P-value and Critical Values	=T.DIST(t, df, 1) =T.DIST.2T(t, df) =T.DIST.RT(t, df) =T.INV.2T(α, df) =T.INV($\frac{\alpha}{2}$, df) or =T.INV(α, df)

12.7 Chi Square Test of Independence

Measure	**Formulas**
Critical Value	$\chi^{2*}_{\alpha,df}$ for right-tailed hypothesis test
Expected Frequency, f_e	$f_e = \frac{(row\ total)(column\ total)}{sample\ size}$
Test Statistic	$\chi^2_{df} = \sum \frac{(f_o - f_e)^2}{f_e}$
Degrees of Freedom	$df = (r-1)\times(c-1)$ where r=number of rows and c=number of columns

Measure	**Excel Formula**
Chi Square Distribution: P-value and Critical Values	$= \mathrm{CHISQ.DIST.RT}(\chi^2, \mathrm{df})$ $= \mathrm{CHISQ.INV.RT}(\alpha, \mathrm{df})$

12.8 Simple Linear Regression

Measure	Formula
Correlation	$r = \frac{covariance}{s_x s_y}$
Slope	$b_1 = r\left(\frac{s_y}{s_x}\right)$
Intercept	$b_0 = \bar{y} - b_1 \times \bar{x}$
Regression Equation	$\hat{y} = b_0 + b_1(x)$
Residual (observed – predicted)	$e = y_i - \hat{y}_i$

Measure	Excel Formula
Covariance	=COVARIANCE.S(<data range>, <data range>)
Correlation	=CORREL(<data range>, <data range>)
Slope	=SLOPE(<y data range>, <x data range>)

12.9 Multiple Linear Regression

Measure	Formula
Regression Equation	$\hat{y} = b_0 + b_1(x_1) + b_2(x_2) + b_3(x_3) \cdots b_k(x_k)$
Residual (observed – predicted)	$e = y_i - \hat{y}_i$
Adjusted R^2	$Adjusted\ R^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-k-1}\right)$

12.10 Dummy Variables with Regression

Measure	Formula
Regression Equation	$\hat{y} = b_0 + b_1(x) + b_2(d_1) + b_3(d_2)$*
Residual (observed – predicted)	$e = y_i - \hat{y}_i$
Reference Category	The category where the value of all the dummy variables (that represent one qualitative variable) are zero. Each category is compared against this category.

*The number of dummy variables depends on the number of qualitative variables and the number of categories represented by each qualitative variable. The number of dummy variables used to represent each variable must be one less than the number of categories.