Empirical Research in Management and Economics
Formula Sheet

January 20, 2022

1 Descriptive Statistics

1.1 Measure of Location

1.1.1 Mean

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

1.1.2 Median

For odd number of elements in a dataset:

$$\tilde{x} = x_{\frac{n+1}{2}} \tag{2}$$

For even number of elements in a dataset:

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\left(\frac{n}{2}+1\right)}}{2} \tag{3}$$

1.1.3 Mode

$$Mo(x) = \max(f(x_i)) \tag{4}$$

1.1.4 Quartile

Measure of percentage of elements less than or equal to a term

1.2 Measure of Spread

1.2.1 Variance

Variance measured on the whole population

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{5}$$

1.2.2 Sample Variance

Variance measured on a sample population

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{6}$$

1.2.3 Standard Deviation and Sample Standard

$$\sigma = \sqrt{\sigma^2} \tag{7}$$

$$s = \sqrt{s^2} \tag{8}$$

1.2.4 Co-efficient of Variance

$$v = \frac{s}{\bar{x}} \tag{9}$$

1.3 Skewness

1.3.1 Types of Skewness

Name	Other Name	Characteristic
Right Skew	Positive Skew	Data concentrated on the lower side
Symmetric Distribution	Normal Distribution	Data distributed evenly
Left Skew	Negative Skew	Data concentrated on the higher side

1.3.2 Measure of Skewness

Skewness is measured by the Moment Co-efficient of Skewness.

$$g_m = \frac{m_3}{s^3}, \text{ where} \tag{10}$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3 \tag{11}$$

Type of Skewness The type of skewness from the value is g_m is:

Value of g_m	Type
$g_m = 0$	Symmetric
$g_m > 0$	Positive Skew
$g_m < 0$	Negative Skew

Degree of Skewness The degree of skewness from the value is g_m is:

Value of g_m	Degree
$\lvert g_m \rvert > 1$	High Skewness
$0.5 < \lvert g_m \rvert \geq 1$	Moderate Skewness
$\lvert g_m \rvert \leq 0.5$	Low Skewness

1.4 Kurtosis

Kurtosis is the measure of peakedness of data. Fisher's kurtosis measure is defined as:

$$\gamma = \frac{m_4}{s^4}, \text{ where} \tag{12}$$

$$m_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4 \tag{13}$$

1.4.1 Type of Kurtosis

The types of kurtosis from the value of γ are:

Value of γ	Type
$\gamma = 0$	Normal Distribution or Mesokurtic
$\gamma < 0$	Flattened or Platykurtic
$\gamma > 0$	Peaked or Lepokurtic

2 Hypothesis Testing

2.1 T-Test

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \tag{14}$$

where:

$$\begin{aligned} \bar{X} &= \text{Sample Mean} \\ \mu &= \text{Assumed Mean} \\ s &= \text{Number of Samples} \\ n &= \text{Number of observations} \end{aligned}$$

If $T < t_c$ the H_0 is not rejected. t_c is a functions of level of significance (α) and degrees of freedom ($v = n - 1$).

2.2 χ^2 Test

$$\chi^2 = \sum_i \sum_j \frac{(h_{ij}^o - h_{ij}^e)^2}{h_{ij}^e} \tag{15}$$

where:

$$\begin{aligned} h_e &= \text{Expected Value} \\ h_o &= \text{Actual Value} \end{aligned}$$

If $\chi^2 < \chi_c^2$ then H_0 is not rejected. χ_c is a functions of level of significance (α) and degrees of freedom ($v = (i-1)(j-1)$).

3 Research and Survey Design

3.1 Population Covariance

$$\text{Cov}(x, y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y) \tag{16}$$

3.2 Sample Covariance

$$\text{Cov}(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \tag{17}$$

3.3 Bravais-Pearson Correlation Co-efficient

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{18}$$

$$= \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x) \cdot \text{Var}(y)}} \tag{19}$$

$$= \frac{\text{Cov}(x, y)}{\sigma_x \cdot \sigma_y} \tag{20}$$

4 Estimation of Regression Function

For the regression functions:

$$Y_i = \beta_0 + \beta_1 X_1 \tag{21}$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_1 \tag{22}$$

$$\tag{23}$$

where Y_i is the observed dependent variable (DV), $\hat{Y}_i$ is the estimated DV, and X_i is the independent variable (IV).

$$u_i = Y_i - \hat{Y}_i \tag{24}$$

$$\Rightarrow Y_i = \hat{Y}_i + u_i \tag{25}$$

$$\Rightarrow Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_1 + u_i \tag{26}$$

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \tag{27}$$

The objective function is:

$$\min_{u_i} \sum u_i = \min \sum_i \left[Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \right]^2$$

Since the regression function passes through: $\left(\bar{X}, \bar{Y}\right)$

$$\beta_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\min_{u_i} \sum u_i = \min \sum_i \left[Y_i - \bar{Y} + \hat{\beta}_1 \bar{X} - \hat{\beta}_1 X_i \right]^2$$

$$= \min \sum_i \left[\left(Y_i - \bar{Y}\right) - \hat{\beta}_1 \left(X_i - \bar{X}\right) \right]^2$$

$$= \min \sum_i \left[\left(Y_i - \bar{Y}\right)^2 - 2 \cdot \left(Y_i - \bar{Y}\right) \cdot \hat{\beta}_1 \left(X_i - \bar{X}\right) + \hat{\beta}_1{}^2 \left(X_i - \bar{X}\right)^2 \right]$$

$$= \min \left[\sum_i \left(Y_i - \bar{Y}\right)^2 - 2 \cdot \hat{\beta}_1 \sum_i \left(Y_i - \bar{Y}\right) \cdot \left(X_i - \bar{X}\right) + \hat{\beta}_1{}^2 \sum_i \left(X_i - \bar{X}\right)^2 \right]$$

$$\Rightarrow u_i^{\beta_1} = -2 \sum_i \left(Y_i - \bar{Y}\right) + 2\hat{\beta}_1 \left(X_i - \bar{X}\right)^2 = 0 \qquad \text{(For optima Conditions)}$$

$$\Rightarrow \hat{\beta}_1 = \boxed{\frac{\sum_i (Y_i - \bar{Y})(X_i - \bar{X})}{\sum_i (X_i - \bar{X})^2}}$$

$$\Rightarrow \hat{\beta}_0 = \boxed{\bar{Y} - \hat{\beta}_1 \bar{X}}$$