Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data

Jan Hauke and Tomasz Kossowski

Adam Mickiewicz University, Poznań, Poland

Abstract

Spearman's rank correlation coefficient (denoted here by r_s) is a non-parametric (distribution-free) rank statistic proposed by Charles Spearman as a measure of the strength of the associations between two variables. It is a measure of a monotone association that is used when the distribution of the data makes Pearson's correlation coefficient undesirable or misleading. Spearman's coefficient is not a measure of the linear relationship between two variables, as some "statisticians" declare. It assesses how well an arbitrary monotonic function can describe the relationship between two variables, without making any assumptions about the frequency distribution of the variables. Unlike Pearson's product-moment correlation coefficient, it does not require the assumption that the relationship between the variables is linear, nor does it require the variables to be measured on interval scales; it can be used for variables measured at the ordinal level.

In principle, r_s is simply a special case of Pearson's product-moment coefficient in which the data are converted to ranks before calculating the coefficient.

Charles Spearman developed his rank correlation in 1904. However, his statistical work was not appreciated by his University College colleague Karl Pearson and there was a long feud between them. Nowadays, the coefficient r_s is widely used in statistical analysis. In the presentation we would like to compare it values and significance for different sets of data (original - for Pearson's coefficient and ranked data for Spearman's coefficient).

What inspired this presentation was a paper by Plata (2006).

Keywords

Pearson's and Spearman's correlation coefficient.

References:

Plata, S. (2006). A note on Fisher's correlation coefficient. *Applied Mathematical Letters 19*, 499–502.

Spearman, Ch. (1904). Proof and measurement of association between two things. American Journal of Psychology *15*, 72–101.