Bounding Pearson's Correlation for Ordinal Data: Theory, Symmetry, and Empirical Consequences

Cees van der Eijk   Scott Moser

Table of contents

0.1 3. Maximizing and Minimizing Pearson's Correlation

To determine the maximum and minimum possible Pearson's correlation coefficients given the marginal distributions, we employ the concept of optimal coupling derived from the Hardy–Littlewood–Pólya rearrangement inequality.

0.1.1 3.1 Maximizing Pearson's Correlation

To maximize Pearson's correlation, we construct a joint distribution that pairs categories from X and Y in a comonotonic (similarly ordered) fashion. Specifically, the optimal joint distribution, denoted as $F_{max}(x,y)$, is constructed by:

1. Sorting the categories of X in descending order according to their values.
2. Independently sorting the categories of Y in descending order according to their values.
3. Pairing the largest values of X with the largest values of Y, the next-largest with the next-largest, and so forth, maintaining consistency with the marginal totals.

Formally, if $x_{(1)} x_{(2)} … x_{(K_X)}$ and $y_{(1)} y_{(2)} … y_{(K_Y)}$ represent the sorted categories of X and Y, respectively, then the joint probabilities are allocated to maximize:

$$\mathbb{E}[XY]_{\max} = \sum_{i,j} x_{(i)} y_{(j)} P_{\max}(X = x_{(i)}, Y = y_{(j)}).$$