Statistical analysis of the tail behaviour of dependent sequences

Brborović, Darko

Doctoral thesis / Disertacija

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:640209

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2025-03-26**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

University of Zagreb

FACULTY OF SCIENCE

DEPARTMENT OF MATHEMATICS

Darko Brborović

Statistical analysis of the tail behaviour of dependent sequences

DOCTORAL DISSERTATION

Zagreb, 2022.

University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF MATHEMATICS

Darko Brborović

Statistical analysis of the tail behaviour of dependent sequences

DOCTORAL DISSERTATION

Supervisor:

Bojan Basrak

Zagreb, 2022.

Sveučilište u Zagrebu

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Darko Brborović

Statistička analiza repnog ponašanja zavisnih nizova

DOKTORSKI RAD

Mentor:

Bojan Basrak

Zagreb, 2022.

Acknowledgements

I am grateful to my mentor, and friend, Bojan Basrak for all guidance and help he has given me over the years. Without his support I could not have completed this thesis.

Also, I am grateful to my family - wife Lorena and kids Nina and Marin - for their support and patience they have shown in allocating me the time needed to finish this thesis.

SUMMARY

In this thesis we analyse the tail behaviour of bivariate random sequences, including the stationary M-dependent sequences, $M \in \mathbb{N}$. Statistical analysis of the asymptotic tail behaviour of random sequences is often done within the theory of point processes. However, if a larger number of tail events is available it is possible to obtain convergence results towards normal distribution, which we do in this thesis.

Motivated by the recent results of DiCiccio and Romano [8] on permutation tests for correlation between random variables, we present and analyse the permutation test of tail dependence for independent and identically distributed bivariate random vectors. Additionally, we present the permutation test of independence for stationary M-dependent random sequences.

SAŽETAK

U ovom radu predstavljamo analizu repnog ponašanja slučajnih nizova dvodimenzionalnih vektora, što uključuje i slučaj stacionarnih M-zavisnih nizova, $M \in \mathbb{N}$. Često se asimptotska analiza repnog ponašanja slučajnih nizova obavlja u okviru teorije točkovnih procesa. Ipak, ukoliko je dostupan veći broj repnih događaja moguće je postići konvergenciju prema normalnoj distribuciji, kao što je to slučaj u ovom radu.

Motivirani relativno novim rezultatima o permutacijskim testovima za korelaciju između dvije slučajne varijable, objavljenima u DiCiccio and Romano [8], u ovom radu predstavljamo i analiziramo test repne zavisnosti za nezavisne i jednako distribuirane dvodimenzinalne slučajne vektore. Dodatno, predstavljamo i permutacijski test nezavisnosti za stacionarne M-zavisne slučajne nizove.

Contents

Introduction

In this thesis we develop permutation tests concerning the joint tail behaviour of some random sequences. Motivation for the analysis of the tail behaviour of random sequences stems from the author's interest in and experience on the financial markets where it is often the case that some measures of central dependence, for example correlations, are insufficient to describe the dependence structure of various assets co-movements. It has been empirically observed that the effect of diversification may diminish in times of crisis as tail events start to dominate the central dependencies' structure (see for example Embrechts et al. [10], Mainik et al. [28], Malevergne and Sornette [29] and references therein). Therefore, it seems desirable to have a simple test for a possible tail dependence between two series of observations. For that purpose we have chosen a permutation test that is straightforward to use and has some desirable properties, such as exactness on finite samples and performing relatively well on small data sizes (see for instance Janssen and Pauls [23] or Chapter 15 in [27]).

Permutation tests are well known and considerable literature exists demonstrating various applications of them. The standard reference is the book by Lehmann and Romano [27]. We present some basic results on the permutation tests in Chapter 1, together with some other results needed in the other two chapters, where we present two permutation tests. In Chapter 2 we present the test for tail dependence of the sequence (Y_i, Z_i), $i = 1, 2, \ldots$ of independent and identically distributed (iid) bivariate random vectors. To formulate the test, we use the model introduced in Ledford and Tawn [26] (see relation 2.2). Of course, the permutation test we develop in Chapter 2 is valid when Y_i and Z_i are independent. In Chapter 3 we formulate the permutation test of independence when Y_i and Z_i are stationary and M-dependent.

Several results that we use in this thesis can be already found in [8]. Their approach, however, does not allow for increasing thresholds (see relation (2.1)). Therefore, the main ideas of the proofs in [8] need to be altered in a nontrivial way to arrive at the main theoretical results we present. In most of the proofs we present, we use some elementary tools like the Chebyshev inequality and the Borel-Cantelli lemma. Despite that, or because of that, some of the proofs are quite lengthy, especially when we treat M-dependent processes in Chapter 3. There we use an idea that is well known in the bootstrap literature, namely, we “erase” from the given sequences blocks of the length M to leave resulting blocks of data that are independent. Note that the main asymptotic results (Theorems 2.1.2 and 3.1.1) in both Chapter 2 and Chapter 3 rely on the Combinatorial Central Limit Theorem proved in Hoeffding [20]. Finally, note that in this thesis we assume that marginal distributions of the bivariate vectors (Y_i, Z_i) are known, that is to say we do not incorporate into our analysis the estimation of their distributions.

1. PREREQUISITES

In this chapter we present results that will be used in later chapters. Besides short overview of the permutation tests (Section 1.1) and the Combinatorial Central Limit Theorem (Section 1.2) we present some measures of dependence that are frequently used in extreme value analysis (Section 1.4). We also recap results on conditional convergence that are needed to formulate our main asymptotic results in Chapters 2 and 3 (Section 1.3).

1.1. PERMUTATION TESTS

For some theoretical examples and an introduction to the background theory of permutation tests, we refer to Lehmann and Romano [27]. Below we give the short overview of some basic facts about permutation tests.

Let $X_i = (Y_i, Z_i)$, $i = 1, 2, \dots$ be a sequence of iid bivariate random vectors with joint distribution P. Assume Y and Z have marginal distributions P_Y and P_Z, respectively. Define $X^n = (X_1, \dots, X_n)$, $Y^n = (Y_1, \dots, Y_n)$ and $Z^n = (Z_1, \dots, Z_n)$, $n \in \mathbb{N}$. P_n is the distribution of the random vector X^n. Suppose we want to test the null hypothesis H_0 that P belongs to a certain family of distributions.

Denote the finite group of permutations of the set $\{1, 2, \dots, n\}$ by $\mathbf{G}_n$. Next introduce the group action of $\mathbf{G}_n$ on $(\mathbb{R}^2)^n = \mathbb{R}^{2n}$ by defining the action of an element $\pi \in \mathbf{G}_n$ as

$$\pi((y_1, z_1) \dots, (y_n, z_n)) = ((y_1, z_{\pi(1)}), \dots, (y_n, z_{\pi(n)})), \tag{1.1}$$

where $((y_1, z_1) \dots, (y_n, z_n)) \in \mathbb{R}^{2n}$.

We say that the randomization hypothesis holds if the distribution of X^n and πX^n are the same for all $\pi \in \mathbf{G}_n$. Note, if G_n is a random element with uniform distribution on the permutation group $\mathbf{G}_n$ independent of X^n, then the randomization hypothesis implies that

G_nX^n and X^n have the same distribution.

Let $T_n(Y^n, Z^n)$ be a test statistic and $\alpha \in (0,1)$ the level of the permutation test. Define the permutation distribution of the statistic T_n as

$$\hat{R}_n(t) = \frac{1}{n!} \sum_{\pi \in \mathbf{G}_n} I_{\{T_n(Y^n, Z^n_\pi) \leq t\}},\ t \in \mathbb{R}, \tag{1.2}$$

where $Z^n_\pi = (Z_{\pi(1)}, \dots, Z_{\pi(n)})$. Its $(1-\alpha)$ quantile is defined as

$$\hat{r}(1-\alpha) = \hat{R}_n^{-1}(1-\alpha) = \inf\{t : \hat{R}_n(t) \geq 1-\alpha\}.$$

The permutation test rejects the null hypothesis if the value of the statistic T_n is greater than $\hat{r}(1-\alpha)$. The proof of consistency of quantiles $\hat{r}(1-\alpha)$ follows from Lemma 11.2.1. in [27]. For the convenience of the reader we restate the lemma below.

Lemma 1.1.1. Let $\alpha \in (0,1)$. The following convergences of quantiles holds:

(i) Let $\{F_n\}$ be a sequence of distribution functions on the real line converging weakly to a distribution function F. Assume F is continuous and strictly increasing at $y = F^{-1}(1-\alpha)$. Then,

$$F_n^{-1}(1-\alpha) \to F^{-1}(1-\alpha),$$

where $F^{-1}(y) = \inf\{x \in \mathbb{R} \mid F(x) \geq y\}$.

(ii) Let $\{\hat{F}_n\}$ be a sequence of random distribution functions satisfying

$$\hat{F}_n(x) \xrightarrow{P} F(x)$$

at all x which are continuity points of some fixed distribution function F. Assume F is continuous and strictly increasing at $y = F^{-1}(1-\alpha)$. Then,

$$\hat{F}_n^{-1}(1-\alpha) \xrightarrow{P} F^{-1}(1-\alpha). \tag{1.3}$$

It is a general goal in the construction of a permutation test to prove

$$\hat{R}_n(t) \xrightarrow{P} R(t). \tag{1.4}$$

The only limiting distribution function R that appears in this thesis is the standard normal cumulative distribution function Φ. In that case and under (1.4) we can use Lemma 1.1.1

(ii) (the randomization hypothesis is not needed here) to conclude that $\hat{r}(1-\alpha)$ converges in probability to the $(1-\alpha)$ quantile of the normal distribution $\Phi^{-1}(1-\alpha)$. Convergence of the $(1-\alpha)$ quantile of the statistic T_n to $\Phi^{-1}(1-\alpha)$ quantile of normal distribution also follows if the conditions of Lemma 1.1.1 (i) are satisfied.

To perform finite sample permutation testing, let $k := n! - \lfloor n!\alpha \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer number smaller than $x \in \mathbb{R}$ (floor). For a given $X^n = x \in \mathbb{R}^{2n}$ compute all $n!$ values of $T_n(\pi x)$ for different permutations π to get the ordered values

$$T_n^{(1)}(x) \leq T_n^{(2)}(x) \leq \cdots \leq T_n^{(n!)}(x).$$

Denote by $M^+(x)$ the number of $T_n^{(i)}(x)$ that are greater than $T_n^{(k)}(x)$ and by $M^0(x)$ the number of $T_n^{(i)}(x)$ that are equal to $T_n^{(k)}(x)$, $i \in \{1, 2 \ldots, n!\}$. Then the permutation test (function) ϕ is given by

$$\phi(x) = \begin{cases} 0, & T_n(x) < T_n^{(k)}(x) \\ \alpha(x), & T_n(x) = T_n^{(k)}(x) \\ 1, & T_n(x) > T_n^{(k)}(x), \end{cases}$$

where $\alpha(x) = (\alpha n! - M^+(x))/M^0(x)$. Suppose that the randomization hypothesis holds. Then the resulting permutation test is exact level α (see Theorem 15.2.1. in [27]), or more precisely

$$E_P[\phi(X^n)] = \alpha,$$

for all probabilities P from the null hypothesis parameter space. Note, $T_n^{(k)}(x)$ are the same as the quantiles $\hat{r}(1-\alpha)$, for realizations x of X^n. Thus, the permutation test rejects the null hypothesis if the value of the statistic is greater than $T_n^{(k)}(x)$, and rejects with the probability $\alpha(x)$ if the value of the statistic is equal to $T_n^{(k)}(x)$. Observe the randomization in ϕ when the test statistic is equal to $T_n^{(k)}(x)$.

Note that the permutation distribution of the general test statistic T_n is a distribution conditional on the set of observations (see, for example, Romano [34] or Janssen and Pauls [23]). It has been known since Hoeffding [21] that it is possible to give sufficient condition for the convergence in (1.4) in the form of unconditional convergence in distribution, but it turns out it is also the necessary condition. That is the content of the following theorem. For the convenience of the reader, we are restating it as it is presented in [11] (Theorem 5.1 in [11]). $\mathscr{X}_n$ in the theorem denotes a general sample space in [11]. In this thesis we will use $\mathscr{X}_n$ of the form $\mathbb{R}^n$.

Theorem 1.1.2. Suppose that X^n has distribution P_n in $\mathscr{X}_n$, and $\mathbf{G}_n$ is a finite group of transformations from $\mathscr{X}_n$ to $\mathscr{X}_n$. Let $\hat{R}_n(\cdot)$ denote the permutation distribution of a statistic T_n. Let G_n and G_n' be independent and uniformly distributed over $\mathbf{G}_n$ (and independent of X^n). Suppose, under P_n

$$\big(T_n(G_nX^n), T_n(G_n'X^n)\big) \xrightarrow{d} (T, T'), \tag{1.5}$$

where T and T' are independent, each with common c.d.f. $R(\cdot)$. Then, for all continuity points t of $R(\cdot)$,

$$\hat{R}_n(t) \xrightarrow{P} R(t). \tag{1.6}$$

Conversely, if (1.6) holds for some limiting c.d.f $R(\cdot)$ whenever t is a continuity point, then (1.5) holds.

We give a few notions before we end this short overview of permutation tests. When applying permutation test procedures, generating all the permutations of a given data set is typically computationally prohibitive. However, this can be avoided by using iid sample of random permutations (see Problem 5.15. in [27]). A similar notion is true for the p-values of a permutation test. Namely, an approximation of the p-value of a permutation test with the permutation group $\mathbf{G}_n$ can be given as

$$\tilde{p} = \frac{1}{B}\Big(1 + \sum_{i=1}^{B-1} I_{\{T(\pi_i X) \geq T(X)\}}\Big), \tag{1.7}$$

where T is the test statistic, $B \in \mathbb{N}$ and $\pi_1, \dots \pi_{B-1}$ are iid and uniformly distributed permutations from $\mathbf{G}_n$. Under the null hypothesis $P(\tilde{p} \leq u) \leq u$, $0 \leq u \leq 1$, so a test that rejects when $\tilde{p} \leq \alpha$ is level α. As B grows approximation $\tilde{p}$ gets asymptotically close in probability to the p-value of the test (for details see [27], section 15.2.1).

If the test statistics' values concentrate on a small number of points, the approximation of the p-value given in (1.7) is actually an upper bound. That is because the number of instances of π_i, such that $T(\pi_i X) = T(X)$, is frequently larger than 1. As suggested in Hemerik and Goeman [19], section 3.4., a randomization of the p-values helps in that case. The authors define the so-called randomized p-value there as

$$p' = \frac{1}{B}\Big(\sum_{i=1}^{B} I_{\{T(\pi_i X) > T(X)\}} + u \cdot \sum_{i=1}^{B} I_{\{T(\pi_i X) = T(X)\}}\Big), \tag{1.8}$$

where u is $U(0,1)$ distributed and independent of X. The randomized p-value p' has the property that $p' \leq \alpha$ if and only if the corresponding permutation test rejects. This property, in particular, implies that p' is uniformly distributed on $[0,1]$ (for details, see [19]).

1.2. THE COMBINATORIAL CENTRAL LIMIT THEOREM

To derive the asymptotic behaviour of the permutation distribution, we will rely on the Combinatorial Central Limit Theorem proved in [20]. Below we give a short overview of that result and its context.

The set up of the problem analyzed in [20] is the following: assume that for each $n \in \mathbb{N}$ we are given $2n$ real numbers $a_n(i), b_n(i)$, $i = 1, 2, \dots, n$, such that neither all instances of $a_n(i)$ nor those of $b_n(i)$ are equal. Denote by $G_n = (G_n(1), \dots, G_n(n))$ a uniformly distributed random element on $\mathbf{G}_n$. Let us define the sum

$$S_n = \sum_{i=1}^{n} a_n(i) b_n(G_n(i)). \tag{1.9}$$

The following theorem gives sufficient conditions for the asymptotic normality of S_n (Theorem 4 in [20]).

Theorem 1.2.1. The distribution of S_n as in (1.9) is asymptotically normal, meaning

$$\lim_{n\to\infty} P(S_n - ES_n \leq x\sqrt{\text{Var } S_n}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy = \Phi(x), \;\; x \in \mathbb{R}, \tag{1.10}$$

if

$$\lim_{n\to\infty} n^{\frac{1}{2}r-1} \frac{\sum_{i=1}^{n}(a_n(i) - \bar{a}_n)^r}{\left(\sum_{i=1}^{n}(a_n(i) - \bar{a}_n)^2\right)^{r/2}} \frac{\sum_{i=1}^{n}(b_n(i) - \bar{b}_n)^r}{\left(\sum_{i=1}^{n}(b_n(i) - \bar{b}_n)^2\right)^{r/2}} = 0, \; r = 3, 4, \dots, \tag{1.11}$$

where

$$\bar{a}_n = \frac{1}{n}\sum_{i=1}^{n} a_n(i), \qquad \bar{b}_n = \frac{1}{n}\sum_{i=1}^{n} b_n(i).$$

Condition (1.11) is satisfied if

$$\lim_{n\to\infty} n \frac{\max_{1\leq i\leq n}(a_n(i) - \bar{a}_n)^2}{\sum_{i=1}^{n}(a_n(i) - \bar{a}_n)^2} \frac{\max_{1\leq i\leq n}(b_n(i) - \bar{b}_n)^2}{\sum_{i=1}^{n}(b_n(i) - \bar{b}_n)^2} = 0. \tag{1.12}$$

The mean and variance of S_n can be expressed explicitly as is shown in Theorem 2 in [20] to get:

$$ES_n = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_n(i) b_n(j), \tag{1.13}$$

$$\mathrm{Var}\,(S_n) = \frac{1}{n-1}\sum_{i=1}^{n}\sum_{j=1}^{n} d_n^2(i,j), \tag{1.14}$$

where

$$d_n(i,j) = a_n(i)b_n(j) - \frac{1}{n}\sum_{g=1}^{n} a_n(g)b_n(j) - \frac{1}{n}\sum_{h=1}^{n} a_n(i)b_n(h) + \frac{1}{n^2}\sum_{g=1}^{n}\sum_{h=1}^{n} a_n(g)b_n(h). \tag{1.15}$$

Remark 1.2.2. Theorem 1.2.1 is referring to asymptotic behaviour of sequences of permuted numbers. Therefore, the only source of randomness is the one from random permutations, i.e. from random elements G_n. Taking into account the structure of probability spaces where G_n, $n \in \mathbb{N}$, are defined we can write probability in relation (1.10) as

$$\lim_{n\to\infty} \frac{1}{n!} \sum_{\pi\in\mathbf{G}_n} I_{\{(S_n^\pi - ES_n)/\sqrt{\mathrm{Var}\,S_n} \le x\}} = \Phi(x), \tag{1.16}$$

where

$$S_n^\pi = \sum_{i=1}^{n} a_n(i)b_n(\pi(i)),\ \pi \in \mathbf{G}_n.$$

More formally, for $n \in \mathbb{N}$, let $\Omega_n = \mathbf{G}_n$. Define σ-algebra $\mathscr{F}_n = \mathscr{P}(\mathbf{G}_n)$ on $\mathbf{G}_n$ and probability P_n that gives equal mass to all permutations form $\mathbf{G}_n$, i.e.

$$P_n(\pi) = \frac{1}{n!},\quad \pi \in \mathbf{G}_n.$$

In such a manner we get the discrete probability space $(\Omega_n, \mathscr{F}_n, P_n)$ on which G_n is defined as identity. With obvious notation for expectation and variance on $(\Omega_n, \mathscr{F}_n, P_n)$ we calculate

$$\begin{aligned} P_n((S_n - E_n(S_n))/\sqrt{\mathrm{Var}_n}(S_n) \le t) &= E_n(I_{\{S_n - E_n(S_n))/\sqrt{\mathrm{Var}_n}(S_n) \le t\}}) \\ &= \int_{\mathbf{G}_n} I_{\{S_n - E_n(S_n))/\sqrt{\mathrm{Var}_n}(S_n) \le t\}} dP_n \\ &= \mathbf{G}_n \text{ is disjoint union of single permutations} \\ &= \frac{1}{n!}\sum_{\pi\in\mathbf{G}_n} I_{\{S_n^\pi - E_n(S_n))/\sqrt{\mathrm{Var}_n}(S_n) \le t\}}. \end{aligned}$$

Extension to probability space that supports all random elements G_n then follows by the Kolmogorov Theorem on Existence of Processes (see Theorem 1 and Remark 2, Section 2.9. in [37]). □

We further note that Hoeffding in [20] actually proves a more general version of its Combinatorial c.l.t. where he looks at numbers of the form $c_n(i,j)$. In that case $d_n(i,j)$ are defined as

$$d_n(i,j) = c_n(i,j) - \frac{1}{n}\sum_{g=1}^{n} c_n(g,j) - \frac{1}{n}\sum_{h=1}^{n} c_n(i,h) + \frac{1}{n^2}\sum_{g=1}^{n}\sum_{h=1}^{n} c_n(g,h). \tag{1.17}$$

and one considers the asymptotic behavior of the sum

$$S_n = \sum_{i=1}^{n} c_n(i, G_n(i)). \tag{1.18}$$

As we can see, the only source of randomness is again random permutation G_n, but now we work with doubly indexed set of numbers $c_n(i,j)$. According to Theorem 2 in [20], the mean and the variance of S_n are

$$ES_n = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} c_n(i,j), \tag{1.19}$$

$$\operatorname{Var}(S_n) = \frac{1}{n-1}\sum_{i=1}^{n}\sum_{j=1}^{n} d_n^2(i,j). \tag{1.20}$$

The following theorem describes the asymptotic behavior of the sum S_n (Theorem 3 in [20]). We additionally assume that $d_n(i,j) \neq 0$ for some pair (i,j) so that $\operatorname{Var}(S_n) > 0$.

Theorem 1.2.3. The distribution of S_n as in (1.18) is asymptotically normal if

$$\lim_{n\to\infty} \frac{\frac{1}{n}\max_{1\le i,j\le n} d_n^r(i,j)}{\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} d_n^2(i,j)\right)^{r/2}} = 0,\ r = 3,4,\ldots \tag{1.21}$$

Condition (1.21) is satisfied if

$$\lim_{n\to\infty} n\frac{\max_{1\le i,j\le n} d_n^2(i,j)}{\sum_{i=1}^{n}\sum_{j=1}^{n} d_n^2(i,j)} = 0. \tag{1.22}$$

1.3. Convergence of conditional distributions

Permutation tests, more precisely permutation distributions (see (1.2)), can be analysed in the context of conditional expectations. In this Section we recap some results regarding conditional distributions. We use the books of Kallenberg [25] and Shiryaev [37] as our main references.

Let ξ and η be random elements on a probability space $(\Omega, \mathscr{F}, P)$ with values in two measurable spaces $(S, \mathscr{S})$ and $(T, \mathscr{T})$ respectively. Suppose that the conditional expectation $E(\xi \mid \eta) = E(\xi \mid \sigma(\eta))$ exists, which is the case if $E|\xi| < \infty$.

Let $F \in \mathscr{F}$ and let $\mathscr{G} \subseteq \mathscr{F}$ be a σ-algebra. The conditional probability of the event F, with respect to $\mathscr{G}$, is the random variable $P(F \mid \mathscr{G}) = E(I_F \mid \mathscr{G})$. From the general properties of conditional expectations we conclude that $P(F \mid \mathscr{G}) = P(F)$ iff F is independent of $\mathscr{G}$, and $P(F \mid \mathscr{G}) = I_F$ iff $F = G$ (a.s), for $G \in \mathscr{G}$. It is also clear that $0 \le P(F \mid \mathscr{G}) \le 1$ (a.s.) and by the monotone convergence theorem for conditional expectation it follows that for the disjoint sets $F_1, F_2, \ldots \in \mathscr{F}$ we have

$$P(\bigcup_i F_i \mid \mathscr{G}) = \sum_i P(F_i \mid \mathscr{G}) \quad \text{(a.s).} \tag{1.23}$$

Despite the validity of the relation (1.23) conditional probability is not in general a measure. The problem is that the equation in (1.23) is satisfied only (a.s.) and exceptional sets on which it does not hold may accumulate to form a set of non-zero P-measure (see Section 2.7. in [37]). The usual remedy for that problem is introduction of regular conditional distributions. We first define a probability kernel (see Chapter 1 in [25]).

Definition 1.3.1. A mapping $\mu : T \times S \to \bar{R}_+$ is the probability kernel from T to S if

i) for $t \in T$ fixed, $B \mapsto \mu(t, B)$, $B \in \mathscr{S}$, is a probability measure on $(S, \mathscr{S})$ and

ii) for $B \in \mathscr{S}$ fixed, $t \mapsto \mu(t, B)$, $t \in T$, is a $\mathscr{T}$-measurable function.

The definition of regular conditional distribution is given next (see Chapter 6 in [25]).

Definition 1.3.2. Let ξ and η be random elements on a probability space $(\Omega, \mathscr{F}, P)$ with values in measurable spaces $(S, \mathscr{S})$ and $(T, \mathscr{T})$, respectively. The regular conditional distribution of ξ, given η, is the function $\mu : T \times S \to \overline{\mathbb{R}}_+$ such that μ is a probability kernel from T to S and

$$\mu(\eta(\omega), B) = P(\xi \in B \mid \eta)(\omega), \ \omega \in \Omega, \ B \in \mathscr{S} \quad \text{(a.s.)} \tag{1.24}$$

It is clear from the definition of the probability kernel that Definition 1.3.2 states that

i) for each $\omega \in \Omega$ the function $B \to \mu(\eta(\omega), B)$, $B \in \mathscr{S}$ is a probability measure on $(S, \mathscr{S})$ and

ii) For each $B \in \mathscr{S}$ the function $\omega \to \mu(\eta(\omega), B)$ is a version of the conditional probability $P(\xi \in B \mid \eta)$, i.e. (1.24) holds.

Note that the regular conditional distribution of ξ, given η, is a random measure on $(S, \mathscr{S})$. If ξ is $\sigma(\eta)$-measurable, then the regular version of $P(\xi \in B \mid \eta)$ is $I_{\{\xi \in B\}}$ and if ξ is independent of $\sigma(\eta)$ the regular version of $P(\xi \in B \mid \eta)$ is $P(\xi \in B)$. Observe that Definition 1.3.2 includes the definition of the regular conditional distribution of ξ, given $\mathscr{G}$, as it is a special case when η is the identity map from $(\Omega, \mathscr{F})$ to $(\Omega, \mathscr{G})$.

The existence of the regular conditional distribution is guaranteed if $(S, \mathscr{S})$ is a Borel space (See Theorem 6.3 in [25] or Theorem 5, Section 2.7. in [37]). The measurable spaces $(S, \mathscr{S})$ and $(T, \mathscr{T})$ are Borel isomorphic if there exists a bijection $f : S \to T$ such that both f and f^{-1} are measurable. $(S, \mathscr{S})$ is a Borel space if it is Borel isomorphic to a Borel subset of $\mathbb{R}$ (see Definition 9, Section 7 in [37]). Note that every Borel space is countably generated, meaning that its σ-algebra is countably generated. Clearly, the prime example of Borel space is $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, but also $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$, $n \in \mathbb{N}$, and $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$. Consequently, for random elements with values in all of those spaces regular conditional distributions exist.

Conditional distributions (we always use regular versions for which (1.24) holds) are useful for calculations of both conditional and unconditional expectations. The next theorem shows how to do that in a fairly general setting (Theorem 6.4. in [25]). It is often referred to as a disintegration theorem as it shows how to disintegrate measures on product space into their one-dimensional components.

Theorem 1.3.3. Let $(S, \mathscr{S})$ and $(T, \mathscr{T})$ be two measurable spaces, $\mathscr{G} \subseteq \mathscr{F}$ σ-subalgebra of $\mathscr{F}$ and ξ a random element in S such that $P(\xi \in \cdot \mid \mathscr{G})$ has a regular version ν. Further consider an $\mathscr{G}$-measurable random element η in T and a measurable function f on $S \times T$ with $E|f(\xi,\eta)| < \infty$. Then

$$E(f(\xi,\eta) \mid \mathscr{G}) = \int f(s,\eta)\nu(ds) \quad \text{(a.s.).} \tag{1.25}$$

If we apply the previous Theorem on the case where $\mathscr{G} = \sigma(\eta)$ and μ is a regular conditional distribution on $T \times S$ such that $P(\xi \in \cdot \mid \eta) = \mu(\eta,\cdot)$ then from (1.25) follows

$$E(f(\xi,\eta) \mid \eta) = \int f(s,\eta)\mu(\eta,ds) \quad \text{(a.s.).} \tag{1.26}$$

Integration of the last equality yields

$$Ef(\xi,\eta) = E\int f(s,\eta)\mu(\eta,ds). \tag{1.27}$$

Note, if ξ and η are independent, then $\mu(\eta,\cdot) = P_\xi(\cdot)$. In that case (1.26) becomes

$$E(f(\xi,\eta) \mid \eta) = \int f(s,\eta)P_\xi(ds) \quad \text{(a.s.).} \tag{1.28}$$

The right hand side of (1.28) is also equal (a.s.) to $E_\xi(f(\xi,\eta))$, where we denote with E_ξ the expectation in respect of the probability measure P_ξ defined on S. Note that the only random element on the right-hand side of (1.28) is η.

Next we give the definition of conditional independence (see Chapter 6 in [25]).

Definition 1.3.4. Let $\mathscr{F}_1, \mathscr{F}_2, \dots$ and $\mathscr{G}$ be σ-subalgebras of $\mathscr{F}$. Then $(\mathscr{F}_i)_{i=1}^\infty$ are conditionally independent, given $\mathscr{G}$, if

$$P\Big(\bigcap_{j=1}^k B_{i_j} \mid \mathscr{G}\Big) = \prod_{j=1}^k P\big(B_{i_j} \mid \mathscr{G}\big) \quad \text{(a.s.),} \quad \text{for all } B_{i_j} \in \mathscr{F}_{i_j},$$

for $j = 1,\dots,k$, for all $k \in \mathbb{N}$ and all k-subsets $\{i_j : 1 \le j \le k\}$ of $\mathbb{N}$. We use the symbol $\perp\!\!\!\perp_{\mathscr{G}}$ to denote (pairwise) conditional independence, given $\mathscr{G}$.

The random elements $(\xi_i)_{i=1}^\infty$ are conditionally independent if induced σ-algebras $\sigma(\xi_i)$, $i \in \mathbb{N}$, are conditionally independent.

Note, if ξ_i in the previous definition are random variables with values in $\mathbb{R}$ then the definition of the conditional independence, given an σ-algebra $\mathscr{G}$, reduces to

$$P\Big(\bigcap_{i=1}^n \{\xi_i < x_i\} \mid \mathscr{G}\Big) = \prod_{i=1}^n P\big(\xi_i < x_i \mid \mathscr{G}\big) \quad \text{(a.s.),} \tag{1.29}$$

for all $(x_1,\dots,x_n)\in\mathbb{R}^n$ (see Theorem 1 and Corollary 1, Section 7.3 in Chow and Teicher [4]).

Further note that any random elements ξ_i that are $\mathscr{G}$-measurable are also conditionally independent, given $\mathscr{G}$. If the ξ_i are independent of $\mathscr{G}$ then their conditional independence, given $\mathscr{G}$, is equivalent to ordinary independence. By using regular conditional distributions some statements from unconditional independence may be translated to the conditional case. Thus, from Lemma 3.8 in [25] it follows that σ-algebras $\mathscr{F}_1,\mathscr{F}_2,\dots$ (contained in $\mathscr{F}$) are conditionally independent, given $\mathscr{G}$, iff

$$(\mathscr{F}_1,\dots,\mathscr{F}_n)\perp\!\!\!\perp_{\mathscr{G}}\mathscr{F}_{n+1},$$

for all $n\in\mathbb{N}$. On the other hand, random variables that are independent can lose their independence under conditioning, but also random variables that are not independent may become independent under conditioning (see Section 7.3. in [4]).

Suppose that (X_n), $n\in\mathbb{N}$, and X are random variables on $(\Omega,\mathscr{F},P)$. By Theorem 6.3. from [25] we conclude that there exists a sequence μ_n, $n\in\mathbb{N}$, and μ of regular conditional distributions of X_n, $n\in\mathbb{N}$, and X, respectively, given $\mathscr{G}$. Therefore, for $\omega\in\Omega$, $\mu_n(\omega,\cdot)$, $n\in\mathbb{N}$, and $\mu(\omega,\cdot)$ are probability measures on $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ and for $B\in\mathscr{B}(\mathbb{R})$ we have

$$\mu_n(\omega,B)=P(X_n\in B\mid\mathscr{G})(\omega),\ n\in\mathbb{N}\quad\text{(a.s.)}\tag{1.30}$$

and

$$\mu(\omega,B)=P(X\in B\mid\mathscr{G})(\omega)\quad\text{(a.s.)}.\tag{1.31}$$

Convergence of the sequence of probability measures $\mu_n(\omega,\cdot)$, $n\in\mathbb{N}$, to $\mu(\omega,\cdot)$ holds if for all μ-continuity Borel sets B, $\mu_n(\omega,B)\to\mu(\omega,B)$ (see Definition 3, Section 3.1. in [37]). Taking into account (1.30) and (1.31) and the fact that Borel σ-algebra is generated by rectangles of the form $(-\infty,x]$, $x\in\mathbb{R}$, we may say that (regular) conditional distributions of X_n converge to the (regular) conditional distribution of X if $P(X_n\le x\mid\mathscr{G})(\omega)\to P(X\le x\mid\mathscr{G})(\omega)$ (a.s.), as $n\to\infty$ for all $x\in\mathbb{R}$ that are continuity points of F_X. For the purpose of easier referencing we will refer to such convergence of conditional distributions as conditional convergence in distribution with respect to $\mathscr{G}$. This is the content of the next definition.

Definition 1.3.5. Let X and $(X_n)_{n=1}^{\infty}$ be random variables on a probability space $(\Omega, \mathscr{F}, P)$ and $\mathscr{G}$ σ-subalgebra of $\mathscr{F}$. The sequence (X_n) converges $\mathscr{G}$-conditionally in distribution to the random variable X if for all continuity points x of the distribution function F of X

$$E(I_{\{X_n \le x\}} \mid \mathscr{G}) \to E(I_{\{X \le x\}} \mid \mathscr{G}) \quad \text{(a.s.)}, \ n \to \infty.$$

We write in that case $X_n \xrightarrow{\mathscr{G}-d} X$. If $\mathscr{G} = \sigma(Y)$, for some random element Y on Ω, we write $X_n|Y \xrightarrow{d} X|Y$.

It is easy to verify that conditional convergence in distribution implies unconditional convergence in distribution. Namely, if F_n are distribution functions of X_n, $n \in \mathbb{N}$, for x a continuity point of F we have, by the dominated convergence theorem,

$$\begin{aligned}\lim_{n\to\infty} F_n(x) &= \lim_{n\to\infty} E(I_{\{X_n \le x\}}) = \lim_{n\to\infty} E(E(I_{\{X_n \le x\}} \mid \mathscr{G})) \\ &= E(\lim_{n\to\infty} E(I_{\{X_n \le x\}} \mid \mathscr{G})) = E(E(I_{\{X \le x\}} \mid \mathscr{G})) = F(x).\end{aligned}$$

For some details on this topic see also Nowak and Zieba [31]. The next two examples show how to put together objects and concepts introduced in this section.

Example 1.3.6. Let $X_1, X_2, \dots$ be a sequence of independent normal random variables, $X_i \sim N(0,1)$, $i \in \mathbb{N}$, and let Y be a random variable with values in $[0,\infty)$ defined on the same probability space and independent of (X_i). Define the new sequence of random variables $V_i = YX_i$, $i \in \mathbb{N}$. By the central limit theorem for i.i.d random variables (Theorem 3, Section 3.3. in [37]) we know that $\sum_i X_i/\sqrt{n}$ converges in distribution to a random variable with the normal cumulative distribution function Φ. We would like to conclude the same for the $\sum_i V_i/\sqrt{n}$ but, since V_i are clearly not independent, we can not use unconditional central limit theorems involving independent random variables.

Although V_i are not independent they are conditionally independent with respect to σ-algebra $\sigma(Y)$. To show that we will use the characterization (1.29) and relation (1.28) with $\xi = (X_1, \dots, X_n)$ and $\eta = Y$. For $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{R}^n$ we have

$$P\big(\bigcap_{i=1}^{n} \{V_i < x_i\} \mid Y\big) = E\big(I_{\bigcap_{i=1}^{n}\{YX_i < x_i\}} \mid Y\big) = \int_{\mathbb{R}^n} \prod_{i=1}^{n} I_{\{s_i Y < x_i\}} P_{\xi}(d(s_1, \dots, s_n)) \quad \text{(a.s.)}.$$

By the Fubini theorem and (1.28) the last integral above is (a.s.) equal to

$$\prod_{i=1}^{n} \int_{\mathbb{R}} I_{\{sY < x_i\}} P_{X_i}(ds) = \prod_{i=1}^{n} P\big(V_i < x_i \mid Y\big).$$

Hence, V_i are conditionally independent with respect to σ-algebra $\sigma(Y)$ and identically distributed with $V_i|Y \sim N(0,Y^2)$.

Next we show $\sum_i V_i/\sqrt{n}$ converges $\sigma(Y)$-conditionally in distribution (see Definition 1.3.5.). Let $n \in \mathbb{N}$ and $x \in \mathbb{R}$ be a continuity point of the distribution function of Y. Use again (1.28) to get

$$E\big(I_{\{\sum_{i=1}^n YX_i/\sqrt{n}\leq x\}} \mid Y\big) = \int_{\mathbb{R}^n} I_{\{Y\sum_{i=1}^n s_i/\sqrt{n}\leq x\}} P_\xi(d(s_1,\dots,s_n)) \quad \text{(a.s.).} \tag{1.32}$$

Note that the only random object on the right hand side of (1.32) is $Y(\omega)$, $\omega \in \Omega$. Take $\omega \in \Omega$ such that $Y(\omega) > 0$. For such ω the integral in (1.32) is equal to

$$\begin{aligned}\int_{\mathbb{R}^n} & I_{\{\sum_{i=1}^n s_i/\sqrt{n}\leq x/Y(\omega)\}} P_\xi(d(s_1,\dots,s_n)) \\ &= P_\xi\Big(\{(x_1,\dots,x_n) \in \mathbb{R}^n \,:\, \sum_{i=1}^n x_i/\sqrt{n} \in (-\infty,\, x/Y(\omega))\}\Big).\end{aligned} \tag{1.33}$$

Let Z be a normal random variable, $Z \sim N(0,1)$, such that $\sum_i X_i/\sqrt{n} \xrightarrow{d} Z$. Choose Z such that it is defined on the same probability space as (X_i) and Y and Z is independent of Y. Then the expression in (1.33) is converging to

$$\Phi\Big(\frac{x}{Y(\omega)}\Big) = P_Z((-\infty,\, x/Y(\omega))),$$

as $n \to \infty$. Use again (1.28) to get

$$P_Z((-\infty,\, x/Y(\omega))) = \int_{\mathbb{R}} I_{\{s\leq x/Y(\omega)\}} P_Z(ds) = E\big(I_{\{ZY\leq x\}} \mid Y\big)(\omega), \tag{1.34}$$

for almost all ω such that $Y(\omega) > 0$. For $\omega \in \Omega$ such that $Y(\omega) = 0$ from (1.32) we conclude that the integral on the right-hand side is converging to δ_0, and the probability measure with all mass concentrated in 0. But then use again (1.28) to get

$$\delta_0 = P_Z(0 \leq x) = \int_{\mathbb{R}} I_{\{0\leq x\}} P_Z(ds) = E\big(I_{\{ZY\leq x\}} \mid Y\big) \tag{1.35}$$

for almost all ω such that $Y(\omega) = 0$. From (1.32) - (1.35) we conclude that

$$E\big(I_{\{\sum_{i=1}^n YX_i/\sqrt{n}\leq x\}}\big) \to E\big(I_{\{ZY\leq x\}} \mid Y\big) \quad \text{(a.s.)}$$

which means

$$\frac{\sum_i YX_i}{\sqrt{n}}|Y \xrightarrow{d} ZY|Y.$$

From the obtained conditional convergence in distribution with respect to Y we conclude that unconditional convergence in distribution of $\sum_i YX_i/\sqrt{n}$ to YZ also holds. □

Example 1.3.7. Let Y be a positive random variable such that $Y \leq M$ (a.s.), $M > 0$. Further, let X_n, $n \in \mathbb{N}$, be a binomial random variable defined conditionally with respect to $\sigma(Y)$, such that $X_n|Y \sim B(n, p_n)$, $p_n := Y/n$. We will determine $\sigma(Y)$-conditional limit in distribution of (X_n) (in the sense of Definition 1.3.5).

Note first that we can choose $n \in \mathbb{N}$, such that $p_n \leq 1$ (a.s.) and then we determine the conditional limit, from that point on, of the sequence (X_n). Alternatively, we can define $p_n = \min\{Y/n, 1\}$ to have a proper situation for the definition of binomial random variables, but we will omit the minimum from the definition of p_n to keep notation simpler.

Let $U_n \sim U(0,1)$, $n \in \mathbb{N}$, be a sequence of independent uniform random variables defined on the same probability space as Y and independent of Y. We define

$$X_n = \sum_{i=1}^{n} I(U_i \leq p_n)$$

and then show $X_n|Y \sim B(n, p_n)$. We denoted by $I(U_i \leq p_n)$ indicator of the set $\{U_i \leq p_n\}$ so $I(U_i \leq p_n) = I_{\{U_i \leq p_n\}}$. First, note that $I(U_i \leq p_n)$ are Bernoulli random variables, conditionally on Y. For $n \in \mathbb{N}$ and $i \in \{1, 2, \ldots, n\}$ we use the relation (1.28) to get

$$\begin{aligned} P(I(U_i \leq p_n) = 0 \mid Y) &= E\big(I(I(U_i \leq p_n) = 0) \mid Y\big) \\ &= \int_{[0,1]} I(I(u \leq Y(\omega)/n) = 0) P_{U_i}(du) \quad \text{(a.s.)}. \end{aligned} \tag{1.36}$$

Similar to the situation in Example 1.3.6, the only random object in the last integral of (1.36) is Y. Therefore (1.36) is equal to

$$\int_0^1 I(I(u \leq Y(\omega)/n) = 0) du = \int_0^{Y(\omega)} du = p_n = Y/n \quad \text{(a.s.)}.$$

In the same way we show $P(I(U_i \leq p_n) = 1 \mid Y) = 1 - p_n$ (a.s.).

Next we show that for $n \in \mathbb{N}$ the random variables $I(U_1 \leq p_n), \ldots, I(U_n \leq p_n)$ are $\sigma(Y)$-conditionally independent (see Definition 1.3.4). We will do that by showing that for all $(x_1, \ldots, x_n) \in \{0,1\}^n$ it holds that

$$P\Big(\bigcap_{i=1}^{n} \{I(U_i \leq p_n) = x_i\} \mid Y\Big) = \prod_{i=1}^{n} P\big(I(U_i \leq p_n) = x_i \mid Y\big) \quad \text{(a.s.)}. \tag{1.37}$$

Let $(x_1, \ldots, x_n) \in \{0,1\}^n$. Then we have by (1.28), the independence of U_i, and Fubini theorem

$$
\begin{aligned}
P\big(\bigcap_{i=1}^{n}\{I(U_i \le p_n) = x_i\} \mid Y\big) &= E\big(I(\bigcap_{i=1}^{n}\{I(U_i \le p_n) = x_i\}) \mid Y\big) \\
&= \int_{[0,1]^n} \prod_{i=1}^{n} I(I(u_i \le Y/n) = x_i) P_{(U_1,\dots,U_n)}(d(u_1,\dots,u_n)) \quad \text{(a.s.)} \\
&= \prod_{i=1}^{n} \int_{[0,1]} I(I(u_i \le Y/n) = x_i) du_i \quad \text{(a.s.)},
\end{aligned}
$$

which proves (1.37).

Finally, we show $X_n|Y \sim B(n, p_n)$ by proving that for $n \in \mathbb{N}$ and $i, j \in \{1,\dots,n\}$ the random variable $I(U_i \le Y/n) + I(U_j \le Y/n)$ is binomial, conditionally on Y. The general case then follows by induction. Let $k \in \{0,1,2\}$. Then we have

$$
\begin{aligned}
P\big(I(U_i \le Y/n) + I(U_j \le Y/n) = k \mid Y\big) &= E\big(I(I(U_i \le Y/n) + I(U_j \le Y/n) = k) \mid Y\big) \\
&= \int_{[0,1]^2} I(I(u_1 \le Y/n) + I(u_2 \le Y/n) = k) P_{U_1,U_2}(d(u_1,u_2)) \quad \text{(a.s.)} \\
&= \int_0^1 \int_0^1 I(I(u_1 \le Y/n) + I(u_2 \le Y/n) = k) du_1 du_2 \quad \text{(a.s.)}. \qquad (1.38)
\end{aligned}
$$

For $k = 0$ the last integral in (1.38) is equal to

$$
\begin{aligned}
&\int_0^1 \int_0^1 I(I(u_1 \le Y/n) = 0, I(u_2 \le Y/n) = 0) du_1 du_2 \\
&\qquad = \int_0^{Y(\omega)/n} \int_0^{Y(\omega)/n} du_1 du_2 = \Big(\frac{Y(\omega)}{n}\Big)^2 = (p_n)^2 \quad \text{(a.s.)}.
\end{aligned}
$$

For $k = 1$ we have two possibilities. The first one is

$$
\begin{aligned}
&\int_0^1 \int_0^1 I(I(u_1 \le Y/n) = 1, I(u_2 \le Y/n) = 0) du_1 du_2 \\
&\qquad = \int_{Y(\omega)/n}^1 \int_0^{Y(\omega)/n} du_1 du_2 = (p_n)(1 - p_n) \quad \text{(a.s.)}
\end{aligned}
$$

and the second one with $I(u_1 \le Y/n) = 0$ and $I(u_2 \le Y/n) = 1$ we treat similarly. Hence

$$
P\big(I(U_i \le Y/n) + I(U_j \le Y/n) = 1 \mid Y\big) = 2p_n(1 - p_n) \ \text{(a.s.)}.
$$

It is clear that for $k = 2$ the last integral in (1.38) is (a.s.) equal to $(1 - p_n)^2$.

We conclude that for $n \in \mathbb{N}$ and $k \in \{1,2,\dots,n\}$ we have

$$
P\big(X_n = k \mid Y\big)(\omega) = \binom{n}{k} \Big(\frac{Y(\omega)}{n}\Big)^k \Big(1 - \frac{Y(\omega)}{n}\Big)^{n-k} \qquad (1.39)
$$

for all $\omega \in A$, $P(A) = 1$. We can choose A such that $Y(\omega) < M$, for $\omega \in A$. Note that $Y(\omega) = \lambda_\omega \in (0, \infty)$ and $np_n(\omega) \to \lambda_\omega$, $n \to \infty$, for $\omega \in A$. By the Poisson theorem (see Section 1.6. in [37]) it follows that for $\omega \in A$

$$\binom{n}{k}\left(\frac{Y(\omega)}{n}\right)^k\left(1-\frac{Y(\omega)}{n}\right)^{n-k} \to \frac{Y(\omega)^k}{k!}e^{-Y(\omega)},\ n \to \infty,$$

for all $k = 0, 1, 2, \ldots$. We conclude $X_n|Y \xrightarrow{d} P|Y$, where P is a (discrete) Poisson random variable with parameter Y. □

Let us now turn our attention back to the permutation distribution $\hat{R}_n(t)$ defined in (1.2). Recall, $\mathbf{G}_n$ is the group of permutations of the set $\{1, 2, \ldots, n\}$. Our aim is to frame the permutation distribution within the concept of conditional distributions.

Note, for $n \in \mathbb{N}$, $(\mathbf{G}_n, \mathscr{P}(\mathbf{G}_n))$ is the Borel space ($\mathbf{G}_n$ is a finite set). Let $\eta = X^n$ be a random vector with values in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $\xi = G_n$ a uniform random variable (discrete) on $\mathbf{G}_n$, independent of (X_n). Theorem 6.3. in [25] implies that then there exists regular conditional distribution of G_n, given X^n. Let $x^n = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\pi = (\pi_n(1), \ldots, \pi_n(n)) \in \mathbf{G}_n$. Suppose that the action of permutation π_n on x^n is given by $\pi_n x^n = (x_{\pi_n(1)}, \ldots, x_{\pi_n(n)})$. Finally, let T_n be a real statistic defined on $\mathbb{R}^n$.

Define function $f_t : \mathbb{R}^n \times \mathbf{G}_n \to \mathbb{R}$, $t \in \mathbb{R}$, by

$$f_t(\pi_n, x^n) = I_{\{T_n(\pi_n x^n) \le t\}}$$

Then, because of the independence between (X^n) and G_n, by (1.28) we get

$$P(T_n(G_n X^n) \le t \mid X^n) = E(I_{\{T_n(G_n X^n) \le t\}} \mid X^n) = \int_{\mathbf{G}_n} I_{\{T_n(gX^n) \le t\}} P_{G_n}(dg) \quad \text{(a.s.)}.$$

P_{G_n} is a discrete probability measure on $\mathbf{G}_n$ such that $P_{G_n}(\pi) = 1/n!$, for $\pi \in \mathbf{G}_n$. Note, $\mathbf{G}_n$ is a disjoint union of individual permutations $\pi \in \mathbf{G}_n$. Therefore

$$P(T_n(G_n X^n) \le t \mid X^n) = \frac{1}{n!} \sum_{\pi \in \mathbf{G}_n} I_{\{T_n(\pi X^n) \le t\}} \quad \text{(a.s.)}. \tag{1.40}$$

We conclude that the permutation distribution $\hat{R}_n(t)$ may be regarded as a conditional distribution of a given statistic with respect to $\sigma(X^n)$. If we can prove, as we will do in the next chapter, that

$$P(T_n(G_n X^n) \le t \mid X^n) \to \Phi(t) \quad \text{(a.s.)},$$

we will get $\sigma(X^n)$-conditional convergence in distribution of (T_n) towards Φ, in the sense of Definition 1.3.5. As already noted, convergence in distribution then follows, and because the limit is deterministic we can conclude that $\hat{R}_n(t) \xrightarrow{P} \Phi(t)$. The obtained convergence in probability implies the convergence of the quantiles $\hat{r}(1-\alpha)$ of $\hat{R}_n(t)$ toward $\Phi^{-1}(1-\alpha)$ (see Lemma 1.1.1 (ii)). Note that stated $\sigma(X^n)$-conditional convergence in distribution is not exactly the same as defined in Definition 1.3.5. because of indexation of $\mathscr{G}$ by n. The difference is notational as we have

$$E(I_{\{T_n(G_nX^n)\leq t\}} \mid X^n) = E(I_{\{T_n(G_nX^n)\leq t\}} \mid \sigma(X)),$$

because of iid assumption on (X^n). $\sigma(X)$ is the σ-algebra generated by the whole process $(X_1, X_2, \ldots)$.

The above considerations show that the permutation distribution $\hat{R}_n(t)$ of the statistic T_n can be regarded as a conditional distribution of the permuted statistic with respect to σ-algebra $\sigma(X^n)$ and the obtained convergence results may be interpreted in the conditional settings. This point of view is quite common in the literature on permutation tests. For example, Romano in [35] shows that the main difference between permutation tests and some bootstrap procedures is in the character of critical values of the related tests. Namely, in the case of bootstrap tests critical values relate to unconditional distributions, whereas in the case of permutation tests critical values relates to conditional distributions. More general treatment of this point of view can be found in Janssen and Pauls [23]. Among other things, the authors there prove the equivalence of the conditional and unconditional tests, that is to say the corresponding critical values, under the condition of convergence of the conditional distribution of the permuted (resampled) statistic towards the distribution of the given statistic (see Lemma 1 in [23]). Note also that conditional interpretation of the permutation distribution shows that Theorem 1.1.2 gives characterization of the conditional convergence in probability in terms of the unconditional (bivariate) convergence in distribution. Similar results were obtained recently by Bücher and Kojadinovic [1] in the bootstrap context.

1.4. TAIL INDEPENDENCE

Let (X_i, Y_i), $i = 1, 2, \ldots$, be a sequence of i.i.d. bivariate random vectors with the joint distribution F and the same marginal distributions $F_X = F_Y$. The (upper) tail dependence coefficient is defined as

$$\chi = \lim_{u \to u^*} P(X_1 > u \mid Y_1 > u)$$

where

$$u^* = \sup\{x \mid F_X(x) < 1\}$$

is the upper limit of the support of F. Sequences (X_i) and (Y_i) are tail independent if $\chi = 0$ and otherwise they are tail dependent. Obviously, tail independence follows from the assumption of independence between X_1 and Y_1.

For the general random vector (X, Y) with marginals F_X and F_Y (not necessarily equal) let $F_X^{\leftarrow}(t) = \inf\{x : F_X(x) \geq t\}$ and $F_Y^{\leftarrow}(t) = \inf\{x : F_Y(x) \geq t\}$, $t \in (0, 1)$. The upper tail dependence coefficient is defined as

$$\chi = \lim_{t \to 1-} P(X > F_X^{\leftarrow}(t) \mid Y > F_Y^{\leftarrow}(t)),$$

whenever this limit exists (see Frahm et al. [13] or Embrechts et al. [10]).

The notions of tail independence and the tail dependence coefficient are often studied within the framework of extremal dependence (see for example Coles et al. [7] or Frahm et al. [13]). In that context it is often the case that further investigations and examples are done within the framework of copulas. Standard references for copula functions are Nelsen [30] and Joe [24].

Following the exposition from [7] for the random vector (X, Y) with the distribution function F and continuous marginal distributions F_X and F_Y we can define uniform random variables $U = F_X(X) = F(X, \infty)$ and $V = F_Y(Y) = F(\infty, Y)$ and conclude that there exists the unique copula function C such that

$$F(x, y) = C(F_X(x), F_Y(y)). \tag{1.41}$$

Recall, the copula is a cumulative distribution function $C : [0, 1] \times [0, 1] \to [0, 1]$ whose margins are uniformly distributed on $[0, 1]$. For continuous F_X and F_Y, the related copula function is unique (Sklar's theorem). Copula function contains information on mutual dependence between X and Y and it is invariant to monotone transformations of their marginal distributions. We can use the copula representation to define tail dependence coefficient after simple calculation. Namely, for $u \in [0,1]$

$$\begin{aligned} P(V > u \mid U > u) &= \frac{P(V > u, U > u)}{P(U > u)} = \frac{1 - 2u + C(u,u)}{1-u} \\ &= 2 - \frac{1 - C(u,u)}{1-u} \sim 2 - \frac{\log C(u,u)}{\log u}, \end{aligned} \tag{1.42}$$

as $u \to 1$. Therefore, for $u \in [0,1]$ we can define function $\chi(u)$ as

$$\chi(u) = 2 - \frac{\log P(U \leq u, V \leq u)}{\log P(U \leq u)}$$

and the tail dependence coefficient as

$$\chi = \lim_{u \to 1} \chi(u).$$

The function $\chi(u)$ can be regarded as a quantile-dependent measure of dependence. The next example, taken from [7], provides a basic insight into the value of the tail dependence coefficient as a measure of dependence in two extreme situations.

Example 1.4.1. Let us first suppose that X and Y are independent. Then we know that $F(x,y) = F_X(x)F_Y(y)$ and so $C(u,v) = uv$ on $[0,1] \times [0,1]$. In this case we know that $\chi = 0$, but we note that $\chi(u) = 0$, for all $u \in [0,1]$.

Next, let us suppose that X and Y are perfectly dependent i.e. $Y = X$ P-$(a.s.)$. Then we have $C(u,v) = \min\{u,v\}$. In this case $\chi(u) = 1$ for all $u \in [0,1]$ and so $\chi = 1$. □

To gain further insight into the characteristics of the dependence function $\chi(u)$ we present some more examples. It turns out that extreme-value copulas offer an interesting class of examples in that regard. We follow discussion from Gudendorf and Segers [15] (similar details can be found in [7]).

Let $X_i = (X_{i1}, \ldots, X_{id})$, $i \in \{1, \ldots, n\}$, be a sample of d-dimensional i.i.d. random vectors with the distribution function F, marginal distribution functions $F_1, \ldots, F_d$ and copula C_F. Suppose F is continuous. We consider vector of componentwise maxima

$$M_n = (M_{n1}, \ldots, M_{nd}), \quad \text{where } M_{nj} = \max\{X_{1j}, \ldots X_{nj}\}, j \in \{1, \ldots, d\}.$$

Then the copula C_n of M_n is given by

$$C_n(u_1,\dots,u_d) = C_F(u_1^{1/n},\dots,u_d^{1/n})^n, \ \ (u_1,\dots,u_d) \in [0,1]^d.$$

We say that the copula C is an extreme-value copula if there exists a copula C_F such that

$$C_F(u_1^{1/n},\dots,u_d^{1/n})^n \to C(u_1,\dots,u_d), \ n \to \infty,$$

for all $(u_1,\dots,u_d) \in [0,1]^d$ (Definition 2.1 in [15]). Closely related to extreme-value copulas are max-stable copulas (Definition 2.2. in [15]): a d-variate copula C is max-stable if

$$C(u_1,\dots,u_d) = C(u_1^{1/m},\dots,u_d^{1/m})^m$$

for every integer $m > 0$ and all $(u_1,\dots,u_d) \in [0,1]^d$. We can state the following theorem (Theorem 2.1. in [15]).

Theorem 1.4.2. A copula is an extreme-value copula if and only if it is max-stable.

The family of extreme-value copulas coincides with the set of copulas of extreme value distributions (for details see [15]). Recall, the d-dimensional probability distribution function G is an extreme value distribution if there exists a sequence of constants $a_{n,i} > 0$ and $b_{n,i} \in \mathbb{R}$, $n \in \mathbb{N}$, $i \in \{1,2,\dots,d\}$, such that G is a limit with non-degenerated margins of the sequence

$$(\frac{M_{n,1}-b_{n,1}}{a_{n,1}},\dots,\frac{M_{n,d}-b_{n,d}}{a_{n,d}}).$$

It is known what the possible extreme value distributions are. We have the following theorem (Theorem 1.1.3. in de Haan and Ferreira [16]) for $d = 1$.

Theorem 1.4.3. The class of extreme value distributions is $G_\gamma(ax+b)$, $x \in \mathbb{R}$, $a > 0$, $b \in \mathbb{R}$, where

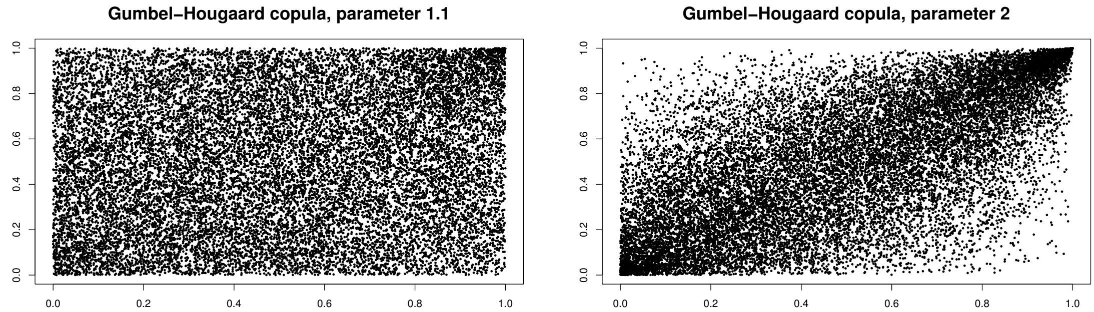
$$G_\gamma(x) = \exp\Big(-(1+\gamma x)^{-1/\gamma}\Big),$$

with $\gamma \in \mathbb{R}$ and where for $\gamma = 0$ the right-hand side is interpreted as $\exp(-e^{-x})$.

Note that the case $\gamma = 0$ yields a Gumbel distribution (or double exponential).

We present the following examples from [7].

Figure 1.1: Results of 20.000 simulations of the Gumbel-Hougaard copula. On the left: copula with parameter 1.1; on the right: copula with parameter 2.



Example 1.4.4. Let us look at the bivariate logistic extreme-value copula or the Gumbel-Hougaard copula defined as

$$C(u,v) = \exp\{-[(-\log u)^{\delta} + (-\log v)^{\delta}]^{1/\delta}\}, \quad \delta \in [1,\infty).$$

The parameter δ determines the strength of dependence. For $\delta = 1$ we get independence, while perfect dependence is attained for $\delta \to \infty$. After an easy calculation we see that $\chi(u) = 2 - 2^{1/\delta}$, which means that the dependence function $\chi(u)$ does not depend on u. It turns out that the dependence function $\chi(u)$ does not depend on u for any other bivariate extreme-value copula (see Section 4.3. in [7] or Section 4 in [15]). Obviously, if the dependence structure is described by the Gumbel-Hougaard copula the tail dependence coefficient χ is equal to $2 - 2^{\alpha}$. In Figure 1.1 we show simulations of the Gumbel-Hougaard copula.

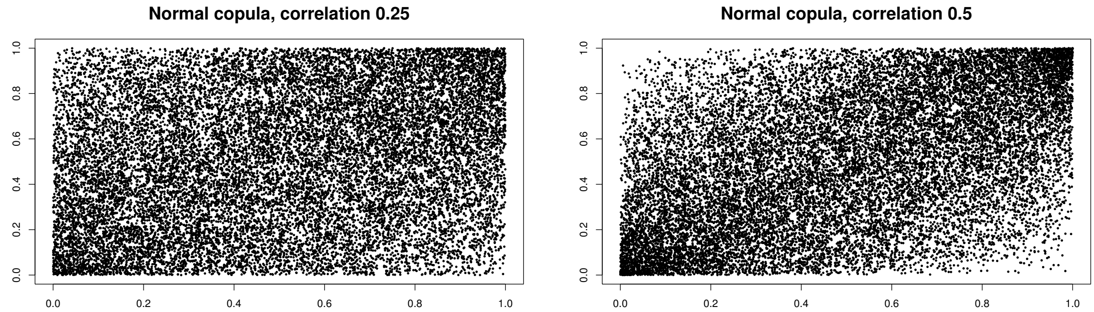
Another example of dependence structure is given by the Gaussian dependence model. Let random vector (X,Y) have bivariate normal distribution with the correlation coefficient ρ. Then marginal distributions of X and Y are also normal and the copula function is given by

$$C(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\{-\frac{1}{2(1-\rho^2)}(s^2 - 2\rho st + t^2)\} ds dt.$$

In Figure 1.2 we show realizations of the Normal (Gaussian) copula.

We note that for $\rho < 1$ random variables X and Y are tail independent, while for $\rho = 1$ they are dependent (see [7]). This copula offers various level of dependence structure within the class of tail independent variables. From the definition of tail independence we

Figure 1.2: Results of 20.000 simulations of the Normal copula. On the left: correlation parameter 0.25; on the right: correlation parameter 0.5.



can conclude that the tail dependence coefficient χ is equal to zero, for all $\rho < 1$. However, as it is shown in [7], the dependence function $\chi(u)$ has strictly positive values for positive ρ up to the neighbourhoods very close to $u = 1$. If the tail dependence coefficient is the only measure of tail dependence, the above-mentioned behaviour of the dependence function $\chi(u)$ for the Gaussian dependence structure may be a problem as any estimate of tail dependence coefficient will be derived from observation for which $u < 1$. $\square$

The rich structure of dependence within tail independence models shows that the tail dependence coefficient χ is unable to provide information on the strength of that dependence. To overcome that problem the authors in [7] introduce new measure of tail dependence $\bar{\chi}$. By analogy with the definition of dependence function $\chi(u)$ we first define

$$\bar{\chi}(u) = \frac{2\log P(U > u)}{\log P(U > u, V > u)} - 1 = \frac{2\log(1-u)}{\log \bar{C}(u,u)} - 1, \quad u \in [0,1]$$

and then the dependence coefficient $\bar{\chi}$ as

$$\bar{\chi} = \lim_{u \to 1} \bar{\chi}(u)$$

for which $-1 < \bar{\chi} \leq 1$. We denoted by $\bar{C}$ the tail copula function. It is related to the copula function by the relation

$$\bar{C}(u,v) = 1 - u - v + C(u,v).$$

Note that $\bar{\chi} = 1$ for all tail dependent random variables, i.e. copulas that describe their dependence structure. That is the content of the next lemma.

Lemma 1.4.5. Let X and Y be continuous random variables and C the copula associated with them. If X and Y are tail dependent then their tail dependence coefficient $\bar{\chi}$ is equal to 1.

Proof. As already discussed, after transforming the marginal distributions of X and Y to uniform distribution the tail dependence coefficient χ is then defined as

$$\chi = \lim_{u\uparrow 1} \frac{\bar{C}(u,u)}{1-u}. \tag{1.43}$$

Suppose $\chi > 0$. We know that (1.43) is equivalent to the fact that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $1-u < \delta$ implies $|\bar{C}(u,u)/(1-u) - \chi| < \varepsilon$, $u \in (0,1)$. Take $\varepsilon = \chi/2$. Then exists $\delta_\chi > 0$ such that for u satisfying $1-u < \delta_\chi$ holds

$$-\frac{\chi}{2} < \frac{\bar{C}(u,u)}{1-u} - \chi < \frac{\chi}{2}$$

i.e.

$$\frac{\chi}{2} < \frac{\bar{C}(u,u)}{1-u} < \frac{3\chi}{2}. \tag{1.44}$$

Note, for $u \in (0,1)$

$$\log \bar{C}(u,u) = \log\left(\bar{C}(u,u)\frac{1-u}{1-u}\right) = \log\frac{\bar{C}(u,u)}{1-u} + \log(1-u).$$

Use the right inequality in (1.44) to conclude that for $u \in (1-\delta_\chi, 1)$ we have

$$\log \bar{C}(u,u) < \log\frac{3\chi}{2} + \log(1-u).$$

By the definition of the tail coefficient $\bar{\chi}$ then we have

$$\bar{\chi} \geq \lim_{u\uparrow 1} \frac{2\log(1-u)}{K + \log(1-u)} - 1, \tag{1.45}$$

where $K = \log(3\chi/2)$ is a constant. As the limit in (1.45) is undetermined use the L'Hospital rule to conclude $\bar{\chi} \geq 1$. Similarly, by using the left inequality in (1.43) we conclude $\bar{\chi} \leq 1$. Hence $\bar{\chi} = 1$. ∎

Example 1.4.6. Let us calculate the value of the coefficient $\bar{\chi}$ for the dependencies we already analysed in the above examples.

If X and Y are perfectly dependent then $\bar{C}(u,v) = 1-u-v+\min\{u,v\}$. Therefore,

$$\bar{\chi}(u) = \frac{2\log(1-u)}{\log(1-u)} - 1 = 1$$

and so $\bar{\chi} = 1$ in this case. When X and Y are independent $\bar{C}(u,v) = 1 - u - v + uv = (1-u)(1-v)$ and so $\bar{\chi} = 0$.

The tail dependence is obtained for the Gumbel-Hougaard copula when the parameter $\delta > 1$. By Lemma 1.4.5 we conclude that $\bar{\chi} = 1$ for such δ. Clearly, $\bar{\chi} = 0$ for $\delta = 1$.

Finally, for the Gaussian copula C with the correlation coefficient ρ we have

$$\bar{C}(u,v) = \int_{\Phi^{-1}(u)}^{\infty} \int_{\Phi^{-1}(v)}^{\infty} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\{-\frac{1}{2(1-\rho^2)}(s^2 - 2\rho st + t^2)\} ds dt.$$

It turns out that in this case $\bar{\chi} = \rho$ (see [7]). Thus, the Gaussian copula offers an example of a parametric model that covers a wide range of dependencies within the class of tail independent models. □

To summarize, the tail dependence coefficient χ is equal to zero in the case of tail independence, while it is in $(0,1]$ in the tail dependence case. Therefore, the tail dependence coefficient χ is useful for detecting tail independence. However, the measure of dependence $\bar{\chi}$ recovers the strength of dependence in the tail independence case as $\bar{\chi} \in (-1,1)$ in that case. In the words of the authors in [7]: "$\{\chi > 0, \bar{\chi} = 1\}$ signifies tail dependence, in which case the value of χ determines a measure of the strength of dependence within the class; alternatively, $\{\chi = 0, \bar{\chi} < 1\}$ signifies tail independence, in which case the value of $\bar{\chi}$ determines the strength of dependence within this class".

Tail independence can be analysed in some other frameworks besides the one presented above. Instead of the transformation of continuous marginals to the uniform distribution one can define, for a random vector (X,Y) with the distribution function F, new random variables $Z = -1/\log F_X(X)$ and $W = -1/\log F_Y(Y)$. In that case Z and W have unit Fréchet distribution, meaning that $P(Z \le z) = \exp(-1/z)$, $z \in (0,\infty)$. For tails of Z and W we have the asymptotic estimate $P(Z > z) = P(W > z) \sim 1/z$, for $z \to \infty$. Ledford and Tawn in [26] consider the model of tail dependence described by

$$P(Z > z, W > z) \sim z^{-1/\eta} L(z)$$

for $z \to \infty$, $\eta \in [1/2, 1]$ the tail dependence coefficient and L a slowly varying function at infinity ($L(tz)/L(z) \to 1$, as $z \to \infty$, and $t > 0$ fixed). Subsequently, the authors in Peng [32] and Draisma et al. [9] provided a theoretical background for the model using a variant of the second order condition imposed on the copula survivor function of Y and Z.

The model describes tail dependence by the parameter $\eta \in (0,1]$ such that

$$P(Y > u, Z > u) \sim L(u)P(Z > u)^{1/\eta}, \tag{1.46}$$

for $u \to \infty$, where L is a slowly varying function at infinity. Observe, for $\chi > 0$, one necessarily has $\eta = 1$. For $\eta = 1/2$, Y and Z are sometimes called near independent (see [26], or Charpentier and Segers [2] in a slightly different context).

Links with coefficients χ and $\bar{\chi}$ are given by (see Heffernan [17])

$$\bar{\chi} = 2\eta - 1$$

and

$$\chi = \begin{cases} c, \;\; \bar{\chi} = 1, \;\; \text{and} \;\; L(t) \to c, \text{ as } t \to \infty \\ 0, \;\; \bar{\chi} = 1, \;\; \text{and} \;\; L(t) \to 0, \text{ as } t \to \infty \\ 0, \;\; \bar{\chi} < 1, \end{cases}$$

where $c > 0$. Calculation of those parameters for various copula functions can be found in Heffernan [17]. A similar list of examples in a slightly different context can be found in [2]. Below we give few examples from that paper.

Example 1.4.7. A copula C is Archimedean if it is of the form

$$C(u_1, \ldots, u_d) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_d)), \;\; (u_1, \ldots, u_d) \in [0,1]^d,$$

where the Archimedean generator $\phi : [0,1] \to [0,\infty]$ is convex, decreasing function with $\phi(0) = 0$ and ϕ^{-1} denotes the generalized inverse of ϕ (see [2]). By using this parametric representation of Archimedean copulas the authors in [2] define slightly different dependence measures. Define the coefficient θ_1 as

$$\theta_1 = -\lim_{s \downarrow 0} \frac{s\phi'(1-s)}{\phi(1-s)}$$

when this limit exists. Then, if $\theta_1 > 1$, the upper tail exhibits tail dependence, while if $\theta_1 = 1$, the upper tail exhibits tail independence. The tail independence is further analysed dependent on the value of $-\phi'(1)$ (for details and connections with other tail dependence measures see [2]).

The Archimedean generator of the Gumbel-Hougaard copula is given by $\phi(t) = (-\log t)^{\theta}$, $\theta \in [1,\infty)$. The tail independent case is attained for $\theta = 1$, which confirms the value of

$-\phi'(1)$, which is equal to 1 in this case. It is worth noting that the Gumbel-Hougaard copula is the only extreme-value copula that is also an Archimedean copula.

Another example of an Archimedean copula is the Joe copula (no. 6 in Table 1 in [2]) with the generator $\phi(t) = -\log(1-(1-t)^{\theta})$, $\theta \in [1,\infty)$. The Joe copula shares tail independence characteristics with the Gumbel-Hougaard copula.

A slightly more delicate member of the Archimedean family of copulas is, for example, the one defined by the Archimedean generator function $\phi(t) = -\log(\theta t + (1-\theta))$, $\theta \in (0,1]$ (no. 7 in Table 1 in [2]). In this case we get near independence (see section 4 in [2]). □

The theory of estimation of tail dependence is very wide and there are many non-parametric estimators of tail dependence (see for example Schmidt and Stadtmüler [36]). Within that theory the diagnostics of tail independence are interesting in their own right. Falk and Michel in [12] argue that testing for tail independence (i.e. tail dependence) is important in data analysis of extreme events. Note, for bivariate extreme-value copulas the marginal random variables are either independent or tail dependent (see the definition of the coefficient of upper tail dependence in [15]).

2. THE PERMUTATION TEST OF TAIL DEPENDENCE - I.I.D. DATA

In this chapter we present a permutation test of tail dependence. The underlying idea is to analyse a natural nonparametric estimator of the parameter χ, which can be constructed based on the number of joint upcrossings of the random variables Y_i and Z_i over a high threshold, that is

$$\sum_{i=1}^{n} I_{\{Y_i > u'_n, Z_i > v'_n\}} \tag{2.1}$$

for some suitable increasing sequences (u'_n) and (v'_n). A similar idea was considered by Schmidt and Stadtmüller in [36] in the case of tail dependence. They show asymptotic normality and strong consistency not only for the estimator of χ but for the whole tail copula under a technical (and often used) second order condition. In Draisma et al. [9] the authors give an estimator of the coefficient η introduced in Ledford and Tawn [26] and construct a test of asymptotic dependence ($\eta = 1$). We refer to [7] for an overview of different measures of tail dependence. Falk and Michel [12] discussed a test for tail independence within a framework of extreme value distributions. Hüsler and Li [22] give a nonparametric test for the asymptotic independence of bivariate random vectors whose distribution function lies in the domain of attraction of an extreme value distribution.

All asymptotic results in this chapter hold under fairly general assumptions. However, to formulate the test for positive tail dependence we will use the model introduced in [26], where the tail dependence is described by a parameter $\eta \in (0,1]$ such that

$$P(Y > u, Z > u) \sim L(u)P(Z > u)^{1/\eta}, \tag{2.2}$$

for $u \to \infty$, where L is a slowly varying function at infinity. Observe, for $\chi > 0$, one necessarily has $\eta = 1$, but for all $\eta > 1/2$ one should expect to observe more joint extremes of Y and Z than if they were independent. However, a similar effect should appear even with $\eta = 1/2$, provided that $\liminf_{x\to\infty} L(x)$ is larger than 1. Therefore, we hope that a test which detects $\eta > 1/2$ can also spot such a subtler form of positive tail association.

In the next section we show that there exist threshold sequences (u_n), (v_n) and a centring and (nondeterministic) scaling, which make the statistic $\hat{T}_n$ given in (2.14), and based on the expression (2.1), asymptotically normal. To justify the application of the permutation test we show that the permutation distribution of the test statistic $\hat{T}_n$ is also asymptotically normal. To formulate the permutation test of tail dependence we suppose the marginal distributions of Y and Z to be continuous while the joint threshold upcrossings of Y and Z follow the model based on (2.2). The null hypothesis then consists of distributions for which $\eta < 1/2$ and some distributions with $\eta = 1/2$, but with the function L not exceeding 1, roughly speaking. See (2.21) in the following section for precise formulation. This allows us to construct a permutation test, which uses the permutation distribution as the null distribution. To perform the test, we

i) calculate the test statistic $\hat{T}_n$ for multiple permutations of the vector $(Z_1,\ldots,Z_n)$ and then

ii) reject the null hypothesis whenever the original test statistic exceeds a predetermined quantile of the empirical permutation distribution.

In the next section, we present our main theoretical results supporting the construction of the permutation test for tail dependence. Section 2.2 presents a simulation study concerning the suggested test's power and an application of the test on financial data. Section 2.3 is dedicated to proofs of the theoretical results from Section 2.1.

2.1. MAIN RESULTS

Let $X_i = (Y_i, Z_i)$, $i = 1,2,\ldots$ be a sequence of iid bivariate random vectors defined on the probability space $(\Omega, \mathscr{F}, P)$. Let $X^n = (X_1,\ldots,X_n)$, $Y^n = (Y_1,\ldots,Y_n)$, $Z^n = (Z_1,\ldots,Z_n)$, $n \in \mathbb{N}$. To simplify the notation, for a generic member of an identically distributed sequence, say (X_i), (Y_i, Z_i), we write X, (Y,Z). Denote by F_Y and F_Z the distribution functions of Y and Z, respectively, and suppose they are known. Let $(m_n)_{n\in\mathbb{N}}$ be an intermediate sequence of integers such that $m_n \to \infty$, for $n \to \infty$, and

$$m_n = O\big(n^{2/3-\tau}\big), \tag{2.3}$$

for some $\tau > 0$. Suppose that there exist two sequences (u_n) and (v_n) of positive real numbers such that $u_n \to \sup\{x : F_Y(x) < 1\}$, $v_n \to \sup\{x : F_Z(x) < 1\}$ and

$$nP(Y > u_n) \to 1, \quad nP(Z > v_n) \to 1,\ n \to \infty. \tag{2.4}$$

Note that the existence of such sequences (u_n) and (v_n) is immediate for continuous random variables.

Remark 2.1.1. Assumption (2.4) is very natural one in the context of regularly varying random variables. Recall that (nonnegative) random variable Y with distribution function F_Y is regularly varying with index α, $\alpha > 0$, if $\bar{F} = 1 - F$ is regularly varying function with index $-\alpha$ meaning $\lim_{t\to\infty} \bar{F}(tx)/\bar{F}(t) = x^{-\alpha}$, for $x > 0$. Then the existence of a sequence (u_n) as in (2.4) follows from Theorem 3.6. in [33]. $\square$

Let $I_{Y,i} = I_{\{Y_i > u_{\sqrt{m_n}}\}}$ and $I_{Z,i} = I_{\{Z_i > v_{\sqrt{m_n}}\}}$. Consider the following auxiliary statistic

$$S_n(X^n) = \frac{m_n}{n}\sum_{i=1}^{n} I_{\{Y_i > u_{\sqrt{m_n}}\}} I_{\{Z_i > v_{\sqrt{m_n}}\}} = \frac{m_n}{n}\sum_{i=1}^{n} I_{Y,i} I_{Z,i}.$$

The unusual choice of thresholds in S_n is motivated by the fact that in the case of independent components of the random vector (Y,Z) we have

$$m_n P(Y > u_{\sqrt{m_n}}, Z > v_{\sqrt{m_n}}) \to 1,\ n \to \infty,$$

because of (2.4).

Let G_n be a random element on Ω with uniform distribution on the permutation group $\mathbf{G}_n$ on the set $\{1,2,\ldots,n\}$. The random element G_n pairs the values $Y_1,\ldots Y_n$ with a permutation of $Z_1,\ldots,Z_n$ (see (1.1) for the definition). Define the sum $S_n^{G_n}$ as

$$S_n^{G_n} = S_n(G_n X^n) := \frac{m_n}{n}\sum_{i=1}^{n} I_{\{Y_i > u_{\sqrt{m_n}}\}} I_{\{Z_{G_n(i)} > v_{\sqrt{m_n}}\}} = \frac{m_n}{n}\sum_{i=1}^{n} I_{Y,i} I_{Z,G_n(i)}.$$

We assume that G_n and X^n are independent in the rest of the text.

Let $h : \mathbf{G}_n \times \mathbb{R}^{2n} \to [0,\infty)$ be a function defined as

$$h(\pi, x^n) = \frac{m_n}{n}\sum_{i=1}^{n} I_{y,i} I_{z,\pi(i)},$$

where $x^n = ((y_1,z_1),\ldots,(y_n,z_n))$, $\pi \in \mathbf{G}_n$. Clearly, $h(G_n,X^n) = S_n^{G_n}$ and, because of (2.3), we have $E|h(G_n,X^n)| \le m_n/n < \infty$. Let P_{G_n} be the probability on $\mathbf{G}_n$ induced by random element G_n. Clearly $P_{G_n}(\pi) = 1/n!$, for $\pi \in \mathbf{G}_n$. Because of the independence between X^n and G_n by (1.28) we have

$$\begin{aligned} E(S_n^{G_n} \mid X^n) &= \int_{\mathbf{G}_n} \frac{m_n}{n} \sum_{i=1}^n I_{Y,i} I_{Z,\pi(i)} P_{G_n}(d\pi) \quad \text{(a.s.)} \\ &= \frac{1}{n!} \sum_{\pi\in\mathbf{G}_n} \frac{m_n}{n} \sum_{i=1}^n I_{Y,i} I_{Z,\pi(i)} \quad \text{(a.s.)} \end{aligned} \tag{2.5}$$

From (2.5) we derive

$$E(S_n^{G_n} \mid X^n) = m_n \bar{I}_Y \bar{I}_Z \ \ \text{(a.s.)} \tag{2.6}$$

and

$$\mathrm{Var}(S_n^{G_n} \mid X^n) = \frac{1}{n-1} \frac{m_n^2}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,j} - \bar{I}_Z\right)^2 \text{ (a.s.)}, \tag{2.7}$$

where

$$\bar{I}_Y = \frac{1}{n} \sum_{i=1}^n I_{Y,i} \ \text{ and } \bar{I}_Z = \frac{1}{n} \sum_{i=1}^n I_{Z,i}.$$

Details are given in Section 2.3 after the proof of Theorem 2.1.2.

Similarly, by using (1.28), we conclude that almost surely

$$\begin{aligned} & P\Big(S_n^{G_n} - E(S_n^{G_n} \mid X^n) \le t\sqrt{\mathrm{Var}(S_n^{G_n} \mid X^n)} \mid X^n\Big) \\ &= \int_{\mathbf{G}_n} I_{\{\frac{m_n}{n}\sum_{i=1}^n I_{Y,i}I_{Z,\pi(i)} - E(S_n^{G_n}|X^n) \le t\sqrt{\mathrm{Var}(S_n^{G_n}|X^n)}\}} P_{G_n}(d\pi) \\ &= \frac{1}{n!} \sum_{\pi\in\mathbf{G}_n} I_{\{\frac{m_n}{n}\sum_{i=1}^n I_{Y,i}I_{Z,\pi(i)} - E(S_n^{G_n}|X^n) \le t\sqrt{\mathrm{Var}(S_n^{G_n}|X^n)}\}},\ t \in \mathbb{R}. \end{aligned} \tag{2.8}$$

The main asymptotic result of this chapter is summarized in the following theorem.

Theorem 2.1.2. Let (Y_i,Z_i), $i \in \mathbb{N}$, be a sequence of iid bivariate random vectors. Suppose that (2.3) and (2.4) hold. Then for $t \in \mathbb{R}$

$$\lim_{n\to\infty} P\big(S_n^{G_n} - E(S_n^{G_n} \mid X^n) \le t\sqrt{\mathrm{Var}(S_n^{G_n} \mid X^n)} \mid X^n\big) = \Phi(t), \ \text{(a.s.)}, \tag{2.9}$$

where Φ is the standard normal cumulative distribution function.

Remark 2.1.3. Denote by $P_{\mathbb{X}}$ the distribution of the sequence (X_i). Then (2.9) can be restated (compare (1.16) and (2.8)), using the expressions in (2.6) and (2.7), as

$$P_{G_n}\Big(S_n^{G_n} - E(S_n^{G_n} \mid X^n) \le t\sqrt{\mathrm{Var}(S_n^{G_n} \mid X^n)}\Big) \to \Phi(t),\ t \in \mathbb{R},\ P_{\mathbb{X}}\text{-a.s.} \tag{2.10}$$

Note that the expression on the left-hand side of (2.9) is still random and dependent on X^n. □

Define the statistic T_n as

$$T_n(X^n) := \sqrt{n-1}\frac{\sum_{i=1}^n I_{Y,i}I_{Z,i} - n\bar{I}_Y\bar{I}_Z}{\sqrt{\sum_{i=1}^n \left(I_{Y,i}-\bar{I}_Y\right)^2}\sqrt{\sum_{j=1}^n \left(I_{Z,j}-\bar{I}_Z\right)^2}}. \tag{2.11}$$

Then we have

$$T_n(G_nX^n) = \sqrt{n-1}\frac{\sum_{i=1}^n I_{Y,i}I_{Z,G_n(i)} - n\bar{I}_Y\bar{I}_Z}{\sqrt{\sum_{i=1}^n \left(I_{Y,i}-\bar{I}_Y\right)^2}\sqrt{\sum_{j=1}^n \left(I_{Z,j}-\bar{I}_Z\right)^2}}.$$

Use the expressions in (2.6) and (2.7) and rearrange slightly the terms of $(S_n^{G_n} - E(S_n^{G_n} \mid X^n))/\sqrt{\mathrm{Var}\,(S_n^{G_n}\mid X^n)}$ to see that

$$T_n(G_nX^n) = \frac{S_n^{G_n} - E(S_n^{G_n}\mid X^n)}{\sqrt{\mathrm{Var}\,(S_n^{G_n}\mid X^n)}}. \tag{2.12}$$

Recall the definition of the permutation distribution $\hat{R}_n(t)$ of the statistic T_n (see (1.2)). From (2.12) we conclude that Theorem 2.1.2 states that the permutation distribution $\hat{R}_n(t)$ of the statistic T_n converges almost surely to the standard normal distribution function Φ. So, for $n \to \infty$ we have

$$\hat{R}_n(t) \to \Phi(t),\; t \in \mathbb{R},\; \text{(a.s.)} \tag{2.13}$$

Remark 2.1.4. Assume that Y and Z are independent, with distributions P_Y and P_Z. Consider the test statistic T_n. Then $P_{X^n} = (P_Y \times P_Z)^n$, and one can easily check that the randomization hypothesis holds. From the definition of the permutation distribution given in (1.2) we conclude that $E(\hat{R}_n(t)) = P(T_n \leq t)$, for $t \in \mathbb{R}$. Take the expectation in (2.13) and use the dominated convergence theorem to get

$$E\big(\lim_{n\to\infty} \hat{R}_n(t)\big) = \lim_{n\to\infty} E(\hat{R}_n(t)) = \lim_{n\to\infty} P(T_n(Y_n,Z_n) \leq t) = \Phi(t).$$

We conclude that the distribution of the test statistic T_n also converges to the standard normal distribution. Hence, the permutation distribution $\hat{R}_n(t)$ asymptotically approximates the true sampling distribution of statistic T_n, enabling us to construct the permutation test. □

In general, when the randomization hypothesis does not hold, one can not conclude from the convergence of the permutation distribution, as in the previous remark, that the same is true for the test statistic. As shown in [8] and Chung and Romano [5], it is helpful to consider the studentization of the test statistic in this context. Define the studentization factor $\hat{\tau}_n$ as

$$\hat{\tau}_n = \hat{\tau}_n(Y^n, Z^n) := \frac{\sqrt{\frac{1}{n}\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2}\sqrt{\frac{1}{n}\sum_{i=j}^n \left(I_{Z,j} - \bar{I}_Z\right)^2}}.$$

and slightly rearrange the test statistic T_n a little to get

$$T_n(Y^n, Z^n) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)\left(I_{Z,i} - \bar{I}_Z\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2}\sqrt{\frac{1}{n}\sum_{j=1}^n \left(I_{Z,j} - \bar{I}_Z\right)^2}}.$$

We replace the factor $\sqrt{n-1}$ by $\sqrt{n}$ in (2.11) for easier notation, making no difference in asymptotic terms. Divide T_n by $\hat{\tau}_n$ to get a new statistic

$$\hat{T}_n(Y^n, Z^n) := \frac{T_n}{\hat{\tau}_n} = \frac{\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)\left(I_{Z,i} - \bar{I}_Z\right)}{\sqrt{\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2}}. \tag{2.14}$$

Denote by $p_{11} = p_{11}(n)$, $p_Y = p_Y(n)$ and $p_Z = p_Z(n)$ the probabilities $P(Y > u_{\sqrt{m_n}}, Z > v_{\sqrt{m_n}})$, $P(Y > u_{\sqrt{m_n}})$ and $P(Z > v_{\sqrt{m_n}})$, respectively. Note that by (2.4), it holds that $p_Y \sim p_Z \sim 1/\sqrt{m_n}$ as $n \to \infty$. The following proposition describes the asymptotic behaviour of the statistic $\hat{T}_n$.

Proposition 2.1.5. Let (Y_i, Z_i), $i \in \mathbb{N}$, be a sequence of iid bivariate random vectors for which (2.3) and (2.4) hold. Then

$$\frac{\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)\left(I_{Z,i} - \bar{I}_Z\right) - n(p_{11} - p_Y p_Z)}{\sqrt{\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2}} \xrightarrow{d} N(0,1),\ n \to \infty. \tag{2.15}$$

We show next that the corresponding permutation distribution $\hat{R}_n^{\tau}$ of $\hat{T}_n$ converges to the standard normal distribution in probability. To do that we adapt the idea of the proof of Theorem 2.3. in [8].

From (2.13) follows $\hat{R}_n(t) \to \Phi(t)$ in probability as $n \to \infty$ and for all $t \in \mathbb{R}$. The necessity part of Theorem 1.1.2 implies that for G_n and G_n' independent and uniformly distributed over $\mathbf{G}_n$ it holds that

$$\left(T_n(G_n X^n), T_n(G_n' X^n)\right) \xrightarrow{d} N(0, I_2),\ n \to \infty, \tag{2.16}$$

where I_2 is a 2×2 unit matrix. The following lemma turns out to be useful in the sequel.

Lemma 2.1.6. With the same assumptions as in Proposition 2.1.5 for G_n uniformly distributed over $\mathbf{G}_n$

$$\hat{\tau}_n(G_n X^n) \xrightarrow{P} 1,\ n \to \infty.$$

By Lemma 2.1.6 and Slutsky's theorem, from (2.16) we obtain

$$\left(T_n(G_n X^n)/\hat{\tau}_n(G_n X^n), T_n(G_n' X^n)/\hat{\tau}_n(G_n' X^n)\right) \xrightarrow{d} N(0, I_2),\ n \to \infty.$$

Now we can apply the sufficiency part of Theorem 1.1.2 to get

$$\hat{R}_n^{\tau}(t) := \frac{1}{n!} \sum_{\pi \in \mathbf{G}_n} I_{\{T_n(Y^n, Z_\pi^n)/\hat{\tau}_n(Y^n, Z_\pi^n) \le t\}} \xrightarrow{P} \Phi(t),\ n \to \infty,\ t \in \mathbb{R}. \tag{2.17}$$

That is, the permutation distribution of $\hat{T}_n$ converges towards the standard normal distribution. By Proposition 2.1.5, the centred statistic $\hat{T}_n$ has the same limiting law. This allows one to detect positive tail dependence between Y and Z. To describe the test, we assume that F_Y and F_Z are continuous and known. In particular, we transform the marginal distributions of Y and Z to the unit Pareto distribution, i.e. we suppose that

$$P(Y > u) = P(Z > u) = \frac{1}{u},\ u \ge 1, \tag{2.18}$$

and we suppose that there exists $\eta \in (0, 1]$ such that

$$P(Y > u, Z > u) = L(u) P(Z > u)^{1/\eta},\ u \ge 1, \tag{2.19}$$

where L is a slowly varying function at infinity. The assumption (2.19) is essentially the model of joint threshold upcrossings given in (2.2) and introduced in [26]. Note that the independence corresponds to the case $\eta = 1/2$ and $L(x) = 1$. On the other hand, positive tail dependence is clearly present if $\liminf_{x \to \infty} L(x)$ is larger than 1, even with $\eta = 1/2$.

Observe that the centring term on the left hand side of (2.15) is asymptotically equivalent to

$$c_n := \sqrt{n} \frac{p_{11} - p_Y^2}{\sqrt{p_{11}}} \tag{2.20}$$

when p_{11} converges to zero more slowly than $m_n^{-3/2}$, which is the consequence of asymptotic equivalence between $\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2$ and np_{11} (see proof of Proposition 2.1.5 for details). Note that stated asymptotic behavior of c_n always holds for $\eta = 1/2$, because of (2.19). Thus, in that case, the conclusion of Proposition 2.1.5 is equivalent to

$$\hat{T}_n - c_n \xrightarrow{d} N(0, 1),\ n \to \infty.$$

Suppose there exists $d > 1/2$ such that

$$\limsup_{u\to\infty} u^d \frac{L(u)-1}{\sqrt{L(u)}} < \infty. \tag{2.21}$$

Note that in the case of unit Pareto random variables, one can take $u_n = n$ in (2.4). Because of assumptions (2.18) and (2.19), for $\eta = 1/2$ we have

$$c_n = \frac{\sqrt{n}}{\sqrt{m_n}} \frac{L(\sqrt{m_n})-1}{\sqrt{L(\sqrt{m_n})}},\ n \in \mathbb{N}. \tag{2.22}$$

It turns out, see Remark 2.1.7, that under (2.21), for $\eta = 1/2$ and (m_n) satisfying (2.3) and $n^{1/(d+1)} = o(m_n)$ we have

$$\limsup_{n\to\infty} c_n \leq 0.$$

The theoretical considerations above allow one to detect whether $\chi > 0$, or more generally whether $\eta > 1/2$. To ensure an asymptotically correct level of the test, under the null hypothesis one can include the case $\eta < 1/2$ (i.e. negative tail dependence), as well as the case $\eta = 1/2$ but under an additional assumption on L in (2.21). Finally, one rejects the null hypothesis when the value of the statistic $\hat{T}_n$ exceeds the corresponding quantile of the permutation distribution $\hat{R}_n^\tau(t)$ and randomizes when they are equal (see the definition of test ϕ in Section 1.1). Denote by $\hat{r}_n(1-\alpha)$ the $1-\alpha$ quantile of $\hat{R}_n^\tau(t)$. From (2.17), we conclude (see Section 1.1 for details) that

$$\hat{r}_n(1-\alpha) \xrightarrow{P} \Phi^{-1}(1-\alpha), \tag{2.23}$$

where $\Phi^{-1}(1-\alpha)$ is a $1-\alpha$ quantile of the standard normal distribution Φ. In particular, the convergence of $1-\alpha$ quantiles of the statistic $\hat{T}_n$ to $\Phi^{-1}(1-\alpha)$ holds if $\eta = 1/2$ and the condition in (2.21) is satisfied.

Remark 2.1.7. Note, for $\eta = 1/2$ and $d > 1/2$ we have

$$c_n = \frac{\sqrt{n}}{\sqrt{m_n}} \frac{1}{\sqrt{m_n}^d} \sqrt{m_n}^d \frac{L(\sqrt{m_n})-1}{\sqrt{L(\sqrt{m_n})}},\ n \in \mathbb{N}.$$

Suppose (2.21) holds. Then $\limsup_{n\to\infty} c_n \leq 0$ if

$$\left(\frac{n}{m_n^{d+1}}\right)^{1/2} \to 0,\ n \to \infty,$$

i.e. if $n^{1/(d+1)} = o(m_n)$. □

Remark 2.1.8. As we already noted,the test statistic in Proposition 2.1.5 has the following asymptotically equivalent form

$$\frac{T_n - d_n}{\hat{\tau}_n} \approx \frac{T_n}{\hat{\tau}_n} - c_n$$

where $d_n = n(p_{11} - p_Y p_Z)$. Therefore, for $z_\alpha \in \mathbb{R}$ the upper α–quantile of the standard normal distribution we have

$$\limsup_n P\left(\frac{T_n}{\hat{\tau}_n} > z_\alpha\right) = \limsup_n P\left(\frac{T_n}{\hat{\tau}_n} - c_n > z_\alpha - c_n\right)$$

If $\limsup_n c_n \leq 0$, the right hand side is bounded by $\lim_n P(T_n/\hat{\tau}_n - c_n \geq z_\alpha)$ which by Proposition 2 equals $1 - \Phi(z_\alpha) = \alpha$. Similar argument shows that for $c_n \to \infty$, the power of the test grows to 1.

If we take specifically that the joint behaviour of Y and Z is described by the Morgenstern copula (see Section 2.2.4 for definition), then $\eta = 1/2$ and $L = 1 + \tilde{\alpha}$, $\tilde{\alpha} \in [-1, 1]$ (see Table 1 in [18]). In that case $p_{11} = (1 + \tilde{\alpha})m_n^{-1}$ and

$$c_n = \frac{\sqrt{n}}{\sqrt{m_n}} \frac{\tilde{\alpha}}{\sqrt{1 + \tilde{\alpha}}}.$$

Therefore, for $\tilde{\alpha} \downarrow 0$ the power of the test converges to the level of the test α. See also the simulation study and the graph on Figure 2.3. □

Remark 2.1.9. Our assumption on the known marginal distributions of Y and Z allows us to determine the threshold sequences $(u_{\sqrt{m_n}})$ and $(v_{\sqrt{m_n}})$. It is hard to justify such an assumption in practical applications. Instead, one would typically use high empirical quantiles. However, the simulation study presented in Section 2.2 indicates that the size and power of the test proposed here are not overly sensitive to the choice of the thresholds, possibly making such a test applicable with a range of possible thresholds. Of course, it would be interesting to incorporate the estimate of marginal distributions into the testing procedure we propose, but that is left for some future research. □

2.2. Simulations and application

In this section, we investigate the behaviour of the test statistic $\hat{T}_n$ defined in (2.14) in a simulation study. We denote the simulated data by $(Y_1, Z_1), \ldots, (Y_n, Z_n)$, $n \in \mathbb{N}$. The threshold levels used to calculate the value of the statistic $\hat{T}_n$ are determined through empirical upper quantiles of the given data. After the calculation of the value of $\hat{T}_n$, the following permuted values of $\hat{T}_n$ were calculated:

$$\hat{T}_n(Y^n, Z^n_\pi) = \frac{\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)\left(I_{Z,\pi(i)} - \bar{I}_Z\right)}{\sqrt{\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,\pi(i)} - \bar{I}_Z\right)^2}}.$$

For the level of the test α the number $k = n - \lfloor n\alpha \rfloor$ is calculated, and then the value of the statistic is compared with the k-th largest value of the calculated permuted values of the statistic $\hat{T}_n$. The permutation test rejects the null hypothesis when the value of the statistic is larger than the k-th largest value of the permuted values of the statistic $\hat{T}_n$ and randomizes when they are equal. We repeated this procedure to get simulated rejection probabilities, which we refer to as empirical rejection probabilities. We also calculated the value of the test statistic $\hat{T}_n$ with exact threshold levels for known marginal distributions and then repeated the whole procedure. We refer to rejection probabilities obtained in such a manner as theoretical rejection probabilities.

2.2.1. Completely independent samples

Consider two iid sequences (Y_i) and (Z_i) from the unit Pareto distribution. Although there is no universal method to determine the threshold level for a given set of observations, one can find some guidance in the assumptions of Theorem 2.1.2. Recall that we assumed that $\sqrt{m_n} P(Y_i > u_{\sqrt{m_n}}) \to 1$ where m_n is an intermediary sequence such that $m_n = O(n^{2/3-\tau})$, $n \to \infty$, $\tau > 0$. Since the marginal distribution is known here, one can use it to select an appropriate threshold (observe that the sequence $u_n = n$ satisfies condition (2.4)). However, from a practical perspective, it seems more interesting to check how our test statistic and corresponding permutation test behave when the threshold level varies, i.e. when we look at different quantiles of data (like 10% or 5% of the most extreme data points).

In Figure 2.1, we present simulation results of the permutation test involving statistic

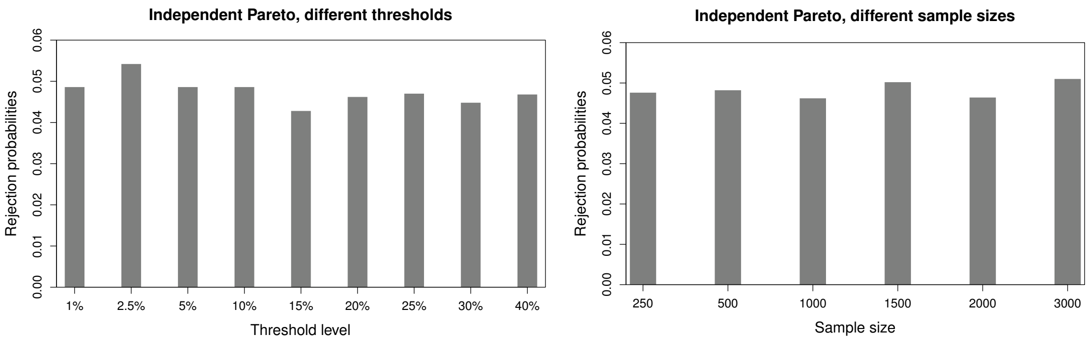


Figure 2.1: On the left: empirical rejection probabilities (y-axis) depending on the threshold level (expressed as the upper quantile) (x-axis) for the sample size $n = 1000$. On the right: empirical rejection probabilities of the permutation test for various sample sizes (x-axis) with the threshold level fixed at the upper 20% quantile. The level of the test is set at 5%, while the number of permutations and repeats equals 5000 in both graphs.

$\hat{T}_n$ for different thresholds and different sample sizes. The rejection probabilities seem quite stable for different threshold levels. However, for very high thresholds (i.e. upper quantiles), one can expect only a very small number of joint exceedances of processes Y and Z. In such cases, the randomization procedure plays an important role in the testing procedure.

Table 2.1 presents empirical rejection probabilities for a threshold level 20% (upper quantile) and a different number of permutations and repeats with the level of the test set at 5%. Similar results were obtained with the threshold level set at 10% and 5% or the level of the test set at 10%. Overall, we have found that the test statistic $\hat{T}_n$ produces permutation tests that maintain the exactness property reasonably well under the null hypothesis of tail independence between iid processes Y and Z and for various threshold levels and sample sizes. Moreover, the results are virtually indistinguishable whether we use empirical or theoretical thresholds, cf. columns 3 and 4 in Table 2.1.

2.2.2. Dependent Pareto-type random variables

Let U_i, V_i and W_i be three independent iid sequences of $U(0,1)$ distributed random variables, $i = 1,2,\ldots$. Define three sequences of Pareto-distributed random variables $X_i^1 = 1/U_i$, $X_i^2 = 1/V_i$ and $X_i^3 = 1/W_i$. Finally let $Y_i = X_i^1 + aX_i^2$ and $Z_i = aX_i^2 + X_i^3$, with $a \geq 0$.

No. of sim.	No. of perm.	Emp. rp	Theo. rp
500	500	0.0560	0.0580
1000	1000	0.0410	0.0410
2000	1000	0.0600	0.0650
1000	2000	0.0450	0.0435
2000	2000	0.0495	0.0590
5000	2000	0.0440	0.0445
2000	5000	0.0514	0.0566
5000	5000	0.0462	0.0488
10000	10000	0.0539	0.0505

Table 2.1: Results of a simulation study in the iid Pareto case for the different number of simulations (column 1) and permutations (column 2). Empirical rejection probabilities for the different number of permutations and number of repeats are shown in column 3 in cases when the threshold is determined using empirical quantiles. Column 4 contains rejection probabilities in cases when the threshold is determined using the known marginal distribution. The test statistic is $\hat{T}_n$, the sample size is 1000, the threshold level is fixed at the upper 20% quantile, and the level of the test is 5%.

Due to the independence and regular variation as $z \to \infty$

$$\begin{aligned} P(Y > z) = P(Z > z) &= P(a \cdot X^2 + X^3 > z) \\ &\sim P(a \cdot X^2 > z) + P(X^3 > z) \sim \left(\frac{z}{a}\right)^{-1} + (z)^{-1}. \end{aligned}$$

Thus, the tails of Y and Z are both of the Pareto type and therefore regularly varying. Obviously, those two random variables are independent if and only if $a = 0$. To measure the level of their dependence, we use the upper tail dependence coefficient χ. For this model, we have

$$\chi = \lim_{z\to\infty} \frac{P(X^1 + a \cdot X^2 > z, a \cdot X^2 + X^3 > z)}{P(a \cdot X^2 + X^3 > z)}.$$

Since

$$P(X^1 + a \cdot X^2 > z, a \cdot X^2 + X^3 > z) \sim P(aX^2 > z) \sim \left(\frac{z}{a}\right)^{-1},$$

one obtains $\chi = a/(a+1)$ in this setting. For $a = 0$, Y and Z are independent and hence $\chi = 0$. In Figure 2.2 (left), simulation results of the permutation test involving the statistic $\hat{T}_n$ are shown. We present corresponding empirical rejection probabilities, with increasing values of the parameter χ (and a). The threshold level was set to 20%, but very similar results were obtained for other thresholds.

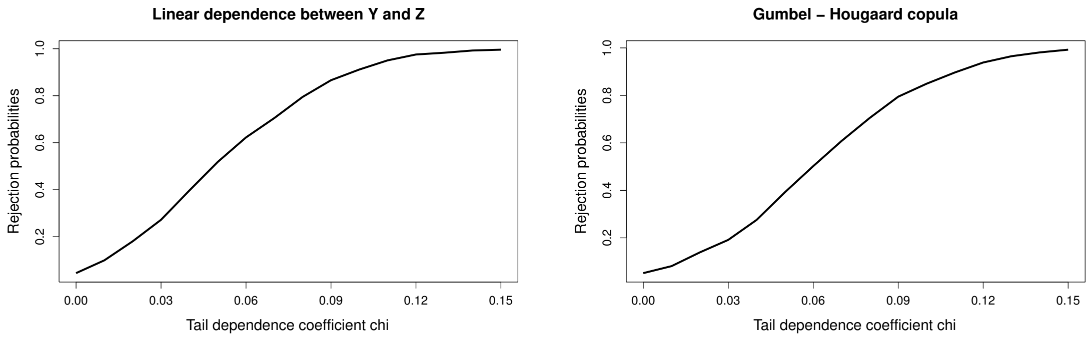


Figure 2.2: Empirical rejection probabilities (y-axis) for various values of the parameter χ (x-axis) for the sample size $n = 1000$ using an upper empirical quantile at level 20%. The test statistic is $\hat{T}_n$. On the left: data generated from the linear dependence model between Y and Z described in subsection 2.2.2. On the right: data generated from the Gumbel-Hougaard copula described in subsection 2.2.3. The level of the test is set at 5%, while the number of permutations and repeats is equal to 2000 in both simulations.

2.2.3. Tail dependent samples

We presented the Gumber-Hougaard copula in Example 1.4.4. Figure 2.2 (right) shows simulation results of the permutation test involving the statistic $\hat{T}_n$ and data generated from the Gumbel-Hougaard copula. The dependence of rejection probabilities on the corresponding tail dependence coefficient χ is shown. We present the results for a threshold level 20%, but similar results were obtained for other thresholds as well. We confirm that very similar overall results were attained when we generated data from some other copulas of a similar type, like the Joe copula (see Table 1 in [18]). As in subsection 2.2.1 the results remain similar whether empirical or theoretical thresholds are used.

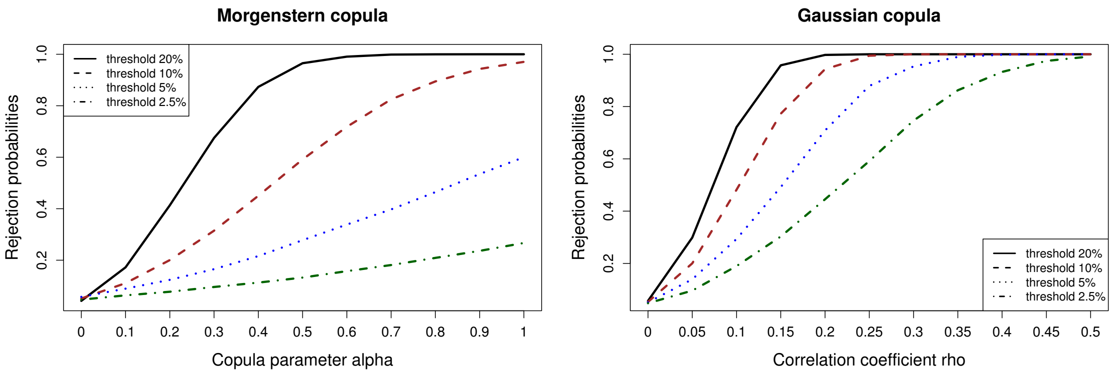


Figure 2.3: Empirical rejection probabilities (y-axis) for various values of the corresponding parameter of the copulas (x-axis): Morgenstern on the left and Gaussian on the right. The test statistic is $\hat{T}_n$. The level of the test is set at 5%, the sample size is $n = 2000$, while the number of permutations and repeats equal 4000. The threshold level varies from 20% through 10% and 5% down to 2.5%.

2.2.4. The Morgenstern and Gaussian copula

It is known that there exist bivariate models with the tail dependence index χ equal to 0 that exhibit considerable preasymptotic tail dependence. For such models, the tail dependence coefficient η determines the strength of dependence within the class of tail independent models. The often studied case of near independence corresponds to $\eta = 1/2$. As we have seen in Section 2.1, that case is more subtle in our procedure as well.

The Morgenstern copula is an example of a copula that falls in the near-independence class of models. It is given by

$$C(u,v) = uv\big[1 + \tilde{\alpha}(1-u)(1-v)\big],\ u,v \in [0,1],\ \tilde{\alpha} \in [-1,1].$$

Another example of a copula with a complex dependence structure is the Gaussian copula described in Example 1.4.4. Exact independence is attained for $\rho = 0$ for the Gaussian copula and for $\tilde{\alpha} = 0$ for the Morgenstern copula. Both copulas are tail independent (see [18]). For the Gaussian copula the associated tail dependence coefficient η is given by the relation $\eta = (1+\rho)/2$ (see Table 1 in [18]).

In Figure 2.3 we present the simulation results for the Morgenstern and the Gaussian copulas with positive dependencies. The threshold level varies from 20% of upper quantile data down to 2.5%. In contrast to the case $\chi > 0$, the choice of threshold levels has greater influence on the empirical rejection probabilities. Unsurprisingly, for fixed thresholds a larger data sample increases the power of the test.

2.2.5. Comparison of the permutation test with an asymptotic test

Convergence of the centred statistic $\hat{T}_n$ towards the standard normal distribution (see Proposition 2.1.5) allows for the construction of another test for tail dependence (apart from the permutation test). It rejects the null hypothesis if the value of the statistic $\hat{T}_n$ is larger than the corresponding quantile of normal distribution. In this section we present comparative results for the two tests, and compare them with a test based on the Hill estimator $\hat{\eta}_2$ of the tail coefficient η suggested by Draisma et al. in [9] (with the null hypothesis $\eta \leq 1/2$).

We first consider two independent iid sequences (Y_i) and (Z_i) from the unit Pareto distribution. In Table 2.2, we compare three tests for those independent samples. The table shows rejection probabilities for different thresholds. For the test based on $\hat{\eta}_2$, the parameter m indicates the number of upper order statistics used to calculate it, see [9]. As can be seen from Table 2.2, the permutation test controls Type 1 error significantly better than the other two tests.

	Thresholds				
Tests	2.5%	5%	10%	20%	30%
Permutation test	0.052	0.050	0.052	0.050	0.052
Asymptotic test	0.003	0.011	0.032	0.049	0.043
Parameter m	25	50	100	200	300
Test based on $\hat{\eta}_2$	0.024	0.039	0.062	0.078	0.102

Table 2.2: Empirical rejection probabilities for independent samples from the unit Pareto distribution and for different thresholds. For the test based on $\hat{\eta}_2$ different thresholds correspond to different values of m. The sample size is 1000, the level of the test is 0.05, and the number of permutations and repeats is 5000.

Next, we compare the performance of the tests on simulated tail dependent data. For that purpose, we use three different copulas and a fixed threshold level of 5%. We first present empirical rejection probabilities for the Gumbel-Hougaard copula with different values of the dependence parameter χ. The permutation test and the test based on $\hat{\eta}_2$ exhibit larger power than the test based on the asymptotic normality of the test statistic $\hat{T}_n$.

	Tail coefficient χ				
Gumbel-Hougaard copula	0	0.04	0.08	0.12	0.16
Permutation test	0.052	0.305	0.655	0.869	0.967
Asymptotic test	0.011	0.131	0.431	0.725	0.910
Test based on $\hat{\eta}_2$	0.039	0.304	0.639	0.848	0.941
	Correlation ρ				
Gaussian copula	0.0	0.1	0.2	0.3	0.4
Permutation test	0.048	0.184	0.484	0.779	0.948
Asymptotic test	0.051	0.062	0.263	0.585	0.872
Test based on $\hat{\eta}_2$	0.038	0.123	0.296	0.526	0.730
	Parameter $\tilde{\alpha}$				
Morgenstern copula	0	0.2	0.4	0.6	0.8
Permutation test	0.048	0.089	0.1428	0.221	0.292
Asymptotic test	0.009	0.022	0.043	0.078	0.125
Test based on $\hat{\eta}_2$	0.038	0.052	0.066	0.080	0.094

Table 2.3: Comparison of empirical rejection probabilities for simulated data from the Gumbel-Hougaard, Gaussian and Morgenstern copula. The level of the test is set at 0.05, and the threshold level of the permutation test is fixed at 5% ($m = 50$ for the test based on $\hat{\eta}_2$). The sample size is 1000 and the number of permutations and repeats is 5000.

Further, we compare the three tests on data generated from the Gaussian copula (see subsection 2.2.4) with different vales of the correlation parameter ρ, and the Morgenstern copula (see subsection 2.2.4) with different values of the parameter $\tilde{\alpha}$. The simulation results obtained are presented in Table 2.3. Observe that the test based on $\hat{\eta}_2$ in [9] was not designed for the purpose of testing dependence exhibited by the Morgenstern copula ($\eta = 1/2$ here). The simulation results obtained are presented in Table 2.3. It appears that the permutation test has better power than the other two in both of these two cases. Simulations with other thresholds confirm that.

2.2.6. Discussion of the simulation study

Overall, the simulation results (some of which are not presented) indicate that the permutation test performs better than the other two test, especially for small data sizes. The advantage over the test based on the asymptotic normality is consistent with general observations on permutation tests, see for instance Janssen and Pauls [23] or Chapter 15 in [27]. The simulation studies above show that permutation test has approximately correct test rejection probabilities at different thresholds and with different sample sizes when Y and Z are independent. They also appear to demonstrate the test's considerable power, especially for larger data sets. The test does not seem overly sensitive to the choice of thresholds except when data are very close to independence.

All the simulations and analysis were done in *R* using the publicly available packages `permute`, `gumbel`, `fMultivar` and `copula`.

2.2.7. Application to financial data

In this subsection we present an application of the permutation test to stock returns. It is known that in certain cases the sample correlation does not indicate significant dependence between two stocks (some assets) but the price movements become very correlated in periods of crisis and rising volatility. Therefore it seems desirable to have a more sensitive test of dependence based, for instance, on the tail behaviour of returns. Additionally, from the diversification perspective, it is useful to test for positive dependence between asset returns, as negatively dependent (or independent) assets may lower the riskiness of a portfolio of assets. The permutation test presented in this article may serve that purpose, if the model of joint threshold exceedances from (2.19) is assumed.

We consider two European stocks, namely Société Générale (the French bank) and Beiersdorf (the German personal care producer). The analysis period includes daily return data for stocks of both between 1 January 2017 and 31 December 2019. Data were taken from the Yahoo Finance platform. To perform tail analysis, we focus on negative returns of the two stocks in order to detect whether they are positively tail dependent. We use the permutation test with our test statistic $\hat{T}_n$ given in (2.14), and we compare the results of that test to the permutation test of independence based on sample correlation (see [8]). In Figure 2.4 are shown plots of daily returns of Société Générale and Beiersdorf stocks (left) and their negative returns only (right). Of course, we have to suppose iid daily returns in order to apply either test.

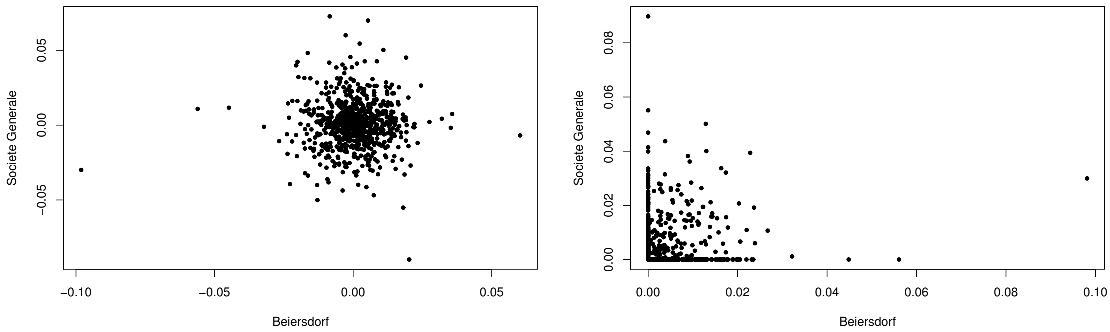


Figure 2.4: On the left side is shown a scatterplot of daily returns of Beiersdorf and Société Générale. On the right side are shown only negative daily returns of those two stocks (more precisely, we looked at max$\{0, -r_B\}$, where r_B, are returns on Beiersdorf stock, and performed the same for daily returns of Société Générale).

We first apply the permutation test of independence based on the sample correlation test statistic $\sqrt{n}\hat{\rho}$ from [8] with the level of the test set to $q = 5\%$ and $20,000$ permutations. In that case, we cannot reject the independence hypothesis as we obtain a test statistic value equal to 0.0231, while the q-th percentile of permuted values of the statistic is equal to 1.6414. The same conclusion is supported by the estimate of the p-value, which is equal to 0.4913. We also apply the studentized version of the correlation permutation test with the statistic S_n from [8] but the conclusion is the same (the estimated p-value is 0.4889, the value of the statistic is 0.5069, while the q-th percentile of permuted values of the the studentized statistic is equal to 44.746).

Next, we apply the permutation test for tail dependence based on the test statistic $\hat{T}_n$ with the level of the test equal to 5% and with $20,000$ permutations, having looked at negative values of daily returns for both stocks. We first choose a fixed threshold level of 5% of available data, and in that case the test statistic $\hat{T}_n$ is equal to 1.79, while the q-th percentile of test statistic permuted values is equal to 1.09. Therefore the test rejects independence of daily returns of those two stocks, with the estimated p-value of the test equal to 0.0017. Similarly, for the threshold level 2.5% the test statistic is equal to 1.5 and the q-th percentile of test statistic permuted values is equal to 1.1 (the estimated p-value is equal to 0.0009). So, in this case the null hypothesis for stocks of Beiersdorf and Société Générale is rejected as well. We report that very similar results were obtained when we used statistics T_n from (2.11) instead of $\hat{T}_n$.

The permutation independence test based on sample correlation and our test based on tail behaviour provide, in many cases, the same conclusion. For example, if we look at the same time interval, the same quantile $q = 5\%$ and the number of permutations set to $20,000$ but choose stocks of Beiersdorf and Siemens (the German industrial conglomerate), then both tests reject the corresponding null hypothesis. More precisely, in the case of the permutation test of independence based on sample correlation we obtain a test statistic value equal to 7.8155, while the q-th percentile of test statistic permuted values is equal to 1.6357. When we apply the permutation test for tail dependence based on the test statistic $\hat{T}_n$ with a threshold level set to 5% we obtain a test statistic value equal to 2.79, while the q-th percentile of test statistic permuted values is equal to 1.1.

2.3. Proofs

In this Section, we present proofs of the results stated above and some auxiliary results used in their proofs. We begin with a lemma that is an immediate consequence of assumptions (2.3) and (2.4).

Lemma 2.3.1. Suppose (2.3) and (2.4) hold. As $n \to \infty$,

$$\frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y)\to 0,$$

almost surely, where $p_Y = P(Y > u_{\sqrt{m_n}})$. An analogous statement holds for sequences (Z_n) and (v_n).

Proof of Lemma 2.3.1. Recall, $I_{Y,i} = I_{\{Y_i > u_{\sqrt{m_n}}\}}$. Let us take $\varepsilon > 0$ arbitrarily chosen. By the Markov inequality we have

$$P\Big(\frac{\sqrt{m_n}}{n}\Big|\sum_{i=1}^{n}(I_{Y,i}-p_Y)\Big| > \varepsilon\Big) \le \frac{m_n^2}{n^4\varepsilon^4}E\Big(\sum_{i=1}^{n}(I_{Y,i}-p_Y)\Big)^4. \tag{2.24}$$

When we take into account that for $i \neq j$ the random variables $I_{Y,i}$ and $I_{Y,j}$ are independent we see that $E(\sum_{i=1}^{n}(I_{Y,i}-p_Y))^4$ is equal to

$$\begin{aligned}
&\sum_{i=1}^{n}E\Big(I_{Y,i}-p_Y\Big)^4 + 4\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{j=1}^{n}E\Big(I_{Y,i}-p_Y\Big)^3E\Big(I_{Y,j}-p_Y\Big)\\
&+3\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{j=1}^{n}E\Big(I_{Y,i}-p_Y\Big)^2E\Big(I_{Y,i}-p_Y\Big)^2\\
&+6\sum_{\substack{i=1\\ i\neq j\neq k}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}E\Big(I_{Y,i}-p_Y\Big)^2E\Big(I_{Y,j}-p_Y\Big)E\Big(I_{Y,k}-p_Y\Big)\\
&+\sum_{\substack{i=1\\ i\neq j\neq k\neq l}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}E\Big(I_{Y,i}-p_Y\Big)E\Big(I_{Y,j}-p_Y\Big)E\Big(I_{Y,k}-p_Y\Big)E\Big(I_{Y,l}-p_Y\Big).
\end{aligned}$$

The second, fourth and fifth sum in the previous expression are all equal to zero because $E(I_{Y,i}-p_Y) = 0$, for all $i \in \{1,2,\dots,n\}$. Furthermore, as $I_{Y,i}$ and p_Y are bounded by 1 it follows $|I_{Y,i}-p_Y| \le 1$ almost surely. Therefore

$$\sum_{i=1}^{n}E\Big(I_{Y,i}-p_Y\Big)^4 \le n.$$

As $E(I_{Y,i}-p_Y)^2 = p_Y-(p_Y)^2$ we have

$$3\sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n} E\Big(I_{Y,i}-p_Y\Big)^2 E\Big(I_{Y,j}-p_Y\Big)^2 \leq 3n^2(p_Y)^2(1-2p_Y+(p_Y)^2).$$

Going back to (2.24) and noting that $p_Y \sim 1/\sqrt{m_n}$, we see that its right-hand side is then bounded (up to a constant) by the following asymptotically dominant term

$$\frac{m_n^2}{n^4\varepsilon^4}n^2(p_Y)^2 \sim \frac{m_n^2}{n^2\varepsilon^4}\frac{1}{m_n} = \frac{1}{\varepsilon^4}\frac{m_n}{n}\frac{1}{n}.$$

Because of assumption (2.3), in (2.24) we have the inequality

$$P\Big(\frac{\sqrt{m_n}}{n}\Big|\sum_{i=1}^{n}(I_{Y,i}-p_Y)\Big| > \varepsilon\Big) \leq \frac{C}{n^{1+\tau}\varepsilon^4},$$

for n large enough, where $\tau > 0$ and $C > 0$ are constant. As the last inequality holds for all $\varepsilon > 0$ we can use the Borel-Cantelli Lemma to conclude

$$\frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y) \to 0$$

almost surely for $n\to\infty$, as we have claimed. ∎

Proof of Theorem 2.1.2. Following the notation used in Section 1.2 we define two triangular arrays of random variables $a_n(i)$ and $b_n(i)$, $i\in\{1,\dots,n\}$, as

$$a_n(i) = \sqrt{\frac{m_n}{n}}I_{Y,i} \quad \text{and} \quad b_n(i) = \sqrt{\frac{m_n}{n}}I_{Z,i}.$$

Next, we define

$$\bar{a}_n := \frac{1}{n}\sum_{i=1}^{n}a_n(i) = \sqrt{\frac{m_n}{n}}\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}$$

and

$$\bar{b}_n := \frac{1}{n}\sum_{i=1}^{n}b_n(i) = \sqrt{\frac{m_n}{n}}\frac{1}{n}\sum_{i=1}^{n}I_{Z,i}.$$

By Theorem 1.2.1 we need to verify

$$\lim_{n\to\infty} n\frac{\max_{1\leq i\leq n}(a_n(i)-\bar{a}_n)^2}{\sum_{i=1}^{n}(a_n(i)-\bar{a}_n)^2}\frac{\max_{1\leq i\leq n}(b_n(i)-\bar{b}_n)^2}{\sum_{i=1}^{n}(b_n(i)-\bar{b}_n)^2} = 0. \tag{2.25}$$

As $a_n(i)$ and $b_n(i)$ are random variables, we need almost sure convergence in the above expression to hold. For simplicity, we focus separately on the numerator and the denominator of the expression in (2.25).

Taking into account previous notation, the numerator in (2.25) is equal to

$$n \max_{1\le i\le n}\Big(\sqrt{\frac{m_n}{n}}I_{Y,i}-\sqrt{\frac{m_n}{n}}\frac{1}{n}\sum_{j=1}^{n}I_{Y,j}\Big)^2\max_{1\le i\le n}\Big(\sqrt{\frac{m_n}{n}}I_{Z,i}-\sqrt{\frac{m_n}{n}}\frac{1}{n}\sum_{j=1}^{n}I_{Z,j}\Big)^2.$$

After a short calculation, we see that we can write it in the form

$$\frac{m_n^2}{n}\max_{1\le i\le n}\Big(I_{Y,i}-\bar{I}_Y\Big)^2\max_{1\le i\le n}\Big(I_{Z,i}-\bar{I}_Z\Big)^2.$$

Both maxima in the numerator are almost surely bounded by 1 as it is true for each $I_{Y,i}$ and $I_{Z,i}$. Therefore, the numerator is almost surely bounded by m_n^2/n. For reasons that will soon become clear, we will bring the factor m_n^2/n to the denominator. Then the expression under the limit in (2.25) is almost surely bounded by

$$\frac{1}{\frac{n}{m_n^2}\sum_{i=1}^{n}(a_n(i)-\bar{a}_n)^2\sum_{i=1}^{n}(b_n(i)-\bar{b}_n)^2}. \tag{2.26}$$

Therefore, we are left with the conclusion that this expression will almost surely converge to zero if the denominator tends to $+\infty$ almost surely. Both sums in the denominator of (2.26) can be treated analogously, so we focus on the first sum. We have:

$$\begin{aligned}\sum_{i=1}^{n}(a_n(i)-\bar{a}_n)^2 &= \sum_{i=1}^{n}\Big(\sqrt{\frac{m_n}{n}}I_{Y,i}-\frac{1}{n}\sum_{j=1}^{n}\sqrt{\frac{m_n}{n}}I_{Y,j}\Big)^2\\ &= \frac{m_n}{n}\sum_{i=1}^{n}I_{Y,i}^2-2\frac{m_n}{n}\sum_{i=1}^{n}I_{Y,i}\frac{1}{n}\sum_{j=1}^{n}I_{Y,j}+\frac{m_n}{n}\sum_{i=1}^{n}\Big(\frac{1}{n}\sum_{j=1}^{n}I_{Y,j}\Big)^2\\ &= \frac{m_n}{n}\sum_{i=1}^{n}I_{Y,i}-m_n\Big(\frac{1}{n}\sum_{j=1}^{n}I_{Y,j}\Big)^2 = m_n\bar{I}_Y(1-\bar{I}_Y)\end{aligned}$$

and analogously

$$\sum_{i=1}^{n}(b_n(i)-\bar{b}_n)^2 = m_n\bar{I}_Z(1-\bar{I}_Z).$$

So, the denominator in (2.26) can be written as

$$\frac{\sqrt{n}}{\sqrt{m_n}}\sqrt{m_n}\,\bar{I}_Y(1-\bar{I}_Y)\frac{\sqrt{n}}{\sqrt{m_n}}\sqrt{m_n}\,\bar{I}_Z(1-\bar{I}_Z).$$

We will analyze the asymptotic behaviour of $\sqrt{m_n}\bar{I}_Y(1-\bar{I}_Y)$ and then the conclusion for the behaviour of $\sqrt{m_n}\bar{I}_Z(1-\bar{I}_Z)$ follows by analogy. Consider first the term $\sqrt{m_n}\bar{I}_Y$. We have

$$\sqrt{m_n}\bar{I}_Y = \frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y+p_Y) = \frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y)+\sqrt{m_n}p_Y,$$

We know, by the assumption (2.4), that $\sqrt{m_n} p_Y \to 1$, for $n \to \infty$. By using Lemma 2.3.1 we conclude that $\sqrt{m_n}\bar{I}_Y$ converges to 1 almost surely. Therefore $\bar{I}_Y \to 0$ almost surely and the same is then true for $\sqrt{m_n}\bar{I}_Y\bar{I}_Y$. So, $\sqrt{m_n}\bar{I}_Y(1-\bar{I}_Y) \to 1$ almost surely, $n \to \infty$.

We conclude that both $\sqrt{m_n}\bar{I}_Y(1-\bar{I}_Y)$ and $\sqrt{m_n}\bar{I}_Z(1-\bar{I}_Z)$ almost surely converge to 1. Then the whole expression in (2.26) converges to zero because of the term $\sqrt{n}/\sqrt{m_n}$ in the denominator, which tends to infinity (recall (2.3)). Then (2.9) follows by Theorem 1.2.1 (compare relations (2.8) and (1.16)). ■

To prove (2.6) and (2.7) let S_n be the sum $\sum_{i=1}^n a_n(i) b_n(G_n(i))$ and note it is a sum as in (1.9) from Section 1.2 (clearly, dependent on ω). On the right-hand side of (2.5) we recognize the expectation of S_n, relative to the probability measure P_{G_n}. From (1.13) we conclude that $E(S_n^{G_n} \mid X^n)$ is almost surely equal to

$$\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n a_n(i)b_n(j) = \frac{1}{n}\frac{m_n}{n}\sum_{i=1}^n\sum_{j=1}^n I_{Y,i}I_{Z,j}$$

and so (2.6) follows. To prove (2.7) first observe that by relation (1.15) we have

$$\begin{aligned} d_n(i,j) &= \frac{m_n}{n} I_{Y,i}I_{Z,j} - \frac{1}{n}\frac{m_n}{n}\sum_{k=1}^n I_{Y,k}I_{Z,j} \\ &- \frac{1}{n}\frac{m_n}{n}\sum_{l=1}^n I_{Y,i}I_{Z,l} + \frac{1}{n^2}\frac{m_n}{n}\sum_{k=1}^n\sum_{l=1}^n I_{Y,k}I_{Z,l} \\ &= \frac{m_n}{n} I_{Z,j}\left(I_{Y,i} - \frac{1}{n}\sum_{k=1}^n I_{Y,k}\right) - \frac{m_n}{n}\frac{1}{n}\sum_{l=1}^n I_{Z,l}\left(I_{Y,i} - \frac{1}{n}\sum_{k=1}^n I_{Y,k}\right) \\ &= \frac{m_n}{n}\left(I_{Y,i} - \bar{I}_Y\right)\left(I_{Z,j} - \bar{I}_Z\right). \end{aligned}$$

Then from (1.14) follows (2.7).

Before we give the proof of Proposition 2.1.5 we state and prove two lemmas that will be needed for it.

Lemma 2.3.2. With the same assumptions as in Proposition 2.1.5

$$\frac{\frac{1}{n}\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2}{E(I_{Y,1} - p_Y)^2 (I_{Z,1} - p_Z)^2} \xrightarrow{P} 1,\ n \to \infty.$$

Proof. We will show that

$$\left| \frac{\frac{1}{n}\sum_{i=1}^n \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2}{E(I_{Y,1} - p_Y)^2 (I_{Z,1} - p_Z)^2} - 1 \right| \xrightarrow{P} 0, \quad \text{for } n \to \infty. \tag{2.27}$$

Let us first calculate the expression in the denominator. We have

$$\begin{aligned}
E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2 &= E(I_{Y,1}-2I_{Y,1}p_Y+(p_Y)^2)(I_{Z,1}-2I_{Z,1}p_Z+(p_Z)^2) \\
&= E\Big(I_{Y,1}I_{Z,1}-2I_{Y,1}I_{Z,1}p_Z+I_{Y,1}(p_Z)^2 \\
&\quad -2I_{Y,1}I_{Z,1}p_Y+4I_{Y,1}I_{Z,1}p_Yp_Z-2I_{Y,1}p_Y(p_Z)^2 \\
&\quad +I_{Z,1}(p_Y)^2-2I_{Z,1}(p_Y)^2p_Z+(p_Y)^2(p_Z)^2\Big) \\
&= p_{11}-2p_{11}p_Z-2p_{11}p_Y+4p_{11}p_Yp_Z \\
&\quad +p_Y(p_Z)^2+p_Z(p_Y)^2-3(p_Y)^2(p_Z)^2 \\
&\sim p_{11}-4p_{11}p_Y+4p_{11}(p_Y)^2+2(p_Y)^3-3(p_Y)^4. \qquad (2.28)
\end{aligned}$$

In the last row of the above expression we used the fact that p_Y and p_Z have the same asymptotic behaviour (see (2.4)). Furthermore, from (2.4) follows $(p_Y)^3 \sim m_n^{-3/2}$. In a similar vein, we see that the numerator of the fraction in (2.27) is equal to

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left(I_{Y,i}-2I_{Y,i}\bar{I}_Y+\bar{I}_Y^2\right)\left(I_{Z,i}-2I_{Z,i}\bar{I}_Z+\bar{I}_Z^2\right) \\
&\quad = \frac{1}{n}\sum_{i=1}^{n}\left(I_{Y,i}I_{Z,i}-2I_{Y,i}I_{Z,i}\bar{I}_Z+I_{Y,i}\bar{I}_Z^2-2I_{Y,i}I_{Z,i}\bar{I}_Y+4I_{Y,i}I_{Z,i}\bar{I}_Y\bar{I}_Z\right. \\
&\quad\left. -2I_{Y,i}\bar{I}_Y\bar{I}_Z^2+I_{Z,i}\bar{I}_Y^2-2I_{Z,i}\bar{I}_Y^2\bar{I}_Z+\bar{I}_Y^2\bar{I}_Z^2\right) \\
&\quad = \frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i}-2\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i}+\bar{I}_Y\bar{I}_Z^2-2\bar{I}_Y\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i} \\
&\quad +4\bar{I}_Y\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i}+\bar{I}_Y^2\bar{I}_Z-3\bar{I}_Y^2\bar{I}_Z^2
\end{aligned}$$

So, in the numerator under the absolute value signs, we get the expression

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i}-2\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i}-2\bar{I}_Y\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i}+4\bar{I}_Y\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,i} \\
&+\bar{I}_Y\bar{I}_Z^2+\bar{I}_Y^2\bar{I}_Z-3\bar{I}_Y^2\bar{I}_Z^2-p_{11}+2p_{11}p_Z+2p_{11}p_Y-4p_{11}p_Yp_Z \\
&-p_Y(p_Z)^2-p_Z(p_Y)^2+3(p_Y)^2(p_Z)^2, \qquad (2.29)
\end{aligned}$$

while in the denominator we have the expression from (2.28). If p_{11} converges to zero more slowly than $m_n^{-3/2}$, $n \to \infty$, then p_{11} is an asymptotically dominant term in (2.28). Otherwise the asymptotically dominant term in (2.28) is p_Y^3. Proof of the lemma is slightly different in those two cases and we present them separately for clarity.

Suppose first that $p_{11}m_n^{3/2} \to \infty$, i.e. p_{11} converges to zero more slowly than $m_n^{-3/2}$, $n \to \infty$. Although the expression in (2.29) looks worrisome, there are some notions that will prove helpful. First of all, let us recall from the proof of Theorem 2.1.2 that both $\sqrt{m_n}\bar{I}_Y$ and $\sqrt{m_n}\bar{I}_Z$ almost surely converge to 1. So, we know $\bar{I}_Y, \bar{I}_Z \sim 1/\sqrt{m_n}$ and therefore $\bar{I}_Y\bar{I}_Z^2, \bar{I}_Y^2\bar{I}_Z \sim 1/m_n^{3/2}$. As we work under the assumption $p_{11}m_n^{3/2} \to \infty$, $n \to \infty$, we can conclude that the terms

$$\bar{I}_Y\bar{I}_Z^2, \bar{I}_Y^2\bar{I}_Z, 3\bar{I}_Y^2\bar{I}_Z^2, 2p_{11}p_Z, 2p_{11}p_Y, 4p_{11}p_Yp_Z, p_Y(p_Z)^2, p_Z(p_Y)^2, 3(p_Y)^2(p_Z)^2$$

divided by the denominator (2.28) all converge to zero (even almost surely). Next, group together the terms $1/n\sum_{i=1}^n I_{Y,i}I_{Z,i}$ and p_{11} to get

$$\frac{1}{n}\sum_{i=1}^n (I_{Y,i}I_{Z,i} - p_{11}).$$

We want to show that the above sum converges to zero in probability when we divide it with the denominator (2.28).

For arbitrary $\varepsilon > 0$, by using Chebyshev's inequality, we get

$$P\Big(\frac{1}{np_{11}}\Big|\sum_{i=1}^n (I_{Y,i}I_{Z,i} - p_{11})\Big| > \varepsilon\Big) \leq \frac{1}{\varepsilon^2 n^2 p_{11}^2} E\Big(\sum_{i=1}^n (I_{Y,i}I_{Z,i} - p_{11})\Big)^2. \tag{2.30}$$

The expectation on the right-hand side in (2.30) is equal to

$$\sum_{i=1}^n E(I_{Y,i}I_{Z,i} - p_{11})^2 + \sum_{i=1}^n \sum_{\substack{j=1\\ i\neq j}}^n E(I_{Y,i}I_{Z,i} - p_{11})(I_{Y,j}I_{Z,j} - p_{11}).$$

The first sum above is equal to $n(p_{11} - p_{11}^2)$ and the second sum is equal to zero, as $I_{Y,i}I_{Z,i}$ and $I_{Y,j}I_{Z,j}$ are independent for $i \neq j$. So, we get

$$P\Big(\frac{1}{np_{11}}\Big|\sum_{i=1}^n (I_{Y,i}I_{Z,i} - p_{11})\Big| > \varepsilon\Big) \leq \frac{1}{\varepsilon^2}\frac{1}{np_{11}^2}(p_{11} - p_{11}^2) \sim \frac{1}{\varepsilon^2}\frac{1}{np_{11}} \to 0,$$

under the assumptions (2.3) and $p_{11}m_n^{3/2} \to \infty$, $n \to \infty$. Therefore, we get

$$\frac{\frac{1}{n}\sum_{i=1}^n (I_{Y,i}I_{Z,i} - p_{11})}{E(I_{Y,1} - p_Y)^2(I_{Z,1} - p_Z)^2} \xrightarrow{P} 0, n \to \infty.$$

We are now almost done as we can similarly see that

$$\bar{I}_Z\frac{1}{n}\sum_{i=1}^n I_{Y,i}I_{,Zi} = \bar{I}_Z\frac{1}{n}\sum_{i=1}^n (I_{Y,i}I_{Z,i} - p_{11}) + \bar{I}_Z p_{11}.$$

The second term, divided by the denominator (2.28), converges trivially to zero (almost surely) because $\bar{I}_Z \to 0$ almost surely, for $n \to \infty$. For the first term, we conclude from the above considerations that, after dividing it with the denominator (2.28), it also converges to zero almost surely. We use similar reasoning to conclude that the terms

$$2\bar{I}_Y \frac{1}{n}\sum_{i=1}^{n} I_{Y,i}I_{Z,i}, 4\bar{I}_Y\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n} I_{Y,i}I_{Z,i}$$

divided by the denominator (2.28) also converge to zero almost surely.

Suppose now that p_{11} converges to zero faster than $m_n^{-3/2}$ and so p_Y^3 is an asymptotically dominant term in (2.28). As before, group together the terms $1/n\sum_{i=1}^{n} I_{Y,i}I_{Z,i}$ and p_{11} and use the Chebyshev inequality to get

$$P\Big(\frac{1}{np_Y^3}\Big|\sum_{i=1}^{n}(I_{Y,i}I_{Z,i} - p_{11})\Big| > \varepsilon\Big) \leq \frac{1}{\varepsilon^2 n^2 p_Y^6} E\Big(\sum_{i=1}^{n}(I_{Y,i}I_{Z,i} - p_{11})\Big)^2. \tag{2.31}$$

Because of the independence of $I_{Y,i}I_{Z,i}$ and $I_{Y,j}I_{Z,j}$, for $i \neq j$, the sum on the right hand side of (2.31) is bounded by

$$\frac{1}{\varepsilon^2}\frac{1}{np_Y^6}(p_{11} - p_{11}^2) \sim \frac{1}{\varepsilon^2}\frac{m_n^{3/2}}{n} m_n^{3/2} p_{11}.$$

Because of the assumption (2.3) $m_n^{3/2}/n$ converges to zero, and because p_{11} converges to zero faster than $m_n^{-3/2}$ the same is true for $m_n^{3/2}p_{11}$. Therefore, the term

$$\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}I_{Z,i} - p_{11})$$

divided by the expression in (2.28) converges to zero in probability. By the same argument as before we conclude the same is true for the terms

$$2\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n} I_{Y,i}I_{,Zi},\ 2\bar{I}_Y\frac{1}{n}\sum_{i=1}^{n} I_{Y,i}I_{Z,i},\ 4\bar{I}_Y\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n} I_{Y,i}I_{Z,i}.$$

It is also straightforward to conclude that the terms $3\bar{I}_Y^2\bar{I}_Z^2$, $2p_{11}p_Z$, $2p_{11}p_Y$, $4p_{11}p_Y p_Z$ and $3(p_Y)^2(p_Z)^2$ divided by the denominator (2.28) all converge to zero. Therefore, we only need to show the same for $\bar{I}_Y\bar{I}_Z^2 - p_Y(p_Z)^2$ and $\bar{I}_Y^2\bar{I}_Z - p_Z(p_Y)^2$. To do that, first note that

$$\bar{I}_Y\bar{I}_Z^2 - p_Y(p_Z)^2 = (\bar{I}_Y - p_Y)\bar{I}_Z^2 + p_Y(\bar{I}_Z^2 - p_Z^2) = K_1 + K_2.$$

Divide K_1 by p_Y^3 to get

$$\frac{K_1}{p_Y^3} = \frac{\bar{I}_Z^2\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i} - p_Y)}{p_Y^3} = \frac{\bar{I}_Z^2}{p_Y^2}\frac{\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i} - p_Y)}{p_Y}.$$

As already noted $\bar{I}_Z^2/p_Y^2 \to 1$ almost surely, while by Lemma 2.3.1 we have

$$\frac{\frac{1}{n}\sum_{i=1}^n (I_{Y,i}-p_Y)}{p_Y} \sim \frac{\sqrt{m_n}}{n}\sum_{i=1}^n (I_{Y,i}-p_Y) \to 0, n\to\infty \text{ (a.s.)}$$

In a similar manner we first write

$$\frac{K_2}{p_Y^3} = \frac{p_Y(\bar{I}_Z^2 - p_Z^2)}{p_Y^3} \sim \frac{(\bar{I}_Z^2 - p_Z^2)}{p_Z^2} = \frac{(\bar{I}_Z - p_Z)}{p_Z}\frac{(\bar{I}_Z + p_Z)}{p_Z},$$

and then use Lemma 2.3.1 again to conclude that $K_2/p_Y^3 \to 0$, for $n\to\infty$.

By using the triangle inequality in both the cases we treated, we conclude that (2.27) is true and therefore the lemma. ■

Lemma 2.3.3. With the same assumptions as in Proposition 2.1.5

$$\frac{\sqrt{n}\big(\bar{I}_Y - p_Y\big)\big(\bar{I}_Z - p_Z\big)}{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2}} \xrightarrow{P} 0,\ n\to\infty. \tag{2.32}$$

Proof. In the proof of Lemma 2.3.2 we have already seen that

$$\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2} \sim \sqrt{p_{11}} \text{ or } \sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2} \sim \sqrt{p_Y^3},$$

depending on the asymptotic behaviour of p_{11}. Let us look more closely at the numerator in (2.32). We have

$$\begin{aligned}\sqrt{n}\big(\bar{I}_Y - p_Y\big)\big(\bar{I}_Z - p_Z\big) &= \sqrt{n}\frac{1}{n}\sum_{i=1}^n (I_{Y,i}-p_Y)\frac{1}{n}\sum_{i=1}^n (I_{Z,i}-p_Z)\\ &= \frac{1}{n^{3/2}}\sum_{i=1}^n\sum_{j=1}^n (I_{Y,i}-p_Y)(I_{Z,j}-p_Z).\end{aligned}$$

To make later analysis easier, let us first look at the case $i=j$ in the above expression. We will first show that

$$\frac{\frac{1}{n^{3/2}}\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z)}{\sqrt{p_{11}}} \xrightarrow{P} 0,\ n\to\infty \tag{2.33}$$

if p_{11} converges to zero more slowly than p_Y^3, i.e. if $p_{11}m_n^{3/2} \to \infty$, for $n\to\infty$. Let $\varepsilon > 0$ be arbitrarily chosen. By using the Markov inequality we get

$$\begin{aligned}P\Big(\frac{1}{n^{3/2}\sqrt{p_{11}}}\Big|\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z)\Big| > \varepsilon\Big)&\\ \le \frac{1}{\varepsilon^2 n^3 p_{11}} E\Big(\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z)\Big)^2.&\end{aligned} \tag{2.34}$$

Use the fact that (Y_i, Z_i) and (Y_k, Z_k) are independent for $i \neq k$. Then the expectation on the right-hand side of (2.34) can be estimated:

$$\sum_{i=1}^{n} E(I_{Y,i} - p_Y)^2 (I_{Z,i} - p_Z)^2 + \sum_{\substack{i=1 \\ i \neq k}}^{n} \sum_{k=1}^{n} E(I_{Y,i} - p_Y)(I_{Z,i} - p_Z) E(I_{Y,k} - p_Y)(I_{Z,k} - p_Z) \leq n + n^2 (p_{11} - p_Y p_Z)^2.$$

Going back to (2.34) we conclude that (2.33) is true. When p_{11} converges to zero faster than p_Y^3 we use the analogous estimates to conclude that

$$\frac{\frac{1}{n^{3/2}} \sum_{i=1}^{n} (I_{Y,i} - p_Y)(I_{Z,i} - p_Z)}{\sqrt{p_Y^3}} \xrightarrow{P} 0, \; n \to \infty.$$

We are left to show that

$$\frac{1}{\sqrt{p_{11}}} \frac{1}{n^{3/2}} \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} (I_{Y,i} - p_Y)(I_{Z,j} - p_Z) \xrightarrow{P} 0, \; n \to \infty, \tag{2.35}$$

when p_{11} converges to zero more slowly than p_Y^3 and to show the analogous statement with $\sqrt{p_{11}}$ replaced by $\sqrt{p_Y^3}$, when p_{11} converges to zero faster than p_Y^3. Use the Chebyshev inequality again, applied on the expression in (2.35). On its right-hand side we get the expression

$$\begin{aligned}
&\frac{1}{\varepsilon^2 n^3 p_{11}} \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} E(I_{Y,i} - p_Y)^2 (I_{Z,j} - p_Z)^2 \\
&\quad + \frac{1}{\varepsilon^2 n^3 p_{11}} \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \neq l, k \neq i}}^{n} \sum_{l=1}^{n} E(I_{Y,i} - p_Y)(I_{Z,j} - p_Z)(I_{Y,k} - p_Y)(I_{Z,l} - p_Z) \\
&\quad + \frac{1}{\varepsilon^2 n^3 p_{11}} \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \neq l, l \neq j}}^{n} \sum_{l=1}^{n} E(I_{Y,i} - p_Y)(I_{Z,j} - p_Z)(I_{Y,k} - p_Y)(I_{Z,l} - p_Z) \\
&\quad = J_1 + J_2 + J_3.
\end{aligned}$$

Note that we have omitted the cases in which $k = i$ or $l = j$ because those two cases boil down to an estimate similar to the estimate of J_1. For J_1 note that $(I_{Y,i} - p_Y)^2 \leq 1$ and $(I_{Z,j} - p_Z)^2 \leq 1$ almost surely so J_1 is bounded by

$$\frac{1}{\varepsilon^2 n^3 p_{11}} n^2 = \frac{1}{\varepsilon^2 n p_{11}}$$

and that expression converges to zero as we have already seen in the proof of Lemma 2.3.2. It is clear that the same conclusion holds when p_{11} converges to zero faster than p_Y^3, because of the assumption (2.3).

As J_2 and J_3 are symmetric, we will only analyze J_2. Let us first note that for the sums in which $l \neq i$ random variables $I_{Y,i}$ are independent of all other $I_{Z,j}, I_{Y,k}$ and $I_{Z,l}$ and therefore expectation under those sums is equal to zero. So, we can write

$$J_2 = \frac{1}{\varepsilon^2 n^3 p_{11}} \sum_{\substack{i=1\\ i\neq j}}^{n} \sum_{j=1}^{n} \sum_{\substack{k=1\\ k\neq i}}^{n} E(I_{Y,i} - p_Y)(I_{Z,i} - p_Z)(I_{Z,j} - p_Z)(I_{Y,k} - p_Y).$$

With similar reasoning we conclude that the sums in which $k \neq j$ also disappear, so we get

$$J_2 = \frac{1}{\varepsilon^2 n^3 p_{11}} \sum_{\substack{i=1\\ i\neq j}}^{n} \sum_{j=1}^{n} E(I_{Y,i} - p_Y)(I_{Z,i} - p_Z)(I_{Z,j} - p_Z)(I_{Y,j} - p_Y).$$

Because (Y_i, Z_i) and (Y_j, Z_j) are independent for $i \neq j$ we get

$$J_2 \leq \frac{1}{\varepsilon^2 n^3 p_{11}} n^2 (p_{11} - p_Y p_Z)^2 = \frac{1}{\varepsilon^2 n p_{11}} (p_{11} - p_Y p_Z)^2$$

and the expression on the right-hand side converges to zero, for $n \to \infty$. We conclude that (2.35) is valid. When p_{11} converges to zero faster than p_Y^3 in a similar manner we get the estimate for J_2 (when we replace $\sqrt{p_{11}}$ by $\sqrt{p_Y^3}$)

$$J_2 \leq \frac{1}{\varepsilon^2 n p_Y^3} (p_{11} - p_Y p_Z)^2,$$

and we again conclude that the expression on the right-hand side converges to zero, for $n \to \infty$, because of the assumption (2.3). Therefore, the statement of the Lemma holds. ■

Proof of Proposition 2.1.5. Multiply

$$\frac{\sum_{i=1}^{n} \left(I_{Y,i} - \bar{I}_Y\right)\left(I_{Z,i} - \bar{I}_Z\right) - n(p_{11} - p_Y p_Z)}{\sqrt{\sum_{i=1}^{n} \left(I_{Y,i} - \bar{I}_Y\right)^2 \left(I_{Z,i} - \bar{I}_Z\right)^2}}$$

by

$$\frac{\sqrt{E(I_{Y,1} - p_Y)^2 (I_{Z,1} - p_Z)^2}}{\sqrt{E(I_{Y,1} - p_Y)^2 (I_{Z,1} - p_Z)^2}}$$

to get

$$\frac{\frac{1}{\sqrt{n}}\left[\sum_{i=1}^n \left(I_{Y,i}-\bar{I}_Y\right)\left(I_{Z,i}-\bar{I}_Z\right)-n(p_{11}-p_Y p_Z)\right]}{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2}} \frac{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2}}{\frac{1}{\sqrt{n}}\sqrt{\sum_{i=1}^n \left(I_{Y,i}-\bar{I}_Y\right)^2\left(I_{Z,i}-\bar{I}_Z\right)^2}}.$$

Use Lemma 2.3.2 to conclude that we are left to show

$$\frac{\frac{1}{\sqrt{n}}\left[\sum_{i=1}^n \left(I_{Y,i}-\bar{I}_Y\right)\left(I_{Z,i}-\bar{I}_Z\right)-n(p_{11}-p_Y p_Z)\right]}{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2}} \xrightarrow{d} N(0,1).$$

We can rewrite the numerator slightly in the above fraction to get

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(I_{Y,i}-p_Y-(\bar{I}_Y-p_Y)\right)\left(I_{Z,i}-p_Z-(\bar{I}_Z-p_Z)\right)\\
&\quad=\frac{1}{\sqrt{n}}\Big[\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z)-(\bar{I}_Z-p_Z)\sum_{i=1}^n (I_{Y,i}-p_Y)\\
&\qquad-(\bar{I}_Y-p_Y)\sum_{i=1}^n (I_{Z,i}-p_Z)+n(\bar{I}_Y-p_Y)(\bar{I}_Z-p_Z)\Big]\\
&\quad=\frac{1}{\sqrt{n}}\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z)-\sqrt{n}(\bar{I}_Y-p_Y)(\bar{I}_Z-p_Z).
\end{aligned}$$

By Lemma 2.3.3 we conclude

$$\frac{\sqrt{n}(\bar{I}_Y-p_Y)(\bar{I}_Z-p_Z)}{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2}} \xrightarrow{P} 0,\ n\to\infty.$$

Therefore, we only need to prove

$$\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z)-\sqrt{n}(p_{11}-p_Y p_Z)}{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2}} \xrightarrow{d} N(0,1) \tag{2.36}$$

to finish the proof of this proposition. We will do that by using the Lindeberg-Feller central limit theorem. Let us define a triangular array (with independent rows)

$$U_{ni} = (I_{Y,i}-p_Y)(I_{Z,i}-p_Z),\ i=1,\dots,n.$$

We see that $|U_{ni}| \leq 1$ (a.s.) and $E(U_{ni}) = (p_{11}-p_Y p_Z)$. As in the proof of Lemma 2.3.2 we calculate

$$E(U_{ni})^2 = (p_{11}-2p_{11}p_Z-2p_{11}p_Y+4p_{11}p_Y p_Z+p_Y(p_Z)^2+p_Z(p_Y)^2-3(p_Y)^2(p_Z)^2)$$

and so

$$\begin{aligned}
\mathrm{Var}(U_{ni}) = (&p_{11}-p_{11}^2-2p_{11}p_Z-2p_{11}p_Y+6p_{11}p_Y p_Z\\
&+p_Y(p_Z)^2+p_Z(p_Y)^2-4(p_Y)^2(p_Z)^2)
\end{aligned}$$

Note that the asymptotically dominant term in $\mathrm{Var}(U_{ni})$ depends on the asymptotic behaviour of p_{11} (see the proof of Lemma 2.3.2). Suppose first that p_{11} converges to zero more slowly than p_Y^3. As $\mathrm{Var}(U_{ni})$ does not depend on i in that case we get

$$s_n^2 = \sum_{i=1}^n \mathrm{Var}(U_{ni}) = n\,\mathrm{Var}(U_{n1}) \sim np_{11}.$$

To apply the Lindeberg-Feller central limit theorem, we need to check the Lindeberg condition

$$L_n := \sum_{i=1}^n \frac{1}{s_n^2} E\left((U_{ni}-EU_{ni})^2 I_{\{|U_{ni}-EU_{ni}|>\varepsilon s_n\}}\right) \to 0,$$

for any $\varepsilon > 0$ and $n \to \infty$. Now we use the fact that $(U_{ni}-EU_{ni})^2 \le 1$ (a.s.) and the Chebyshev inequality to get

$$\begin{aligned}
L_n &\sim \frac{1}{np_{11}} \sum_{i=1}^n E\left((U_{ni}-EU_{ni})^2 I_{\{|U_{ni}-EU_{ni}|>\varepsilon s_n\}}\right) \\
&\le \frac{1}{np_{11}} \sum_{i=1}^n E\left(I_{\{|U_{ni}-EU_{ni}|>\varepsilon s_n\}}\right) \\
&= \frac{1}{np_{11}} nP(|U_{n1}-EU_{n1}| > \varepsilon s_n) \\
&\le \frac{1}{p_{11}} \frac{1}{\varepsilon^2} \frac{1}{s_n^2} \mathrm{Var}(U_{n1}) \\
&\sim \frac{1}{p_{11}} \frac{1}{\varepsilon^2} \frac{1}{np_{11}} p_{11} = \frac{1}{\varepsilon^2} \frac{1}{np_{11}} \to 0
\end{aligned}$$

for $n \to \infty$, under the assumption (2.3). The same conclusion follows if p_{11} converges to zero faster than p_Y^3 (in that case $s_n^2 \sim np_Y^3 \sim nm_n^{-3/2}$). So, the Lindeberg condition is satisfied in both cases and we can conclude

$$\frac{\sum_{i=1}^n U_{ni} - \sum_{i=1}^n EU_{ni}}{s_n} \xrightarrow{d} N(0,1).$$

We are now over with the proof, as $(\sum_{i=1}^n U_{ni} - \sum_{i=1}^n EU_{ni})/s_n$ is equal to

$$\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n (I_{Y,i}-p_Y)(I_{Z,i}-p_Z) - \sqrt{n}(p_{11}-p_Y^2)}{\sqrt{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2 - (p_{11}^2 - 2p_{11}p_Y p_Z + (p_Y)^2(p_Z)^2)}}.$$

We can overcome the slight difference in the denominator of the above expression by noting that

$$\frac{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2 - (p_{11}^2 - 2p_{11}p_Y p_Z + (p_Y)^2(p_Z)^2)}{E(I_{Y,1}-p_Y)^2(I_{Z,1}-p_Z)^2} \to 1, \ \ n \to \infty$$

because of the expansion in (2.28). Note that for n large enough we have

$$\frac{\sum_{i=1}^{n} EU_{ni}}{s_n} \sim \frac{\sqrt{n}(p_{11}-p_Y^2)}{\sqrt{p_{11}}} \sim \frac{\sqrt{n}(p_{11}-1/m_n)}{\sqrt{p_{11}}}$$

when p_{11} converges to zero more slowly than p_Y^3, and

$$\frac{\sum_{i=1}^{n} EU_{ni}}{s_n} \sim \frac{\sqrt{n}(p_{11}-p_Y^2)}{\sqrt{p_Y^3}} \sim \sqrt{n}\, m_n^{3/4}(p_{11}-1/m_n) \sim -\sqrt{n} m_n^{-1/4},$$

when p_{11} converges to zero faster than p_Y^3. ∎

Proof of Lemma 2.1.6. Multiply $\hat{\tau}_n(G_n X^n)^2$ by

$$\frac{E(I_{Y,1}-p_Y)^2 E(I_{Z,1}-p_Z)^2}{E(I_{Y,1}-p_Y)^2 E(I_{Z,1}-p_Z)^2}.$$

Then first observe

$$\frac{\frac{1}{n}\sum_{i=1}^{n}\left(I_{Y,i}-\bar{I}_Y\right)^2 \frac{1}{n}\sum_{j=1}^{n}\left(I_{Z,i}-\bar{I}_Z\right)^2}{E(I_{Y,1}-p_Y)^2 E(I_{Z,1}-p_Z)^2} \to 1 \quad \text{(a.s.)}. \tag{2.37}$$

To prove (2.37) it is enough to show

$$\left|\frac{\frac{1}{n}\sum_{i=1}^{n}\left(I_{Y,i}-\bar{I}_Y\right)^2}{E(I_{Y,1}-p_Y)^2} - 1\right| \to 0 \quad \text{(a.s.)}.$$

After a small calculation, we see that the above expression can be bounded

$$\frac{|\bar{I}_Y - \bar{I}_Y^2 - p_Y + (p_Y)^2|}{p_Y-(p_Y)^2} \le \frac{|\bar{I}_Y - p_Y|}{p_Y-(p_Y)^2} + \frac{|\bar{I}_Y^2-(p_Y)^2|}{p_Y-(p_Y)^2}.$$

For the first term on the right-hand side of the above inequality recall (2.4) and then we get

$$\frac{|\bar{I}_Y - p_Y|}{p_Y-(p_Y)^2} = \frac{\left|\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y)\right|}{p_Y-(p_Y)^2} \sim \left|\frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y)\right| \to 0 \ \text{(a.s.)},$$

as we have already shown in Lemma 2.3.1. For the second term, we can use a similar procedure to get

$$\begin{aligned}\frac{|\bar{I}_Y^2-(p_Y)^2|}{p_Y-(p_Y)^2} &= \frac{|(\bar{I}_Y-p_Y)|\cdot|(\bar{I}_Y+p_Y)|}{p_Y-(p_Y)^2} \\ &\le 2\frac{\left|\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y)\right|}{p_Y-(p_Y)^2} \sim \left|\frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}(I_{Y,i}-p_Y)\right| \to 0 \ \text{(a.s.)}\end{aligned}$$

and that proves the almost sure convergence stated.

To finish the proof of the lemma, all that remains is to show

$$\frac{\frac{1}{n}\sum_{i=1}^{n}\left(I_{Y,i}-\bar{I}_Y\right)^2\left(I_{Z,G_n(i)}-\bar{I}_Z\right)^2}{E(I_{Y,1}-p_Y)^2E(I_{Z,1}-p_Z)^2}\xrightarrow{P}1. \tag{2.38}$$

Note first that in the denominator in (2.38) we have the expression

$$p_Yp_Z-p_Z(p_Y)^2-p_Y(p_Z)^2+(p_Y)^2(p_Z)^2$$

and p_Yp_Z is the asymptotically dominant term ($\sim 1/m_n$). So, we need to prove that

$$\left|\frac{\frac{1}{n}\sum_{i=1}^{n}\left(I_{Y,i}-\bar{I}_Y\right)^2\left(I_{Z,G_n(i)}-\bar{I}_Z\right)^2}{p_Yp_Z}-1\right|\xrightarrow{P}0. \tag{2.39}$$

Next we calculate the expression in the numerator under the absolute sign to get

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,G_n(i)}-2\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,G_n(i)}-2\bar{I}_Y\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,G_n(i)}\\&+4\bar{I}_Y\bar{I}_Z\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,G_n(i)}-3\bar{I}_Y^2\bar{I}_Z^2+\bar{I}_Y\bar{I}_Z^2+\bar{I}_Y^2\bar{I}_Z-p_Yp_Z.\end{aligned}$$

We will show that the above-written expression divided by p_Yp_Z converges to zero in probability. First, we note that

$$\frac{1}{n}\sum_{i=1}^{n}I_{Y,i}I_{Z,G_n(i)}-p_Yp_Z=\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}I_{Z,G_n(i)}-p_Yp_Z).$$

Then we use the Markov inequality to get, for arbitrary $\varepsilon>0$,

$$\begin{aligned}&P\Big(\frac{1}{np_Yp_Z}\Big|\sum_{i=1}^{n}(I_{Y,i}I_{Z,G_n(i)}-p_Yp_Z)\Big|>\varepsilon\Big)\\&\qquad\le\frac{1}{\varepsilon^2n^2(p_Y)^2(p_Z)^2}E\Big(\sum_{i=1}^{n}(I_{Y,i}I_{Z,G_n(i)}-p_Yp_Z)\Big)^2.\end{aligned} \tag{2.40}$$

The expectation on the right-hand side of (2.40) is equal to

$$\begin{aligned}&\sum_{i=1}^{n}E(I_{Y,i}I_{Z,G_n(i)}-p_Yp_Z)^2\\&+\sum_{i=1}^{n}\sum_{\substack{j=1\\i\neq j}}^{n}E\Big((I_{Y,i}I_{Z,G_n(i)}-p_Yp_Z)(I_{Y,j}I_{Z,G_n(j)}-p_Yp_Z)\Big)=J_1+J_2.\end{aligned}$$

Let $h:\mathbf{G}_n\times\mathbb{R}^{2n}\to[0,\infty)$ be a function defined as

$$h(\pi,x^n)=\sum_{i=1}^{n}(I_{y,i}I_{z,\pi(i)}-p_Yp_Z)^2.$$

Similar to the calculation in (2.5) we get

$$E(h(G_n,X^n)\mid X^n)=\frac{1}{n!}\sum_{\pi\in\mathbf{G}_n}\sum_{i=1}^{n}(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)^2\quad\text{(a.s.)}$$

Take the expectation in the last equation to get

$$J_1=\sum_{i=1}^{n}\frac{1}{n!}\sum_{\pi\in\mathbf{G}_n}E(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)^2.$$

Fix $i\in\{1,2,\ldots,n\}$ and define $A_n=A_n(i):=\{\pi\in\mathbf{G}_n\mid\pi(i)=i\}$ and $B_n=A_n^{\mathrm{c}}$. A_n has $(n-1)!$ elements. Then we have

$$\begin{aligned}\frac{1}{n!}\sum_{\pi\in\mathbf{G}_n}E(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)^2&=\frac{1}{n!}\sum_{\pi\in A_n}E(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)^2\\&\quad+\frac{1}{n!}\sum_{\pi\in B_n}E(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)^2=J_{11}+J_{12}.\end{aligned}$$

By the definition of A_n we have

$$J_{11}=\frac{1}{n}(p_{11}-2p_{11}p_Yp_Z+(p_Y)^2(p_Z)^2)$$

and similarly

$$J_{12}=\frac{n-1}{n}(p_Yp_Z-(p_Y)^2(p_Z)^2).$$

Note that J_{11} converges to zero at least as fast as $1/(n\sqrt{m_n})$ ($p_{11}\sim 1/\sqrt{m_n}$ for $Y=Z$). As $p_Yp_Z\sim 1/m_n$ we conclude that J_1 can be estimated as

$$n\Big(\frac{1}{n\sqrt{m_n}}+\frac{n-1}{n}p_Yp_Z\Big)\sim\frac{1}{\sqrt{m_n}}+\frac{n}{m_n}\sim\frac{n}{m_n}$$

and its contribution in (2.40) is then asymptotically bounded by

$$\frac{m_n^2}{\varepsilon^2n^2}\frac{n}{m_n}\sim\frac{m_n}{\varepsilon^2 n}.$$

The last fraction above converges to zero because of the assumption (2.3).

To estimate J_2 we use reasoning similar to that above to get

$$J_2=\sum_{\substack{i=1\\i\neq j}}^{n}\sum_{j=1}^{n}\frac{1}{n!}\sum_{\pi\in\mathbf{G}_n}E(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)(I_{Y,j}I_{Z,\pi(j)}-p_Yp_Z).$$

Now fix $i,j\in\{1,2,\ldots,n\}$, $i\neq j$. Let A_n be a set of permutations $\pi\in\mathbf{G}_n$ such that $\pi(i)\neq i$, $\pi(i)\neq j$, $\pi(j)\neq i$ and $\pi(j)\neq j$. Let $B_n=A_n^c$. For the permutations $\pi\in A_n$ we have

$$E(I_{Y,i}I_{Z,\pi(i)}-p_Yp_Z)(I_{Y,j}I_{Z,\pi(j)}-p_Yp_Z)=0,$$

because $I_{Y,i}I_{Z,\pi(i)}$ and $I_{Y,j}I_{Z,\pi(j)}$ are independent, but also $I_{Y,i}$ and $I_{Z,\pi(i)}$ are independent as well as $I_{Y,j}$ and $I_{Z,\pi(j)}$. For $\pi \in B_n$ we have $\pi(i) = i$ or $\pi(i) = j$ or $\pi(j) = i$ or $\pi(j) = j$ and in each of these four cases there are $(n-1)!$ permutations (some of them are counted more than once). Therefore, the number of permutations in B_n is less than $4(n-1)!$. On the other hand, for the permutation $\pi \in B_n$ we have various possibilities for the values of

$$E(I_{Y,i}I_{Z,\pi(i)} - p_Y p_Z)(I_{Y,j}I_{Z,\pi(j)} - p_Y p_Z),$$

but in all cases, the asymptotically dominant terms are p_{11}^2 or $p_{11}p_Y p_Z$. As p_{11} converges to zero at least at the rate $1/\sqrt{m_n}$, J_2 is asymptotically bounded by

$$n^2 \frac{4}{n} \frac{1}{m_n}.$$

Going back to the relation (2.40) we see that the contribution of J_2 is then asymptotically bounded by

$$\frac{1}{\varepsilon^2 n (p_Y)^2 (p_Z)^2} \frac{4n}{m_n} \sim \frac{m_n}{\varepsilon^2 n},$$

for any $\varepsilon > 0$. The last expression converges to zero because of the assumption (2.3).

We conclude hat the right-hand side in (2.40) converges to zero and therefore

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}I_{Z,G_n(i)} - p_Y p_Z)}{p_Y p_Z} \xrightarrow{P} 0.$$

Now we can finish the proof of this lemma similarly as we have done in the proof of Lemma 2.3.2. For example, we can write

$$2\bar{I}_Z \frac{1}{n}\sum_{i=1}^{n} I_{Y,i}I_{Z,G_n(i)} = 2\bar{I}_Z \frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}I_{Z,G_n(i)} - p_Y p_Z) + 2p_Y p_Z \bar{I}_Z.$$

The first term on the right-hand side of the above equality, divided by $p_Y p_Z$, converges to zero, as we have just shown that for the term

$$\frac{1}{n}\sum_{i=1}^{n}(I_{Y,i}I_{Z,\pi(i)} - p_Y p_Z).$$

Recall that $\bar{I}_Z \sim 1/\sqrt{m_n}$ (a.s.). The second term, divided by $p_Y p_Z$, also converges to zero almost surely. We can use similar reasoning for the other terms under the absolute sign in (2.39), which finishes the proof of this lemma.

■

3. The Permutation test of independence for M-dependent data

In this chapter we formulate the permutation test of independence when Y_i and Z_i are stationary and M-dependent. To prove the results needed we use the idea that is quite common in the bootstrap literature: we remove blocks of length M from the original sequences to leave independent blocks of resulting data. Otherwise, the ideas we use are similar to those presented and used in Chapter 2.

3.1. Asymptotic results

Let $Y_i, Z_i,\ i = 1, 2, \dots$ be two strictly stationary M-dependent processes defined on the probability space $(\Omega, \mathscr{F}, P)$. Recall, the process $\{Y_i, i \in \mathbb{N}\}$ is M-dependent, $M \in \mathbb{N}$, if for all $j \in \mathbb{N}$, the vector $(Y_1, \dots, Y_j)$ is independent of $(Y_{j+k}, Y_{j+k+1} \dots)$ whenever $k > M$ and it is strictly stationary if the joint distributions of $(Y_{i_1}, \dots, Y_{i_k})$ and $(Y_{i_1+h}, \dots, Y_{i_k+h})$ are the same for all positive integers $i_1, \dots, i_k$, k and h. Note, the strict stationarity is equivalent to the statement that the distributions of $(Y_1, \dots, Y_k)$ and $(Y_{1+h}, \dots, Y_{k+h})$ are the same for all positive integers k and h. If $\{Y_i, i \in \mathbb{N}\}$ is strictly stationary, it immediately follows that all Y_i, $i \in \mathbb{N}$, have the same distribution.

As in the previous section we choose an intermediate sequence of integers m_n such that $m_n \to \infty$, for $n \to \infty$, but here we will need the stronger assumption

$$m_n = O\big(n^{1/4-\tau}\big), \tag{3.1}$$

for some $\tau > 0$. Suppose that there exist two sequences (u_n) and (v_n) of positive real numbers such that $u_n \to \sup\{x : F_Y(x) < 1\}$, $v_n \to \sup\{x : F_Z(x) < 1\}$ and

$$nP(Y > u_n) \to 1, \quad nP(Z > v_n) \to 1, \; n \to \infty. \tag{3.2}$$

Note, as before, for the generic member Y_i, Z_i of the processes $\{Y_i, i \in \mathbb{N}\}$ and $\{Z_i, i \in \mathbb{N}\}$ we write Y, Z. This is justified by the stationarity of those processes.

Let us fix $n \in \mathbb{N}$. We will suppose that $n = m_n \cdot N_n + M \cdot N_n$, where $N = N_n$ depends on n. That assumption makes no difference in asymptotic terms, but it will make our calculations and notation easier. With that assumption in mind we can divide both sequences of random variables $Y_1, \ldots, Y_n$ and $Z_1, \ldots, Z_n$ into N_n blocks of m_n subsequent random variables and separated by M blocks of them. In other words we are considering the sequences

$$Y_1, \ldots, Y_m, Y_{m+M+1}, \ldots, Y_{2m+M}, Y_{2m+2M+1}, \ldots, Y_{3m+2M}, \ldots, Y_n$$

and

$$Z_1, \ldots, Z_m, Z_{m+M+1}, \ldots, Z_{2m+M}, Z_{2m+2M+1}, \ldots, Z_{3m+2M}, \ldots, Z_n$$

and we form the random vectors $Y^1, \ldots Y^N$ and $Z^1, \ldots Z^N$ out of them. $Y^1, \ldots Y^N$ are independent as well as $Z^1, \ldots Z^N$ because we assumed $\{Y_i, i \in \mathbb{N}\}$ and $\{Z_i, i \in \mathbb{N}\}$ are M-dependent processes. Additionally, because of stationarity assumption, $Y^1, \ldots Y^N$ are equally distributed. The same is true for $Z^1, \ldots Z^N$. Note, although m_n and N_n depend on n we sometimes write them without the subscript n to ease notation.

We will use the hypothesis testing setup we have previously described. For that purpose we form the new sequence of vectors X^i where $X^i = (Y^i, Z^i)$, for $i \in \{1, \ldots, N_n\}$. Let $\hat{X}^N = (X^1, \ldots, X^N)$. As Y^i and Z^i are $\mathbb{R}^m$-valued vectors, the vector X^i has values in $\mathbb{R}^{2m}$. In terms of components we see that, for example, $X^1 = (Y_1, \ldots, Y_m, Z_1, \ldots, Z_m)$. We immediately conclude that $X^1, \ldots, X^N$ are independent as in each vector X^i are only the random variables Y_j and Z_k from blocks Y^i and Z^i and those are independent of the other blocks of original data, $i = 1, \ldots, N_n$. Also, all $X^1, \ldots, X^N$ have the same distribution, and we can denote it with P^X, as it is true for the random vectors Y^i and Z^i.

Denote the finite group of permutations of the set $\{1, 2, \ldots, N\}$ by $\mathbf{G}_N$. The group action of $\mathbf{G}_N$ on $(\mathbb{R}^{2m})^N = \mathbb{R}^{2mN}$ is defined by the action of the element $\pi \in \mathbf{G}_N$ as

$$\pi((y^1, z^1), \ldots (y^N, z^N)) = ((y^1, z^{\pi(1)}), \ldots (y^N, z^{\pi(N)})), \tag{3.3}$$

where $((y^1,z^1),\dots(y^N,z^N)) \in \mathbb{R}^{2mN}$. Let G_N be a random element on Ω with uniform distribution on the permutation group $\mathbf{G}_N$. We assume that G_N and $\hat{X}^N$ are independent in the rest of this chapter.

We will use several abbreviations in the sequel to make notation easier. As before, let $p_Y = p_Y(n) := P(Y_1 > u_{\sqrt{m_n}})$ and $p_Z = p_Z(n) := P(Z_1 > u_{\sqrt{m_n}})$ and

$$p^Y_{kl} = P(Y_k > u_{\sqrt{m_n}}, Y_l > u_{\sqrt{m_n}}) \quad \text{and} \quad p^Z_{kl} = P(Z_k > v_{\sqrt{m_n}}, Z_l > V_{\sqrt{m_n}}),$$

where $k,l \in \{1,2,\dots,m_n\}$. Next, let

$$I_{Y,ik} := I_{\{Y_{(i-1)(m_n+M)+k} > u_{\sqrt{m_n}}\}} \quad \text{and} \quad I_{Z,jk} := I_{\{Z_{(j-1)(m_n+M)+k} > u_{\sqrt{m_n}}\}}, \tag{3.4}$$

where $i,j \in \{1,\dots,N_n\}$ and $k \in \{1,\dots,m_n\}$. For $i \neq j$ $I_{Y,ik}$ and $I_{Y,jk}$ are independent as well as $I_{Y,ik}$ and $I_{Y,il}$, for $k-l > M$, $k,l \in \{1,\dots,m_n\}$. Finally, let

$$\bar{I}^Y_{n,k} := \frac{1}{N_n}\sum_{g=1}^{N_n} I_{\{Y_{(g-1)(m_n+M)+k} > u_{\sqrt{m_n}}\}} \quad \text{and} \quad \bar{I}^Z_{n,k} := \frac{1}{N_n}\sum_{h=1}^{N_n} I_{\{Z_{(h-1)(m_n+M)+k} > u_{\sqrt{m_n}}\}}.$$

Both $\{\bar{I}^Y_{n,k}\}_{k=1}^{m_n}$ and $\{\bar{I}^Z_{n,k}\}_{k=1}^{m_n}$ are triangular arrays of bounded random variables because

$$|\bar{I}^Y_{n,k}| \le \frac{1}{N_n}\sum_{g=1}^{N_n} |I_{\{Y_{(g-1)(m_n+M)+k} > u_{\sqrt{m_n}}\}}| \le \frac{1}{N_n} N_n = 1,$$

with the obvious analogy for $\bar{I}^Z_{n,k}$. Because of the stationarity of the processes Y and Z mean of both triangular arrays does not depend on k, as we have

$$E(\bar{I}^Y_{n,k}) = p_Y = p_Y(n) \ \text{ and } \ E(\bar{I}^Z_{n,k}) = p_Z = p_Z(n).$$

Note that for different $k,l \in \{1,\dots,m_n\}$ sums $\bar{I}^Y_{n,k}$ and $\bar{I}^Y_{n,l}$ contain members of the process $\{Y_i\}$ that are in a different block of the size m_n and separated by $k-l$ indices. Therefore, for $k-l > M$, $\bar{I}^Y_{n,k}$ and $\bar{I}^Y_{n,l}$ are independent, meaning that $\{\bar{I}^Y_{n,k}\}_1^{m_n}$ is an M-dependent triangular array. Obviously, the same applies to $\{\bar{I}^Z_{n,k}\}_1^{m_n}$.

On several occasions in rest of the chapter we will use the following asymptotic result.

$$\frac{m_n^3}{N_n} = \frac{m_n^3(m_n+M)}{N_n(m_n+M)} \sim \frac{m_n^4}{n} \to\ 0,\ n \to \infty. \tag{3.5}$$

In (3.5) we used the assumption (3.1) and the fact that, as n grows, M becomes negligible compared to both n and m_n.

Consider now the following auxiliary statistic

$$S_n = S_n(\hat{X}^N) := \frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{\{Y_{(i-1)(m_n+M)+k}>u_{\sqrt{m_n}}\}} I_{\{Z_{(i-1)(m_n+M)+k}>u_{\sqrt{m_n}}\}} = \frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,ik}.$$

Define the sum $S_n^{G_N}$ as

$$S_n(G_N\hat{X}^N) := \frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,G_N(i)k}.$$

Let $n \in \mathbb{N}$ such that $n = 2(m_n+M)N_n$, where $M, m_n, N_n \in \mathbb{N}$ have the same interpretation as before. Let $y_1,\dots,y_n, z_1,\dots,z_n \in \mathbb{R}$ and let $\hat{x}^n = ((y^1,z^1),\dots,(y^N,z^N)) \in \mathbb{R}^{2m_nN_n}$, where $y^i = (y_{(i-1)(m_n+M)+1},\dots,y_{(i-1)(m_n+M)+m_n})$, $z^i = (z_{(i-1)(m_n+M)+1},\dots,z_{(i-1)(m_n+M)+m_n}) \in \mathbb{R}^{m_n}$, $i = 1,\dots,N_n$. Next, define the function $h : \mathbf{G}_{N_n} \times \mathbb{R}^{2m_nN_n} \to [0,\infty)$ by

$$h(\pi,\hat{x}^n) = \frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{y,ik}I_{z,\pi(i)k},$$

Clearly, $h(G_{N_n},\hat{X}^{N_n}) = S_n^{G_{N_n}}$ and, because of (3.1), we have $E|h(G_{N_n},\hat{X}^{N_n})| \le m_n/n < \infty$ (a.s.). Let P_{G_N} be the probability on $\mathbf{G}_{N_n}$ induced by the random element G_{N_n}. Clearly $P_{G_N}(\pi) = 1/N_n!$, for $\pi \in \mathbf{G}_{N_n}$. Because of the independence between $\hat{X}^{N_n}$ and G_{N_n} by (1.28) we have

$$\begin{aligned} E(S_n^{G_{N_n}} \mid \hat{X}^{N_n}) &= \int_{\mathbf{G}_{N_n}} \frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,\pi(i)k}P_{G_N}(d\pi) \quad \text{(a.s.)} \\ &= \frac{1}{N_n!}\sum_{\pi\in\mathbf{G}_{N_n}} \frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,\pi(i)k} \quad \text{(a.s.)} \end{aligned} \tag{3.6}$$

From (3.6) we derive

$$E(S_n^{G_{N_n}} \mid \hat{X}^{N_n}) = \frac{m_n}{m_n+M}\sum_{k=1}^{m_n} \bar{I}_{nk}^Y \bar{I}_{nk}^Z \quad \text{(a.s.)} \tag{3.7}$$

and

$$\operatorname{Var}(S_n^{G_{N_n}} \mid \hat{X}^{N_n}) = \frac{1}{N_n-1}\frac{m_n^2}{n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Z,jk}-\bar{I}_{n,k}^Z)\Big)^2 \quad \text{(a.s.)}, \tag{3.8}$$

To see how, observe that in the notation of Section 1.2, and using (1.19), we get

$$E(S_n^{G_{N_n}} \mid \hat{X}^{N_n}) = \frac{1}{N_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n} c_{N_n}(i,j),$$

where

$$c_{N_n}(i,j) = \frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,jk},\ i,j = 1,\dots,N_n. \tag{3.9}$$

Then, $E(S_n^{G_{N_n}} \mid \hat{X}^{N_n})$ is equal to

$$\frac{m_n}{m_n+M}\frac{1}{N_n}\sum_{i=1}^{N_n}\frac{1}{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,jk} = \frac{m_n}{m_n+M}\sum_{k=1}^{m_n}\bar{I}_{nk}^{Y}\bar{I}_{nk}^{Z}.$$

Similarly, by (1.20) and Lemma 3.3.1 we get (3.8).

Furthermore, by using (1.28), we conclude that almost surely

$$\begin{aligned} P\Big(S_n^{G_{N_n}} - E(S_n^{G_{N_n}} \mid \hat{X}^{N_n}) \le t\sqrt{\mathrm{Var}\,(S_n^{G_{N_n}} \mid \hat{X}^{N_n})} \mid \hat{X}^{N_n}\Big) \\ = \int_{\mathbf{G}_{N_n}} I_{\{\frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,\pi(i)k} - E(S_n^{G_{N_n}}|\hat{X}^{N_n}) \le t\sqrt{\mathrm{Var}\,(S_n^{G_{N_n}}|\hat{X}^{N_n})}\}} P_{G_{N_n}}(d\pi) \\ = \frac{1}{N_n!}\sum_{\pi\in\mathbf{G}_{N_n}} I_{\{\frac{m_n}{n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,\pi(i)k} - E(S_n^{G_{N_n}}|\hat{X}^{N_n}) \le t\sqrt{\mathrm{Var}\,(S_n^{G_{N_n}}|\hat{X}^{N_n})}\}},\ t\in\mathbb{R}. \end{aligned} \tag{3.10}$$

We will again use the context of the Hoeffding Central Limit Theorem, specifically Theorem 1.2.3, to prove the asymptotic result needed to formulate the permutation test of independence for M-dependent processes. Below we give the analogue of Theorem 2.1.2 applied to the present case of M-dependent processes (Y_i) and (Z_i).

Theorem 3.1.1. Let $(Y_i)_{i=1}^{\infty}$ and $(Z_i)_{i=1}^{\infty}$ be two stationary and M-dependent sequences of random variables. Suppose they are mutually independent and conditions (3.1) and (3.2) hold. Then for $t\in\mathbb{R}$

$$\lim_{n\to\infty} P\Big(S_n^{G_{N_n}} - E(S_n^{G_{N_n}} \mid \hat{X}^{N_n}) \le t\sqrt{\mathrm{Var}\,(S_n^{G_{N_n}} \mid \hat{X}^{N_n})} \mid \hat{X}^{N_n}\Big) = \Phi(t) \quad \text{(a.s.)} \tag{3.11}$$

Convergence in (3.11) is the same as in the conclusion of Theorem 2.1.2.

The proof of Theorem 3.1.1 is given in Section 3.3.

Define the statistic T_n as

$$T_n(\hat{X}^N) = \sqrt{N_n - 1}\,\frac{\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,ik} - N_n\sum_{k=1}^{m_n}\bar{I}_{nk}^{Y}\bar{I}_{nk}^{Y}}{\sqrt{\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^{Y})(I_{Z,jk}-\bar{I}_{n,k}^{Z})\Big)^2}}. \tag{3.12}$$

Then

$$T_n(G_N\hat{X}^N) = \sqrt{N_n - 1}\,\frac{\sum_{i=1}^{N_n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,G_N(i)k} - N_n\sum_{k=1}^{m_n}\bar{I}_{nk}^{Y}\bar{I}_{nk}^{Y}}{\sqrt{\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^{Y})(I_{Z,jk}-\bar{I}_{n,k}^{Z})\Big)^2}}. \tag{3.13}$$

Use (3.7) and (3.8) to conclude

$$T_n(G_N\hat{X}^N) = \frac{S_n^{G_N} - E(S_n^{G_N} \mid \hat{X}^N)}{\sqrt{\mathrm{Var}\,(S_n^{G_N} \mid \hat{X}^N)}} \tag{3.14}$$

Suppose the null hypothesis, that H_0 : (Y_i) and (Z_i) are independent, holds. Then, because $X^1, \dots, X^N$ are independent and equally distributed, we have

$$P_{\hat{X}^N} = P_{X^1} \dots P_{X^N} = (P_{Y^1} \times P_{Z^1})^N.$$

Therefore, the randomization hypothesis holds. Let

$$\hat{R}_{N_n}(t) = \frac{1}{N_n!} \sum_{\pi \in \mathbf{G}_{N_n}} I_{\{T_n(G_{N_n}\hat{X}^{N_n}) \leq t\}},\ t \in \mathbb{R}. \tag{3.15}$$

Theorem 3.1.1 implies that the permutation distribution $\hat{R}_{N_n}(t)$ of T_n converges towards a standard normal random variable almost surely and therefore in probability too. Then, as in Remark 2.1.8, it follows that the test statistic T_n converges in distribution towards a standard normal random variable. Hence, we can perform the permutation test of the independence of the M-dependent processes (Y_i) and (Z_i) by using the test statistic $T_n(\hat{X}^N)$.

Remark 3.1.2. We supposed that $Y_i, Z_i,\ i = 1, 2, \dots$ are M-dependent, with the same $M \in \mathbb{N}$. The proof of Theorem 3.1.1 shows that its conclusion remains the same, if (Y_i) is M_1-dependent and (Z_i) is M_2-dependent, for $M = \max\{M_1, M_2\}$.

3.2. SIMULATIONS

In this section, we investigate the behaviour of the test statistic T_n defined in (3.12) in a simulation study. We denote the simulated data by $(Y_1,Z_1),\dots,(Y_n,Z_n)$, $n \in \mathbb{N}$. The threshold levels used to calculate the value of the statistic T_n are determined through empirical upper quantiles of the given data.

Unlike the case described in Chapter 2, here the test statistic depends on n through N_n and m_n which makes the testing procedure more complex. Additionally, there may be different dependencies contained in the class of M-dependent process. Therefore, we will test the behaviour of the test statistic for different values of N and m.

3.2.1. Independent iid samples

Consider two iid sequences (Y_i) and (Z_i) from the unit Pareto distribution. In Table 3.1 we present empirical rejection probabilities of the test for various combinations of N and m.

N	m	Thresh. 0.2	Thresh. 0.1	Thresh 0.05
25	25	0.045	0.040	0.047
10	50	0.055	0.044	0.058
10	100	0.048	0.048	0.053
50	25	0.049	0.054	0.054
100	25	0.049	0.049	0.043

Table 3.1: Results of a simulation study in the iid Pareto case for various values of N (column 1) and m (column 2). Empirical rejection probabilities for the different threshold levels are shown in columns 3-5. The level of the test is set at 5%, while the number of permutations and repeats is equal to 2000 in all simulations.

Although our asymptotic results indicate that N should be, at least asymptotically, much larger than m, simulation results on independent iid samples indicate that performance of the permutation test is not so sensitive to the choice of N and m. Simulations have been performed for various choices of M, but the same conclusion holds. That is not surprising, as we are dealing here with iid sequences. Also, it appears that the suggested permutation test retains approximate exactness for various choices of the threshold level.

3.2.2. Dependent iid samples

In this subsection we present the simulation results for three copulas we have already encountered in Chapter 2: the Gumbel-Hougaard, the Morgenstern and the Gaussian copula. Here the choice of N and m matter more for the power of the test as somewhat better results are achieved when N is larger or equal to m. Therefore, we present only the simulation results for such choices of N and m. As in the case of independent iid data, the choice of M does not influence the simulation results and therefore we do not report simulations for the different values of M.

The results presented in this section were simulated with the values of N equal to 31, 50, 100 and 200 and the values of m equal to 31, 20, 10 and 5, respectively. In Figure 3.1 we present the empirical rejection probabilities for various values of the parameter χ with $N = 100$ and $m = 10$. Very similar results were obtained for the combination $N = 200$ and $m = 5$. As can be seen, the obtained results are very similar to the results

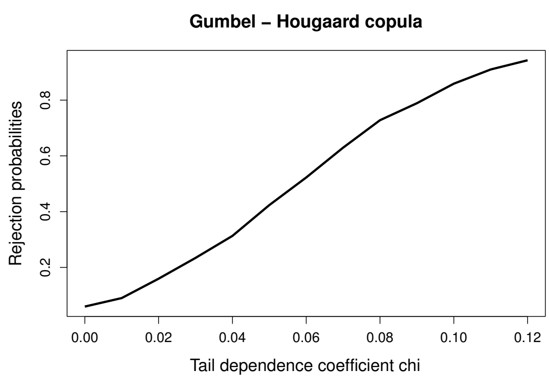


Figure 3.1: Empirical rejection probabilities (y-axis) for various values of the parameter χ (x-axis) for the sample size $n = 1000$, with $N = 100$ and $m = 10$, using an upper empirical quantile at a level of 20%. Data were generated from the Gumbel-Hougaard copula. The level of the test is set at 5%, while the number of permutations and repeats is equal to 2000.

presented in Section 2.2.3. The power of the test was quite stable for various empirical quantiles (thresholds). Also, the test is not overly sensitive to various combinations of N and m as can be seen from Table 3.2, where we present empirical rejection probabilities for $\chi = 0.06$. Note that approximately on that level the differences are the largest.

N	31	50	100	200
m	31	20	10	5
	Empirical rejection probabilities			
Threshold 20%	0.0505	0.0502	0.0523	0.0528
Threshold 10%	0.0477	0.0506	0.0528	0.0542

Table 3.2: Empirical rejection probabilities for data generated from the Gumbel-Hougaard copula with the parameter $\chi = 0.06$, for various combinations of the vales of N and m and for different thresholds. The level of the test is 0.05, and the number of permutations and repeats is 2000.

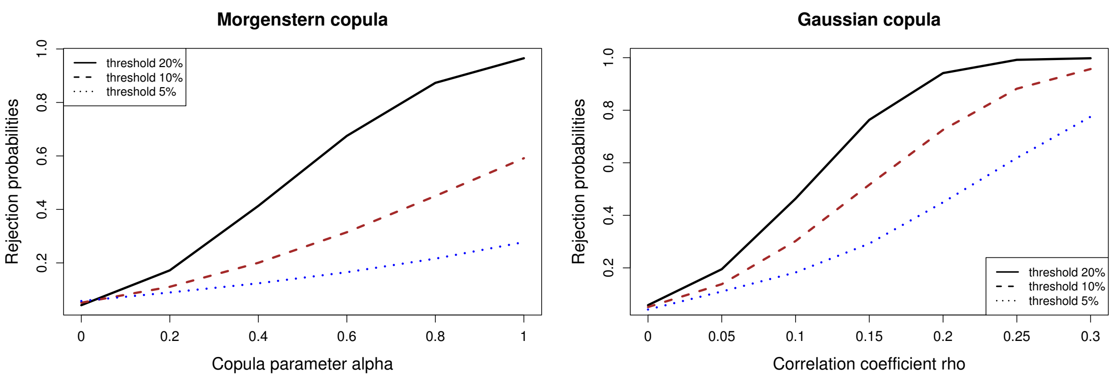


Figure 3.2: Empirical rejection probabilities (y-axis) for various values of the corresponding parameter of the copulas (x-axis): Morgenstern on the left and Gaussian on the right. The level of the test is set at 5%, the sample size is $n = 1000$, with $N = 100$ and $m = 10$, while the number of permutations and repeats equal 2000. The threshold level varies from 20% through 10% down to 5%.

Next we present simulation results for the data generated from the Morgenstern and the Gaussian copula. In Figure 3.2 we present empirical rejection probabilities with $N =$

100 and $m = 10$. Similar to the results from Section 2.2.4, the power of the test for the data generated from those two copulas depends on the choice of the threshold level. On the other hand, the differences between empirical rejection probabilities for various choices of of N and m are even smaller than those observed for the data generated from the Gumbel-Hougaard copula.

3.2.3. Samples of M-dependent data that are mutually independent

Let X_i^1 and X_i^2, $i = 1,2,\ldots$, be two independent sequences of Pareto distributed random variables. Define the new independent sequences $Y_i = X_i^1 + X_{i+1}^1$ and $Z_i = X_i^2 + X_{i+1}^2$, $i = 1,2,\ldots$. Clearly, (Y_i) and (Z_i) are 1-dependent sequences. We simulated the sequences Y_i and Z_i and tested for independence with the testing statistic T_n for different combinations of N, m and M. In Table 3.3 we present some of the results obtained.

Threshold 10%	$M=0$	$M=1$	$M=2$
$N=50, m=20$	0.0430	0.0495	0.0595
$N=100, m=10$	0.0520	0.0575	0.0570
$N=200, m=5$	0.0540	0.0450	0.0515
Threshold 5%	$M=0$	$M=1$	$M=2$
$N=50, m=20$	0.0470	0.0510	0.0565
$N=100, m=10$	0.0585	0.0485	0.0520
$N=200, m=5$	0.0540	0.0410	0.0500

Table 3.3: Empirical rejection probabilities for data generated from two 1-dependent sequences Y_i and Z_i, for various combinations of the vales of N and m and for different thresholds. The level of the test is 0.05, and the number of permutations and repeats is 2000.

The results of the simulations seem to indicate that approximate exactness is achieved for various values of M. To confirm that, we defined new 2-dependent sequences $Y_i' = X_i^1 + X_{i+1}^1 + X_{i+2}^1$ and $Z_i' = X_i^2 + X_{i+1}^2 + X_{i+2}^2$, $i = 1,2,\ldots$, and repeated the simulation procedure. We also expanded the number of permutations and repeats to 4000. Some of the obtained results are presented in Table 3.4

Threshold 10%	$M = 0$	$M = 1$	$M = 2$	$M = 3$
$N = 45, m = 45$	0.0555	0.0565	0.0518	0.0480
$N = 100, m = 20$	0.0473	0.0525	0.0473	0.0500
$N = 200, m = 10$	0.0568	0.0435	0.0550	0.0493
$N = 400, m = 5$	0.0598	0.0515	0.0518	0.0500
Threshold 5%	$M = 0$	$M = 1$	$M = 2$	$M = 3$
$N = 45, m = 45$	0.0500	0.0508	0.0473	0.0550
$N = 100, m = 20$	0.0533	0.0493	0.0510	0.0455
$N = 200, m = 10$	0.0495	0.0500	0.0538	0.0545
$N = 400, m = 5$	0.0593	0.0483	0.0555	0.0555

Table 3.4: Empirical rejection probabilities for data generated from two 2-dependent sequences Y_i' and Z_i', for various combinations of the vales of N and m and for different thresholds. The level of the test is 0.05, and the number of permutations and repeats is 4000.

The simulation results again do not distinguish between any particular choice of M, although visual inspection of qq-plots of the test statistic T_n seems to indicate that better stability is achieved for $M = 2$. Overall, we conclude that the presented permutation test achieves approximate exactness for various combinations of N, m and M.

3.2.4. Samples of M-dependent data that are mutually dependent

Let X_i^1 and X_i^2, $i = 1,2,\ldots$, be two independent sequences of Pareto distributed random variables and define new 2-dependent sequences $Y_i' = X_i^1 + a \cdot X_{i+1}^1 + X_{i+2}^1$ and $Z_i' = X_i^2 + a \cdot X_{i+1}^1 + X_{i+2}^2$, $i = 1,2,\ldots$, where $a \geq 0$. Sequences Y_i and Z_i are now dependent and the parameter a determines the strength of dependence. We simulated the sequences Y_i and Z_i and applied our permutation to test the hypothesis of independence for various values of parameter a.

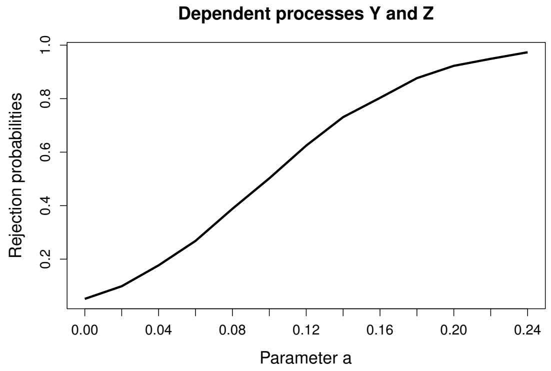


Figure 3.3: Empirical rejection probabilities (y-axis) for various values of the parameter a (x-axis) for for $N = 100$, $m = 20$ and $M = 2$, using an upper empirical quantile at a level of 10%. Data were generated from the sequences Y_i and Z_i defined in this subsection. The level of the test is set at 5%, while the number of permutations and repeats is equal to 4000.

$a = 0.07$	$M = 0$	$M = 1$	$M = 2$
$N = 45, m = 45$	0.3113	0.3313	0.3175
$N = 100, m = 20$	0.3303	0.3220	0.3280
$N = 200, m = 10$	0.3335	0.3305	0.3338
$N = 400, m = 5$	0.3555	0.3518	0.3435
$a = 0.14$	$M = 0$	$M = 1$	$M = 2$
$N = 45, m = 45$	0.7165	0.7183	0.7048
$N = 100, m = 20$	0.7123	0.7108	0.7163
$N = 200, m = 10$	0.7223	0.7150	0.7310
$N = 400, m = 5$	0.7358	0.7445	0.7465

Table 3.5: Empirical rejection probabilities for data generated from two 2-dependent sequences Y_i' and Z_i', for various combinations of the vales of N and m and for different thresholds. The level of the test is 0.05, and the number of permutations and repeats is 4000.

Simulation results are presented in Figure 3.3 for data generated by Y_i and Z_i. We show simulation results for $N = 200$, $m = 10$ and $M = 2$ as the permutation test shows the best power for the larger values of N (similar results were obtained for $N = 400$, $m = 5$ and $M = 2$). The power of the test was weaker for other combinations of N and m (for example, $N = 65$ and $m = 30$ or $N = 100$ and $m = 20$). In Table 3.5 we present some results of the simulations for various values of M. We note that the simulation results indicate that the test shows the largest power when $M = 2$, although the simulation results do not show clear trend of increased power related to larger value of M.

Overall, we conclude that the permutation test we present shows somewhat better performance for combinations of N and m in which N is larger.

3.3. PROOFS

Before we give the proof of Theorem 3.1.1 we will state and prove four lemmas. The proofs of the last three are quite technical and lengthy, but otherwise elementary as they rely on the Markov/Chebyshev inequality and the Borel-Cantelli lemma.

Lemma 3.3.1. With the same assumptions as in Theorem 3.1.1 let

$$c_{N_n}(i,j) := \frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,jk}.$$

Then for $d_{N_n}(i,j)$ (see (1.17) for the definition) we have

$$d_{N_n}(i,j) = \frac{m_n}{n}\sum_{k=1}^{m_n}(I_{Z,jk}-\bar{I}^Z_{n,k})(I_{Y,ik}-\bar{I}^Y_{n,k}). \tag{3.16}$$

Proof. Recall that

$$d_{N_n}(i,j) = c_{N_n}(i,j) - \frac{1}{N_n}\sum_{g=1}^{N_n} c_{N_n}(g,j) - \frac{1}{N_n}\sum_{h=1}^{N_n} c_{N_n}(i,h) + \frac{1}{N_n^2}\sum_{g=1}^{N_n}\sum_{h=1}^{N_n} c_{N_n}(g,h). \tag{3.17}$$

where $i,j \in \{1,\ldots,N_n\}$. The sum of the first and the third addend in (3.17) is equal to

$$\frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,jk} - \frac{1}{N_n}\sum_{h=1}^{N_n}\frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,hk} = \frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,ik}\big(I_{Z,jk}-\bar{I}^Z_{n,k}\big). \tag{3.18}$$

In the same vein we calculate the sum of the second and the fourth addend in (3.17):

$$\begin{aligned}
&-\frac{1}{N_n}\sum_{g=1}^{N_n}\frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,gk}I_{Z,jk} + \frac{1}{N_n^2}\sum_{g=1}^{N_n}\sum_{h=1}^{N_n}\frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,gk}I_{Z,hk}\\
&= -\frac{m_n}{n}\sum_{k=1}^{m_n}\frac{1}{N_n}\sum_{g=1}^{N_n} I_{Y,gk}I_{Z,jk} + \frac{m_n}{n}\sum_{k=1}^{m_n}\frac{1}{N_n}\sum_{g=1}^{N_n} I_{Y,gk}\frac{1}{N_n}\sum_{h=1}^{N_n} I_{Z,hk}\\
&= -\frac{m_n}{n}\sum_{k=1}^{m_n}\frac{1}{N_n}\sum_{g=1}^{N_n} I_{Y,gk}\big(I_{Z,jk}-\bar{I}^Z_{n,k}\big)
\end{aligned} \tag{3.19}$$

From (3.18) and (3.19) we get (3.16). ■

Lemma 3.3.2. With same assumptions as in Theorem 3.1.1 we have:

a)

$$\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{j=1}^{N_n}(I_{Z,jk}-p_Z) \to 0,\ n\to\infty\ \text{(a.s.)}$$

b)

$$\frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Z,jk}-p_Z)\to 0,\ n\to\infty\ \text{(a.s.)}$$

c)

$$\frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}(I_{Z,ik}-p_Z)(I_{Z,jk}-p_Z)\to 0,\ n\to\infty\ \text{(a.s.)}$$

d)

$$\frac{1}{N_n^3}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Z,jk}-p_Z)(I_{Z,qk}-p_Z)\to 0,\ n\to\infty\ \text{(a.s.)}.$$

Because of the symmetry between Y and Z we also have

a')

$$\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y)\to 0,\ n\to\infty\ \text{(a.s.)}$$

c')

$$\frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{\substack{i=1\\ i\neq j}}^{N_n}\sum_{j=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Y,jk}-p_Y)\to 0,\ n\to\infty\ \text{(a.s.)}.$$

Proof of Lemma 3.3.2. For the proof of all given almost sure convergences we will use the Markov inequality and the Borell-Cantelli lemma. The main idea of the proofs is to use the fact that we centred random variables $I_{Y,ik}$ and $I_{Z,jk}$ and, because of the few independence arguments, many multiple sums we will encounter will just vanish.

a') Let us take $\varepsilon > 0$. Then we have by Markov inequality

$$P\Big(\frac{1}{N_n}\Big|\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y)\Big|>\varepsilon\Big)\le\frac{1}{\varepsilon^4N_n^4}E\Big(\sum_{i=1}^{N_n}\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\Big)^4. \tag{3.20}$$

The expectation on the right hand side of the previous inequality is equal to

$$\sum_{i=1}^{N_n} E\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\Big)^4 + 4\sum_{\substack{i=1\\ i\neq j}}^{N_n}\sum_{j=1}^{N_n} E\Big(\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^3\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)\Big)$$
$$+3\sum_{\substack{i=1\\ i\neq j}}^{N_n}\sum_{j=1}^{N_n} E\Big(\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)^2\Big)$$
$$+6\sum_{\substack{i=1\\ i\neq j\neq p}}^{N_n}\sum_{j=1}^{N_n}\sum_{p=1}^{N_n} E\Big(\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)\big(\sum_{k=1}^{m_n}(I_{Ypk}-p_Y)\big)\Big)$$
$$+\sum_{\substack{i=1\\ i\neq j\neq p\neq q}}^{N_n}\sum_{j=1}^{N_n}\sum_{p=1}^{N_n}\sum_{q=1}^{N_n} E\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\sum_{k=1}^{m_n}(I_{Ypk}-p_Y)\sum_{k=1}^{m_n}(I_{Yqk}-p_Y)\Big).$$

For different i and j sums $\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)$ and $\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)$ are independent (they contain random variables from different blocks separated by M places) and so the second, the fourth and the fifth sum in the above expression vanish. Therefore we get

$$E\Big(\sum_{i=1}^{N_n}\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\Big)^4 = \sum_{i=1}^{N_n} E\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\Big)^4$$
$$+3\sum_{\substack{i=1\\ i\neq j}}^{N_n}\sum_{j=1}^{N_n} E\Big(\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)^2\Big). \tag{3.21}$$

By using a similar decomposition for the fourth power of a sum as we used above, we see that the asymptotically dominant term of the first sum in (3.21) is

$$\sum_{i=1}^{N_n}\sum_{\substack{k=1\\ k\neq l\neq r\neq s}}^{m_n}\sum_{l=1}^{m_n}\sum_{r=1}^{m_n}\sum_{s=1}^{m_n} E\big((I_{Y,ik}-p_Y)(I_{Y,il}-p_Y)(I_{Y,ir}-p_Y)(I_{Y,is}-p_Y)\big). \tag{3.22}$$

Note that $I_{Y,ik}$ and $I_{Y,il}$ are independent for $|k-l|>M$ and the same observation holds for indices r and s. Therefore, for $s\in\{1,\dots,m_n\}$ such that $|s-k|>M$, $|s-l|>M$ and $|s-r|>M$ the sum in (3.22) is equal to zero as $E(I_{Y,is}-p_Y)=0$. If that is not true for at least one index, let say $|s-r|\leq M$, then we can not conclude that the sum in (3.22) is equal to zero, but then it can be bounded by the asymptotically dominant term $2MN_nm_n^3p_Y$, i.e. we get

$$\sum_{i=1}^{N_n} E\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\Big)^4 \leq CN_nm_n^3p_Y.$$

For the second sum in (3.21) we first note that, as $i \neq j$, we have

$$E\Big(\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)^2\Big) = E\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2 E\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)^2$$

Next, we get estimate

$$\begin{aligned} E\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2 &= \sum_{k=1}^{m_n}E(I_{Y,ik}-p_Y)^2 + \sum_{\substack{k=1 \\ k\neq l}}^{m_n}\sum_{l=1}^{m_n}E(I_{Y,ik}-p_Y)(I_{Y,il}-p_Y) \\ &= \sum_{k=1}^{m_n}\big(E(I_{Y,ik})-(p_Y)^2\big) + \sum_{\substack{k=1 \\ |k-l|\leq M}}^{m_n}\sum_{l=1}^{m_n}E(I_{Y,ik}-p_Y)E(I_{Y,il}-p_Y) \\ &\leq m_n p_Y + 2Mm_n p_Y = Cm_n p_Y. \end{aligned}$$

To estimate the second sum above we used the same reasoning we employed to estimate the sum in (3.22). Also, since $0 \leq I_{Y,ik}, I_{Z,jk} \leq 1$, for $i,j \in \{1,\dots,N_n\}$, $k,l \in \{1,\dots,m_n\}$, we have the estimate $E(I_{Y,ik}I_{Y,il}) \leq E(I_{Y,ik}) = p_Y$. Hence, we get estimate

$$3\sum_{\substack{i=1 \\ i\neq j}}^{N_n}\sum_{j=1}^{N_n}E\Big(\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\big)^2\big(\sum_{k=1}^{m_n}(I_{Y,jk}-p_Y)\big)^2\Big) \leq CN_n^2 m_n^2 (p_Y)^2.$$

Taking into account obtained estimates we conclude from (3.21) that

$$E\Big(\sum_{i=1}^{N_n}\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)\Big)^4 \leq CN_n m_n^3 p_Y + CN_n^2 m_n^2 (p_Y)^2,$$

where $C > 0$ is a constant. Finally, use the assumption (3.1) to conclude

$$\frac{CN_n m_n^3}{N_n^4} = Cp_y \frac{m_n^3}{N_n^3}\cdot\frac{(m_n+M)^3}{(m_n+M)^3} \sim p_Y \frac{m_n^6}{n^3} \sim \frac{m_n^3}{n}\frac{m_n^{5/2}}{n}\frac{1}{n} < \frac{1}{n^{1+\eta}}$$

and

$$\frac{N_n^2 m_n^2 p_Y p_Z}{N_n^4} = \frac{p_Y p_Z m_n^2}{N_n^2}\frac{m_n^2}{m_n^2} \sim (m_n p_Y p_Z)\frac{m_n^3}{n^2} \sim 1\cdot\frac{m_n^3}{n}\frac{1}{n} < \frac{1}{n^{1+\eta}},$$

where $\eta > 0$. Going back into (3.20) we see that we have the estimate

$$P\Big(\frac{1}{N_n}\Big|\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y)\Big| > \varepsilon\Big) \leq \frac{1}{\varepsilon^4 n^{1+\eta}}$$

and then we use the Borel-Cantelli lemma to conclude that almost surely

$$\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y) \to 0, \;\; n\to\infty.$$

That prove parts $a)$ and $a')$ of the lemma.

b) We would like to show that

$$B_n = \frac{1}{N_n^2} \sum_{k=1}^{m_n} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)$$

is converging to zero almost surely. Let $\varepsilon > 0$ be arbitrarily chosen. Then by the Chebyshev inequality we have

$$P(|B_n| > \varepsilon) \leq \frac{1}{N_n^4 \varepsilon^2} E\Big(\sum_{k=1}^{m_n} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)\Big)^2. \tag{3.23}$$

Expectation on the right hand side of (3.23) is equal to

$$\begin{aligned}
&\sum_{i=1}^{N_n} E\Big(\sum_{j=1}^{N_n} \sum_{k=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)\Big)^2 \\
&+ \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{q=1}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E\Big((I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z) \cdot (I_{Y,pl} - p_Y)(I_{Z,ql} - p_Z)\Big) \\
&= B_1 + B_2.
\end{aligned}$$

Recall that (Y_i) and (Z_i) are independent, as well as $I_{Y,ik}$ and $I_{Y,pl}$, because $i \neq p$. Therefore

$$B_2 = \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{q=1}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Y,ik} - p_Y) E(I_{Y,pl} - p_Y) E\big((I_{Z,jk} - p_Z)(I_{Z,ql} - p_Z)\big) = 0.$$

Using the similar reasoning we get

$$\begin{aligned}
B_1 &= \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} E\Big(\sum_{k=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)\Big)^2 \\
&+ \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j \neq q}}^{N_n} \sum_{q=1}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E\big((I_{Y,ik} - p_Y)(I_{Y,il} - p_Y)(I_{Z,jk} - p_Z)(I_{Z,ql} - p_Z)\big) \\
&= B_{11} + B_{12}.
\end{aligned}$$

Because Y and Z are independent and $j \neq q$ we see that the expectation in B_{12} is equal to

$$E\big((I_{Y,ik} - p_Y)(I_{Y,il} - p_Y)\big) E(I_{Z,jk} - p_Z) E(I_{Z,ql} - p_Z) = 0$$

and so $B_{12} = 0$. Finally,

$$B_{11} = \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E\big((I_{Y,ik} - p_Y)(I_{Y,il} - p_Y)\big) E\big((I_{Z,jk} - p_Z)(I_{Z,jl} - p_Z)\big)$$

The asymptotically dominant term in B_{11} is

$$\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{l=1}^{m_n} E(I_{Y,ik}I_{Y,il})E(I_{Z,jk}I_{Z,jl})$$

and it can be bounded by $m_n^2 N_n^2 p_Y p_Z$. We used estimates $E(I_{Y,ik}I_{Y,il}) \leq p_Y$ and $E(I_{Z,ik}I_{Z,il}) \leq p_Z$. Note that other terms in B_{12} contain multiple appearances of p_Y and p_Z that converges to zero by rate $1/\sqrt{m_n}$ according to the assumption (3.1). Going back to (3.23) we conclude that

$$P(|B_n| > \varepsilon) \leq \frac{p_Y p_Z m_n^2 N_n^2}{N_n^4 \varepsilon^2} = \frac{1}{\varepsilon^2}\frac{(m_n p_Y p_Z)m_n}{N_n^2} \sim \frac{1}{\varepsilon^2}\frac{m_n}{N_n^2}\cdot\frac{m_n^2}{m_n^2} \sim \frac{1}{\varepsilon^2}\frac{m_n^3}{n^2}.$$

Because of the assumption (3.1) we conclude

$$P(|B_n| > \varepsilon) < \frac{1}{\varepsilon^2}\frac{1}{n^{1+\eta}},$$

for some $\eta > 0$. Use the Borel-Cantelli lemma to conclude that B_n converges to zero almost surely. That proves $b)$ part of the lemma.

c) Note first that for $i = j$ we have

$$\Big|\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n}(I_{Z,ik}-p_Z)(I_{Z,ik}-p_Z)\Big| \leq \frac{m_n N_n}{N_n^2} = \frac{m_n}{N_n} \to 0,\ n\to\infty \quad \text{(a.s.)}$$

Therefore, we need to prove

$$C_n = \frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{\substack{j=1\\ i\neq j}}^{N_n}(I_{Z,ik}-p_Z)(I_{Z,jk}-p_Z)$$

is converging to zero almost surely. We will use very similar arguments as we have used in proving $b)$ part of this lemma. Choose $\varepsilon > 0$ arbitrarily. Then by the Chebyshev inequality we have

$$P(|C_n| > \varepsilon) \leq \frac{1}{N_n^4\varepsilon^2}E\Big(\big(\sum_{i=1}^{N_n}\sum_{\substack{j=1\\ i\neq j}}^{N_n}\sum_{k=1}^{m_n}(I_{Z,ik}-p_Z)(I_{Z,jk}-p_Z)\big)\Big)^2. \tag{3.24}$$

The expectation on the right hand side of (3.24) is equal to

$$\sum_{i=1}^{N_n}E\Big(\sum_{\substack{j=1\\ j\neq i}}^{N_n}\sum_{k=1}^{m_n}(I_{Z,ik}-p_Z)(I_{Z,jk}-p_Z)\Big)^2$$
$$+\sum_{i=1}^{N_n}\sum_{\substack{p=1\\ p\neq i}}^{N_n}\sum_{\substack{j=1\\ j\neq i}}^{N_n}\sum_{\substack{q=1\\ q\neq j}}^{N_n}\sum_{k=1}^{m_n}\sum_{l=1}^{m_n}E(I_{Z,ik}-p_Z)(I_{Z,pl}-p_Z)(I_{Z,jk}-p_Z)(I_{Z,ql}-p_Z)$$
$$= C_1 + C_2.$$

For sums in C_2 for which $q \neq i$ we get $i \neq p$, $i \neq j$ and $i \neq q$ and in those cases the expectation in C_2 is equal to

$$E(I_{Z,ik} - p_Z)E(I_{Z,pl} - p_Z)(I_{Z,jk} - p_Z)(I_{Z,ql} - p_Z) = 0$$

and so $C_2 = 0$. So, we can suppose $q = i$ and then

$$C_2 = \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z)E(I_{Z,pl} - p_Z)(I_{Z,jk} - p_Z),$$

because $p \neq i$ and $j \neq i$. For the sums in which $p \neq j$ we again conclude $C_2 = 0$. Therefore, we can suppose $p = j$ to get

$$C_2 = \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z)E(I_{Z,jl} - p_Z)(I_{Z,jk} - p_Z).$$

Now note that for $k,l \in \{1,2,\ldots,m_n\}$ such that $|k-l| > M$, indicators $I_{Z,ik}$ and $I_{Z,il}$ are independent and therefore $E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z) = E(I_{Z,ik} - p_Z)E(I_{Z,il} - p_Z) = 0$, which again leads to conclusion $C_2 = 0$. Finally, we conclude that

$$C_2 = \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z)E(I_{Z,jl} - p_Z)(I_{Z,jk} - p_Z).$$

The asymptotically dominant term in C_2 is

$$\sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} E(I_{Z,ik}I_{Z,il})E(I_{Z,jl}I_{Z,jk})$$

and it can be bounded by $2Mm_nN_n^2p_Z^2$.

By using similar arguments we see that

$$\begin{aligned} C_1 &= \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} E\Big(\sum_{k=1}^{m_n} (I_{Z,ik} - p_Z)(I_{Z,jk} - p_Z)\Big)^2 \\ &+ \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{\substack{q=1 \\ q\neq j, q\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z)(I_{Z,jk} - p_Z)(I_{Z,ql} - p_Z) \\ &= C_{11} + C_{12}. \end{aligned}$$

As $q \neq j$ and $q \neq i$ we conclude that $C_{12} = 0$ because $I_{Z,ql}$ is independent to all $I_{Z,ik}$, $I_{Z,il}$ and $I_{Z,jk}$. On the other hand,

$$C_{11} = \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z)E(I_{Z,jk} - p_Z)(I_{Z,jl} - p_Z).$$

Note, again, that for $|k-l| > M$ $C_{11} = 0$ and so

$$C_{11} = \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} E(I_{Z,ik} - p_Z)(I_{Z,il} - p_Z)E(I_{Z,jk} - p_Z)(I_{Z,jl} - p_Z).$$

The asymptotically dominant term in C_{11} is the same as in C_2. Therefore, C_1 is asymptotically bounded by $2Mm_nN_n^2p_Z^2$. We plug that estimate into (3.24) and then we get

$$P(|C_n| > \varepsilon) \leq \frac{2Mm_nN_n^2p_Z^2}{N_n^4\varepsilon^2} \sim \frac{1}{\varepsilon^2 N_n^2}. \tag{3.25}$$

Using the same arguments as in the proofs of the a) and b) parts of this lemma, we conclude that $C_n \to 0$ almost surely, for $n \to \infty$. By the same argument c') holds.

d) Using the similar ideas as in the proofs of b) and c) parts of this lemma we will show that

$$D_n = \frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq q}}^{N_n} \sum_{q=1}^{N_n} \sum_{k=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)(I_{Z,qk} - p_Z)$$

is converging to zero almost surely, as $n \to \infty$. Choose $\varepsilon > 0$. Then by the Chebyshev inequality we have

$$P(|D_n| > \varepsilon) \leq \frac{1}{N_n^6\varepsilon^2} E\Big(\sum_{i=1}^{N_n} \sum_{\substack{j=1 \\ j\neq q}}^{N_n} \sum_{q=1}^{N_n} \sum_{k=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)(I_{Z,qk} - p_Z)\Big)^2. \tag{3.26}$$

Expectation on the right hand side of (3.26) is equal to

$$\sum_{i=1}^{N_n} E\Big(\sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ q\neq j}}^{N_n} \sum_{k=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)(I_{Z,qk} - p_Z)\Big)^2$$
$$+ \sum_{\substack{i=1 \\ i\neq p}}^{N_n} \sum_{p=1}^{N_n} \sum_{\substack{j=1 \\ j\neq q}}^{N_n} \sum_{q=1}^{N_n} \sum_{\substack{r=1 \\ r\neq s}}^{N_n} \sum_{s=1}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E\Big((I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)(I_{Z,qk} - p_Z)$$
$$\cdot (I_{Y,pl} - p_Y)(I_{Z,rl} - p_Z)(I_{Z,sl} - p_Z)\Big)$$
$$= D_1 + D_2.$$

Although the term D_2 looks quite complex at first, we note that $I_{Y,ik}$ and $I_{Y pl}$ are independent, for $i \neq p$, and both are independent to process (Z_i). Therefore, the expectation under those sums in D_2 is equal to

$$E(I_{Y,ik}-p_Y)E(I_{Y,pl}-p_Y)E(I_{Z,jk}-p_Z)(I_{Z,qk}-p_Z)(I_{Z,rl}-p_Z)(I_{Z,sl}-p_Z)$$

which is equal to zero and so, D_2 is equal to zero. We continue in the same spirit with the calculation of D_1. We get

$$\begin{aligned}
D_1 &= \sum_{i=1}^{N_n}\sum_{j=1}^{N_n} E\Big(\sum_{\substack{q=1\\q\neq j}}^{N_n}\big(\sum_{k=1}^{m_n}(I_{Y,ik}-p_Y)(I_{Z,jk}-p_Z)(I_{Z,qk}-p_Z)\big)\Big)^2 \\
&+ \sum_{i=1}^{N_n}\sum_{\substack{j=1\\r\neq j}}^{N_n}\sum_{r=1}^{N_n}\sum_{\substack{q=1\\q\neq j}}^{N_n}\sum_{\substack{s=1\\s\neq r}}^{N_n}\sum_{k=1}^{m_n}\sum_{l=1}^{m_n} E(I_{Y,ik}-p_Y)(I_{Y,il}-p_Y) \\
&\qquad\qquad \cdot E(I_{Z,jk}-p_Z)(I_{Z,qk}-p_Z)(I_{Z,rl}-p_Z)(I_{Z,sl}-p_Z) \\
&= D_3 + D_4.
\end{aligned}$$

If we look at D_4 we see that $r \neq j$ and $q \neq j$. If we additionally suppose that $s \neq j$ $I_{Z,jk}$ will be independent with all $I_{Z,qk}$, $I_{Z,rl}$ and $I_{Z,sl}$. In that case D_4 will be just zero. So, we can suppose that $s = j$ and in that case D_4 is equal to

$$\begin{aligned}
&\sum_{i=1}^{N_n}\sum_{\substack{j=1\\r\neq j}}^{N_n}\sum_{r=1}^{N_n}\sum_{\substack{q=1\\q\neq j}}^{N_n}\sum_{k=1}^{m_n}\sum_{l=1}^{m_n} E(I_{Y,ik}-p_Y)(I_{Y,il}-p_Y) \\
&\qquad\qquad \cdot E(I_{Z,jk}-p_Z)(I_{Z,jl}-p_Z)(I_{Z,qk}-p_Z)(I_{Z,rl}-p_Z).
\end{aligned}$$

Applying similar arguments we see that for $r \neq q$ the whole sum is equal to zero. For $r = q$ we get

$$\begin{aligned}
D_4 &= \sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{q=1\\q\neq j}}^{N_n}\sum_{k=1}^{m_n}\sum_{l=1}^{m_n} E(I_{Y,ik}-p_Y)(I_{Y,il}-p_Y) \\
&\qquad\qquad \cdot E(I_{Z,jk}-p_Z)(I_{Z,jl}-p_Z)E(I_{Z,qk}-p_Z)(I_{Z,ql}-p_Z).
\end{aligned}$$

The asymptotically dominant terms in the above expression is

$$\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{q=1\\q\neq j}}^{N_n}\sum_{k=1}^{m_n}\sum_{l=1}^{m_n} E(I_{Y,ik}I_{Y,il})E(I_{Z,jk}I_{Z,jl})E(I_{Z,qk}I_{Z,ql}),$$

and it is bounded by $m_n^2 N_n^3 p_Y p_Z^2$. For the sum D_3 and apply the same idea to get

$$D_3 = \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ q\neq j}}^{N_n} E\Big(\big(\sum_{k=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Z,jk} - p_Z)(I_{Zqk} - p_Z) \big) \Big)^2$$

$$+ \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ q\neq j}}^{N_n} \sum_{\substack{r=1 \\ r\neq j, r\neq q}}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,il} - p_Y)$$

$$\cdot E(I_{Z,jk} - p_Z)(I_{Z,jl} - p_Z)(I_{Z,qk} - p_Z)(I_{Z,rl} - p_Z).$$

In the second sum the term $I_{Z,rl}$ is independent to all $I_{Z,jk}$, $I_{Z,jl}$ and $I_{Z,qk}$. Hence, the whole sum is equal to zero. Therefore, we have

$$D_3 = \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ q\neq j}}^{N_n} \sum_{k=1}^{m_n} \sum_{l=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,il} - p_Y)$$

$$\cdot E(I_{Z,jk} - p_Z)(I_{Z,jl} - p_Z) E(I_{Z,qk} - p_Z)(I_{Z,ql} - p_Z)$$

The asymptotically dominant term in the sum D_3 is the same for the sum D_4. We conclude that the whole expression on the right hand side of (3.26) is asymptotically dominated by $\sqrt{m_n}/(\varepsilon^2 N_n^3)$ and then we easily apply the Borel-Cantelli lemma to prove D_n converges to zero almost surely, for $n \to \infty$. That proves part $d)$ of the lemma. ■

Lemma 3.3.3. With same assumptions as in Theorem 3.1.1 we have:

a)

$$\frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|\leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y)(I_{Z,jk} I_{Z,jl} - p_{kl}^Z) \to 0,\ n \to \infty \text{ (a.s.)}$$

b)

$$\frac{1}{N_n^2} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{k=1 \\ |k-l|\leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y) p_{kl}^Z \to 0,\ n \to \infty \text{ (a.s.)}$$

c)

$$\frac{1}{N_n^2} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|\leq M}}^{m_n} \sum_{l=1}^{m_n} p_Y (I_{Y,ik} - p_Y)(I_{Z,jk} I_{Z,jl} - p_{kl}^Z) \to 0,\ n \to \infty \text{ (a.s.)}$$

d)

$$\frac{1}{N_n}\sum_{i=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n} p_Y p_{kl}^Z (I_{Y,ik}-p_Y) \to 0,\ n\to\infty \ \text{(a.s.)}.$$

e)

$$\frac{1}{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n} p_Y^2 (I_{Z,jk}I_{Z,jl}-p_{kl}^Z) \to 0,\ n\to\infty \ \text{(a.s.)}.$$

Proof of Lemma 3.3.3. **a)** We will show that

$$K_1 = \frac{1}{N_n^3}\sum_{i=1}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n}(I_{Y,ik}-p_Y)(I_{Y,pl}-p_Y)(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)$$

is converging to zero almost surely. Let $\varepsilon > 0$ be arbitrarily chosen. Then by the Chebyshev inequality we have

$$P(|K_1|>\varepsilon) \le \frac{1}{N_n^6\varepsilon^2}E\Big(\sum_{i=1}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n}(I_{Y,ik}-p_Y)(I_{Y,pl}-p_Y)(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)\Big)^2. \tag{3.27}$$

Expectation on the right hand side of (3.27) is equal to

$$\sum_{i=1}^{N_n}E\Big(\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n}(I_{Y,ik}-p_Y)(I_{Y,pl}-p_Y)(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)\Big)^2$$
$$+\sum_{i=1}^{N_n}\sum_{\substack{t=1\\t\neq i}}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{u=1\\u\neq t}}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|\le M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Y,ik}-p_Y)(I_{Y,tr}-p_Y)(I_{Y,pl}-p_Y)(I_{Y,us}-p_Z)$$
$$\cdot E(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)(I_{Z,qr}I_{Z,qs}-p_{rs}^Z)$$
$$=K_{12}+K_{13}.$$

Recall that (Y_i) and (Z_i) are independent. Note first that for $q\neq j$

$$E(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)(I_{Z,qr}I_{Z,qs}-p_{rs}^Z)=E(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)E(I_{Z,qr}I_{Z,qs}-p_{rs}^Z)=0$$

and then $K_{13}=0$. Therefore

$$K_{13}=\sum_{i=1}^{N_n}\sum_{\substack{t=1\\t\neq i}}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{u=1\\u\neq t}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\le M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|\le M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Y,ik}-p_Y)(I_{Y,tk}-p_Y)(I_{Y,pl}-p_Y)(I_{Y,ul}-p_Z)$$
$$\cdot E(I_{Z,jk}I_{Z,jl}-p_{kl}^Z)(I_{Z,jr}I_{Z,js}-p_{rs}^Z).$$

Next, note that $i \neq t$ and $i \neq p$. If additionally $i \neq u$ we conclude that $I_{Y,ik}$ is independent to all $I_{Y,tk}$, $I_{Y,pl}$ and $I_{Y,ul}$. Hence, $E(I_{Y,ik}-p_Y)(I_{Y,tr}-p_Y)(I_{Y,pl}-p_Y)(I_{Y,us}-p_Z)=0$ and the whole K_{13} is equal to zero. Let $u=i$. Then

$$K_{13}=\sum_{i=1}^{N_n}\sum_{\substack{t=1\\t\neq i}}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|\leq M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Y,ik}-p_Y)(I_{Y,ul}-p_Z)E(I_{Y,tk}-p_Y)(I_{Y,pl}-p_Y)$$
$$\cdot E(I_{Z,jk}I_{Z,jl}-p^Z_{kl})(I_{Z,jr}I_{Z,js}-p^Z_{rs}).$$

For $p\neq t$ we again conclude $K_{13}=0$. Therefore,

$$K_{13}=\sum_{i=1}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|\leq M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Y,ik}-p_Y)(I_{Y,ul}-p_Z)E(I_{Y,pk}-p_Y)(I_{Y,pl}-p_Y)$$
$$\cdot E(I_{Z,jk}I_{Z,jl}-p^Z_{kl})(I_{Z,jr}I_{Z,js}-p^Z_{rs}).$$

The asymptotically dominant term in K_{13} is

$$\sum_{i=1}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|\leq M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Y,ik}I_{Y,ul})E(I_{Y,pk}I_{Y,pl})E(I_{Z,jk}I_{Z,jl}I_{Z,jr}I_{Z,js})$$

and it can be bounded by $N_n^3 4M^2 m_n^2 p_Y 2p_Z$.

In the same vein we calculate

$$K_{12}=\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}E\Big(\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{k=1\\|k-l|\leq M}}^{m_n}\sum_{l=1}^{m_n}(I_{Y,ik}-p_Y)(I_{Y,pl}-p_Y)(I_{Z,jk}I_{Z,jl}-p^Z_{kl})\Big)^2$$
$$+\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{q=1\\q\neq j}}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{u=1\\u\neq i}}^{N_n}\sum_{\substack{k=1\\|k-l|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|\leq M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Y,ik}-p_Y)(I_{Y,ir}-p_Y)(I_{Y,pl}-p_Y)(I_{Y,us}-p_Z)$$
$$\cdot E(I_{Z,jk}I_{Z,jl}-p^Z_{kl})(I_{Z,qr}I_{Z,qs}-p^Z_{rs})$$
$$=K_{14}+K_{15}.$$

As now $q\neq j$ we conclude

$$E(I_{Z,jk}I_{Z,jl}-p^Z_{kl})(I_{Z,qr}I_{Z,qs}-p^Z_{rs})=E(I_{Z,jk}I_{Z,jl}-p^Z_{kl})E(I_{Z,qr}I_{Z,qs}-p^Z_{rs})=0$$

and so $K_{15} = 0$. For K_{14} we have

$$K_{14} = \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} E\Big(\sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y)(I_{Z,jk} I_{Z,jl} - p^Z_{kl}) \Big)^2$$
$$+ \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{u=1 \\ u \neq i, u \neq p}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s| \leq M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,ir} - p_Y)(I_{Y,pl} - p_Y)(I_{Y,us} - p_Z)$$
$$\cdot E(I_{Z,jk} I_{Z,jl} - p^Z_{kl})(I_{Z,jr} I_{Z,js} - p^Z_{rs})$$
$$= K_{15} + K_{16}.$$

We can again conclude that $K_{16} = 0$, because $I_{Y,us}$ is independent to all $I_{Y,ik}$, $I_{Y,ir}$ and $I_{Y,pl}$. On the other hand K_{15} is bounded by $N_n^3 4M^2 m_n^2 p_Y 2 p_Z$ (note the same bound as for the sum K_{13}). Because of (3.27), we have

$$P(|K_1| > \varepsilon) \leq \frac{N_n^3 4M^2 m_n^2 p_Y^2 p_Z}{\varepsilon^2 N_n^6} \sim \frac{m_n^{1/2}}{N_n^3} \cdot \frac{m_n^2}{m_n^2} \sim \frac{1}{\varepsilon^2} \frac{m_n^{5/2}}{n^2} < \frac{1}{\varepsilon^2 n^{1+\eta}}, \; n \to \infty,$$

for some $\eta > 0$. By the Borel-Cantelli lemma then follows **a)**.

b) We will show that

$$K_2 = \frac{1}{N_n^2} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y) p^Z_{kl},$$

is converging to zero almost surely. Choose $\varepsilon > 0$ arbitrarily. Then by the Chebyshev inequality we have

$$P(|K_2| > \varepsilon) \leq \frac{1}{N_n^4 \varepsilon^2} E\Big(\sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y) p^Z_{kl} \Big)^2. \tag{3.28}$$

Expectation on the right hand side of (3.28) is equal to

$$\sum_{i=1}^{N_n} E\Big(\sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y) p^Z_{kl} \Big)^2$$
$$+ \sum_{i=1}^{N_n} \sum_{\substack{t=1 \\ t \neq i}}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{u=1 \\ u \neq t}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s| \leq M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,tr} - p_Y)(I_{Y,pl} - p_Y)(I_{Y,us} - p_Z) p^Z_{kl} p^Z_{rs}$$
$$= K_{22} + K_{23}.$$

For sums in K_{23} for which $u \neq i$ we have $i \neq t$, $i \neq p$ and $i \neq u$ and then we conclude that $I_{Y,ik}$ is independent to all $I_{Y,tk}$, $I_{Y,pl}$ and $I_{Y,ul}$. Hence, $E(I_{Y,ik} - p_Y)(I_{Y,tr} - p_Y)(I_{Y,pl} - p_Y)(I_{Y,us} - p_Z) = 0$ and the whole K_{23} is equal to zero. For $u = i$ we get

$$K_{23} = \sum_{i=1}^{N_n} \sum_{\substack{t=1 \\ t\neq i}}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} \sum_{r=1}^{m_n} \sum_{\substack{s=1 \\ |r-s|\leq M}}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,is} - p_Z)E(I_{Y,tr} - p_Y)(I_{Y,pl} - p_Y)p^Z_{kl}p^Z_{rs}.$$

If $t \neq p$ we again get $K_{23} = 0$. Therefore,

$$K_{23} = \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} \sum_{r=1}^{m_n} \sum_{\substack{s=1 \\ |r-s|\leq M}}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,is} - p_Z)E(I_{Y,pr} - p_Y)(I_{Y,pl} - p_Y)p^Z_{kl}p^Z_{rs}.$$

The asymptotically dominant term in K_{23} is then

$$\sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} \sum_{r=1}^{m_n} \sum_{\substack{s=1 \\ |r-s|\leq M}}^{m_n} E(I_{Y,ik}I_{Y,is})E(I_{Y,pr}I_{Y,pl})p^Z_{kl}p^Z_{rs}$$

and it can be bounded by $N_n^2 4M^2 m_n^2 p_Y^2 p_Z^2$.

Let us calculate K_{22} further. We get

$$\begin{aligned} K_{22} = &\sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} E\Big(\sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y)p^Z_{kl}\Big)^2 \\ &+ \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{u=1 \\ u\neq i, u\neq p}}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} \sum_{r=1}^{m_n} \sum_{\substack{s=1 \\ |r-s|\leq M}}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,ir} - p_Y)(I_{Y,pl} - p_Y)(I_{Y,us} - p_Z)p^Z_{kl}p^Z_{rs} \\ &= K_{24} + K_{25}. \end{aligned}$$

We use the fact that $I_{Y,us}$ is independent to all $I_{Y,ik}$, $I_{Y,ir}$ and $I_{Y,pl}$ to conclude that $K_{25} = 0$. For K_{24} we get the same bound as for K_{23}. Plug that bound back to (3.28) to get

$$P(|K_2| > \varepsilon) \leq \frac{N_n^2 4M^2 m_n^2 p_Z^2 p_Y^2}{\varepsilon^2 N_n^4} \sim \frac{1}{\varepsilon^2 N_n^2} \cdot \frac{m_n^2}{m_n^2} \sim \frac{1}{\varepsilon^2} \frac{m_n^2}{n^2} < \frac{1}{\varepsilon^2 n^{1+\eta}},\ n \to \infty,$$

for some $\eta > 0$. By the Borel-Cantelli lemma we prove **b)** part of this lemma. In the same way we can prove **c)** part of the lemma so we skip that proof. Also, proofs of **d)** and **e)** parts of the lemma are almost the same so we only present the proof for the **d)** part.

d) We will show that

$$K_3 = \frac{1}{N_n} \sum_{i=1}^{N_n} \sum_{k=1}^{m_n} \sum_{\substack{l=1 \\ |k-l|\leq M}}^{m_n} p_Y p^Z_{kl}(I_{Y,ik} - p_Y)$$

is converging to zero almost surely. Choose $\varepsilon > 0$ arbitrarily. Then by the Chebyshev inequality we have

$$P(|K_3| > \varepsilon) \leq \frac{1}{N_n^2 \varepsilon^2} E\Big(\sum_{i=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y) p_Y p_{kl}^Z \Big)^2. \tag{3.29}$$

Expectation on the right hand side of (3.29) is equal to

$$\begin{aligned} &\sum_{i=1}^{N_n} E\Big(\sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y) p_Y p_{kl}^Z \Big)^2 \\ &+ \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s| \leq M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,pr} - p_Y) p_Y^2 p_{kl}^Z p_{rs}^Z \\ &= K_{32} + K_{33}. \end{aligned}$$

Clearly, $K_{33} = 0$ because $I_{Y,ik}$ and $I_{Y,pr}$ are independent. Then

$$K_{32} = p_Y^2 \sum_{i=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s| \leq M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,ir} - p_Y) p_{kl}^Z p_{rs}^Z.$$

For k and r such that $|k-r| > M$ we conclude that $I_{Y,ik}$ and $I_{Y,ir}$ are independent and therefore $K_{32} = 0$. So, we may write

$$K_{32} = p_Y^2 \sum_{i=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M \\ |k-r| \leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s| \leq M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Y,ik} - p_Y)(I_{Y,ir} - p_Y) p_{kl}^Z p_{rs}^Z.$$

Note that all indices l, r and s depend on k. Therefore, we have at most $16M^3 m_n$ members of the sums over k, l, r and s (for each i). Furthermore, the asymptotically dominant terms of K_{32} are $E(I_{Y,ik} I_{Y,ir})$ and they can be bounded by p_Y. Hence, we may asymptotically bound the whole K_{32} by $N_n 16 M^3 m_n p_Y^3 p_Z^2$. Consequently, in (3.29) we get the inequality

$$P(|K_3| > \varepsilon) \leq \frac{N_n 16 M^3 m_n p_Y^3 p_Z^2}{N_n^2 \varepsilon^2} \sim \frac{p_Y p_Z^2}{\varepsilon^2 N_n} \sim \frac{1}{\varepsilon^2} \frac{1}{N_n m_n} \frac{1}{\sqrt{m_n}} \sim \frac{1}{\varepsilon^2} \frac{1}{n} \frac{1}{\sqrt{m_n}} < \frac{1}{\varepsilon^2 n^{1+\eta}},$$

for some $\eta > 0$, $n \to \infty$. Use again the Borel-Cantelli lemma to conclude that part **d)** of the lemma holds. ■

Lemma 3.3.4. With the same assumptions as in the Theorem 3.1.1 we have:

a)
$$\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\to 0,\ n\to\infty\ \text{(a.s.)}$$

b)
$$\frac{1}{N_n^3}\sum_{i=1}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Z,jk}I_{Z,ql}-p_Z^2\right)\to 0,\ n\to\infty\ \text{(a.s.)}$$

c)
$$\frac{1}{N_n^3}\sum_{\substack{i=1\\i\neq p}}^{N_n}\sum_{p=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,pl}-p_Y^2\right)\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\to 0,\ n\to\infty\ \text{(a.s.)}$$

d)
$$\frac{1}{N_n^4}\sum_{\substack{i=1\\i\neq p}}^{N_n}\sum_{p=1}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,pl}-p_Y^2\right)\left(I_{Z,jk}I_{Z,ql}-p_Z^2\right)\to 0,\ n\to\infty\ \text{(a.s.)}.$$

Proof of Lemma 3.3.4. Proof is similar to the proof of previous two lemmas. We will use the Chebyshev inequality and the Borel-Cantelli lemma to validate stated almost sure convergences.

a) Choose $\varepsilon > 0$ arbitrarily. Let

$$L_1=\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)$$

Than we have, by the Chebyshev inequality,

$$P\Big(\big|L_1\big|>\varepsilon\Big)\le\frac{1}{\varepsilon^2N_n^4}E\Big(\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\Big)^2 \tag{3.30}$$

The expectation on the right hand side of (3.30) is equal to

$$\begin{aligned}
&\sum_{i=1}^{N_n}E\Big(\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\Big)^2\\
&+\sum_{i=1}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n}E\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Y,pr}I_{Y,ps}-p_Y^2\right)\\
&\qquad\cdot E\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\left(I_{Z,qr}I_{Z,qs}-p_Z^2\right)\\
&=L_{12}+L_{13}.
\end{aligned}$$

We note that (Y_i) and (Z_i) are independent by assumption and for $i \neq p$ $I_{Y,ik}I_{Y,il}$ and $I_{Y,pr}I_{Y,ps}$ are independent too. Therefore, $L_{13} = 0$ We continue calculations with L_{12} to see it is equal to

$$\begin{aligned}
&\sum_{i=1}^{N_n}\sum_{j=1}^{N_n} E\Big(\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\Big)^2\\
&+\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{q=1\\q\neq j}}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n} E\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Y,ir}I_{Y,is}-p_Y^2\right)\\
&\qquad\qquad\cdot E\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\left(I_{Z,qr}I_{Z,qs}-p_Z^2\right)\\
&= L_{14}+L_{15}.
\end{aligned}$$

As $q \neq j$ in L_{15}, we conclude that $L_{15} = 0$. Finally,

$$L_{14} = \sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n} E\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Y,ir}I_{Y,is}-p_Y^2\right) \cdot E\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\left(I_{Z,jr}I_{Z,js}-p_Z^2\right) \tag{3.31}$$

Observe, if $|k-r| > M$, $|k-s| > M$, $|l-r| > M$ and $|l-s| > M$, then $I_{Y,ik}I_{Y,il}$ and $I_{Y,ir}I_{Y,is}$ are independent and therefore $L_{14} = 0$. Otherwise, without loss of generality, let us suppose $|k-r| \leq M$. Then

$$L_{14} = \sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M\\|k-r|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n} E\left(I_{Y,ik}I_{Y,il}-p_Y^2\right)\left(I_{Y,ir}I_{Y,is}-p_Y^2\right) \cdot E\left(I_{Z,jk}I_{Z,jl}-p_Z^2\right)\left(I_{Z,jr}I_{Z,js}-p_Z^2\right) \tag{3.32}$$

The asymptotically dominant term in (3.32) is

$$\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M\\|k-r|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n} E\left(I_{Y,ik}I_{Y,il}I_{Y,ir}I_{Y,is}\right)E\left(I_{Z,jk}I_{Z,jl}I_{Z,jr}I_{Z,js}\right). \tag{3.33}$$

The other terms contain factors of the form p_Y^2 or p_Z^2 that make convergence of those other terms to zero faster than the term in (3.33). We use only modest estimates

$$E\left(I_{Y,ik}I_{Y,il}I_{Y,ir}I_{Y,is}\right) \leq p_Y \quad \text{and} \quad E\left(I_{Z,jk}I_{Z,jl}I_{Z,jr}I_{Z,js}\right) \leq p_Z$$

to conclude that L_{14} may be asymptotically bounded by $2MN_n^2m_n^3p_Yp_Z$. Apply that bound in (3.30) to get the asymptotic bound for the sum on the right hand side of (3.30):

$$\frac{2MN_n^2m_n^3p_Yp_Z}{\varepsilon^2N_n^4} \sim \frac{m_n^2}{\varepsilon^2N_n^2}\cdot\frac{m_n^2}{m_n^2} \sim \frac{1}{\varepsilon^2}\frac{m_n^4}{n}\cdot\frac{1}{n} < \frac{1}{\varepsilon^2}\frac{1}{n^{1+\eta}},$$

for some $\eta > 0$, because of the assumption (3.1). Together with Borel-Cantelli lemma that proves part $a)$ of the lemma.

d) Choose $\varepsilon > 0$ arbitrarily. Let

$$L_4 = \frac{1}{N_n^4}\sum_{\substack{i=1\\i\neq p}}^{N_n}\sum_{p=1}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\big(I_{Y,ik}I_{Y,pl}-p_Y^2\big)\big(I_{Z,jk}I_{Z,ql}-p_Z^2\big).$$

By the Chebyshev inequality we have

$$P\Big(\big|L_4\big| > \varepsilon\Big) \leq \frac{1}{\varepsilon^2N_n^8}E\Big(\sum_{\substack{i=1\\i\neq p}}^{N_n}\sum_{p=1}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\big(I_{Y,ik}I_{Y,pl}-p_Y^2\big)\big(I_{Z,jk}I_{Z,ql}-p_Z^2\big)\Big)^2. \quad (3.34)$$

The expectation on the right hand side of the above inequality is equal to

$$\begin{aligned}&\sum_{i=1}^{N_n}E\Big(\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\big(I_{Y,ik}I_{Y,pl}-p_Y^2\big)\big(I_{Z,jk}I_{Z,ql}-p_Z^2\big)\Big)^2\\
&+\sum_{i=1}^{N_n}\sum_{\substack{t=1\\t\neq i}}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{u=1\\u\neq t}}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{v=1}^{N_n}\sum_{\substack{w=1\\w\neq v}}^{N_n}\sum_{\substack{k=1\\|k-l|>M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Yik}I_{Ypl}-p_Y^2)(I_{Ytr}I_{Yus}-p_Y^2)\\
&\qquad\cdot E(I_{Zjk}I_{Zql}-p_Z^2)(I_{Zvr}I_{Zws}-p_Z^2)\\
&=L_{41}+L_{42}.\end{aligned}$$

Observe, for the sums in which $t \neq p$, $u \neq i$ and $u \neq p$, indicators $I_{Yik}I_{Ypl}$ and $I_{Ytr}I_{Yus}$ are independent. In that case $L_{42} = 0$. Without loss of generality suppose $t = p$. If additionally $|k-r| > M$, $|k-s| > M$, $|l-r| > M$ and $|l-s| > M$, then $I_{Yik}I_{Ypl}$ and $I_{Ytr}I_{Yus}$ are independent and so $L_{42} = 0$. Again, without loss of generality suppose $|k-r| \leq M$. Then

$$\begin{aligned}L_{42} = &\sum_{i=1}^{N_n}\sum_{\substack{p=1\\p\neq i}}^{N_n}\sum_{\substack{u=1\\u\neq p}}^{N_n}\sum_{\substack{j=1\\j\neq q}}^{N_n}\sum_{q=1}^{N_n}\sum_{v=1}^{N_n}\sum_{\substack{w=1\\w\neq v}}^{N_n}\sum_{\substack{k=1\\|k-l|>M\\|k-r|\leq M}}^{m_n}\sum_{l=1}^{m_n}\sum_{\substack{r=1\\|r-s|>M}}^{m_n}\sum_{s=1}^{m_n}E(I_{Yik}I_{Ypl}-p_Y^2)(I_{Ypr}I_{Yus}-p_Y^2)\\
&\qquad\cdot E(I_{Zjk}I_{Zql}-p_Z^2)(I_{Zvr}I_{Zws}-p_Z^2).\end{aligned}$$

If, additionally, $j \neq v$, $j \neq w$, $q \neq v$ and $q \neq w$, then $I_{Zjk}I_{Zql}$ and $I_{Zvr}I_{Zws}$ are independent and therefore $L_{42} = 0$. So, without loss of generality suppose $w = q$. Then,

$$L_{42} = \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{u=1 \\ u\neq p}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ j\neq q}}^{N_n} \sum_{\substack{v=1 \\ v\neq j}}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M \\ |k-r|\leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s|>M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Yik}I_{Ypl} - p_Y^2)(I_{Ypr}I_{Yus} - p_Y^2) \cdot E(I_{Zjk}I_{Zql} - p_Z^2)(I_{Zvr}I_{Zqs} - p_Z^2).$$

The asymptotically dominant term in L_{42} is

$$\sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{u=1 \\ u\neq p}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ j\neq q}}^{N_n} \sum_{\substack{v=1 \\ v\neq j}}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M \\ |k-r|\leq M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s|>M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Yik}I_{Ypl}I_{Ypr}I_{Yus})E(I_{Zjk}I_{Zql}I_{Zvr}I_{Zqs}).$$

We use estimates

$$E\big(I_{Yik}I_{Ypl}I_{Ypr}I_{Yus}\big) \leq p_Y^2 \quad \text{and} \quad E\big(I_{Zjk}I_{Zql}I_{Zvr}I_{Zqs}\big) \leq p_Z^2$$

to conclude that L_{42} may be asymptotically bounded by $2Mm_n^3N_n^6p_Y^2p_Z^2$.

Let us calculate L_{41} further. We get

$$\begin{aligned} L_{41} = &\sum_{i=1}^{N_n} \sum_{j=1}^{N_n} E\Big(\sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{q=1 \\ j\neq q}}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(I_{Y,ik}I_{Y,pl} - p_Y^2\big)\big(I_{Z,jk}I_{Z,ql} - p_Z^2\big)\Big)^2 \\ &+ \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{u=1 \\ u\neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ j\neq q}}^{N_n} \sum_{\substack{v=1 \\ v\neq j}}^{N_n} \sum_{\substack{w=1 \\ w\neq v}}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s|>M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Yik}I_{Ypl} - p_Y^2)(I_{Yir}I_{Yus} - p_Y^2) \\ &\qquad \cdot E(I_{Zjk}I_{Zql} - p_Z^2)(I_{Zvr}I_{Zws} - p_Z^2) \\ = &\, L_{43} + L_{44}. \end{aligned}$$

For $j \neq w$, $q \neq v$ and $q \neq w$, $L_{44} = 0$. Without loss of generality suppose $v = q$. Similar reasoning employed in estimation of the asymptotic behavior of L_{42} allows for the additional assumption, without loss of generality, $|k-r| \leq M$. Therefore,

$$L_{44} = \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p\neq i}}^{N_n} \sum_{\substack{u=1 \\ u\neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{q=1 \\ j\neq q}}^{N_n} \sum_{\substack{w=1 \\ w\neq q}}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \sum_{\substack{r=1 \\ |r-s|>M}}^{m_n} \sum_{s=1}^{m_n} E(I_{Yik}I_{Ypl} - p_Y^2)(I_{Yir}I_{Yus} - p_Y^2) \cdot E(I_{Zjk}I_{Zql} - p_Z^2)(I_{Zvr}I_{Zws} - p_Z^2)$$

We conclude that we can asymptotically bound sum L_{44} by the same bound we obtained for L_{42}. Note, if we continue the same procedure to calculate L_{43} we will get less and less sums with the range of the indices between 1 and N_n. Such sums are asymptotically inferior to sums L_{42} and L_{44}. Therefore, we can asymptotically bound the sum on the right hand side of (3.34) by $2Mm_n^3N_n^6p_Y^2p_Z^2$, i.e. we get

$$P\Big(|L_4| > \varepsilon\Big) \leq \frac{2Mm_n^3N_n^6p_Y^2p_Z^2}{\varepsilon^2N_n^8} \sim \frac{m_n}{\varepsilon^2N_n^2}\cdot\frac{m_n^2}{m_n^2} \sim \frac{1}{\varepsilon^2}\frac{m_n^3}{n}\frac{1}{n} < \frac{1}{\varepsilon^2n^{1+\eta}},\ n\to\infty,$$

for some $\eta > 0$. By the Borel-Cantelli lemma we conclude part **d)** of the lemma holds.

Parts **b)** and **c)** of the lemma can be proven in a very similar way so we skip their proofs. ■

Proof of Theorem 3.1.1. The proof is similar to the proof we gave in the i.i.d. case. Additional complexity stems from the fact that we are now dealing with sums (over k) of dependent random variables. Instead of previously defined a_n and b_n from Theorem 2.1.2 in this case we define

$$c_{N_n}(i,j) := \frac{m_n}{n}\sum_{k=1}^{m_n} I_{Y,ik}I_{Z,jk}, \tag{3.35}$$

where $i,j = 1,\dots,N_n$. We see that $c_{N_n}(i,j)$ counts joint up crossings of the Y and Z over level $u_{\sqrt{m_n}}$ in blocks numbered by i and j. Recall the assumption $n = N_n(m_n+M)$ to see c_{N_n} depends on N_n. According to Theorem 1.2.3 we need to check the condition

$$N_n\frac{\max_{1\leq i,j\leq N_n} d_{N_n}^2(i,j)}{\sum_{i=1}^{N_n}\sum_{j=1}^{N_n} d_{N_n}^2(i,j)} \to 0 \quad \text{(a.s.)} \tag{3.36}$$

where $d_{N_n}(i,j)$ are defined in (1.17), and $N_n\to\infty$. We will prove (3.36) for $n\to\infty$ in which case both N_n and m_n are growing to infinity. By Lemma 3.3.1 from Section 3.3 we get

$$d_{N_n}(i,j) = \frac{m_n}{n}\sum_{k=1}^{m_n}(I_{Z,jk}-\bar{I}_{n,k}^Z)(I_{Y,ik}-\bar{I}_{n,k}^Y).$$

As $n = N_n(m_n+M)$ and $m_n/(m_n+M) \leq 1$ we get

$$\frac{m_n}{n} \leq \frac{1}{N_n}. \tag{3.37}$$

Because all $I_{Y,ik}$, $I_{Z,jk}$, $\bar{I}_{n,k}^Y$ and $\bar{I}_{n,k}^Z$ are almost surely less or equal to 1 the same is true for $|I_{Z,jk}-\bar{I}_{n,k}^Z|$ and $|I_{Y,ik}-\bar{I}_{n,k}^Y|$. Therefore,

$$|d_{N_n}(i,j)| \leq \frac{1}{N_n}\sum_{k=1}^{m_n}|I_{Z,jk}-\bar{I}_{n,k}^Z|\cdot|I_{Y,ik}-\bar{I}_{n,k}^Y| \leq \frac{m_n}{N_n}$$

and so

$$d_{N_n}^2(i,j) \leq \frac{m_n^2}{N_n^2}.$$

We conclude that the whole numerator in (3.36) is bounded by m_n^2/N_n and that bound is converging to zero as we show in (3.5). Summarizing the above considerations we conclude that the numerator in (3.36) is converging to zero almost surely, for $n \to \infty$.

Let us look at the denominator in (3.36). Use Lemma 3.3.1 to conclude that the sum $\sum_{i=1}^{N_n}\sum_{j=1}^{N_n} d_{N_n}^2(i,j)$ is equal to

$$\frac{m_n^2}{n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Z,jk}-\bar{I}_{n,k}^Z)\Big)^2.$$

We use again the fact that $m_n/(m_n+M) \to 1$, as $n \to \infty$, to conclude that asymptotically m_n/n is actually behaving like $1/N_n$. Therefore, we have asymptotic estimate for the sum $\sum_{i=1}^{N_n}\sum_{j=1}^{N_n} d_{N_n}^2(i,j)$ in the form

$$\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Z,jk}-\bar{I}_{n,k}^Z)\Big)^2. \tag{3.38}$$

Formally, we can choose n large enough such that the denominator is arbitrarily close to the last expression, whether that expression converges or diverges to ∞. For the reason of simplicity we will analyse that expression in what follows instead of the expression with the factor m_n^2/n^2. Let us look closer what happens when we square the sum over k in (3.38). We get

$$\begin{aligned}\Big(\sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Z,jk}-\bar{I}_{n,k}^Z)\Big)^2 &= \sum_{k=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)^2(I_{Z,jk}-\bar{I}_{n,k}^Z)^2\\ &+\sum_{\substack{k=1\\k\neq l}}^{m_n}\sum_{l=1}^{m_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Z,jk}-\bar{I}_{n,k}^Z)(I_{Y,il}-\bar{I}_{n,l}^Y)(I_{Zjl}-\bar{I}_{n,l}^Z).\end{aligned}$$

Therefore, from (3.38) we see that $\sum_{i=1}^{N_n}\sum_{j=1}^{N_n} d_{N_n}^2(i,j)$ is asymptotically behaving as

$$\begin{aligned}&\sum_{k=1}^{m_n}\Big(\frac{1}{N_n}\sum_{i=1}^{N_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)^2\Big)\Big(\frac{1}{N_n}\sum_{j=1}^{N_n}(I_{Z,jk}-\bar{I}_{n,k}^Z)^2\Big)\\ &+\sum_{\substack{k=1\\k\neq l}}^{m_n}\sum_{l=1}^{m_n}\frac{1}{N_n}\sum_{i=1}^{N_n}(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Y,il}-\bar{I}_{n,l}^Y)\frac{1}{N_n}\sum_{j=1}^{N_n}(I_{Z,jk}-\bar{I}_{n,k}^Z)(I_{Zjl}-\bar{I}_{n,l}^Z)\\ &=I_1+I_2.\end{aligned} \tag{3.39}$$

We will show that $I_1 \to 1$ almost surely, for $n \to \infty$. For other terms we will either show they are nonnegative or they converge to zero almost surely.

Let us turn our attention to the sum we denoted by I_1. We use the following short calculation

$$\begin{aligned}\frac{1}{N_n}\sum_{i=1}^{N_n}(I_{Y,ik}-\bar{I}^Y_{n,k})^2 &= \frac{1}{N_n}\sum_{i=1}^{N_n}I^2_{Y,ik}-2\bar{I}^Y_{n,k}\frac{1}{N_n}\sum_{i=1}^{N_n}I_{Y,ik}+(\bar{I}^Y_{n,k})^2\\ &= \frac{1}{N_n}\sum_{i=1}^{N_n}I_{Y,ik}-(\bar{I}^Y_{n,k})^2\\ &= \bar{I}^Y_{n,k}-(\bar{I}^Y_{n,k})^2\end{aligned}$$

to see that

$$I_1 = \sum_{k=1}^{m_n}\bar{I}^Y_{n,k}\bar{I}^Z_{n,k}(1-\bar{I}^Y_{n,k})(1-\bar{I}^Z_{n,k}). \tag{3.40}$$

Although this expression is written in compact form we note that it contains multiple sums that has range between 1 and N_n. Let us remind that N_n is the number of blocks we divided our original sequences (Y_i) and (Z_j) and the length of each block is m_n. Blocks are separated by M elements from the original sequences. Because of the assumption of M dependence it means that all random variables from different blocks are independent. Moreover, $\bar{I}^Y_{n,k}$ and $\bar{I}^Y_{n,l}$ are also independent for $|k-l|>M$.

To analyse the asymptotic behavior of I_1 note first that

$$\begin{aligned}\sum_{k=1}^{m_n}\bar{I}^Y_{n,k}\bar{I}^Z_{n,k} &= \frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}I_{Y,ik}\bar{I}^Z_{n,k}\\ &= \frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y)\bar{I}^Z_{n,k}+\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}p_Y\bar{I}^Z_{n,k}\\ &= J_1+J_2.\end{aligned}$$

For J_2 we see that it is equal to

$$\begin{aligned}p_Y\sum_{k=1}^{m_n}\bar{I}^Z_{n,k} &= p_Y\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{j=1}^{N_n}(I_{Z,jk}-p_Z)+p_Y\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{j=1}^{N_n}p_Z\\ &= p_Y\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{j=1}^{N_n}(I_{Z,jk}-p_Z)+m_np_Yp_Z.\end{aligned}$$

By the assumption (3.2) we conclude that $m_np_Yp_Z \to 1$, for $n \to \infty$. On the other hand, by Lemma 3.3.2 $a)$ and the fact that $p_Y \to 0$, for $n \to \infty$, we conclude that

$$p_Y\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{j=1}^{N_n}(I_{Z,jk}-p_Z) \to 0 \tag{3.41}$$

almost surely, as $n \to \infty$. Therefore, J_2 converges to 1 almost surely. Let us turn our attention to J_1. It is equal to

$$\frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}(I_{Y,ik}-p_Y)I_{Z,jk} = \frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Z,jk}-p_Z) + p_Z\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y).$$

First term in the above equation converges to zero almost surely by Lemma 3.3.2 $b)$ and the second term also converges to zero almost surely by $a')$ part of the same lemma. Therefore

$$\sum_{k=1}^{m_n}\bar{I}_{nk}^Y\bar{I}_{nk}^Z \to 1 \quad \text{(a.s.)}.$$

as $n \to \infty$.

Let us look now at $\sum_{k=1}^{m_n}\bar{I}_{nk}^Y(\bar{I}_{nk}^Z)^2$, with the notion that we can treat $\sum_{k=1}^{m_n}\bar{I}_{nk}^Z(\bar{I}_{nk}^Y)^2$ in a similar manner. We would like to show that sum converges to zero almost surely. First we note that

$$\begin{aligned}\sum_{k=1}^{m_n}\bar{I}_{nk}^Y(\bar{I}_{nk}^Z)^2 &= \frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y)(\bar{I}_{nk}^Z)^2 + \frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}p_Y(\bar{I}_{nk}^Z)^2 \\ &= \frac{1}{N_n^3}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}(I_{Y,ik}-p_Y)I_{Z,jk}I_{Z,qk} + p_Y\sum_{k=1}^{m_n}(\bar{I}_{nk}^Z)^2 \\ &= K_1 + K_2.\end{aligned}$$

Let us turn our attention to K_2 first. We have

$$\begin{aligned}K_2 &= p_Y\frac{1}{N_n^2}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}I_{Zik}I_{Z,jk} \\ &= p_Y\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}(I_{Z,ik}-p_Z)(I_{Z,jk}-p_Z) + p_Y\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}p_Z(I_{Z,ik}-p_Z) \\ &+ p_Y\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}p_Z(I_{Z,jk}-p_Z) - p_Y\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}p_Z^2.\end{aligned}$$

The first sum above converges to zero almost surely by Lemma 3.3.2 $c)$. For the second sum note that

$$p_Y p_Z\frac{1}{N_n}\sum_{i=1}^{N_n}\sum_{k=1}^{m_n}(I_{Z,ik}-p_Z) \to 0,\ n\to\infty \ \text{(a.s.)}$$

because of Lemma 3.3.2 $a)$. The same conclusion holds for the third sum. Finally, the last term in the expansion of K_2 also converges to zero as it is equal to $p_Y(\sqrt{m_n}p_Z)^2 \sim p_Y$.

Clearly, $p_Y \to 0$, because of the assumption (3.2), for $n \to \infty$. Hence, we have proved that the whole term K_2 converges to zero almost surely. We will prove the same is true for K_1. We have

$$\begin{aligned} K_1 = &\frac{1}{N_n^3}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Z,jk}-p_Z)(I_{Z,qk}-p_Z) \\ &+\frac{1}{N_n^3}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Z,qk}-p_Z)p_Z \\ &+\frac{1}{N_n^3}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}(I_{Y,ik}-p_Y)(I_{Z,jk}-p_Z)p_Z+\frac{1}{N_n^3}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{q=1}^{N_n}(I_{Y,ik}-p_Y)p_Z^2 \end{aligned}$$

The first sum above converges to zero almost surely by Lemma 3.3.2 d) while the second and the third sum converge to zero almost surely by Lemma 3.3.2 b), for $n \to \infty$. The last sum is equal to

$$p_Z^2\frac{1}{N_n}\sum_{k=1}^{m_n}\sum_{i=1}^{N_n}(I_{Y,ik}-p_Y)$$

and it converges to zero by a') part of Lemma 3.3.2. Therefore, we have shown that both

$$\sum_{k=1}^{m_n}\bar{I}_{nk}^Y(\bar{I}_{nk}^Z)^2 \quad \text{and} \quad \sum_{k=1}^{m_n}\bar{I}_{nk}^Z(\bar{I}_{nk}^Y)^2$$

converge to zero almost surely. By noting that

$$\sum_{k=1}^{m_n}(\bar{I}_{nk}^Y)^2(\bar{I}_{nk}^Z)^2 \le \sum_{k=1}^{m_n}\bar{I}_{nk}^Y(\bar{I}_{nk}^Z)^2 \to 0 \quad \text{(a.s.)}$$

we see that we actually proved that I_1 converges to 1 almost surely, as $n \to \infty$.

To get better insight into the asymptotic behavior of the sum I_2 we will first calculate its expectation. Let $k,l \in \{1,2,\ldots,m_n\}$. Because (Y_i) and (Z_i) are independent we have

$$\begin{aligned} &E(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Y,il}-\bar{I}_{n,l}^Y)(I_{Z,jk}-\bar{I}_{n,k}^Z)(I_{Z,jl}-\bar{I}_{n,l}^Z) \\ &\qquad = E(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Y,il}-\bar{I}_{n,l}^Y)E(I_{Z,jk}-\bar{I}_{n,k}^Z)(I_{Z,jl}-\bar{I}_{n,l}^Z). \end{aligned}$$

Next we calculate $E(I_{Y,ik}-\bar{I}_{n,k}^Y)(I_{Y,il}-\bar{I}_{n,l}^Y)$. It is equal to

$$E\Big(I_{Y,ik}I_{Y,il}-\frac{1}{N_n}\sum_{g=1}^{N_n}I_{Y,il}I_{Y,gk}-\frac{1}{N_n}\sum_{h=1}^{N_n}I_{Y,ik}I_{Y,hl}+\frac{1}{N_n^2}\sum_{g=1}^{N_n}\sum_{h=1}^{N_n}I_{Y,gk}I_{Y,hl}\Big) \tag{3.42}$$

and after few steps we see it is equal to

$$p_{kl}^Y-p_Y^2-\frac{1}{N_n}(p_{kl}^Y-p_Y^2).$$

Observe, for $|k-l|>M$ that expression is equal to zero and so the whole expectation of I_2 is equal to zero if k and l are separated by more than M places. Otherwise, the expectation of I_2 is equal to

$$\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}(p_{kl}^Y-p_Y^2-\frac{1}{N_n}(p_{kl}^Y-p_Y^2))(p_{kl}^Z-p_Z^2-\frac{1}{N_n}(p_{kl}^Z-p_Z^2)).$$

If $Y_k=Y_l$ for some k and l, then $p_{kl}^Y=p_Y$ (same for Z) and asymptotically dominant term is p_Y^2. Hence, the asymptotically dominant term in I_2 may be $m_n p_Y p_Z\sim 1$, for $n\to\infty$.

Because of the distinctly different behavior of the sum I_2, depending on the distance between k and l, we will divide the analysis of the asymptotic behavior of I_2 on two cases. Note that I_2 may be written as

$$I_2=\sum_{k=1}^{m_n}\sum_{\substack{l=1\\k\neq l}}^{m_n}\Big(\frac{1}{N_n}\sum_{i=1}^{N_n}(I_{Y,ik}I_{Y,il})-\bar{I}_{n,k}^Y\bar{I}_{n,l}^Y\Big)\Big(\frac{1}{N_n}\sum_{j=1}^{N_n}(I_{Z,jk}I_{Z,jl})-\bar{I}_{n,k}^Z\bar{I}_{n,l}^Z\Big). \tag{3.43}$$

Let us first consider the sum

$$I_2'=\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}\Big(\frac{1}{N_n}\sum_{i=1}^{N_n}(I_{Y,ik}I_{Y,il})-\bar{I}_{n,k}^Y\bar{I}_{n,l}^Y\Big)\Big(\frac{1}{N_n}\sum_{j=1}^{N_n}(I_{Z,jk}I_{Z,jl})-\bar{I}_{n,k}^Z\bar{I}_{n,l}^Z\Big).$$

Multiplicate the terms in I_2' to get

$$\begin{aligned}I_2'&=\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}I_{Y,ik}I_{Y,il}I_{Z,jk}I_{Z,jl}-\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}I_{Y,ik}I_{Y,il}\bar{I}_{n,k}^Z\bar{I}_{n,l}^Z\\&-\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}\bar{I}_{n,k}^Y\bar{I}_{n,l}^Y I_{Z,jk}I_{Z,jl}+\frac{1}{N_n^2}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}\bar{I}_{n,k}^Y\bar{I}_{n,l}^Y\bar{I}_{n,k}^Z\bar{I}_{n,l}^Z\\&=J_1-J_2-J_3+J_4.\end{aligned}$$

Note that J_1 and J_4 are almost surely nonnegative. As J_2 and J_3 are symmetric we will only analyse the sum J_3. Observe that

$$J_3=-\frac{1}{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}\bar{I}_{n,k}^Y\bar{I}_{n,l}^Y I_{Z,jk}I_{Z,jl}=-\frac{1}{N_n^3}\sum_{i=1}^{N_n}\sum_{p=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}I_{Y,ik}I_{Y,pl}I_{Z,jk}I_{Z,jl}.$$

For $i=p$ we conclude

$$|J_3|\le\frac{1}{N_n^3}\sum_{i=1}^{N_n}\sum_{j=1}^{N_n}\sum_{k=1}^{m_n}\sum_{\substack{l=1\\|k-l|\le M}}^{m_n}|I_{Y,ik}I_{Y,il}I_{Z,jk}I_{Z,jl}|\le\frac{2Mm_n}{N_n}\to 0,\ n\to\infty\quad\text{(a.s.)}$$

Let

$$J = \frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} I_{Y,ik} I_{Y,pl} I_{Z,jk} I_{Z,jl}.$$

We will show that J converges to zero almost surely. Write J as

$$J = \frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} ((I_{Y,ik} - p_Y) + p_Y)((I_{Y,pl} - p_Y) + p_Y)((I_{Z,jk} I_{Z,jl} - p_{kl}^Z) + p_{kl}^Z).$$

After multiplication we conclude

$$J = K_1 + K_2 + K_3 + K_4 + K_5 + K_6 + K_7 + K_8,$$

where

$$K_1 = \frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y)(I_{Z,jk} I_{Z,jl} - p_{kl}^Z),$$

$$K_2 = \frac{1}{N_n^2} \sum_{i=1}^{N_n} \sum_{\substack{p=1 \\ p \neq i}}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} (I_{Y,ik} - p_Y)(I_{Y,pl} - p_Y) p_{kl}^Z,$$

$$K_3 = \frac{1}{N_n^2} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} p_Y (I_{Y,ik} - p_Y)(I_{Z,jk} I_{Z,jl} - p_{kl}^Z),$$

$$K_4 = \frac{1}{N_n} \sum_{i=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} p_Y p_{kl}^Z (I_{Y,ik} - p_Y),$$

$$K_5 = \frac{1}{N_n^2} \sum_{p=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} p_Y (I_{Y,pk} - p_Y)(I_{Z,jk} I_{Z,jl} - p_{kl}^Z),$$

$$K_6 = \frac{1}{N_n} \sum_{p=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} p_Y p_{kl}^Z (I_{Y,pk} - p_Y),$$

$$K_7 = \frac{1}{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} p_Y^2 (I_{Z,jk} I_{Z,jl} - p_{kl}^Z),$$

and

$$K_8 = p_Y^2 \sum_{\substack{k=1 \\ |k-l| \leq M}}^{m_n} \sum_{l=1}^{m_n} p_{kl}^Z.$$

For the sum K_8 we immediately get

$$K_8 \leq p_Y^2 2Mm_n p_Z \sim p_Z \to 0,\ n \to \infty.$$

Note the symmetry between sums K_3 and K_5 and between sums K_4 and K_6. Then by Lemma 3.3.3 follows that the sum J converges to zero almost surely. We conclude that $I_1 + I_2'$ converges to 1 almost surely.

Let us turn our attention on $I_2'' = I_2 - I_2'$. By (3.43) we have

$$I_2'' = \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(\frac{1}{N_n} \sum_{i=1}^{N_n} (I_{Y,ik} I_{Y,il}) - \bar{I}_{n,k}^Y \bar{I}_{n,l}^Y\big)\big(\frac{1}{N_n} \sum_{j=1}^{N_n} (I_{Z,jk} I_{Z,jl}) - \bar{I}_{n,k}^Z \bar{I}_{n,l}^Z\big).$$

We will use slightly different strategy to prove $I_2'' \to 0$ almost surely, for $n \to \infty$. We first rewrite the sum I_2'' slightly

$$I_2'' = \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(\frac{1}{N_n} \sum_{i=1}^{N_n} (I_{Y,ik} I_{Y,il}) - p_Y^2 + p_Y^2 - \bar{I}_{n,k}^Y \bar{I}_{n,l}^Y\big)$$

$$\cdot \big(\frac{1}{N_n} \sum_{j=1}^{N_n} (I_{Z,jk} I_{Z,jl}) - p_Z^2 + p_Z^2 - \bar{I}_{n,k}^Z \bar{I}_{n,l}^Z\big).$$

After multiplication and some rearrangement we see that I_2'' can be written as the sum of the following four sums:

$$L_1 = \frac{1}{N_n^2} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(I_{Y,ik} I_{Y,il} - p_Y^2\big)\big(I_{Z,jk} I_{Zjl} - p_Z^2\big) \tag{3.44}$$

$$L_2 = -\frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \sum_{q=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(I_{Y,ik} I_{Y,il} - p_Y^2\big)\big(I_{Z,jk} I_{Zql} - p_Z^2\big) \tag{3.45}$$

$$L_3 = -\frac{1}{N_n^3} \sum_{i=1}^{N_n} \sum_{p=1}^{N_n} \sum_{j=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(I_{Y,ik} I_{Ypl} - p_Y^2\big)\big(I_{Z,jk} I_{Zjl} - p_Z^2\big) \tag{3.46}$$

and

$$L_4 = \frac{1}{N_n^4} \sum_{i=1}^{N_n} \sum_{p=1}^{N_n} \sum_{j=1}^{N_n} \sum_{q=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \big(I_{Y,ik} I_{Ypl} - p_Y^2\big)\big(I_{Z,jk} I_{Zql} - p_Z^2\big). \tag{3.47}$$

The sum L_1 is converging to zero almost surely by Lemma 3.3.4 *a*). In the treatment of the sum L_2 we can separate our analysis on two cases. When $j = q$ then we see that

$L_2 = -L_1/N_n$ and we proved that L_1 is converging to zero almost surely. For $i \neq j$ we use Lemma 3.3.4 $b)$ to conclude that L_2 is converging to zero almost surely. Because of the symmetry between Y and Z we make same conclusion for L_3 where we use Lemma 3.3.4 part $c)$.

Finally, when we look at L_4, note that for $i = p$ we get that $L_4 = -L_2/N_n$. As we proved that L_2 is converging to zero almost surely we conclude that in this case L_4 is converging to zero almost surely. Similarly, when $j = q$ we get $L_4 = -L_3/N_n$ and we make same conclusion. Finally when $i \neq p$ and $j \neq q$ we get

$$L_4 = \frac{1}{N_n^4} \sum_{\substack{i=1 \\ i \neq p}}^{N_n} \sum_{p=1}^{N_n} \sum_{\substack{j=1 \\ j \neq q}}^{N_n} \sum_{q=1}^{N_n} \sum_{\substack{k=1 \\ |k-l|>M}}^{m_n} \sum_{l=1}^{m_n} \left(I_{Y,ik} I_{Ypl} - p_{kl}^Y\right)\left(I_{Z,jk} I_{Zql} - p_{kl}^Z\right).$$

We apply Lemma 3.3.4 $d)$ to conclude that L_4 converges to zero almost surely. Therefore, we proved that the whole I_2 converges to zero almost surely. Taken all results together, we proved that $I_1 + I_2$ converges to 1 almost surely, while the numerator in (3.36) converges to zero. Hence, condition (3.36) holds and the theorem is proved. ■

Conclusion

We present in this thesis two simple permutation tests. The first is used for testing tail dependence in an iid bivariate sample and the second for testing the independence of two stationary M-dependent sequences. The asymptotic analysis of the test for tail dependence under the null hypothesis is derived using a studentization idea similar to those presented in [8] and [5]. The simulation studies for both tests indicate the considerable power of the tests to reject the null hypothesis when it is false.

Although our background assumption of known distributions seems hard to justify in practical applications, our simulations indicate that the tests are quite robust if we employ empirical quantiles instead.

For practical purposes, it would be very useful to see if one could adjust the testing procedures above, especially the one for M-dependent sequences, for processes Y and Z, which are merely stationary and weakly dependent. It is worth noting that the theoretical justification of the analysis of such a procedure remains an open and, we believe, practically very relevant problem.

BIBLIOGRAPHY

[1] Bücher, Axel and Ivan Kojadinovic: *A note on conditional versus joint unconditional weak convergence in bootstrap consistency results*. Journal of Theoretical Probability, 32, September 2019. ↑ 20.

[2] Charpentier, Arthur and Johan Segers: *Tails of multivariate archimedean copulas*. Journal of Multivariate Analysis, 100(7):1521–1537, 2009. ↑ 28, 29.

[3] Chaudhuri, Kausik, Rituparna Sen, and Zheng Tan: *Testing extreme dependence in financial time series*. Economic Modelling, 73:378–394, 2018.

[4] Chow, Yuan Shih and Henry Teicher: *Independence, Interchangeability, Martingales*. Springer, New York, NY, 3rd edition, 1997. ↑ 14.

[5] Chung, EunYi and Joseph P. Romano: *Exact and asymptotically robust permutation tests*. The Annals of Statistics, 41(2):484 – 507, 2013. ↑ 35, 106.

[6] Chung, EunYi and Joseph P. Romano: *Multivariate and multiple permutation tests*. Journal of Econometrics, 193:76–91, 2016.

[7] Coles, Stuart, Janet Heffernan, and Jonathan Tawn: *Dependence measures for extreme value analyses*. Extremes, 2:339–365, 1999. ↑ 21, 22, 23, 24, 25, 27, 30.

[8] DiCiccio, Cyrus J. and Joseph P. Romano: *Robust permutation tests for correlation and regression coefficients*. Journal of the American Statistical Association, 112:1211 – 1220, 2017. ↑ ii, iii, 1, 2, 35, 47, 106.

[9] Draisma, Gerrit, Holger Drees, Ana Ferreira, and Laurens De Haan: *Bivariate tail estimation: dependence in asymptotic independence*. Bernoulli, 10(2):251 – 280, 2004. ↑ 27, 30, 44, 46.

[10] Embrechts, Paul, Alexander J. McNeil, and Daniel Straumann: *Correlation and Dependence in Risk Management: Properties and Pitfalls*, page 176–223. Cambridge University Press, 2002. ↑ 1, 21.

[11] EunYi Chung, Joseph P. Romano: *Exact and asymptotically robust permutation tests*. The Annals of Statistics, 41:484 – 507, 2013. ↑ 6.

[12] Falk, Michael and René Michel: *Testing for tail independence in extreme value models*. Ann Inst Stat Math, 58:261–290, 2006. ↑ 29, 30.

[13] Frahm, Gabriel, Markus Junker, and Rafael Schmidt: *Estimating the tail-dependence coefficient: Properties and pitfalls*. Insurance: Mathematics and Economics, 37(1):80–100, 2005. ↑ 21.

[14] Genest, Christian and Johan Segers: *Rank-based inference for bivariate extreme-value copulas*. The Annals of Statistics, 37(5B):2990–3022, 2009.

[15] Gudendorf, Gordon and Johan Segers: *Extreme-value copulas*. In Jaworski, Piotr, Fabrizio Durante, Wolfgang Karl Härdle, and Tomasz Rychlik (editors): *Copula Theory and Its Applications*, pages 127–145. Springer Berlin Heidelberg, 2010. ↑ 22, 23, 24, 29.

[16] Haan, Laurens de and Ana Ferreira: *Testing Statistical Hypotheses*. Springer, New York, NY, 1st edition, 2006. ↑ 23.

[17] Heffernan, Janet E.: *A directory of coefficients of tail dependence*. Extremes, 3:279–290, 2000. ↑ 28.

[18] Heffernan, Janet E.: *A directory of coefficients of tail dependence*. Extremes, 3:279–290, 2000. ↑ 38, 42, 43.

[19] Hemerik, Jesse and Jelle Goeman: *Exact testing with random permutations*. TEST, 27:811–825, 2018. ↑ 6, 7.

[20] Hoeffding, Wassily: *A combinatorial central limit theorem*. The Annals of Mathematical Statistics, 22:558–566, 1951. ↑ 2, 8, 10.

[21] Hoeffding, Wassily: *Probability inequalities for sums of bounded random variables*. Journal of the American Statistical Association, 58(301):13–30, 1963. ↑ 5.

[22] Hüsler, Jürg and Deyuan Li: *Testing asymptotic independence in bivariate extremes*. Journal of Statistical Planning and Inference, 139(3):990–998, 2009. ↑ 30.

[23] Janssen, Arnold and Thorsten Pauls: *How do bootstrap and permutation tests work?* The Annals of Statistics, 31(3):768 – 806, 2003. ↑ 1, 5, 20, 46.

[24] Joe, Harry: *Dependence Modeling with Copulas*. Chapman and Hall/CRC, 1st edition, 2014. ↑ 21.

[25] Kallenberg, Olav: *Foundations of Modern Probability*. Springer-Verlag, 2nd edition, 2002. ↑ 11, 12, 13, 14, 19.

[26] Ledford, Anthony W. and Jonethan A. Tawn: *Statistics for near independence in multivariate extreme values*. Biometrika, 83(1):169–187, March 1996. ↑ 1, 27, 28, 30, 36.

[27] Lehmann, Erich L. and Joseph P. Romano: *Testing Statistical Hypotheses*. Springer-Verlag New York, 3rd edition, 2005. ↑ 1, 3, 4, 5, 6, 46.

[28] Mainik, Georg, Georgi Mitov, and Ludger Rüschendorf: *Portfolio optimization for heavy-tailed assets: Extreme risk index vs. markowitz*. Journal of Empirical Finance, 32:115–134, 2015, ISSN 0927-5398. ↑ 1.

[29] Malevergne, Yannick and Didier Sornette: *Extreme financial risks: From dependence to risk management*. Springer-Verlag, Berlin Heidelberg, 1st edition, 2006. ↑ 1.

[30] Nelsen, Roger B.: *An Introduction to Copulas*. Springer, New York, NY, 2nd edition, 2006. ↑ 21.

[31] Nowak, Wioletta and Wiesław Zieba: *Types of conditional convergence*. Annales Universitatis Mariae Curie-Skłodowska. Sectio A, Mathematica., 59:97–105, 2005. ↑ 15.

[32] Peng, L.: *Estimation of the coefficient of tail dependence in bivariate extremes*. Statistics & Probability Letters, 43(4):399–409, 1999. ↑ 27.

[33] Resnick, Sidney I.: *Heavy-Tail Phenomena*. Springer, New York, NY, 1st edition, 2007. ↑ 32.

[34] Romano, Joseph P.: *Bootstrap and Randomization Tests of some Nonparametric Hypotheses*. The Annals of Statistics, 17(1):141 – 159, 1989. ↑ 5.

[35] Romano, Joseph P.: *Bootstrap and randomization tests of some nonparametric hypotheses*. The Annals of Statistics, 17:141–159, 1989. ↑ 20.

[36] Schmidt, Rafael and Ulrich Stadtmüler: *Non-parametric estimation of tail dependence*. Scandinavian Journal of Statistics, 33(2):307–335, 2006. ↑ 29, 30.

[37] Shiryaev, Albert N.: *Probability-1*. Springer, New York, NY, 3rd edition, 2016. ↑ 9, 11, 12, 14, 15, 19.

CURRICULUM VITAE

Darko Brborović was born in Pula on 27 May 1970. He finished elementary and high school in Pula, before graduating from the Faculty of Science (PMF) of the University of Zagreb, Department of Mathematics in 1995. He was a teaching assistant there until 1998, when he left to work for Zagrebačka banka. In 2002 he became portfolio manager at the pension fund AZ and from 2005 until 2021 he was the CEO of OTP Invest, an asset management company.

In 2010 Darko Brborović finished his masters degree with the thesis ”Portfolio optimization with the drawdown constraint”. He has lectured on asset management for the Department of Mathematics since then. He also lectures at Juraj Dobrila University in Pula.