Cramér's V

Measures the association between two categorical variables

Formula:

$$V = \sqrt{\frac{\chi^2}{n \min(r-1, c-1)}}$$

Terms Explanation:

- χ^2: Chi-square statistic based on the contingency table

- n: Total sample size

- r, c: Number of rows and columns of the contigency table, respectively

@AIinMinutes

Kendall's Tau-b (τ_b)

Measures strength and direction of association between ordinal variables

Formula:

$$\tau_b = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}$$

Terms Explanation:

- n_c: Number of concordant pairs
- n_d: Number of discordant pairs
- $n_0 = \frac{n(n-1)}{2}$: Total pairs
- n_1, n_2: Number of tied pairs in each variable

@AlinMinutes

Spearman's Rho (ρ)

Measures monotonic relationship between ranked variables

Formula:

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

Terms Explanation:

- d_i: Difference between ranks of corresponding values

- n: Number of observations

@AIinMinutes

Pearson's Correlation Coefficient (r)

Measures strength and direction of linear relationship between two variables

Formula:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Terms Explanation:

- X_i, Y_i: Individual data points
- $\bar{X}$, $\bar{Y}$: Mean values of X and Y

@AlinMinutes

Autocorrelation at Lag k (r_k)

Measures correlation of a time series with a lagged version of itself

Formula:

$$r_k = \frac{\sum_{t=k+1}^{n}(X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}$$

Terms Explanation:

- X_t: Value of time series at time t
- X_{t-k}: Value at lag k
- $\bar{X}$: Mean of time series
- n: Total observations

@AIinMinutes