

相关系数算法

目录

SPSSAU 数据格式..... 2
计算公式..... 2
 1. Pearson 时..... 2
 2. Spearman 时..... 2
 3. Kendall 时..... 3
参考文献..... 3

相关系数研究两两变量之间的相关关系情况，在 SPSSAU 中支持：

- ✓ 批量对‘标题’进行分析；
- ✓ 支持区分 X 和 Y；
- ✓ 提供 Pearson（默认）、Spearman 和 Kendall 三种相关系数。

The image shows a user interface for performing a correlation analysis. At the top left is a blue button labeled '开始分析' (Start Analysis). To its right is a trash icon and a dropdown menu currently set to 'Pearson相关系数' (Pearson Correlation Coefficient). Below these are two large, light gray rectangular input boxes. The top box is labeled '分析项Y(定量)' (Analysis Item Y (Quantitative)). The bottom box is labeled '分析项X(定量) [可选]' (Analysis Item X (Quantitative) [Optional]).

如果分析时有明确区分 X 和 Y，可将 X 和 Y 分别放入对应框中，反之把所有分析项放入‘分析项 Y’框中即可。

SPSSAU 数据格式

Title1	Title2	Title3
2.870334	5.16743	3.555898
6.360674	8.517929	8.62671
0.007093	5.183305	4.023125
5.587655	1.797669	9.950338
9.820287	8.482111	8.612621
6.424085	1.681026	3.197264
5.934659	7.22094	4.261088
0.431016	1.250926	3.265462
9.494304	7.255518	5.732024
9.823583	2.262717	5.380164
0.392169	9.28603	3.265742
0.086585	2.759854	2.597386

比如图中包括 3 个标题，可直接拖拽至右侧框进行分析。

计算公式

1. Pearson 时

$$r = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E(X)^2} \sqrt{E(Y^2) - E(Y)^2}}$$

$$cov(X, Y) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

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

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

其中：

\bar{X} 和 \bar{Y} 分别是 X 和 Y 的平均值

n 表示样本量

2. Spearman 时

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

其中：

d_i 是每对观测值的秩差（即两变量的秩次之差）

$$d_i = R(X_i) - R(Y_i)$$

$R(X_i)$ 和 $R(Y_i)$ 分别是 X 和 Y 的秩次（排名）

n 是样本量

3. Kendall 时

$$\tau = \frac{C - D}{\frac{1}{2}n(n - 1)}$$

其中：

C 是顺序对的数量（即两个变量同时增加的对数）

D 是逆序对的数量（即一个变量增加而另一个变量减少的对数）

n 是样本容量

✓ 顺序对 (Concordant Pairs)

顺序对是指在两个变量中，两个观测值的相对顺序一致。对于一对观测值 (X_i, Y_i) 和 (X_j, Y_j) ，如果 $X_i > X_j$ 时 $Y_i > Y_j$ ，或者 $X_i < X_j$ 时 $Y_i < Y_j$ ，则称这对观测值为顺序对。

✓ 逆序对 (Discordant Pairs)

逆序对是指在两个变量中，两个观测值的相对顺序不一致。对于一对观测值 (X_i, Y_i) 和 (X_j, Y_j) ，如果 $X_i > X_j$ 时 $Y_i < Y_j$ ，或者 $X_i < X_j$ 时 $Y_i > Y_j$ ，则称这对观测值为逆序对。

参考文献

【1】 The SPSSAU project (2024). SPSSAU. (Version 24.0) [Online Application Software]. Retrieved from <https://www.spssau.com>.

【2】 周俊,马世澎. SPSSAU 科研数据分析方法与应用.第 1 版[M]. 电子工业出版社,2024.