Parametric and Non-Parametric Tests of Independence

Download Files

Download PDF | Download Word | Download PowerPoint

Learning Module 9: Parametric and Non-Parametric Tests of Independence

---

Pearson Correlation (or Bivariate Correlation)

$$
r_{XY} = \frac{s_{XY}}{s_X s_Y} \tag{1}
$$

Where:

- r_{XY}: Pearson correlation coefficient between variables X and Y
- s_{XY}: sample covariance between X and Y
- s_X: sample standard deviation of X
- s_Y: sample standard deviation of Y

View Markdown Source

```
## Pearson Correlation (or Bivariate Correlation)

$$
r_{XY} = \frac{s_{XY}}{s_X s_Y} \tag{1}
$$

Where:
```

```
* $r_{XY}$: Pearson correlation coefficient between variables $X$ and $Y$
* $s_{XY}$: sample covariance between $X$ and $Y$
* $s_X$: sample standard deviation of $X$
* $s_Y$: sample standard deviation of $Y$
```

---

t-Test Statistic for Pearson Correlation

$$
t = \frac{r \sqrt{n - 2}}{\sqrt{1 - r^2}} \tag{2}
$$

Where:

- t: test statistic for hypothesis testing of the correlation coefficient
- r: sample Pearson correlation coefficient
- $n - 2$ degrees of freedom
- n: sample size

View Markdown Source

```
## t-Test Statistic for Pearson Correlation

$$
t = \frac{r \sqrt{n - 2}}{\sqrt{1 - r^2}} \tag{2}
$$

Where:

* $t$: test statistic for hypothesis testing of the correlation coefficient
* $r$: sample Pearson correlation coefficient
* $n - 2$ degrees of freedom
* $n$: sample size
```

Spearman Rank Correlation

$$
r_s = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{3}
$$

Where:

- r_s: Spearman rank correlation coefficient
- d_i: difference between the ranks of paired observations for item i
- n: sample size

View Markdown Source

```
## Spearman Rank Correlation

$$
r_s = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{3}
$$

Where:

* $r_s$: Spearman rank correlation coefficient
* $d_i$: difference between the ranks of paired observations for item $i$
* $n$: sample size
```

Chi-Square Test Statistic for Independence

$$
\chi^2 = \sum_{i=1}^{m} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{4}
$$

Where:

- χ^2: chi-square test statistic
- m = the number of cells in the table, which is the number of groups in the first class multiplied by the number of groups in the second class;
- O_{ij} = the number of observations in each cell of row i and column j (i.e., observed frequency); and
- E_{ij} = the expected number of observations in each cell of row i and column j, assuming independence (i.e., expected frequency).

View Markdown Source

```
## Chi-Square Test Statistic for Independence

$$
\chi^2 = \sum_{i=1}^{m} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{4}
$$

Where:

* $\chi^2$: chi-square test statistic
* $m$ = the number of cells in the table, which is the number of groups in the first class mu
* $O_{ij}$ = the number of observations in each cell of row $i$ and column $j$ (i.e., observe
* $E_{ij}$ = the expected number of observations in each cell of row $i$ and column $j$, assu
```

Calculating Expected number of ETFs (E_{ij})

$$
E_{ij} = \frac{(\text{Total row } i) \times (\text{Total column } j)}{\text{Overall total}} \tag{5}
$$

Where:

- E_{ij}: The expected number of ETFs
- Total row i: sum of observed frequencies in row i
- Total column j: sum of observed frequencies in column j
- Overall total: total number of observations in the table

View Markdown Source

```
## Calculating Expected number of ETFs $(E_{ij})$

$$
E_{ij} = \frac{(\text{Total row } i) \times (\text{Total column } j)}{\text{Overall total}} \
$$

Where:

* $E_{ij}$: The expected number of ETFs
* $\text{Total row } i$: sum of observed frequencies in row $i$
* $\text{Total column } j$: sum of observed frequencies in column $j$
* $\text{Overall total}$: total number of observations in the table
```

Standardized Residual (also referred to as a Pearson residual)

$$
\text{Standardized residual} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}} \tag{6}
$$

Where:

- O_{ij}: the number of observations in each cell of row i and column j (i.e., observed frequency); and
- E_{ij} = the expected number of observations in each cell of row i and column j, assuming independence (i.e., expected frequency).

View Markdown Source

```
## Standardized Residual (also referred to as a Pearson residual)

$$
\text{Standardized residual} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}} \tag{6}
$$

Where:

* $O_{ij}$:  the number of observations in each cell of row $i$ and column $j$ (i.e., observe
* $E_{ij}$ = the expected number of observations in each cell of row $i$ and column $j$, assu
```
