

subtypelist

# 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 . . . . .	2
0.1.1	3.1 Maximizing Pearson's Correlation . . . . .	2
0.1.2	3.2 Minimizing Pearson's Correlation . . . . .	2
0.2	3. Maximizing and Minimizing Pearson's Correlation . . . . .	3
0.2.1	3.1 Maximizing Pearson's Correlation . . . . .	3
0.2.2	3.2 Minimizing Pearson's Correlation . . . . .	3
0.2.3	3.3 Computing Pearson's Correlation Bounds . . . . .	4
0.3	4. Connection to Fréchet–Hoeffding Bounds . . . . .	4
0.3.1	4.1 Fréchet–Hoeffding Bounds . . . . .	4
0.3.2	4.2 Achievability of Bounds . . . . .	4
0.3.3	4.3 Role of Asymmetry and Ties . . . . .	5
0.3.4	4.4 Practical Implications . . . . .	5
0.4	5. Symmetry and Asymmetry in Correlation Bounds . . . . .	5
0.4.1	5.1 Defining Symmetry in Correlation Bounds . . . . .	5
0.4.2	5.2 Conditions for Symmetry . . . . .	6
0.4.3	5.3 Conditions for Asymmetry . . . . .	6
0.4.4	5.4 Visualizing Symmetry and Asymmetry . . . . .	6
0.4.5	5.5 Role of Entropy in Measuring Asymmetry . . . . .	6
0.4.6	5.6 Empirical and Practical Considerations . . . . .	6
1	References . . . . .	8

## 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  $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(K_X)}$  and  $y_{(1)} \geq y_{(2)} \geq \dots \geq y_{(K_Y)}$  represent the sorted categories of  $X$  and  $Y$ , respectively, then the joint probabilities are allocated

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

### 0.1.2 3.2 Minimizing Pearson's Correlation

Conversely, to minimize Pearson's correlation, we construct the joint distribution  $F_{\min}(x,y)$  by pairing categories of  $X$  and  $Y$  in a countermonotonic (oppositely ordered) fashion. Specifically:

1. Sorting the categories of  $X$  in descending order according to their values.
2. Independently sorting the categories of  $Y$  in ascending order according to their values.
3. Pairing the largest values of  $X$  with the smallest values of  $Y$ , and so forth, while respecting marginal totals.

The minimum expected value under this joint distribution is:

$$\mathbb{E}[XY]_{\min} = \sum_{i,j} x_{(i)} y_{(j)}$$

## 0.2 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.

### 0.2.1 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  $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(K_X)}$  and  $y_{(1)} \geq y_{(2)} \geq \dots \geq y_{(K_Y)}$  represent the sorted categories of  $X$  and  $Y$ , respectively, then the joint probabilities are allocated

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

### 0.2.2 3.2 Minimizing Pearson's Correlation

Conversely, to minimize Pearson's correlation, we construct the joint distribution  $F_{\min}(x,y)$  by pairing categories of  $X$  and  $Y$  in a countermonotonic (oppositely ordered) fashion. Specifically:

1. Sorting the categories of  $X$  in descending order according to their values.
2. Independently sorting the categories of  $Y$  in ascending order according to their values.
3. Pairing the largest values of  $X$  with the smallest values of  $Y$ , and so forth, while respecting marginal totals.

The minimum expected value under this joint distribution is:

$$\mathbb{E}[XY]_{\min} = \sum_{i,j} x_{(i)} y_{(K_Y-j+1)} P_{\min}(X = x_{(i)}, Y = y_{(K_Y-j+1)}).$$

### 0.2.3 3.3 Computing Pearson's Correlation Bounds

Given the optimal pairings, the maximum and minimum Pearson's correlations are calculated as:

$$r_{\max} = \frac{\mathbb{E}[XY]_{\max} - \mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{\text{Var}(X)\text{Var}(Y)}}, \quad r_{\min} = \frac{\mathbb{E}[XY]_{\min} - \mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

These bounds precisely characterize the feasible range of correlation coefficients given the fixed marginal distributions, providing a rigorous basis for interpreting correlations derived from ordinal data.

## 0.3 4. Connection to Fréchet–Hoeffding Bounds

The theoretical bounds on Pearson's correlation given fixed marginals are closely related to the classical Fréchet–Hoeffding bounds. These bounds define the feasible region for joint distributions of two random variables given their marginal distributions.

### 0.3.1 4.1 Fréchet–Hoeffding Bounds

The Fréchet–Hoeffding bounds specify the limits of joint cumulative distribution functions (CDFs) given marginals  $F_X(x)$  and  $F_Y(y)$ :

$$\max\{F_X(x) + F_Y(y) - 1, 0\} \leq F_{XY}(x, y) \leq \min\{F_X(x), F_Y(y)\}.$$

These bounds represent the minimum and maximum possible joint probabilities for each pair of categories  $(x, y)$ .

### 0.3.2 4.2 Achievability of Bounds

The maximum Pearson correlation, as described in Section 3, is achieved precisely when the joint distribution aligns with the upper Fréchet–Hoeffding bound (comonotonic arrangement). Conversely, the minimum correlation corresponds to the lower Fréchet–Hoeffding bound (countermonotonic arrangement).

- **Upper Bound (comonotonic):** Always achievable as probabilities can be consistently allocated by aligning cumulative marginals.
- **Lower Bound (countermonotonic):** Often not achievable in practice due to structural constraints like ties or discrete probability masses, making it impossible to perfectly align cumulative marginals inversely.

### 0.3.3 4.3 Role of Asymmetry and Ties

Asymmetry in the marginal distributions or the presence of ties significantly impacts the achievability and symmetry of these bounds:

- Marginal distributions with equal or symmetrical frequencies typically yield symmetric correlation bounds (i.e.,  $r_{\min} \approx -r_{\max}$ ).
- Marginal distributions with strong asymmetry or significant ties typically yield asymmetric bounds ( $r_{\min} \neq -r_{\max}$ ), reflecting structural constraints in the joint probability allocation.

The asymmetry phenomenon underscores the importance of considering distribution shape and entropy in interpreting correlation results.

### 0.3.4 4.4 Practical Implications

Recognizing these connections provides crucial insights for applied researchers:

- **Interpretation:** Observed correlation coefficients should be contextualized within their theoretically permissible range.
- **Diagnostics:** Calculation of Fréchet–Hoeffding bounds offers a diagnostic tool for evaluating whether reported correlations in empirical research are substantively meaningful or artifacts of marginal constraints.

In subsequent sections, we explore specific conditions under which asymmetry arises, providing explicit analytical conditions and empirical illustrations to guide practical interpretation.

## 0.4 5. Symmetry and Asymmetry in Correlation Bounds

A critical insight from the analysis of correlation bounds is that the achievable maximum and minimum Pearson correlations are not always symmetric around zero. This section explores conditions under which this symmetry holds or breaks and introduces entropy as a measure of marginal asymmetry.

### 0.4.1 5.1 Defining Symmetry in Correlation Bounds

Symmetry in correlation bounds occurs when:

$$r_{\min} = -r_{\max}$$

This equality implies that the magnitudes of the highest achievable positive and negative correlations are identical.

### 0.4.2 5.2 Conditions for Symmetry

Symmetry typically arises under conditions of balanced or symmetric marginal distributions. More explicitly:

- Marginal distributions with uniform frequencies or probabilities.
- Marginal distributions symmetric about their median category.

Under these conditions, the cumulative probability masses can be aligned or inversely aligned without structural discrepancies, maintaining symmetric correlation bounds.

### 0.4.3 5.3 Conditions for Asymmetry

Asymmetry ( $r_{\min} \neq -r_{\max}$ ) commonly arises under the following scenarios :

- Strongly skewed marginal distributions: When one or both marginal distributions are heavily skewed, the allocation of joint probabilities inherently favors one direction of correlation over the other.
- Presence of ties: Significant ties in marginal distributions limit the flexibility of inverse alignment in cumulative distributions, causing structural constraints that prevent the achievement of perfectly inverse pairings.

### 0.4.4 5.4 Visualizing Symmetry and Asymmetry

Visualization strategies, such as heatmaps and entropy-based diagrams, illustrate how marginal distributions transition from symmetric to asymmetric and the corresponding

impact on achievable correlation bounds. Such visualizations aid researchers in understanding and interpreting empirical correlation results within their permissible theoretical ranges.

#### 0.4.5 5.5 Role of Entropy in Measuring Asymmetry

Entropy provides a useful measure of distributional symmetry or asymmetry. Higher entropy typically indicates a more uniform (symmetrical) distribution, while lower entropy signals concentrated (asymmetrical) probabilities. Formally, entropy ( $H$ ) is given by:

$$H(X) = - \sum_{i=1}^{K_X} p_X(x_i) \log p_X(x_i), \quad H(Y) = - \sum_{j=1}^{K_Y} p_Y(y_j) \log p_Y(y_j).$$

As entropy decreases, indicating greater skewness or concentration in probability masses, the bounds on correlations typically become more asymmetric.

#### 0.4.6 5.6 Empirical and Practical Considerations

Empirically, assessing symmetry or asymmetry in correlation bounds provides researchers with important context for interpreting results:

- Symmetric bounds imply similar interpretative weight for positive and negative correlations.
- Asymmetric bounds highlight the necessity for cautious interpretation, especially regarding the relative magnitude and substantive significance of observed correlations.

Recognizing symmetry conditions and their practical implications thus plays an essential role in robustly interpreting correlations in empirical analyses of ordinal data.

## 1 References