



## 1 | PO11Q

Statistic	Formula
Confidence Interval	$Pr(\bar{y} - t_{\alpha/2} \cdot se \leq \mu \leq \bar{y} + t_{\alpha/2} \cdot se) = 1 - \alpha$
Deviation	$d = y_i - \bar{y}$
Mean	$\bar{y} = \frac{\sum y_i}{n}$ $\mu = \sum y P(y) = E(y)$
Range	$y_{range} = y_{max} - y_{min}$
Standard Deviation	$s = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$
Standard Error	$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}$ $se = \frac{s}{\sqrt{n}}$
t-test	$t = \frac{\bar{y} - \mu_0}{se}$
Variance	$s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1}$
z-score	$z = \frac{y - \mu}{\sigma}$

Table 1: Formulae for PO11Q

Statistic	Formula
<b>Crosstabulations</b>	
$\chi^2$ -Test Statistic	$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$
<b>Two Sample Tests</b>	
Standard Error of Difference	$se = \sqrt{(se_1)^2 + (se_2)^2}$
<i>Means</i>	
Standard Error of Difference	$se = \sqrt{\frac{se_1^2}{n_1} + \frac{se_2^2}{n_2}}$
Standard Error (pooled)	$se_0 = \sqrt{s_0^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$
Variance (pooled)	$s_0^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$
<i>Proportions</i>	
Standard Error	$se = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$
Standard Error of Difference	$se = \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$
Standard Error (pooled)	$se_0 = \sqrt{\hat{\pi}(1-\hat{\pi}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$
<b>Regression</b>	
Adjusted $R^2$	$\bar{R}^2 = 1 - \frac{\sum \hat{e}_i^2 / (n-k-1)}{\sum (y_i - \bar{y})^2 / (n-1)}$
Coefficient of Determination	$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum \hat{e}_i^2}{\sum (y_i - \bar{y})^2}$
Estimated Variance	$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{n-2}$
Estimator of $\beta_0$	$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$
Estimator of $\beta_1$	$\hat{\beta}_1 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}$
Explained Sum of Squares	$\sum (\hat{y}_i - \bar{y})^2$
Residual Sum of Squares	$\sum \hat{e}_i^2 = \sum (\hat{y}_i - y_i)^2$
Standard Error of $\beta_0$	$se(\hat{\beta}_0) = \sqrt{\frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2}} \sigma$
Standard Error of $\beta_1$	$se(\hat{\beta}_1) = \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}$
Total Sum of Squares	$\sum (y_i - \bar{y})^2$

Table 2: Formulae for PO12Q

### 3 | Notation

Symbol	Explanation
<b>P011Q</b>	
$d$	Deviation
$n$	Sample Size
$s$	Standard Deviation
$s^2$	Variance
$\bar{y}$	Mean
$y_i$	Observation $i$
$f$	(Absolute) Frequency
$cf$	Cumulative (Absolute) Frequency
$rf$	Relative Frequency
$crf$	Cumulative Relative Frequency
$E[x]$	The expected value of $x$
$\mu$	Mean of the Population
$se$	Standard Error (with $s$ of sample)
$\sigma$	Standard Deviation of the Population
$\sigma_{\bar{y}}$	Standard Error (with $\sigma$ of population)
$t$	t-value
$z$	z-value
<b>P012Q</b>	
$\chi^2$	Chi-Squared for test of independence
$\pi$	Population Proportion
$\hat{\pi}$	Sample Proportion
$\bar{\pi}$	Pooled Proportion
$se_0$	Standard Error under the Null Hypothesis
$\hat{\beta}$	Estimated Regression Coefficient
$\epsilon$	Error Term
$\hat{\epsilon}$	Estimated Error Term / Residual
$A^{-1}$	Inverse of Matrix $A$
$A'$	Transpose of Matrix $A$
$I$	Identity Matrix
$R^2$	R-Squared / Model Fit
$\sigma^2$	Mean Squared Error
$k$	Number of Slope Coefficients
$\bar{R}^2$	Adjusted R-Squared
$\log(x)$	Logarithm of variable $x$
$\alpha_i$	Regression coefficients of secondary regression models
$P$	Total Number of Regression Coefficients, including the Intercept
$VcV$	Variance-Covariance Matrix
$\Omega$	Is equal to $\sigma^2 I$ , where $\sigma^2$ represents the mean squared error
$t$	Time period in time-series data

Table 3: Notation