

Product-Moment Correlation Coefficient

$$r = \frac{\sum(xy)}{\sqrt{(\sum x^2)(\sum y^2)}}$$

r = product – moment correlation coefficient

x = first set of data

y = second set of data

Population Correlation Coefficient

$$\rho = \frac{1}{N} * \sum \left\{ \left[\frac{(X_i - \mu_X)}{\sigma_X} \right] * \left[\frac{(Y_i - \mu_Y)}{\sigma_Y} \right] \right\}$$

ρ = population correlation coefficient

N = population size

X_i = the ith element of the first population

Y_i = the ith element of the second population

μ_X = mean of the first population

μ_Y = mean of the second population

σ_X = the standard deviation of the first population

σ_Y = the standard deviation of the second population

Sample Correlation Coefficient

$$r = \left[\frac{1}{(n - 1)} \right] * \sum \left\{ \left(\frac{x_i - \bar{x}}{s_x} \right) * \left(\frac{y_i - \bar{y}}{s_y} \right) \right\}$$

r = sample correlation coefficient

n = sample size

x_i = the ith element of the first sample

y_i = the ith element of the second sample

̄x = mean of the first sample

̄y = mean of the second sample

s_x = first sample's standard deviation

s_y = second sample's standard deviation

Population Regression Line

$$Y = B_0 + B_1X$$

Y = value of the dependent variable

B₀ = a constant

$B_1 = \text{regression coefficient}$
 $X = \text{value of the independent variable}$

Estimate of Population Regression Line

$$\hat{y} = b_0 + b_1x$$

$\hat{y} = \text{predicted value of the dependent variable}$
 $b_0 = \text{a constant}$
 $b_1 = \text{regression coefficient}$
 $x = \text{value of the independent variable}$

Regression Coefficient

$$b_1 = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sum[(x_i - \bar{x})^2]}$$

$b_1 = \text{regression coefficient}$
 $\Sigma = \text{summation}$
 $x_i = \text{the } i\text{th element of the first sample}$
 $y_i = \text{the } i\text{th element of the second sample}$
 $\bar{x} = \text{mean of the first sample}$
 $\bar{y} = \text{mean of the second sample}$

Constant in the Regression Equation

$$b_0 = \bar{y} - b_1 * \bar{x}$$

$b_0 = \text{constant}$
 $\bar{y} = \text{mean of sample } y$
 $b_1 = \text{regression coefficient}$
 $\bar{x} = \text{mean of sample } x$

Coefficient of Determination

$$R^2 = \left\{ \left(\frac{1}{N} \right) * \frac{\sum[(x_i - \bar{x}) * (y_i - \bar{y})]}{\sigma_x * \sigma_y} \right\}^2$$

$R^2 = \text{the coefficient of determination}$
 $N = \text{population size}$
 $\bar{x} = \text{mean of the first sample}$
 $\bar{y} = \text{mean of the second sample}$
 $x_i = \text{the } i\text{th element of the first sample}$

$y_i =$ the i th element of the second sample

Standard Deviation of Population X

$$\sigma_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}}$$

$\sigma_x =$ the standard deviation of the first sample

$\bar{x} =$ mean of sample x

$N =$ population size

$x_i =$ the i th element of the first sample

Residual

$$e = y - \hat{y}$$

$e =$ residual

$y =$ observed value

$\hat{y} =$ predicted value

Exponential Model

$$\log(y) = b_0 + b_1x$$

$\log =$ log base 10

$y =$ dependent variable

$b_0 =$ constant

$b_1 =$ regression coefficient

$x =$ independent variable

Quadratic Model

$$\sqrt{y} = b_0 + b_1x$$

$y =$ dependent variable

$b_0 =$ constant

$b_1 =$ regression coefficient

$x =$ independent variable

Reciprocal Model

$$\frac{1}{y} = b_0 + b_1x$$

$y =$ dependent variable

$b_0 = \text{constant}$
 $b_1 = \text{regression coefficient}$
 $x = \text{independent variable}$

Logarithmic Model

$$y = b_0 + b_1 \log(x)$$

$\log = \text{log base 10}$
 $y = \text{dependent variable}$
 $b_0 = \text{constant}$
 $b_1 = \text{regression coefficient}$
 $x = \text{independent variable}$

Power Model

$$\log(y) = b_0 + b_1 \log(x)$$

$\log = \text{log base 10}$
 $y = \text{dependent variable}$
 $b_0 = \text{constant}$
 $b_1 = \text{regression coefficient}$
 $x = \text{independent variable}$

$$SE = s_{b1} = \frac{\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n - 2}}}{\sqrt{\sum(x_i - \bar{x})^2}}$$

$SE = \text{standard error}$
 $s_{b1} = \text{standard deviation of } b_1$
 $\bar{x} = \text{mean of sample } x$
 $\hat{y} = \text{predicted value}$
 $x_i = \text{the } i\text{th element of the first sample}$
 $y_i = \text{the } i\text{th element of the second sample}$

Standard Error

$$t = \frac{b_1}{SE}$$

$b_1 = \text{regression coefficient}$
 $SE = \text{standard error}$

t = test statistic for t norms

Degrees of Freedom (Linear Regression)

$$DF = n - 2$$

DF = degrees of freedom,

n = sample size