PO91Q - Fundamentals in Quantitative Research Methods

Formula Collection and Notation

Dr Florian Reiche
F.Reiche@warwick.ac.uk

1 | Formulae

Statistic	Formula
Descriptives	
Deviation	$d = y_i - \bar{y}$
Mean	$\bar{y} = \frac{\sum y_i}{n}$
	$\mu = \sum y P(y) = E[y]$
Range	$y_{\text{range}} = y_{\max} - y_{\min}$
Standard Deviation	$s = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$
Inference	
Confidence Interval	$Pr\left(\bar{y} - t_{\alpha/2} \cdot se \leq \mu \leq \bar{y} + t_{\alpha/2} \cdot se\right) = 1 - \alpha$
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	$z = \frac{y - \mu}{\sigma}$
Crosstabulations	
χ^2-Test Statistic	$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$

continues on next page

Statistic	Formula
Correlation	
Pearson's Correlation	$r_p = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$
Regression	
Adjusted R^2	$\bar{R}^2 = 1 - \frac{\sum \hat{\epsilon}_i^2 / (n - k - 1)}{\sum(y_i - \bar{y})^2 / (n - 1)}$
Coefficient of Determination	$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum \hat{\epsilon}_i^2}{\sum(y_i - \bar{y})^2}$
Estimated Variance	$\hat{\sigma}^2 = \frac{\sum \hat{\epsilon}_i^2}{n - p}$
Estimator of β_0	$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$
Estimator of β_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	$\sum \hat{\epsilon}_i^2 = \sum(y_i - \hat{y}_i)^2$
Standard Error of β_0	$se(\hat{\beta}_0) = \sqrt{\frac{\sum x_i^2}{n \sum(x_i - \bar{x})^2}}\ \sigma$
Standard Error of β_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 PO91Q

2 | Notation

Symbol	Explanation
Descriptives	
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
Inference	
$E[x]$	The expected value of x
μ	Mean of the Population
se	Standard Error (with s of sample)
σ	Standard Deviation of the Population
$\sigma_{\bar{y}}$	Standard Error (with σ of population)
t	t-value
z	z-value
Bivariate Methods	
χ^2	Chi-Squared for test of independence
r	Correlation Coefficient
r_p	Pearson's Product-Moment Correlation Coefficient
se_0	Standard Error under the Null Hypothesis
Regression	
β	Regression Coefficient
$\hat{\beta}$	Estimated Regression Coefficient
ϵ	Error Term
$\hat{\epsilon}$	Estimated Error Term / Residual
R^2	R-Squared / Model Fit
σ^2	Mean Squared Error
p	Total Number of Regression Coefficients, including the Intercept
k	Number of Slope Coefficients
$\bar{R}^2$	Adjusted R-Squared
log(x)	Logarithm of variable x

Table 3: Notation