

FORMULAE FOR STATISTICS

Categorical Data(Proportions – p)	Quantitative Data(Random variables – y)
Population proportion = p	Population mean = μ S.D. = σ
Sample proportion = \hat{p}	Sample mean = \bar{y} Sample standard deviation = s
Standard deviation of sample proportions = $s.d(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$	Standard deviation of sample means = $s.d(\bar{y}) = \frac{\sigma}{\sqrt{n}}$
Standard error of sample proportions = $s.e(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	Standard error of sample means = $s.e(\bar{y}) = \frac{s}{\sqrt{n}}$
z- score for distribution of sample proportions = $z = \frac{\hat{p} - p}{s.d(\hat{p})}$	z- score for distribution of sample means = $z = \frac{\bar{y} - \mu}{s.d(\bar{y})}$
	t-score: $t = \frac{\bar{y} - \mu_0}{s.e(\bar{y})} = \frac{\sqrt{n}(\bar{y} - \mu_0)}{s}$, df = n-1
Confidence interval = $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ z^* corresponds to the confidence level	Confidence interval = $\bar{y} \pm t^* \frac{s}{\sqrt{n}}$ t^* corresponds to the confidence level
Standard error of difference $s.e(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	Standard error of difference(unpooled) $s.e.(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ df = $n_1 + n_2 - 2$
Format for Confidence interval = Estimate $\pm z^*$ (Standard Error of Estimate) (For quantitative data use t^* instead) \Rightarrow	Standard error of difference(paired) $s.e(\bar{d}) = \frac{s_d}{\sqrt{n}}$, df = n-1, $\bar{d} = \bar{x}_1 - \bar{x}_2$, and $d = x_1 - x_2$
	Standard deviation $s = \sqrt{\frac{\sum (y - \bar{y})^2}{n-1}}$
	Correlation $r = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n-1)S_x \cdot S_y}$
	Equation of linear regression: $\hat{y} = mx + b$ $m = \frac{rS_y}{S_x}$