

Elastic Bounds of Pearson's r : Theoretical and Empirical Insights for Ordinal Variables

Cees van der Eijk Scott Moser

Table of contents

1	Preamble	4
2	Introduction	4
2.1	On ordinal-categorical variables (ordcats), their level of measurement and their common use in quantitative analyses	5
2.2	Ordinal categorical variables (ordcats) and ‘constrained’ correlations	6
3	Background and Definitions	7
4	Theoretical Framework	7
4.1	Optimal Coupling and Correlation Bounds	7
4.2	Symmetry and the Role of Ties	13
4.3	Special Cases and Boundary Conditions	13
4.4	Relation to Copula Theory (Optional)	14
5	Empirical Illustration and Practical Relevance	14
5.1	Effects of Marginal Shapes	14
5.2	Symmetry Breaking in Correlation Bounds	14
5.3	Applications and Implications	14
6	Extensions and Open Questions	14
6.1	To Do	15
7	Conclusion	15

References	16
8 Appendix A: Mathematical Proofs	17
9 Appendix B: Software Tools and Code Listings	17
10 Appendix C: Additional Tables and Simulations	17

Algorithm Name

Step 1: Do this

Step 2: Do that

Algorithm: My Algorithm

1. Do something
2. Do something else

Just a sample algorithmn

Algorithm 1: While loop with If/Else condition

Result: Write here the result

Input : Write here the input

Output: Write here the output

```

1 while While condition do
2   | instructions
3   | if condition then
4   | | instructions1
5   | | instructions2
6   | else
7   | | instructions3
8   | end
9 end

```

Abstract

The use of product-moment correlations to assess relationships between ordered-categorical variables is widespread and generally tolerated in spite of the non-metric character of such variables. Potential problems of this practice have been identified in the literature. This article focusses on one of these, namely that correlation coefficients are poorly comparable across pairs of ordered-categorical variables because their ranges are constrained in non-uniform ways. This not only affects the validity of correlation coefficients for descriptive purposes, but also as bases for subsequent analyses, as in, e.g., principal component analysis, factor analysis, and structural equation modelling. This article focusses on this problem of ‘constrained’ correlations. It explains the problem and explores conditions under which it is most severe. It provides illustrations based on both fabricated and actual empirical data, the latter from a mass-survey that ubiquitously contains many ordered-categorical variables that are commonly analysed with product-moment correlations. The article is accompanied by a software tool in R for analysing the attainable upper- and lower bounds of correlations in actual data.

Pearson’s correlation coefficient r , when applied to ordinal data, is subject to elasticity due to constraints from marginal distributions. This paper develops theoretical bounds for r , based on Boole–Fréchet–Hoeffding inequalities and rearrangement inequalities, and derives analytic expressions for their limits. We investigate when symmetry in the bounds holds or breaks, and explore practical implications for correlation-based methods like PCA and SEM. An accompanying R package allows users to compute bounds and perform hypothesis tests given fixed marginals.

1 Preamble

- The widespread use of Pearson's r with ordinal data.
- The problem of elasticity: how max/min bounds of r vary with marginals.
- Contributions:
 1. Theory of optimal bounds
 2. Symmetry-breaking conditions
 3. Empirical consequences \rightsquigarrow hypo testing and significance
 4. Supporting tools

The 'Point':

1. *Alert* applied researchers of 'elastic bounds'
 1. Problems of ignoring the above \rightsquigarrow errors in interpretation (e.g. Cohen and Cohen)
2. *Provide* tools to assess if problems exist in own data/ research
3. *Clarify* that traditional hypothesis testing (e.g. t-test of $H_0 : r = 0$) are incorrect and provide correct hypo testing (permutation/ randomization tests)

2 Introduction

In this article we elaborate why correlations involving ordered-categorical variables are generally 'constrained', which means that their attainable maximum and minimum values are not +1 and -1, as is often thought, but between values that have to be calculated for each pair of variables. We assess the magnitude of these constraints in fictitious as well as empirical data and analyse under which conditions they are most severe. The article is accompanied by a software tool in R that allows applied researchers to easily assess the extent of the problem in their own data. The use of product-moment correlations (also known as Pearson correlations) to assess relationships between ordered-categorical variables –henceforth referred to as *ordcats*– is widespread and widely condoned in empirical studies, in spite of such variables not being metric while the product-moment correlation assumes that they are. We do not focus on this problem from an obsessive desire for methodological purity, were it only because many statistical procedures are quite robust for violations of their assumptions [fn: examples?]. Instead, we draw attention to a problematic consequence of this violation of assumptions that, as far as we know, has been rarely recognised in the applied literature although it is well-known in the theoretical statistical literature. This is the problem

that correlation coefficients involving ordcats are constrained in their range which may make them mutually incomparable and may lead to further problems downstream the analytical process. This article proceeds as follows. We start by briefly elaborating the character of ordinal categorical variables (ordcats), why they are generally not regarded as being metric, and why, in spite of that, they are nevertheless very frequently used in statistical procedures that assume metric level data. We then proceed to show why correlations based on ordcats tend to be constrained, which implies that their range is smaller than from -1 to +1, and why this is problematic in applied research. This raises the question how to determine the actual range of correlations between ordcats, which we address in the subsequent section. In section 5 we assess the magnitude of these constraining effects on 7503 correlations between 123 ordcats from a mass survey that typically gives rise ordered categorical variables (ordcats), the British Election Study. We use the same data to explore conditions that influence the magnitude of the constraining effect. [Other sections in between?] In the concluding section we discuss implications of our findings, suggestions for dealing with this problem for applied researchers, and directions for further research.

2.1 On ordinal-categorical variables (ordcats), their level of measurement and their common use in quantitative analyses

Ordcats are ordinal variables with a relatively small number of categories (or ranks) compared to the number of cases, thus resulting in partial orderings of cases. Although they can be found in all field of quantitative inquiry they are ubiquitous in data from standardised social science surveys. Likert items and rating scales [describe in fn Likert items and rating scales, add that there are yet other formats such as semantic-differential scales that also yield ordcats] are among the most favoured ways to elicit responses in such surveys. These survey-question formats have most often between 3 and 11 ordinal categories [fn (or not?) about ANES ‘feeling thermometer’ which has 101 categories but in view of the clustering of responses around only some values can perhaps better be regarded as a 13-point ordcat]. The numerical values used to distinguish the categories of ordcats are most often consecutive integers, starting with either 0 or 1, and in applied survey data analyses it is common to interpret these values as scores. A frequently issued objection against this so-called equal-distance scoring practice [fn: add references] is that it implies that the difference in meaning between, e.g., categories 1 and 2 is equal to that between categories 3 and 4, etc., without clear justification. In other words, it implies an interval-level interpretation by fiat. [add fn: In principle it is quite possible to assign values to survey-based ordcats that have a clear empirical justification, for example by calibrating against psychophysical measurements (cf., Lodge; Tillie; others), or against the results of psychometric/linguistic analysis of verbal response labels (cf Sönning 2024). However, such procedures are rarely integrated in applied survey research because of their high costs.] The consequence of this objection is that ordcats should be considered ordinal variables and that they should not be analysed with statistical procedures requiring metric data. In spite of this objection we find that ordcats are very commonly analysed with product-moment correlations or with procedures that build upon such correlations. We find this clearly exhibited in published articles in high-ranking academic journals [fn: explain how we did this] and in highly regarded textbooks where

examples are given of the use of correlations between ordcats as input for procedures such as PCA, FA or SEM [fn: examples: Kline; Byrne; xxx]. This practice is so established that it is occasionally rationalised by referring to ordcats as being ‘quasi-interval level’ measures [fn: reference]. One can, however, also find some justifications in the literature for this practice, which emphasises the robustness of correlation coefficients for violations of variables involved. Labovitz (1967, 1970) in particular has compared correlations between ordcats using the equal-scoring distance with a monotonic random scoring (a procedure that only assumes ordinality between categories). He claims that the differences are negligible, which suggests that even if the theoretical objection against unfounded interval level interpretations of ordcats would be justified, that objection would nevertheless have such minor practical consequences as to make it irrelevant in most situations. Not surprisingly, Labovitz’ work has, in turn, been criticised itself on various grounds, most particularly of (inadvertently) mainly covering situations where this robustness holds, and not very well other situations (cf O’Brien 1979; xxx). At this place we will not focus on discussions about the equal-distance scoring practice. Our main reason for referring to it is to demonstrate the existence of lively debates in the applied literature about potential issues when using ordcats in procedures requiring metric data. The existence of these debates makes it remarkable that another important issue has – to the best of our knowledge – remained mainly overlooked in the applied literature while it has been well-known in the theoretical statistical literature. This is the problem that correlations between ordcats are generally ‘constrained’, which means that they do not range between -1 and +1, but between minima and maxima that have to be calculated separately for each pair. The nature of this issue is the subject of the next section.

2.2 Ordinal categorical variables (ordcats) and ‘constrained’ correlations

- Simple example of 4x4 crosstab
- Conclusion: for every given distribution of a variable, there are only two distributions of the other variable which do not constrain r : the same, and the inverse. Every other distribution does constrain r
- In case of constraintment (‘shrinkage?’), the relationship becomes partially nonlinear, by necessity (does this mean that the problem of constraintment does not exist for eta?) Is there a measure for nonlinearity? (polynomial multiple correlation?)

Determining the minimum and maximum attainable values of r for given marginal distributions

- Explain different strategies which will all yield the same answers: separate sorting and pairing / Frechet-Hoeffding bounds/ etc
- Proofs in appendix?

How severely are correlations constrained – or: how long is a piece of string?

- Explain strategy of using BES data and resulting dataset of 7503 item-pairs

- Descriptive analyses:
 - Define a measure for extent of constraintment
 - Describe univariate distribution of this measure; asymmetry etc. -Probability of reaching the r-max and r-min (and implications: does this mean that the constraintment is even more severe?)

Factors influencing the degree to which correlations are constrained

- Analyses of the same dataset of 7503 item pairs
- Other characteristics of these include
 - Characteristics of each of the two vars separately (their variance / skew / kurtosis / mean etc), as well as dyadic characteristics (difference of means / skew/ etc)
 - Hypotheses/expectations:
 - * Effect of difference of means and of difference of skew
 - * Perhaps effect of number of categories (constraintment less severe the larger the number of categories)

Consequences of constrained correlations in applied research

- Descriptive interpretation of correlations (Cohen's criteria) is undermined
- It is also undermined by every correlation being expressed on its own unique scale/unique calibration
- Need a nice example of 2 correlations with $r(a,b) \leq r(b,c)$ and where $r(a,b)_{\max}$ is $< r(b,c)$: is it smaller because the linear relationship is weaker, or because the constraintment is stronger?

3 Background and Definitions

- Ordinal data and assumptions of interval-scale correlation
- Existing alternatives: Spearman, polychoric, polyserial; their strengths and limitations
- Coupling and rearrangement in discrete settings

4 Theoretical Framework

4.1 Optimal Coupling and Correlation Bounds

We formalize the problem of identifying the joint distribution (coupling) of two ordinal variables with fixed marginals that maximizes or minimizes the Pearson correlation.

Proposition 1. *Let X and Y be discrete ordinal random variables, each with finite ordered support:*

$$X : x_1 < x_2 < \dots < x_{K_X}, \quad Y : y_1 < y_2 < \dots < y_{K_Y}$$

with marginal probability distributions $p_X(x_i) = P(X = x_i)$, $p_Y(y_j) = P(Y = y_j)$, respectively.

Then, the maximum and minimum values of the Pearson correlation coefficient r_{XY} , consistent with the fixed marginals, are obtained as follows:

- *Maximum r_{XY} : achieved by pairing largest X -values with largest Y -values (comonotonic arrangement).*
- *Minimum r_{XY} : achieved by pairing largest X -values with smallest Y -values (anti-comonotonic arrangement).*

Proof. The Pearson correlation coefficient is given by:

$$r_{XY} = \frac{E[XY] - E[X]E[Y]}{\sigma_X \sigma_Y}.$$

Given fixed marginals $p_X(x_i)$, $p_Y(y_j)$, the expectations $E[X]$, $E[Y]$ and standard deviations σ_X , σ_Y are constant. Thus, to maximize or minimize r_{XY} , we only need to maximize or minimize the expectation $E[XY]$:

$$E[XY] = \sum_{i=1}^{K_X} \sum_{j=1}^{K_Y} x_i y_j P(X = x_i, Y = y_j).$$

This proof relies on a rearrangement inequality of [hard52-inequalities] (Theorem 368):

For real sequences $a_1 \leq a_2 \leq \dots \leq a_n$ and $b_1 \leq b_2 \leq \dots \leq b_n$, and for any permutation π of indices $\{1, \dots, n\}$ we have the following:

$$a_1 b_1 + a_2 b_2 + \dots + a_n b_n \geq a_1 b_{\pi(1)} + a_2 b_{\pi(2)} + \dots + a_n b_{\pi(n)}$$

and

$$a_1 b_n + a_2 b_{n-1} + \dots + a_n b_1 \leq a_1 b_{\pi(1)} + a_2 b_{\pi(2)} + \dots + a_n b_{\pi(n)}$$

Equality occurs only if the permutation π results in a monotonically increasing sequence (for maximum) or monotonically decreasing sequence (for minimum).

To explicitly connect this to our problem, enumerate observations explicitly from sorted marginals:

$$X_{\downarrow} : x_{[K_X]} \geq x_{[K_X-1]} \geq \dots \geq x_{[1]}, \quad Y_{\downarrow} : y_{[K_Y]} \geq y_{[K_Y-1]} \geq \dots \geq y_{[1]}.$$

(Sorted in descending order for maximization)

- Expand discrete probability distributions into explicit observation-level sequences. If $K_X \neq K_Y$, extend the shorter sequence with zero-probability categories so both sequences have equal length, which does not affect marginal probabilities or correlation.

Hence, to maximize $E[XY]$: pair the largest ranks of X with the largest ranks of Y . Explicitly, start from the top (highest values) and move downward:

- Pair as many observations of the largest X -category as possible with the largest Y -category, then next-largest, etc., until all observations are paired.

- By the Hardy–Littlewood–Pólya Rearrangement Inequality, this explicitly constructed pairing yields the maximum possible sum of products, hence maximizing $E[XY]$.

Likewise, to minimize $E[XY]$ pair the largest X -categories with the smallest Y -categories, then second-largest X -categories with second-smallest Y -categories, and so forth (anti-comonotonic ordering). By the rearrangement inequality, this arrangement explicitly yields the minimal sum of products, hence minimizing $E[XY]$.

□

In words, given fixed marginals, we have shown:

- **Maximum correlation** r_{\max} is achieved by comonotonic (DEFINE THIS!) arrangement:

$$P_{\max}(X = x_{[i]}, Y = y_{[i]}) \quad \text{in either descending or ascending order.}$$

- **Minimum correlation** r_{\min} is achieved by anti-comonotonic arrangement:

$$P_{\min}(X = x_{[i]}, Y = y_{[n+1-i]}) \quad \text{pairing descending } X \text{ with ascending } Y, \text{ or viceversa.}$$

With this explicit construction we can simply calculate r_{\max} and r_{\min} when the values of X and Y are ranks.

Proposition 2 (Corollary: Analytic Forms for r_{\min} and r_{\max}). *Let two ordinal variables X and Y have values as follows:*

* X with K_X ordinal levels: $1, 2, \dots, K_X$

* Y with K_Y ordinal levels: $1, 2, \dots, K_Y$

with absolute frequencies:

- $n_X(i)$, number of observations at level i of X .

- $n_Y(j)$, number of observations at level j of Y .

Then the maximum Pearson correlation is:

$$r_{max} = \frac{E[XY]_{max} - \bar{X}\bar{Y}}{\sigma_X \sigma_Y}.$$

Where

$$E[XY]_{max} = \frac{1}{N} \sum_{i=1}^{K_X} \sum_{j=1}^{K_Y} x_i y_j M_{i,j}$$

where explicitly defined frequencies $M_{i,j}$ pair observations in a comonotonic manner, calculated explicitly via the greedy iterative algorithm as follows.

Define the matrix $M_{i,j}$, the frequency with which the level x_i pairs with the level y_j . Clearly, we must have:

$$\sum_j M_{i,j} = n_X(i), \quad \sum_i M_{i,j} = n_Y(j)$$

****Algorithm for explicit calculation of $M_{i,j}$ **** (greedy method):

Algorithm 2: While loop with If/Else condition

Result: Write here the result

Input : Marginal frequencies $n_X(i), i \in \{1, 2, \dots, K_X\}$ and $n_Y(j) j \in \{1, 2, \dots, K_Y\}$

Output: A K_X by K_Y matrix, M

```

1 while While condition do
2   instructions
3   if condition then
4     instructions1
5     instructions2
6   else
7     instructions3
8   end
9 end
10 Initialize available frequencies explicitly:  $n'_Y(j) = n_Y(j)$  for each  $j$ 
11 - For each level  $x_i$ , from smallest  $i = 1$  to largest:
12 - For each level  $y_j$ , from smallest  $j = 1$  to largest:
13 - Pair as many observations as possible:
14
```

$$M_{i,j} = \min \left(n_X(i) - \sum_{s=1}^{j-1} M_{i,s}, n'_Y(j) \right)$$

15 Next, update explicitly:

$$n'_Y(j) \leftarrow n'_Y(j) - M_{i,j}$$

Similarly, the minimum Pearson correlation (r_{min}) achieved by anti-comonotonic alignment (pairing largest ranks of X with smallest ranks of Y) is explicitly:

$$r_{min} = \frac{E[XY]_{min} - \bar{X}\bar{Y}}{\sigma_X\sigma_Y}.$$

Proof. Let the total number of observations be N , thus:

$$\sum_{i=1}^{K_X} n_X(i) = N, \quad \sum_{j=1}^{K_Y} n_Y(j) = N.$$

Means and standard deviations clearly become:

Means:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{K_X} i \cdot n_X(i), \quad \bar{Y} = \frac{1}{N} \sum_{j=1}^{K_Y} j \cdot n_Y(j).$$

Standard deviations:

$$\sigma_X = \sqrt{\frac{1}{N} \sum_{i=1}^{K_X} (i - \bar{X})^2 n_X(i)}, \quad \sigma_Y = \sqrt{\frac{1}{N} \sum_{j=1}^{K_Y} (j - \bar{Y})^2 n_Y(j)}.$$

To achieve the maximum expected product $E[XY]_{max}$, do the following: First *sort* ranks from largest to smallest independently for both X and Y ; then *pair* ranks directly from highest to lowest available observations.

****Minimum Expectation****

Define explicitly a matrix $R_{i,j}$ giving frequencies of pairings $X = x_i$ with $Y = y_j$. For the minimal expectation, pair largest available X -values explicitly with smallest available Y -values first (greedy algorithm):

Initialize explicitly available frequencies:

$$n'_Y(j) = n_Y(j) \quad \text{for each } j$$

- For i from largest to smallest ($i = K_X$ down to 1):

- For j from smallest to largest ($j = 1$ up to K_Y): - Pair explicitly as many observations as possible:

$$R_{i,j} = \min \left(n_X(i) - \sum_{s=1}^{j-1} R_{i,s}, n'_Y(j) \right)$$

- Update explicitly available frequencies:

$$n'_Y(j) \leftarrow n'_Y(j) - R_{i,j}$$

Then explicitly the minimal expectation is:

$$E[XY]_{\min} = \frac{1}{N} \sum_{i=1}^{K_X} \sum_{j=1}^{K_Y} x_i y_j R_{i,j}$$

This explicit equation applies Whitt's rearrangement theorem (@whit76-bivariate), ensuring proper maximal pairing of ranks.

The formula for $E[XY]_{\min}$ is derived in a similar maner.

□

Discussion

We want to construct a joint distribution table M such that:

- The **row sums** match the marginal counts for X : $\sum_j M_{i,j} = n_X(i)$
- The **column sums** match the marginal counts for Y : $\sum_i M_{i,j} = n_Y(j)$

To do this, we go **greedily**, pairing the highest available X values with the **smallest available** Y values, one cell at a time.

To avoid **exceeding** the available marginal totals in Y , we must track how much of $n_Y(j)$ is **still unassigned** at each step — and that's what $n'_Y(j)$ represents.

After assigning $M_{i,j}$ units to the cell (i, j) , we subtract that amount from the remaining pool of level y_j values, like this: $n'_Y(j) \leftarrow n'_Y(j) - M_{i,j}$. This ensures we don't **overfill** any column of the matrix.

To maximize the expectation $E[XY]$, the rearrangement inequality states we must pair values of X and Y from smallest-to-smallest, second-smallest-to-second-smallest, and so forth (comonotonic alignment).

Thus, explicitly:

- Sort X : lowest-to-highest:

$$X_{\text{sorted}} = (\underbrace{x_1, \dots, x_1}_{n_X(1)}, \underbrace{x_2, \dots, x_2}_{n_X(2)}, \dots, \underbrace{x_{K_X}, \dots, x_{K_X}}_{n_X(K_X)})$$

Sorted Y -values (ascending order for max, descending for min):

$$Y_{\text{sorted}(\max)} = (\underbrace{y_1, \dots, y_1}_{n_Y(1)}, \underbrace{y_2, \dots, y_2}_{n_Y(2)}, \dots, \underbrace{y_{K_Y}, \dots, y_{K_Y}}_{n_Y(K_Y)})$$

$$Y_{\text{sorted}(\min)} = (\underbrace{y_{K_Y}, \dots, y_{K_Y}}_{n_Y(K_Y)}, \underbrace{y_{K_Y-1}, \dots, y_{K_Y-1}}_{n_Y(K_Y-1)}, \dots, \underbrace{y_1, \dots, y_1}_{n_Y(1)})$$

The maximum and minimum expectations are simply:

Maximum expectation (pair element-by-element):

$$E[XY]_{\max} = \frac{1}{N} \sum_{k=1}^N X_{\text{sorted}}(k) \cdot Y_{\text{sorted}(\max)}(k)$$

- **Minimum expectation:**

$$E[XY]_{\min} = \frac{1}{N} \sum_{k=1}^N X_{\text{sorted}}(k) \cdot Y_{\text{sorted}(\min)}(k)$$

These classical bounds define feasible joint distributions given marginals and constrain the attainable values of r . There is an interesting connection to Fréchet–Hoeffding and Boole–Fréchet Inequalities, which we note but leave for future research.¹

4.2 Symmetry and the Role of Ties

- Conditions under which $r_{\min} = -r_{\max}$
- How ties in marginal distributions affect the attainability of the lower bound

4.3 Special Cases and Boundary Conditions

- Equal vs. unequal numbers of categories
- Symmetric vs. asymmetric marginal distributions

¹Fréchet–Hoeffding and Boole–Fréchet Inequalities

This is the second paragraph. Indent each new paragraph with four spaces or a tab, just like you would for a multi-paragraph list item.

You can also include code blocks:

```
print("Hello, footnote!")
```

Or even more paragraphs and content as needed.

4.4 Relation to Copula Theory (Optional)

Although our setting is discrete, the problem is conceptually linked to copula-based modeling of dependence with fixed marginals. The connection is discussed as a direction for future work.

5 Empirical Illustration and Practical Relevance

5.1 Effects of Marginal Shapes

- Shape-based elasticity of correlation bounds
- Influence of skewness, modality, and entropy

5.2 Symmetry Breaking in Correlation Bounds

- When and why $r_{\min} \neq -r_{\max}$
- Empirical indicators of bound asymmetry

When we talk about “asymmetry” in this context, we’re referring to how different the distribution of Y is from the distribution of X . Since X is kept uniform, any deviation of Y from uniformity creates asymmetry between the distributions. The *total variation distance* (TV) quantifies this asymmetry:

- It measures how different Y is from being symmetric around its midpoint
- When $TV = 0$: Y is symmetric (like X)
- When TV is large: Y is highly asymmetric compared to X

5.3 Applications and Implications

- Impact on PCA, factor analysis, and SEM using ordinal data
- Practical risks: over-factoring, biased structural coefficients
- Recommendations for interpreting Pearson’s r in applied contexts

6 Extensions and Open Questions

- Role of entropy and joint uncertainty in bounding behavior

- Probabilistic modeling over the $r \in [r_{\min}, r_{\max}]$ interval
- Possibility of adjusting or normalizing r

6.1 To Do

- Can $r_{\min} > 0$?
- Are there systematic biases in testing $H_0 : r = 0$ using t-test? (Answer with Monte Carlo simulated data?)
- Add: “ r_m is invariant to ‘problem 1’ (value assignment. So using ranks is actually w.l.o.g.)

7 Conclusion

- Summary: Theoretical bounds, symmetry-breaking conditions, and implications for applied research
- An R package accompanies this paper, providing tools to compute max/min bounds and perform randomization-based hypothesis tests for fixed marginals. Details are available in the appendix and online.
- Future work: deeper integration with copula models, correction strategies, and generalizations to higher dimensions

References

8 Appendix A: Mathematical Proofs

- Formal derivations of optimal coupling and analytical bounds

9 Appendix B: Software Tools and Code Listings

- Documentation and examples for the R package
- Description of randomization testing features

10 Appendix C: Additional Tables and Simulations

- Supplementary empirical examples, simulation results