Mathematical Statistics
Stockholm University

Statistical Properties of Bernstein Copulae with Applications in Multiple Testing

André Neumann
Taras Bodnar
Dietmar Pfeifer
Thorsten Dickhaus

Research Report 2016:13

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se

Mathematical Statistics
Stockholm University
Research Report **2016:13**,
http://www.math.su.se

Statistical Properties of Bernstein Copulae with Applications in Multiple Testing

June 2016

André Neumann[a1], Taras Bodnar[b], Dietmar Pfeifer[c], Thorsten Dickhaus[a]

[a] *Institute for Statistics, University of Bremen, Bibliothekstraße 1, D-28359 Bremen, Germany*

[b] *Department of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden*

[c] *Institute of Mathematics, Carl von Ossietzky University of Oldenburg, D-26111 Oldenburg, Germany*

Abstract

A general way to estimate continuous functions consists of approximations by means of Bernstein polynomials. Sancetta and Satchell (2004) proposed to apply this technique to the problem of approximating copula functions. The resulting so-called Bernstein copulae are nonparametric copula estimates with some desirable mathematical features like smoothness. In the present paper, we extend previous statistical results regarding bivariate Bernstein copulae to the multivariate case and study their impact on multiple tests. In particular, we utilize them to derive asymptotic confidence regions for the family-wise error rate (FWER) of simultaneous test procedures which are empirically calibrated by making use of Bernstein copulae approximations of the dependency structure among the test statistics. This extends a similar approach by Stange at al. (2015) in the parametric case. A simulation study quantifies the gain in FWER level exhaustion and, consequently, power which can be achieved by exploiting the dependencies, in comparison with common threshold calibrations like the Bonferroni or the Šidák correction. Finally, we demonstrate an application of the proposed methodology to real-life data from insurance.

Keywords: Asymptotic oscillation behavior, Bonferroni correction, family-wise error rate, p-value, risk management, Šidák correction, simultaneous test procedure

[1]Corresponding author. E-mail address: neumann@uni-bremen.de.

1 Introduction

Copula-based modeling of dependency structures has become a standard tool in applied multivariate statistics and quantitative risk management; see, e. g., [23], [17], [14], [12], and Chapter 5 of [22]. The estimation of an unknown copula is key to a variety of modern multivariate statistical methods. In particular, applications of copulae to the calibration and the analysis of multiple tests have been considered in [10], [3], [31], [6], [28], and [27]; see also Sections 2.2.4 and 4.4 in [9]. Specifically, the copula-based construction of simultaneous test procedures (STPs) in the sense of [13] developed in [10] and [31] under parametric assumptions regarding the type of dependencies among test statistics considerably extends previous approaches as in [15] which are confined to asymptotic Gaussianity and, consequently, linear dependencies.

In the case of a parametric copula, generic estimation techniques like the (generalized) method of moments or maximum likelihood estimation are established notions; cf. Section 3.2 of [31] and references therein. The empirical copula as well as its asymptotic properties as a nonparametric estimator have been studied, among others, in [25], [8], [33], and, more recently, in [5], [29], and [4], to mention only a few references. However, similarly as multivariate histogram estimators, the empirical copula in dimension m has some undesirable properties. For example, it is discontinuous, and it typically assigns zero mass to large subsets of $[0, 1]^m$, even if the sample size n is large, due to the concentration of measures phenomenon. One way to tackle these issues consists of smoothing of the empirical copula. In particular, in [26] smoothing by Bernstein polynomials has been proposed, leading to so-called Bernstein copulae. Approximation theory for Bernstein copulae has been derived in [7], and asymptotic statistical properties of Bernstein copula estimators in the bivariate case ($m = 2$) have been proven in [16] and [2]. Applications of Bernstein copulae to modeling dependencies in non-life insurance have been considered in [11].

In the present work, we contribute to theory and applications of Bernstein copulae in the case of a general dimension $m \geq 2$. In Section 2, we extend the asymptotic theory regarding the Bernstein copula estimator by proving its rate of convergence in infinity norm as well as its asymptotic normality, together with explicit expressions for the limit variance for arbitrary m. Section 3 is then devoted to applications of Bernstein copulae in multiple testing, in particular to the calibration of multivariate STPs for control of the family-wise error rate (FWER), avoiding restrictive parametric dependency assumptions. The application of the central limit theorem derived in Section 2 allows for a precise quantification of the uncertainty about the realized FWER in the case that the copula of test statistics is pre-estimated prior to calibrating the significance threshold of the STP. This extends the results in [31] to the case of nonparametric copula pre-estimation. Section 4 demonstrates by means of a simulation study that the latter pre-estimation approach leads to a better exhaustion of the FWER level and thus enhances the power of the STP compared with traditional approaches which only take univariate marginal distributions of test statistics into account. Finally, we apply the proposed multiple testing methodology to real-life data from insurance (Section 5), and we conclude with a discussion in Section 6. Lengthy proofs and some auxiliary results are deferred to Section 7.

2 Oscillation behavior of empirical Bernstein copulae

Let $\mathbf{X} = (X_1, \ldots, X_m)^\top$ be a random vector taking values in the probability space $(\mathcal{X}, \mathcal{F}, P)$, where $\mathcal{X} \subseteq \mathbb{R}^m$, $\mathcal{F}$ is a σ-field over $\mathcal{X}$, and P denotes the (joint) distribution of $\mathbf{X}$. The univariate marginal cumulative distribution functions (cdfs) of $\mathbf{X}$ we denote by F_j, $j = 1, \ldots, m$, whereas C stands for the copula related to the distribution P.

Assume that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are stochastically independent and identically distributed (i.i.d.) random vectors with $\mathbf{X}_1 \sim P$. Then, the empirical copula $\hat{C}_n$ pertaining to $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is given by

$$\hat{C}_n(\mathbf{u}) = \hat{H}_n\left(\hat{\mathbf{F}}_n^{\leftarrow}(\mathbf{u})\right)$$

with $\hat{\mathbf{F}}_n^{\leftarrow}(\mathbf{u}) = (\hat{F}_{1,n}^{\leftarrow}(u_1), ..., \hat{F}_{m,n}^{\leftarrow}(u_m))^\top$. In this, $\hat{F}_{j,n}^{\leftarrow}$ denotes the generalized inverse of the marginal empirical cumulative distribution function (ecdf) in coordinate $1 \leq j \leq m$, given by $\hat{F}_{j,n}(x_j) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, x_j]}(X_{i,j})$, and

$$\hat{H}_n(\mathbf{x}) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\boldsymbol{\infty}, \mathbf{x}]}(\mathbf{X}_i).$$

The symbol $\mathbb{1}_{\mathcal{A}}$ denotes the indicator function of set $\mathcal{A}$ and $(-\boldsymbol{\infty}, \mathbf{x}] = (-\infty, x_1] \times \ldots \times (-\infty, x_m]$. An analogous bold-face notation for vectors will be used throughout the remainder.

The Bernstein copula estimation is based on the Bernstein polynomial approximation, which is for a fixed copula C given by

$$B_{\mathbf{K}}(\mathbf{u}) := \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} C(\mathbf{k}/\mathbf{K}) \prod_{j=1}^m P_{k_j, K_j}(u_j),$$

where $\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} := \sum_{k_1=0}^{K_1} \cdots \sum_{k_m=0}^{K_m}$, $\mathbf{k}/\mathbf{K} := \left(\frac{k_1}{K_1}, \ldots, \frac{k_m}{K_m}\right)$,

$$P_{k,K}(u) := \binom{K}{k} u^k (1-u)^{K-k},$$

and $K_1, \dots, K_m$ are given positive integers. It has been proved in [7, Corollary 3.1] that $\lim_{\mathbf{K}\to\infty} B_{\mathbf{K}}(\mathbf{u}) = C(\mathbf{u})$. The Bernstein copula estimator for C is then given by

$$\hat{B}_{n,\mathbf{K}}(\mathbf{u}) := \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \hat{C}_n(\mathbf{k}/\mathbf{K}) \prod_{j=1}^{m} P_{k_j,K_j}(u_j).$$

Theorem 2.1 establishes the consistency of $\hat{B}_{n,\mathbf{K}}$.

Theorem 2.1 **(Chung-Smirnov consistency rate).** *Let* $\mathbf{K} = \mathbf{K}(n) \longrightarrow \infty$ *such that* $\sum_{j=1}^{m} K_j^{-1/2} = O\left(n^{-1/2}(\log\log n)^{1/2}\right)$. *Then*

$$\left\|\hat{B}_{n,\mathbf{K}} - C\right\|_\infty = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \quad a.s. \text{ for } \; n \longrightarrow \infty,$$

where $\|g\|_\infty := \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} |g(\mathbf{u})|$ *for* $g : [\mathbf{0},\mathbf{1}] \longrightarrow \mathbb{R}$.

Proof. By the triangle inequality, it holds that

$$\left\|\hat{B}_{n,\mathbf{K}} - C\right\|_\infty \le \left\|\hat{B}_{n,\mathbf{K}} - B_{\mathbf{K}}\right\|_\infty + \left\|B_{\mathbf{K}} - C\right\|_\infty. \tag{1}$$

From Lemma 7.2 we get that $\|B_{\mathbf{K}} - C\|_\infty = O\left(n^{-1/2}(\log\log n)^{1/2}\right)$.

For the first summand in (1), we get

$$\begin{aligned}\left\|\hat{B}_{n,\mathbf{K}} - B_{\mathbf{K}}\right\|_\infty &\le \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \left|\hat{C}_n(\mathbf{k}/\mathbf{K}) - C(\mathbf{k}/\mathbf{K})\right| \prod_{j=1}^{m} P_{k_j,K_j}(u_j) \\ &\le \max_{\mathbf{k}\in\{\mathbf{0},\dots,\mathbf{K}\}} \left|\hat{C}_n(\mathbf{k}/\mathbf{K}) - C(\mathbf{k}/\mathbf{K})\right|,\end{aligned}$$

where $\{\mathbf{0},\dots,\mathbf{K}\} := \{0,\dots,K_1\} \times \dots \times \{0,\dots,K_m\}$. Let $\tilde{F}_{j,n}$ denote the marginal ecdf of $U_{i,j} := F_j(X_{i,j})$ for $j = 1,\dots,m$ and $i = 1,\dots,n$ and let $\tilde{H}_n$ stand for the ecdf of $\mathbf{U}_1,\dots,\mathbf{U}_n$. Application of the identity (see, e.g. [34, Section 3]) $\tilde{F}_{j,n}^{\leftarrow}(u_j) = F_j\left(F_{j,n}^{\leftarrow}(u_j)\right)$ leads to $\hat{C}_n(\mathbf{k}/\mathbf{K}) = \tilde{H}_n\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right)$ and

$$\begin{aligned}\left\|\hat{B}_{n,\mathbf{K}} - B_{\mathbf{K}}\right\|_\infty &\le \max_{\mathbf{k}\in\{\mathbf{0},\dots,\mathbf{K}\}} \left|\tilde{H}_n\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right) - C(\mathbf{k}/\mathbf{K})\right| \\ &\le \max_{\mathbf{k}\in\{\mathbf{0},\dots,\mathbf{K}\}} \left|\tilde{H}_n\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right) - C\left(\tilde{\mathbf{F}}_n^{\leftarrow}(\mathbf{k}/\mathbf{K})\right)\right| \qquad (2) \\ &\quad + \sum_{j=1}^{m} \max_{k_j\in\{0,\dots,K_j\}} \left|\tilde{F}_{j,n}^{\leftarrow}\left(\frac{k_j}{K_j}\right) - \frac{k_j}{K_j}\right|. \qquad (3)\end{aligned}$$

From [18, Theorem 2] we get that the summand in (2) is of order $O\left(n^{-1/2}(\log\log n)^{1/2}\right)$ as well as that each summand in (3) is of order $O\left(n^{-1/2}(\log\log n)^{1/2}\right)$. This completes the proof. □

In Theorem 2.2 we prove a pointwise central limit theorem for the Bernstein copula estimator.

Theorem 2.2 (**Asymptotic normality**). *Assume bounded second order partial derivatives for* C *on* $(\mathbf{0},\mathbf{1})$*. If* $n^{1/2}\sum_{j=1}^{m} K_j^{-1/2} \to 0$*,* $n\to\infty$*, then it holds for all* $\mathbf{u}\in(\mathbf{0},\mathbf{1})$ *and* sequences $(\mathbf{u}_n)_{n\in\mathbb{N}}$ *in* $(\mathbf{0},\mathbf{1})$ with $\lim_{n\to\infty}\mathbf{u}_n=\mathbf{u}$ *that*

$$n^{1/2}\cdot\left(\hat{B}_{n,\mathbf{K}}\left(\mathbf{u}_n\right)-C\left(\mathbf{u}_n\right)\right)\xrightarrow{d}\mathcal{N}\left(0,\sigma^2\left(\mathbf{u}\right)\right),$$

where

$$\begin{aligned}\sigma^2\left(\mathbf{u}\right):=&C\left(\mathbf{u}\right)\cdot\left(1-C\left(\mathbf{u}\right)\right)+\sum_{j=1}^{m}\left(\partial_j C\left(\mathbf{u}\right)\right)^2 u_j\left(1-u_j\right)\\&-2\sum_{j=1}^{m}\partial_j C\left(\mathbf{u}\right)C\left(\mathbf{u}\right)\left(1-u_j\right)\\&+2\sum_{j=1}^{m}\sum_{j'=j+1}^{m}\partial_j C\left(\mathbf{u}\right)\partial_{j'}C\left(\mathbf{u}\right)\left(\mathbb{P}\left[U_{i,j}\le u_j,U_{i,j'}\le u_{j'}\right]-u_j u_{j'}\right).\end{aligned}$$

Proof. It holds that

$$n^{1/2}\left(\hat{B}_{n,\mathbf{K}}\left(\mathbf{u}_n\right)-C\left(\mathbf{u}_n\right)\right)=n^{1/2}\left(\hat{B}_{n,\mathbf{K}}\left(\mathbf{u}_n\right)-B_{\mathbf{K}}\left(\mathbf{u}_n\right)\right)+n^{1/2}\left(B_{\mathbf{K}}\left(\mathbf{u}_n\right)-C\left(\mathbf{u}_n\right)\right).$$

From Lemma 7.2 we have

$$n^{1/2}\left(B_{\mathbf{K}}\left(\mathbf{u}_n\right)-C\left(\mathbf{u}_n\right)\right)=O\left(n^{1/2}\sum_{j=1}^{m}K_j^{-1/2}\right)=o\left(1\right)$$

by our assumption on $\mathbf{K}$.

Let $U_{i,j}=F_j\left(X_{i,j}\right)$. Since the remainder term in Lemma 7.3 does not depend on $\mathbf{u}$, we conclude that

$$n^{1/2}\cdot\left(\hat{B}_{n,\mathbf{K}}\left(\mathbf{u}_n\right)-B_{\mathbf{K}}\left(\mathbf{u}_n\right)\right)=n^{-1/2}\cdot\sum_{i=1}^{n}Y_{i,\mathbf{K}}\left(\mathbf{u}_n\right)+O\left(n^{-1/4}\cdot\left(\log n\right)^{1/2}\cdot\left(\log\log n\right)^{1/4}\right),$$

where

$$\begin{aligned}Y_{i,\mathbf{K}}\left(\mathbf{u}\right):=\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\Bigg(&\mathbb{1}_{(-\infty,\mathbf{k}/\mathbf{K}]}\left(\mathbf{U}_i\right)-C\left(\mathbf{k}/\mathbf{K}\right)\\&-\sum_{j=1}^{m}\partial_j C\left(\mathbf{k}/\mathbf{K}\right)\left(\mathbb{1}_{(-\infty,k_j/K_j]}\left(U_{i,j}\right)-\frac{k_j}{K_j}\right)\Bigg)\cdot\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right).\end{aligned}$$

From Lemma 7.5 we get

$$n^{-1/2}\sum_{i=1}^{n}Y_{i,\mathbf{K}}\left(\mathbf{u}_n\right)\xrightarrow{\mathrm{d}}\mathcal{N}\left(0,\sigma^2\left(\mathbf{u}\right)\right),$$

completing the proof. □

Remark. If the marginal distributions of $\mathbf{X}$ are known, then the Bernstein copula estimator can be constructed by replacing $\hat{C}_n(\mathbf{u})$ with $\hat{H}_n(\mathbf{F}^{\leftarrow}(\mathbf{u}))$. In this case, the $Y_{i,\mathbf{K}}$ can be simplified to

$$Y_{i,\mathbf{K}}(\mathbf{u}) := \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \left(\mathbb{1}_{(-\infty,\mathbf{k}/\mathbf{K}]}(\mathbf{U}_i) - C(\mathbf{k}/\mathbf{K})\right) \prod_{j=1}^{m} P_{k_j,K_j}(u_j)$$

with variance

$$\mathbb{V}[Y_{i,\mathbf{K}}(\mathbf{u})] = C(\mathbf{u}) \cdot (1 - C(\mathbf{u})) + O\left(\sum_{j=1}^{m} K_j^{-1/2}\right),$$

and it holds

$$n^{1/2} \cdot \left(\hat{B}_{n,\mathbf{K}}(\mathbf{u}) - B_{\mathbf{K}}(\mathbf{u})\right) = n^{-1/2} \cdot \sum_{i=1}^{n} Y_{i,\mathbf{K}}(\mathbf{u}).$$

3 Calibration of multivariate multiple test procedures

In this section, we assume that we have uncertainty about the distribution of $\mathbf{X}$. We thus consider a statistical model of the form $(\mathcal{X}, \mathcal{F}, (P_{\boldsymbol{\vartheta},C} : \boldsymbol{\vartheta} \in \Theta, C \in \mathcal{C}))$. The probability measure $P_{\boldsymbol{\vartheta},C}$ is indexed by two parameters. The parameter C denotes the copula of $\mathbf{X}$, and $\boldsymbol{\vartheta}$ is a vector of marginal parameters which refer to $F_1, \ldots, F_m$. The model for the i.i.d. sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ consequently reads as $(\mathcal{X}^n, \mathcal{F}^{\otimes n}, (\mathbb{P}_{\boldsymbol{\vartheta},C} : \boldsymbol{\vartheta} \in \Theta, C \in \mathcal{C}))$, where $\mathbb{P}_{\boldsymbol{\vartheta},C} = P_{\boldsymbol{\vartheta},C}^{\otimes n}$.

Based on this model, we consider multiple test problems of the form $(\mathcal{X}^n, \mathcal{F}^{\otimes n}, (\mathbb{P}_{\boldsymbol{\vartheta},C} : \boldsymbol{\vartheta} \in \Theta, C \in \mathcal{C}), \mathcal{H})$, where $\mathcal{H} = \{H_1, \ldots, H_m\}$ with $\emptyset \neq H_j \subset \Theta$ for all $1 \leq j \leq m$ denotes a family of m null hypotheses regarding the parameter $\boldsymbol{\vartheta}$. The copula C is not the primary target of statistical inference, but a nuisance parameter in the sense that it does not depend on $\boldsymbol{\vartheta}$. This is a common setup in multiple test theory. We will mainly consider a semi-parametric situation, where Θ is of finite dimension, while $\mathcal{C}$ is a function space.

Remark 3.1. The assumption that the number of tests equals the dimension of $\mathbf{X}$ is only made for notational convenience. The case that these two quantities differ can be treated with obvious modifications.

A multiple test for $\mathcal{H}$ is a measurable mapping $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_m) : (\mathcal{X}^n, \mathcal{F}^{\otimes n}) \to \{0,1\}^m$, where $\varphi_j(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 1$ means rejection of the j-th null hypothesis H_j in favor of the alternative $K_j = \Theta \setminus H_j$, $1 \leq j \leq m$. We restrict our attention to multiple tests which are such that $\varphi_j = \mathbb{1}_{(c_j,\infty)}(T_j)$, where $\mathbf{T} = (T_1, \ldots, T_m)^\top : \mathcal{X}^n \to \mathbb{R}^m$ denotes a vector of real-valued test statistics which tend to larger values under alternatives, and $\mathbf{c} = (c_1, \ldots, c_m)^\top$ are the corresponding critical values. In many problems of practical interest, T_j will only use the data $(x_{i,j})_{1 \leq i \leq n}$, for every $1 \leq j \leq m$. In particular, this typically holds true if $\boldsymbol{\vartheta} = (\boldsymbol{\vartheta}_1^\top, \ldots, \boldsymbol{\vartheta}_m^\top)^\top$, $\boldsymbol{\vartheta}_j$ represents a set of functionals of F_j, and H_j only concerns $\boldsymbol{\vartheta}_j$, for every $1 \leq j \leq m$.

For the calibration of $\mathbf{c}$, we aim at controlling the FWER in the strong sense. For given $\boldsymbol{\vartheta} \in \Theta$ and $C \in \mathcal{C}$, the FWER is defined as the probability for at least one false rejection (type I error) of $\boldsymbol{\varphi}$ under $\mathbb{P}_{\boldsymbol{\vartheta},C}$, i. e.,

$$\mathrm{FWER}_{\boldsymbol{\vartheta},C}(\boldsymbol{\varphi}) = \mathbb{P}_{\boldsymbol{\vartheta},C}\left(\bigcup_{j \in I_0(\boldsymbol{\vartheta})} \{\varphi_j = 1\}\right),$$

where $I_0(\boldsymbol{\vartheta}) = \{1 \leq j \leq m : \boldsymbol{\vartheta} \in H_j\}$ denotes the index set of true null hypotheses under $\boldsymbol{\vartheta}$. The multiple test $\boldsymbol{\varphi}$ is said to control the FWER at level $\alpha \in [0, 1]$, if

$$\sup_{\boldsymbol{\vartheta} \in \Theta, C \in \mathcal{C}} \mathrm{FWER}_{\boldsymbol{\vartheta},C}(\boldsymbol{\varphi}) \leq \alpha.$$

Notice that, although the trueness of the null hypotheses is determined by $\boldsymbol{\vartheta}$ alone, the FWER depends on $\boldsymbol{\vartheta}$ and C, because the dependency structure in the data typically influences the distribution of $\boldsymbol{\varphi}$ when regarded as a statistic with values in $\{0, 1\}^m$.

Throughout the remainder, we assume that the following set of conditions is fulfilled.

Assumption 3.2.

(a) Letting $H_0 = \bigcap_{j=1}^m H_j$ denote the global hypothesis of $\mathcal{H}$, there exists a $\boldsymbol{\vartheta}^* \in H_0$ such that

$$\forall C \in \mathcal{C} : \forall \boldsymbol{\vartheta} \in \Theta : FWER_{\boldsymbol{\vartheta},C}(\boldsymbol{\varphi}) \leq FWER_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi}).$$

Notice that this assumption can be weakened by considering closed test procedures, where our proposed methodology is applied to every non-empty intersection hypothesis in $\mathcal{H}$; cf. Remark 1 in [32] for details.

(b) The vector of marginal cdfs of $\mathbf{T} = (T_1, \ldots, T_m)^\top$ depends on $\boldsymbol{\vartheta}$ only, and is (at least asymptotically as $n \to \infty$) known under $\boldsymbol{\vartheta}^*$. We denote the vector of marginal cdfs of $\mathbf{T} = (T_1, \ldots, T_m)^\top$ under $\boldsymbol{\vartheta}^*$ by $\mathbf{G} = (G_1, \ldots, G_m)^\top$.

(c) Letting $C_{\boldsymbol{\vartheta}^*}$ denote the copula of $\mathbf{T}$ under $\boldsymbol{\vartheta}^*$, there exists a continuously differentiable function $h : [0, 1] \to [0, 1]$ such that $C_{\boldsymbol{\vartheta}^*}(\mathbf{u}) = h(C(\mathbf{u}))$ for all $\mathbf{u} \in [0, 1]^m$. The function h may be unknown. Notice that, if T_j only uses the data $(x_{i,j})_{1 \leq i \leq n}$, for every $1 \leq j \leq m$, then $C_{\boldsymbol{\vartheta}^*} \equiv C_{\mathbf{T}}$ is independent of $\boldsymbol{\vartheta}^*$.

Before we start to explain the proposed method for the calibration of $\mathbf{c}$, let us illustrate prototypical example applications of our general setup.

Example 3.3.

(a) Let $\Theta = \mathbb{R}^m$ and assume that $\vartheta_j \in \mathbb{R}$ is the expected value of X_j for every $1 \leq j \leq m$. The j-th null hypothesis may be the one-sided null hypothesis $H_j = \{\vartheta_j \leq 0\}$ with corresponding alternative $K_j = \{\vartheta_j > 0\}$. Assume that the scale parameter of the marginal distribution of each X_j is known and without loss of generality equal to one. A suitable test statistic T_j is then given by $T_j(\mathbf{X}_1, \ldots, \mathbf{X}_n) = \sum_{i=1}^n X_{i,j}/\sqrt{n}$. Under $\boldsymbol{\vartheta}^* = \mathbf{0} \in \mathbb{R}^m$, we have that $G_j = \Phi$ (the cdf of the standard normal law on $\mathbb{R}$) is the cdf of the (asymptotic) null distribution of T_j for every $1 \leq j \leq m$. If the considered copula family $\mathcal{C}$ consists of multivariate stable copulae (meaning that the observables follow a multivariate stable distribution), then the copula $C_{\boldsymbol{\vartheta}^*}$ is of the same type as C, hence all parts of Assumption 3.2 are fulfilled.

(b) Let $\mathcal{X} = [0, \infty)$ and assume that the stochastic representations $X_j \stackrel{d}{=} \vartheta_j Z_j$ with $\vartheta_j > 0$ hold true for all $1 \leq j \leq m$, where Z_j is a random variable taking values in $[0, 1]$. The parameter of interest in this problem is $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_m)^\top \in \Theta = (0, \infty)^m$. For each coordinate j, we consider the pair of hypotheses $H_j : \{\vartheta_j \leq \vartheta_j^*\}$ versus $K_j : \{\vartheta_j > \vartheta_j^*\}$, for a given vector $\boldsymbol{\vartheta}^* = (\vartheta_1^*, \ldots, \vartheta_m^*)^\top \in (0, \infty)^m$ of hypothesized upper bounds for the supports (or right end-points of the distributions) of the X_j's. This has applications in the context of stress testing in actuarial science and financial mathematics (cf., e. g., [21]). Suitable test statistics are given by the component-wise maxima of the observables, i. e., $T_j(\mathbf{X}_1, \ldots, \mathbf{X}_n) = \max_{1 \leq i \leq n} X_{i,j}/\vartheta_j^*$, $1 \leq j \leq m$. Assuming that the tail behavior of each X_j is known such that the marginal (limiting) extreme value distribution of T_j under $\boldsymbol{\vartheta}^*$ can be derived and letting $\mathcal{C}$ consist of max-stable copulae, all parts of Assumption 3.2 are fulfilled here, too.

Let us remark here that these two examples have been treated under the restrictive assumption of one-parametric copula families $\mathcal{C}$ in [31].

The following lemma is taken from [10] and connects $\mathrm{FWER}_{\boldsymbol{\vartheta},C}(\boldsymbol{\varphi})$ with $C_{\boldsymbol{\vartheta}^*}$.

Lemma 3.4. *Let Assumption 3.2 be fulfilled and assume that $\boldsymbol{\varphi}$ is a simultaneous test procedure (STP) in the sense of [13], meaning that all m critical values are identical, say $c_j = c(\alpha)$ for all $1 \leq j \leq m$. Then we have that*

$$FWER_{\boldsymbol{\vartheta},C}(\boldsymbol{\varphi}) \leq 1 - C_{\boldsymbol{\vartheta}^*}(\mathbf{1} - \boldsymbol{\alpha}_{loc}),$$

where $\alpha_{loc}^{(j)} = 1 - G_j(c(\alpha))$ *denotes a local significance level for the* j*-th marginal test problem. In practice, it is convenient to carry out the STP in terms of p-values* $p_j = 1 - G_j(T_j)$ *such that* $\varphi_j = \mathbb{1}_{[0,\alpha_{loc}^{(j)})}(p_j)$*.*

Lemma 3.4 shows that the problem of calibrating the local significance levels corresponding to $c(\alpha)$ is equivalent to the problem of estimating the contour line of $C_{\boldsymbol{\vartheta}^*}$ at contour level $1-\alpha$. Any point on that contour line defines a valid set of local significance levels. Thus, one may weight the m hypotheses for importance by choosing particular points on the contour line. If all m hypotheses are equally important it is natural to choose equal local levels $\alpha_{loc}^{(j)} \equiv \alpha_{loc}$ for all $1 \leq j \leq m$. This amounts to finding the point of intersection of the contour line of $C_{\boldsymbol{\vartheta}^*}$ at contour level $1-\alpha$ and the "main diagonal" in the m-dimensional unit hypercube.

Recall that we assume that C and, consequently, $C_{\boldsymbol{\vartheta}^*}$ are unknown. Based on our investigations in Section 2 and making use of Assumption 3.2.(c), we thus propose to calibrate $\boldsymbol{\varphi}$ empirically. If h is known, this can be done by solving the equation

$$h\left(\hat{B}_{n,\mathbf{K}}\left(\mathbf{1} - \boldsymbol{\alpha}_{loc}\right)\right) = 1 - \alpha \tag{4}$$

for $\boldsymbol{\alpha}_{loc}$, for fixed weights. If for a given α the solution of (4) with the chosen weights is not unique, one should choose the smallest set of local significance levels such that (4) holds. To facilitate the argumentation, let us from now on assume that we choose equal local levels. We denote the solution of (4) with $\alpha_{loc}^{(j)} \equiv \alpha_{loc}$ by $\hat{\alpha}_{loc,n}$. This leads to the definition

$$\hat{\alpha}_{loc,n} := 1 - \hat{B}_{n,\mathbf{K}}^{\leftarrow}\left(h^{\leftarrow}(1-\alpha)\right),$$

where $\hat{B}_{n,\mathbf{K}}^{\leftarrow}(p) := \inf\left\{u \in [0,1] \,\middle|\, \hat{B}_{n,\mathbf{K}}(u,\ldots,u) \geq p\right\}$ for $p \in (0,1)$. Since $\hat{B}_{n,\mathbf{K}}$ depends on the data, $\hat{\alpha}_{loc,n}$ is a random variable and

$$\widehat{\mathrm{FWER}}_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi}) = 1 - C_{\boldsymbol{\vartheta}^*}\left(1 - \hat{\alpha}_{loc,n}, \ldots, 1 - \hat{\alpha}_{loc,n}\right)$$

is a random variable, too, which is distributed around the target FWER level α. The following theorem is the main result of this section and quantifies the uncertainty about the realized FWER if the empirical calibration of $\boldsymbol{\varphi}$ is performed via (4).

Theorem 3.5. *Let Assumption 3.2 be fulfilled. Then* $\widehat{FWER}_{\boldsymbol{\vartheta}^*,C}$ *has the following properties.*

a) Consistency:

$$\forall C \in \mathcal{C} : \widehat{FWER}_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi}) \to \alpha \text{ almost surely.}$$

b) Asymptotic Normality:

$$\forall C \in \mathcal{C} : \sqrt{n}\left(\widehat{FWER}_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi}) - \alpha\right) \xrightarrow{d} \mathcal{N}(0, \sigma_\alpha^2),$$

where

$$\sigma_\alpha^2 = \frac{\sigma\left(C_{\boldsymbol{\vartheta}^*}^{\leftarrow}(1-\alpha), \ldots, C_{\boldsymbol{\vartheta}^*}^{\leftarrow}(1-\alpha)\right)}{\tilde{C}'\left(C_{\boldsymbol{\vartheta}^*}^{\leftarrow}(1-\alpha)\right)} \cdot \left(\tilde{C}'_{\boldsymbol{\vartheta}^*}\left(C_{\boldsymbol{\vartheta}^*}^{\leftarrow}(1-\alpha)\right)\right)^2$$

and $\tilde{C}(u) := C(u, \ldots, u), \tilde{C}_{\boldsymbol{\vartheta}^*}(u) := C_{\boldsymbol{\vartheta}^*}(u, \ldots, u)$.

c) Asymptotic Confidence Region:

$$\forall \delta \in (0,1) : \forall C \in \mathcal{C} : \lim_{n\to\infty} \mathbb{P}_{\boldsymbol{\vartheta}^*,C}\left(\sqrt{n}\frac{\widehat{FWER}_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi}) - \alpha}{\hat{\sigma}_n} \leq z_{1-\delta}\right) = 1 - \delta\,,$$

where $\hat{\sigma}_n^2 : \mathcal{X}^n \to (0,\infty)$ *is a consistent estimator of the asymptotic variance* σ_α^2. *In this,* $z_\beta = \Phi^{-1}(\beta)$ *denotes the* β*-quantile of the standard normal distribution on* $\mathbb{R}$.

Proof.

(a) Let $C \in \mathcal{C}$ be arbitrary, but fixed. Since h is continuously differentiable, h is also Lipschitz-continuous with Lipschitz constant $L > 0$. Therefore, with Theorem 2.1 we get

$$\begin{aligned}\left|\widehat{\mathrm{FWER}}_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi}) - \alpha\right| &= \left|1 - \alpha - C_{\vartheta^*}\left(\mathbf{1} - \hat{\boldsymbol{\alpha}}_{loc,n}\right)\right| \\ &= \left|h\left(\hat{B}_{n,\mathbf{K}}\left(\mathbf{1} - \hat{\boldsymbol{\alpha}}_{loc,n}\right)\right) - h\left(C\left(\mathbf{1} - \hat{\boldsymbol{\alpha}}_{loc,n}\right)\right)\right| \\ &\leq \left\|h\left(\hat{B}_{n,\mathbf{K}}\right) - h(C)\right\|_\infty \\ &\leq L \cdot \left\|\hat{B}_{n,\mathbf{K}} - C\right\|_\infty \\ &= O\left(n^{-1/2}(\log\log n)^{1/2}\right) \;\; a.s.\end{aligned}$$

(b) Letting $p := h^{\leftarrow}(1-\alpha)$, Lemma 7.6 yields that

$$\begin{aligned}\sqrt{n}\left(1 - \alpha_{loc,n} - C^{\leftarrow}(p)\right) &= \sqrt{n}\left(\hat{B}_{n,\mathbf{K}}^{\leftarrow}(p) - C^{\leftarrow}(p)\right) \\ &\xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma\left(C^{\leftarrow}(p), \ldots, C^{\leftarrow}(p)\right)}{\tilde{C}'\left(C^{\leftarrow}(p)\right)}\right).\end{aligned}$$

Therefore, applying the Delta Method to $\tilde{C}_{\boldsymbol{\vartheta}^*}$, we have that

$$\begin{aligned}\sqrt{n}\left(\widehat{\mathrm{FWER}}_{\boldsymbol{\vartheta}^*,C}(\boldsymbol{\varphi})-\alpha\right) &= -\sqrt{n}\left(\tilde{C}_{\boldsymbol{\vartheta}^*}\left(1-\hat{\alpha}_{loc,n}\right)-(1-\alpha)\right)\\ &= -\sqrt{n}\left(\tilde{C}_{\boldsymbol{\vartheta}^*}\left(1-\hat{\alpha}_{loc,n}\right)-\tilde{C}_{\boldsymbol{\vartheta}^*}\left(C^{\leftarrow}\left(p\right)\right)\right)\\ &\xrightarrow{d}\mathcal{N}\left(0,\frac{\sigma\left(C^{\leftarrow}\left(p\right),\ldots,C^{\leftarrow}\left(p\right)\right)}{\tilde{C}'\left(C^{\leftarrow}\left(p\right)\right)}\cdot\left(\tilde{C}'_{\boldsymbol{\vartheta}^*}\left(C^{\leftarrow}\left(p\right)\right)\right)^2\right).\end{aligned}$$

The result follows from the definition of p.

(c) Follows directly from part b) using Slutsky's Theorem.

□

If the function h is unknown, one may approximate the value of $\hat{\alpha}_{loc,n}$ with high precision by a Monte Carlo simulation for a given number M of Monte Carlo repetitions. To this end, generate $M \times n$ pseudo-random vectors which follow the estimated (joint) distribution of $\mathbf{X}$ under $\boldsymbol{\vartheta}^*$, by combining $\hat{B}_{n,\mathbf{K}}$ and the marginal cdfs $F_1, \ldots, F_m$ of $X_1, \ldots, X_m$ under the global hypothesis. From these, calculate a pseudo-sample $\mathbf{T}_1, \ldots, \mathbf{T}_M$ from the distribution of $\mathbf{T}$ under $\boldsymbol{\vartheta}^*$. Then, $\mathbf{G}(\mathbf{T}_1), \ldots, \mathbf{G}(\mathbf{T}_M)$ constitutes a pseudo-random sample from the estimator of $C_{\boldsymbol{\vartheta}^*}$, and the empirical equi-coordinate $(1-\alpha)$-quantile of this pseudo-sample approximates $\hat{\alpha}_{loc,n}$. Since the number M of pseudo-random vectors to be generated is in principle unlimited, Theorem 3.5 continues to hold true if this strategy is pursued.

4 Simulation study

In this section we report the results of a simulation study regarding the FWER and the power of multiple tests which are empirically calibrated as proposed in Section 3. Assume w.l.o.g. that $I_0(\boldsymbol{\vartheta}) := \{1, ..., m_0\}$ and let $m_1 := m - m_0$. The empirical FWER and the empirical power, respectively, are given by

$$\widehat{\mathrm{FWER}}_{\vartheta,C}(\boldsymbol{\varphi}) := L^{-1}\sum_{\ell=1}^{L}\mathbb{1}_{\bigcup_{j=1}^{m_0}\left\{\varphi_j^{(\ell)}=1\right\}}\left(\mathbf{x}_1^{(\ell)},\ldots,\mathbf{x}_n^{(\ell)}\right)$$

and

$$\widehat{\mathrm{power}}(\boldsymbol{\varphi}) := L^{-1}\sum_{\ell=1}^{L}\left(m_1^{-1}\sum_{j=m_0+1}^{m}\mathbb{1}_{\left\{\varphi_j^{(\ell)}=1\right\}}\left(\mathbf{x}_1^{(\ell)},\ldots,\mathbf{x}_n^{(\ell)}\right)\right),$$

where $\left(\mathbf{x}_1^{(\ell)},\ldots,\mathbf{x}_n^{(\ell)}\right) \in \mathcal{X}^n$ denotes the pseudo-sample in the ℓ-th simulation run. In this simulation we use the setting given in [31, Section 4.3]. This means that we are in the situation of Example 3.3.(b) except that the copula C of $\mathbf{X}$ is not max-stable. The distribution of $\mathbf{Z}$ is given by an m-variate t_4-copula with Beta$(3,4)$ marginal distributions, where $\Sigma_{i,j} := \mathbb{1}_{\{i=j\}} + \rho\mathbb{1}_{\{i\neq j\}}$, $1 \leq i,j \leq m$. The extreme value copula of the multivariate t-distribution has been derived in [24]. It has an analytically intractable form. In [31], this extreme value copula has been approximated by a Gumbel-Hougaard copula with an estimated copula parameter. In contrast, we approximated this extreme value copula nonparametrically with the Bernstein copula estimator. The calculation of Bernstein copulae has been performed as in [7, Example 4.2], using $K_j := n$ for all $j \in \{1,\ldots,m\}$.

Since the function h is intractable, we calibrate the test with the following algorithm which was outlined at the end of Section 3.

Algorithm 4.1.

(a) Choose a number M of Monte Carlo repetitions.

(b) For each $b = 1,\ldots,M$ draw a sample $\mathbf{U}_1^{\#b},\ldots,\mathbf{U}_n^{\#b}$ of $\hat{B}_{n,\mathbf{K}}$ and calculate

$$X_{i,j}^{\#b} = F_j^{\leftarrow}\left(U_{i,j}^{\#b}\right), 1 \leq i \leq n, 1 \leq j \leq m.$$

(c) For all $1 \leq j \leq m$, compute $T_j^{\#b} = T_j\left(\mathbf{X}_1^{\#b},\ldots,\mathbf{X}_n^{\#b}\right)$ and obtain the pseudo-sample

$$V_j^{\#b} = F_{Z_j}^n\left(T_j^{\#b}\right)$$

from the copula of $\mathbf{T}$. Notice that, due to the stochastic independence of the observables, the marginal cdfs of $\mathbf{T}$ are given by $F_{Z_j}^n$.

(d) Finally, calibrate $\hat{\boldsymbol{\alpha}}_{loc,n}$ by solving

$$\#\left\{\mathbf{V}^{\#b} \leq \mathbf{1} - \hat{\boldsymbol{\alpha}}_{loc,n}\right\} = \lceil (1-\alpha)\,M \rceil\,. \tag{5}$$

Notice that in Algorithm 4.1, we implicitly weight the hypotheses. This means that the weights corresponding to the obtained $\hat{\boldsymbol{\alpha}}_{loc,n}$ depend on the simulation data, for convenience of implementation. In comparison, the classical Bonferroni and Šidák corrected local significance levels are given by

$$\alpha_{loc}^{(j)} = \frac{\alpha}{m} \text{ and } \alpha_{loc}^{(j)} = 1 - (1-\alpha)^{1/m}, 1 \leq j \leq m,$$

respectively.

$m = 6$						
	$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
	$m_0 = 2$	$m_0 = 4$	$m_0 = 2$	$m_0 = 4$	$m_0 = 2$	$m_0 = 4$
Empirical FWER						
Bonferroni	0.0172	0.0280	0.0156	0.0220	0.0120	0.0172
Šidák	0.0176	0.0284	0.0156	0.0228	0.0120	0.0172
Bernstein	0.0232	0.0404	0.0216	0.0340	0.0192	0.0280
Empirical power						
Bonferroni	0.4904	0.4902	0.4913	0.4838	0.4892	0.4778
Šidák	0.4920	0.4912	0.4930	0.4846	0.4906	0.4794
Bernstein	0.5038	0.5034	0.5048	0.5014	0.5040	0.4946

$m = 12$						
	$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
	$m_0 = 5$	$m_0 = 10$	$m_0 = 5$	$m_0 = 10$	$m_0 = 5$	$m_0 = 10$
Empirical FWER						
Bonferroni	0.0188	0.0268	0.0156	0.0184	0.0112	0.0132
Šidák	0.0188	0.028	0.0156	0.0184	0.0116	0.0132
Bernstein	0.0360	0.062	0.0268	0.0476	0.0208	0.0320
Empirical power						
Bonferroni	0.4477	0.4490	0.4481	0.4568	0.4499	0.4510
Šidák	0.4488	0.4498	0.4495	0.4580	0.4513	0.4522
Bernstein	0.4694	0.4686	0.4709	0.4770	0.4685	0.4738

Table 1: Comparison of empirical FWER and power regarding Bonferroni, Šidák and Bernstein corrections with $m \in \{6, 12\}$, varying m_0, $\rho \in \{0.2, 0.5, 0.8\}$, $\alpha = 0.05$, $L = 2500$, $M = 400$, and $n = 100$.

With the notation introduced in Example 3.3.(b), we put $\boldsymbol{\vartheta}^* = (2, ..., 2)^\top$ and

$$\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_m)^\top \quad \text{with} \quad \vartheta_j = \begin{cases} 2, & \text{for } j \leq m_0, \\ 2 + U, \ U \sim \text{UNI}(0, 0.5], & \text{otherwise.} \end{cases}$$

The results of the simulation study for different values of m, m_0, and ρ are displayed in Table 1. It reveals that the Bernstein correction performs better in all cases,

i. e., its empirical FWER is closer to α and its empirical power is higher than those of the generic calibrations. Of course, one could increase the difference in empirical powers between the Bernstein calibration and the Bonferroni or Šidák calibrations by choosing other values for the ϑ_j under the alternatives.

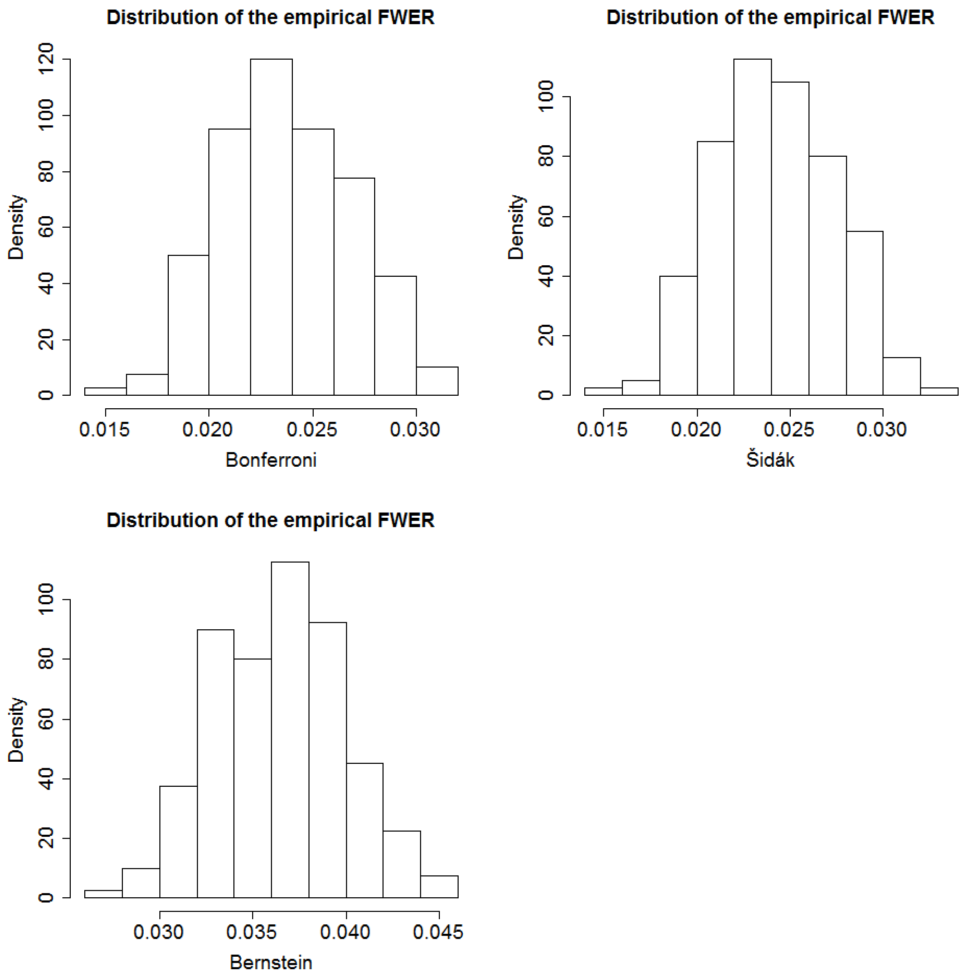

Figure 1: Distribution of the empirical FWER regarding Bonferroni, Šidák and Bernstein correction with $m = 6$, $m_0 = 4$, $\rho = 0.5$, $\alpha = 0.05$, $L = 2500$, $M = 400$, and $n = 100$.

Figure 1 and Figure 2 provide a more detailed view on the distribution of the empirical FWER and the empirical power, respectively, of the three concurring calibrations from Table 1. Namely, we performed the entire simulation algorithm comprising L simulation runs with M Monte Carlo repetitions for the approximation of h in each of these runs 200 times. In the figures, we display results for one particular parameter configuration (m, m_0, ρ). For other parameter configurations, the results are very similar.

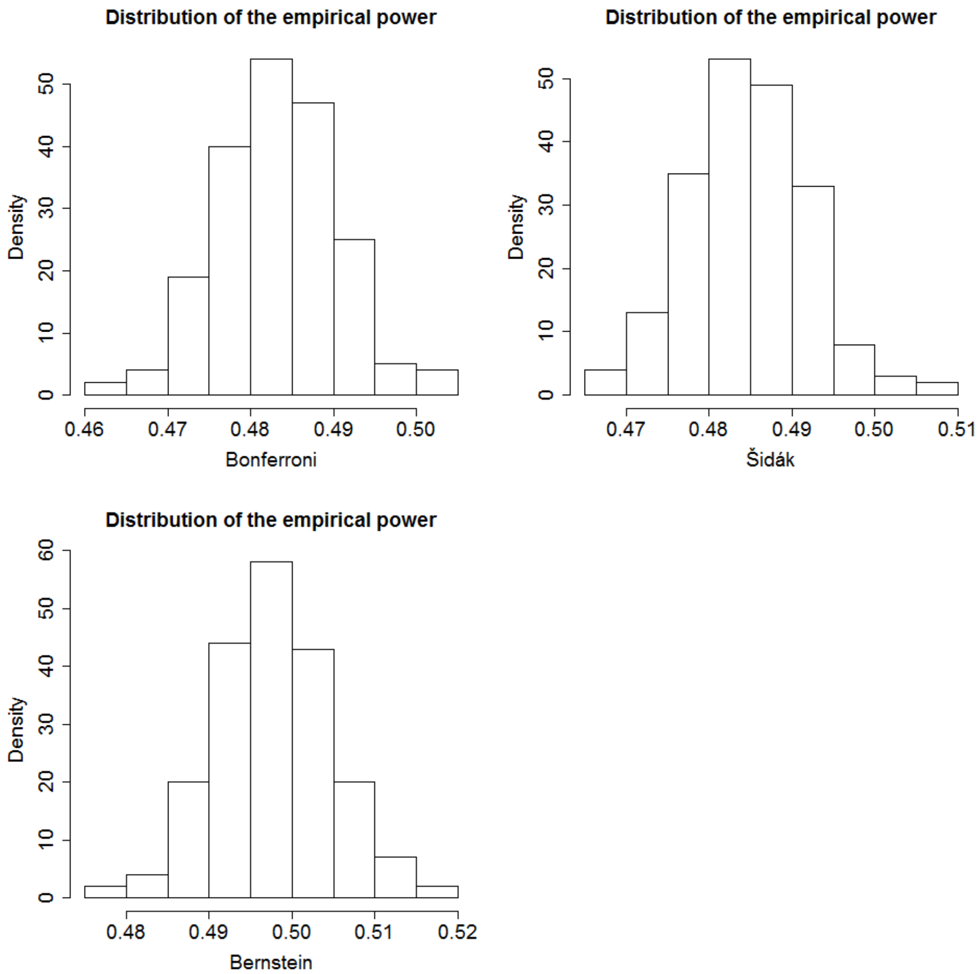

Figure 2: Distribution of the empirical power regarding Bonferroni, Šidák and Bernstein correction with $m = 6$, $m_0 = 4$, $\rho = 0.5$, $\alpha = 0.05$, $L = 2500$, $M = 400$, and $n = 100$.

As expected from Theorem 3.5, the histograms look normal, each with a small variance.

5 Application

In this section, we analyze insurance claim data from $m = 19$ adjacent geographical regions (see Table 4). For every region $j \in \{1, \ldots, 19\}$ these claims have, for confidentiality reasons, been adjusted to a neutral monetary scale. The claim amounts and types have been aggregated to full years, such that temporal dependencies are considered negligible. However, strong non-linear spatial dependencies are likely to be present in the data. Hence, we treat each of the $n = 20$ rows in Table 4 as an independent repetition $\mathbf{X}_i = \mathbf{x}_i$ of an m-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_m)^\top$, where $1 \leq i \leq 20$ is the time index in years and $m = 19$ refers to the regions.

An important quantity for regulators and risk managers is the region-specific value-at-risk (VaR). The VaR at level p for region j is defined as the p-quantile of the (marginal) distribution of X_j, i. e.,

$$\mathrm{VaR}_j(p) := F^{\leftarrow}_{X_j}(p).$$

In insurance mathematics, typically considered values of p are close to one. Here, we chose $p = 0.995$. Our goal is to derive multiplicity-corrected confidence intervals for $\vartheta_j = \mathrm{VaR}_j(0.995)$, $1 \leq j \leq m = 19$ which are compatible with (i. e., dual to) the Bonferroni, Šidák and Bernstein copula-based correction methods discussed before. To this end, let auxiliary point hypotheses be defined as $H_{\vartheta_j^*} : \left\{\vartheta_j = \vartheta_j^*\right\}$ for fixed $\vartheta_j^* > 0$. According to the Extended Correspondence Theorem (see [9, Section 1.3]), the set of all values ϑ_j^* for which $H_{\vartheta_j^*}$ is retained by a multiple test at FWER level α (leading to a local significance level $\alpha_{loc}^{(j)}$ in coordinate j) constitutes a confidence region at simultaneous confidence level $1 - \alpha$ for ϑ_j, $1 \leq j \leq m$. We set $\alpha = 5\%$.

In quantitative risk management, it is common practice to model the excess distribution of X_j over some given threshold u_j by a generalized Pareto distribution (GPD) (cf., e. g., [22, Section 7.2.2]).

Definition 5.1 ([22, Definition 7.16])**.** For shape parameter $\xi \in \mathbb{R}$ and scale parameter $\beta > 0$, the cdf of the GPD is given by

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi} & , \xi \neq 0, \\ 1 - \exp(-x/\beta) & , \xi = 0, \end{cases}$$

where $x \geq 0$ if $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ if $\xi < 0$.

In the remainder, we make the following assumption.

Assumption 5.2. *For every $1 \leq j \leq m = 19$ there exists a threshold u_j and parameter values ξ_j and β_j such that*

$$\mathbb{P}[X_j - u_j \leq x \,|\, X_j > u_j] \approx G_{\xi_j,\beta_j}(x)$$

for all $x \in \mathbb{R}$.

Under Assumption 5.2, an approximation of the VaR at level p for region j is given by

$$\mathrm{VaR}_{\xi_j,\beta_j}(p) \approx u_j + \frac{\beta_j}{\xi_j}\left(\left(\frac{1-p}{1-F_{X_j}(u_j)}\right)^{-\xi_j} - 1\right) =: q_j(\xi_j, \beta_j),$$

provided that $p \geq F_{X_j}(u_j)$. For ease of notation, we let $\vartheta_j = q_j(\xi_j, \beta_j)$ in the sequel.

For computational convenience, we carried out the test for $H_{\vartheta_j^*}$ as a confidence-region test in the sense of [1] based on the family

$$\left(H_{\xi_j^*,\beta_j^*} : \left\{\xi_j = \xi_j^*, \beta_j = \beta_j^*\right\} | \beta_j^* > 0, \xi_j^* \in \mathbb{R}\right) \tag{6}$$

of point hypotheses. Namely, the test procedure works as follows.

Algorithm 5.3.

1. *Test each $H_{\xi_j^*,\beta_j^*}$ by an arbitrary level $\alpha_{loc}^{(j)}$ test, where $\alpha_{loc}^{(j)}$ denotes a multiplicity-corrected significance level based on the Bonferroni, Šidák or Bernstein copula calibration, respectively.*
2. *Let a confidence region $C_{\xi_j,\beta_j}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ at confidence level $1 - \alpha_{loc}^{(j)}$ for (ξ_j, β_j) be defined as the set of all parameter values (ξ_j^*, β_j^*) for which $H_{\xi_j^*,\beta_j^*}$ is retained.*
3. *Reject $H_{\vartheta_j^*}$ at level $\alpha_{loc}^{(j)}$, if the set $\{(\xi_j^*, \beta_j^*) : q_j\left(\xi_j^*, \beta_j^*\right) = \vartheta_j^*\}$ has an empty intersection with $C_{\xi_j,\beta_j}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.*

Due to Algorithm 5.3, it suffices to construct point hypothesis tests for (6). A standard technique for testing parametric hypotheses is to perform a likelihood ratio test. In the risk management context, this method is described in [22, A.3.5]. Define the random variable $N_{u_j} := \#\{1 \leq i \leq n \,|\, X_{i,j} > u_j\}$ and let $\tilde{X}_{1,j}, \ldots, \tilde{X}_{N_{u_j},j}$ denote the corresponding sub-sample for region j. Then the excesses $Y_{1,j}, \ldots, Y_{N_{u_j},j}$ over u_j are defined by

$$Y_{i,j} := \tilde{X}_{i,j} - u_j.$$

The test statistic for testing $H_{\xi_j^*,\beta_j^*}$ is then given by

$$T_j\left(Y_{1,j}, \ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right) := -2\log\Lambda\left(Y_{1,j}, \ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right),$$

where the likelihood ratio Λ is defined by

$$\Lambda\left(Y_{1,j}, \ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right) := \frac{L\left(Y_{1,j}, \ldots, Y_{N_{u_j},j}; \xi_j^*, \beta_j^*\right)}{\sup_{(\xi,\beta)} L\left(Y_{1,j}, \ldots, Y_{N_{u_j},j}; \xi, \beta\right)}$$

with log-likelihood function

$$\log L\left(Y_{1,j},\ldots,Y_{N_{u_j},j};\xi,\beta\right) = -N_{u_j}\log\beta - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{N_{u_j}}\log\left(1+\xi\frac{Y_{i,j}}{\beta}\right).$$

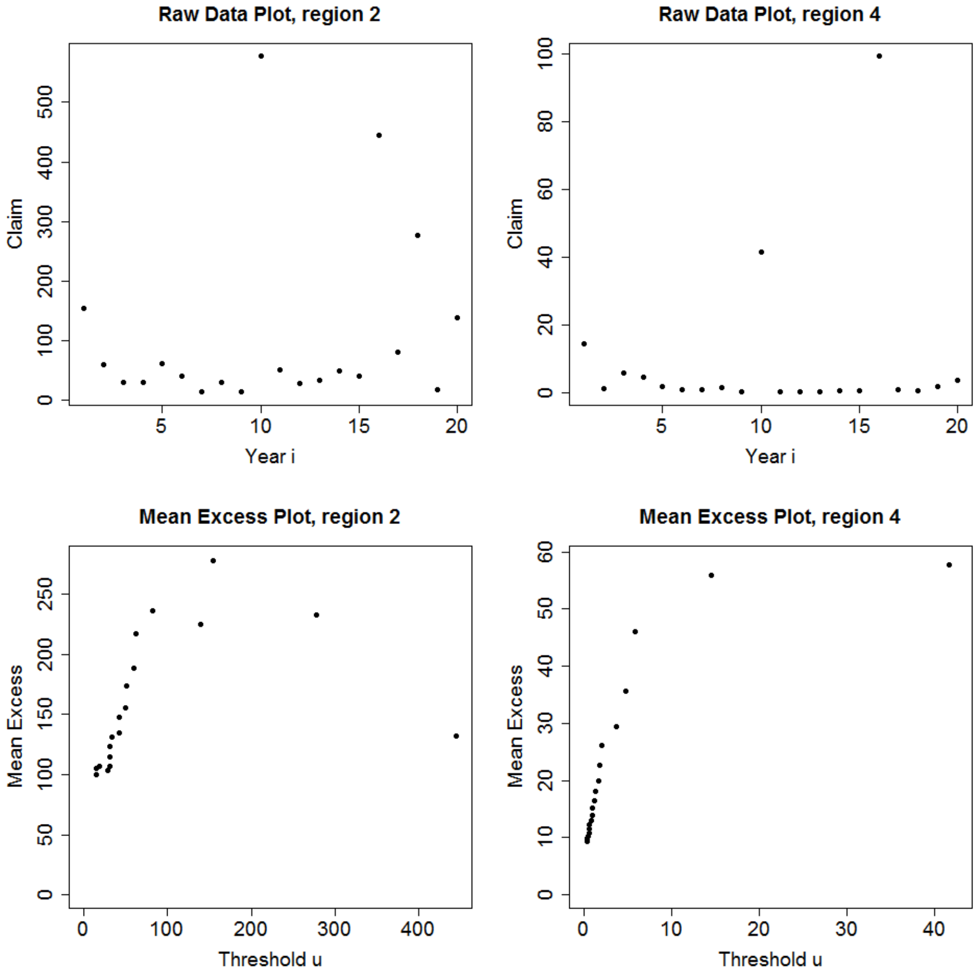

Figure 3: Raw data and mean excess plots for regions 2 and 4. The graphs in the upper panel display the data from Table 4 for $j \in \{2,4\}$, respectively. The graphs in the lower panel show the corresponding mean excess plots.

Under $H_{\xi_j^*,\beta_j^*}$, T_j is asymptotically χ^2-distributed with two degrees-of-freedom. This means that the (asymptotic) confidence interval $C_{\xi_j,\beta_j}(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ in the second step of Algorithm 5.3 is given by

$$C_{\xi_j,\beta_j}(\mathbf{X}_1,\ldots,\mathbf{X}_n)=\left\{\left(\xi_j^*,\beta_j^*\right):T_j\left(Y_{1,j},\ldots,Y_{N_{u_j},j};\xi_j^*,\beta_j^*\right)\le F_{\chi_2^2}^{-1}\left(1-\alpha_{loc}^{(j)}\right)\right\}. \tag{7}$$

Utilizing (7), the confidence region $[\vartheta_j^{\text{lower}},\vartheta_j^{\text{upper}}]$ for ϑ_j based on the third step of Algorithm 5.3 is constructed by finding the minimum value $\vartheta_j^{\text{lower}}=\min q_j(\xi_j^*,\beta_j^*)$ and the maximum value $\vartheta_j^{\text{upper}}=\max q_j(\xi_j^*,\beta_j^*)$, where (ξ_j^*,β_j^*) are located on the boundary of $C_{\xi_j,\beta_j}(\mathbf{x}_1,\ldots,\mathbf{x}_n)$.

A graphical method for the determination of a suitable threshold u_j is based on the mean excess plot in coordinate j; see [22, Section 7.2.2] for details. Namely, all possible values u of u_j are plotted against the mean of the values of $Y_{1,j},\ldots,Y_{N_u,j}$. If the GPD model is appropriate, the plot should yield an approximately linear graph for arguments exceeding u_j. Usually the few largest values of u are ignored, because they lead to very small values of N_u.

For example, Figure 3 shows the mean excess plots for the two regions 2 and 4. The mean excess plot for region 2 is approximately linear when ignoring the three smallest and the four largest values of u. This means that a suitable threshold u_2 would be between 18.815 and 28.316. Similarly, the mean excess plot for region 4 is approximately linear when ignoring the two largest values of u, hence $u_4<0.321$. Based on such considerations, we chose the thresholds $\mathbf{u}=(u_1,\ldots,u_{19})^\top$ given by

$$\mathbf{u}:=(1.0,28.0,9.0,0.3,0.2,0.4,2.6,1.2,0.4,1.1,0.1,0.2,22.5,1.6,3.2,0.2,12.5,1.2,0.5)^\top.$$

Finally, it remains to determine the local significance levels $\left(\alpha_{loc}^{(j)}\right)_{1\le j\le 19}$. In the case of the Bonferroni or the Šidák method, this is trivial. To calibrate the local significance levels with the Bernstein method, we employed a modified version of Algorithm 4.1 based on the empirical excess distribution. Algorithm 5.4 yields a resampling-based approximation of the copula of the vector $\mathbf{T}=(T_1,\ldots,T_m)^\top$ of the region-specific likelihood ratio test statistics.

Algorithm 5.4 (Calibration of the local significance levels for FWER level α)**.**

(a) For every $1\le j\le m$, estimate the parameters ξ_j and β_j of the excess distribution of X_j via maximum likelihood and calculate N_{u_j}.

(b) Choose a number M of Monte Carlo repetitions.

(c) For each $1\le b\le M$ draw a pseudo sample $\mathbf{U}_1^{\#b},\ldots,\mathbf{U}_n^{\#b}$ from the (empirical) Bernstein copula $\hat{B}_{n,\mathbf{K}}$ and calculate the corresponding GPD excesses

$$Y_{i,j}^{\#b}=G_{\hat{\xi}_j,\hat{\beta}_j}^{\leftarrow}\left(U_{(i),j}^{\#b}\right),1\le i\le N_{u_j},1\le j\le m,$$

where $U_{(i),j}^{\#b}$ denotes the i-th reverse order statistic of $\left(U_{i,j}^{\#b}\right)_{1\le i\le n}$.

(d) *For each* $1 \le j \le m$*, compute* $T_j^{\#b} = T_j\left(Y_{1,j}^{\#b}, \ldots, Y_{N_{u_j},j}^{\#b}; \hat{\xi}_j, \hat{\beta}_j\right)$*, and obtain the pseudo-sample*

$$V_j^{\#b} = \hat{G}_{j,M}\left(T_j^{\#b}\right), 1 \le j \le m$$

from the copula of $\mathbf{T}$*.*

(e) *Finally, calibrate* $\hat{\boldsymbol{\alpha}}_{loc,n}$ *by solving*

$$\#\left\{\mathbf{V}^{\#b} \le \mathbf{1} - \hat{\boldsymbol{\alpha}}_{loc,n}\right\} = \lceil (1-\alpha)\, M \rceil\, .$$

Table 2 displays the parameter estimates for the region-specific GPD models, and Table 3 displays the lower bounds $\left(\vartheta_j^{\text{lower}}\right)_{1 \le j \le m}$ of the region-specific confidence intervals for the 99.5% VaR obtained by the Bonferroni, Šidák and Bernstein copula calibration, respectively.

$\hat{\xi}_j$	0.41	1.17	0.75	1.43	0.87	1.51	1.10	0.30	0.49	0.79
	0.56	0.98	1.00	0.73	0.47	0.81	1.08	0.60	0.89	
$\hat{\beta}_j$	19.59	22.21	18.41	0.82	1.10	1.56	4.57	9.75	2.91	6.46
	0.64	0.99	5.12	3.42	20.34	4.52	6.98	1.96	1.64	

Table 2: Estimated parameters $\hat{\xi}_j$, $\hat{\beta}_j$, $1 \le j \le 19$, for the region-specific GPD models. Estimation has been performed via maximum likelihood.

Bonferroni	89.08	283.30	126.20	19.41	10.00	36.68	62.57	39.45	14.62	51.14
	3.74	10.13	53.74	25.43	101.62	37.11	84.99	12.79	14.80	
Šidák	89.22	284.03	126.46	19.48	10.03	36.81	62.75	39.51	14.64	51.25
	3.75	10.15	53.82	25.47	101.78	37.20	85.20	12.81	14.83	
Bernstein	91.59	287.32	127.61	19.81	10.13	37.37	63.54	38.82	14.73	51.74
	3.78	10.25	52.89	26.27	99.90	37.58	82.71	12.91	14.98	

Table 3: Lower confidence bounds $\vartheta_j^{\text{lower}}$ for the 99.5% VaR, $1 \le j \le 19$, obtained by the Bonferroni, the Šidák and the Bernstein copula method, respectively.

Similarly as in Algorithm 4.1, an implicit weighting has been employed for the determination of the local significance levels $\left(\alpha_{loc}^{(j)}\right)_{1 \le j \le m}$ in Algorithm 5.4. Therefore, the confidence bounds obtained with the Bernstein copula method are not guaranteed to be more informative (i. e., larger) than the ones obtained by the Bonferroni or the Šidák methods for all regions. However, we observe improvements in almost all regions j. It is remarkable that this expected behavior of the Bernstein copula calibration can already be verified for the rather moderate sample size of $n = 20$, because the likelihood ratio tests and the Bernstein copula calibration are both based on asymptotic considerations.

Raw data $x_{i,j}$	region j							
year i	1	2	3	4	5	6	7	8
1	23.664	154.664	40.569	14.504	10.468	7.464	22.202	17.682
2	1.080	59.545	3.297	1.344	1.859	0.477	6.107	7.196
3	21.731	31.049	55.973	5.816	14.869	20.771	3.580	14.509
4	28.990	31.052	30.328	4.709	0.717	3.530	6.032	6.512
5	53.616	62.027	57.639	1.804	2.073	4.361	46.018	22.612
6	29.950	41.722	12.964	1.127	1.063	4.873	6.571	11.966
7	3.474	14.429	10.869	0.945	2.198	1.484	4.547	2.556
8	10.020	31.283	21.116	1.663	2.153	0.932	25.163	3.222
9	5.816	14.804	128.072	0.523	0.324	0.477	3.049	7.791
10	170.725	576.767	108.361	41.599	20.253	35.412	126.698	71.079
11	21.423	50.595	4.360	0.327	1.566	64.621	5.650	1.258
12	6.380	28.316	3.740	0.442	0.736	0.470	3.406	7.859
13	124.665	33.359	14.712	0.321	0.975	2.005	3.981	4.769
14	20.165	49.948	17.658	0.595	0.548	29.350	6.782	4.873
15	78.106	41.681	13.753	0.585	0.259	0.765	7.013	9.426
16	11.067	444.712	365.351	99.366	8.856	28.654	10.589	13.621
17	6.704	81.895	14.266	0.972	0.519	0.644	8.057	18.071
18	15.550	277.643	26.564	0.788	0.225	1.230	26.800	64.538
19	10.099	18.815	9.352	2.051	1.089	6.102	2.678	4.064
20	8.492	138.708	46.708	3.680	1.132	1.698	165.600	7.926

year i	9	10	11	12	13	14	15	16	17	18	19
1	12.395	18.551	1.842	4.100	46.135	14.698	44.441	7.981	35.833	10.689	7.299
2	1.436	3.720	0.429	1.026	7.469	7.058	4.512	0.762	14.474	9.337	0.740
3	17.175	87.307	0.209	2.344	22.651	4.117	26.586	3.920	13.804	2.683	3.026
4	0.682	3.115	0.521	0.696	31.126	1.878	29.423	6.394	18.064	1.201	0.894
5	1.581	11.179	2.715	1.327	40.156	4.655	104.691	28.579	17.832	1.618	3.402
6	15.676	24.263	4.832	0.701	16.712	11.852	29.234	7.098	17.866	5.206	5.664
7	0.456	1.137	0.268	0.580	11.851	2.057	11.605	0.282	16.925	2.082	1.008
8	1.581	5.477	0.741	0.369	3.814	1.869	8.126	1.032	14.985	1.390	1.703
9	4.079	7.002	0.524	6.554	5.459	3.007	8.528	1.920	5.638	2.149	2.908
10	21.762	64.582	9.882	6.401	106.197	44.912	191.809	90.559	154.492	36.626	36.276
11	0.626	3.556	1.052	8.277	22.564	8.961	19.817	16.437	25.990	2.364	6.434
12	0.894	3.591	0.136	0.364	28.000	7.574	3.213	1.749	12.735	1.744	0.558
13	2.006	1.973	1.990	15.176	57.235	23.686	110.035	17.373	7.276	2.494	0.525
14	2.921	6.394	0.630	0.762	25.897	3.439	8.161	3.327	24.733	2.807	1.618
15	2.180	3.769	0.770	15.024	36.068	1.613	6.127	8.103	12.596	4.894	0.822
16	9.589	19.485	0.287	0.464	24.211	38.616	51.889	1.316	173.080	3.557	11.627
17	5.515	13.163	0.590	2.745	16.124	2.398	20.997	2.515	5.161	2.840	3.002
18	2.637	80.711	0.245	0.217	12.416	4.972	59.417	3.762	24.603	7.404	19.107
19	2.373	2.057	0.415	0.351	10.707	2.468	10.673	1.743	27.266	1.368	0.644
20	2.972	5.237	0.566	0.708	22.646	6.652	14.437	63.692	113.231	7.218	2.548

Table 4: Insurance claim data from 19 adjacent geographical regions over 20 years.

We omitted the values of $\left(\vartheta_j^{\text{upper}}\right)_{1 \leq j \leq m}$, because they are uninformative (extremely large). This is in line with the fact that all scale parameter estimates $\hat{\xi}_j$ in Table 2 are positive. For $\xi \geq 0$, the GPD has infinite support, thus the modeled 99.5% VaR tends to be very large.

6 Discussion

We have derived a nonparametric approach to the calibration of multivariate STPs which take the joint distribution of test statistics into account. In contrast to previous approaches which were restricted to cases with low-dimensional copula parameters, the Bernstein copula-based approximation of the local significance levels proposed in the present work can be applied under almost no assumptions regarding the dependency structures among test statistics or p-values, respectively. This makes the proposed methodology an attractive choice for data which have not been explicitly modeled prior to the statistical analysis. Furthermore, our empirical results on simulated as well as on real-life data indicate the gain in power which is possible by an explicit consideration of the dependency structure among test statistics in the calibration of the multiple test. This is particularly important for modern applications with high dimensionality of, but also pronounced dependencies in the data.

On the other hand, Theorem 3.5 provides a precise asymptotic performance guarantee for the empirically calibrated multiple test, meaning that a sharp upper bound for its realized FWER can be obtained, at least asymptotically for large sample sizes. This is in contrast to most of the existing resampling-based multiple test procedures like the 'max T' and 'min P' tests proposed in [35], which are obvious competitors of our approach.

Future work shall explore the case that some qualitative assumptions regarding the dependency structure are at hand. For example, it will be interesting to quantify the uncertainty of the FWER of an STP which is calibrated by assuming an Archimedean p-value copula as in [3]. In this case, nonparametric estimation of the copula generator function as for instance proposed in [19] will lead to an empirical calibration of the multiple test.

7 Appendix

In this section several lemmas are formulated and proved which are used in the proofs of Theorem 2.1 and Theorem 2.2.

Lemma 7.1. *It holds that*

a) $\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]}\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left(\frac{k_j}{K_j}-u_j\right)^2\cdot\prod_{j=1}^{m}P_{k_j,K_j}(u_j)\leq\frac{1}{2}\sum_{j=1}^{m}K_j^{-1}$

b) $\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]}\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left|\frac{k_j}{K_j}-u_j\right|\cdot\prod_{j=1}^{m}P_{k_j,K_j}(u_j)\leq\frac{1}{2}\sum_{j=1}^{m}K_j^{-1/2}$

c) $\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]}\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left|\frac{k_j\wedge k_j'}{K_j}-u_j\right|\cdot\prod_{j=1}^{m}P_{k_j,K_j}(u_j)P_{k_j',K_j}(u_j)\leq\sum_{j=1}^{m}K_j^{-1/2}$

Proof.

a) Using the fact that $P_{k_j,K_j}(u_j)$ is the probability function of the binomial distribution for each $u_j\in[0,1]$ and $j=1,\ldots,m$, we get

$$\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]}\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left(\frac{k_j}{K_j}-u_j\right)^2\cdot\prod_{j=1}^{m}P_{k_j,K_j}(u_j)=\sum_{j=1}^{m}\sup_{u_j\in[0,1]}\sum_{k_j=0}^{K_j}\left(\frac{k_j}{K_j}-u_j\right)^2P_{k_j,K_j}(u_j)$$
$$=\sum_{j=1}^{m}\sup_{u_j\in[0,1]}\frac{u_j(1-u_j)}{K_j}=\frac{1}{4}\sum_{j=1}^{m}K_j^{-1}.$$

b) Similarly, with the Jensen inequality we get

$$\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]}\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left|\frac{k_j}{K_j}-u_j\right|\cdot\prod_{j=1}^{m}P_{k_j,K_j}(u_j)=\sum_{j=1}^{m}\sup_{u_j\in[0,1]}\sum_{k_j=0}^{K_j}\left|\frac{k_j}{K_j}-u_j\right|P_{k_j,K_j}(u_j)$$
$$\leq\sum_{j=1}^{m}\sup_{u_j\in[0,1]}\left(\sum_{k_j=0}^{K_j}\left(\frac{k_j}{K_j}-u_j\right)^2P_{k_j,K_j}(u_j)\right)^{1/2}$$
$$=\sum_{j=1}^{m}\sup_{u_j\in[0,1]}\left(\frac{u_j(1-u_j)}{K_j}\right)^{1/2}=\frac{1}{2}\sum_{j=1}^{m}K_j^{-1/2}.$$

c) Finally, we obtain

$$\sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]}\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left|\frac{k_j\wedge k_j'}{K_j}-u_j\right|\cdot\prod_{j=1}^{m}P_{k_j,K_j}(u_j)P_{k_j',K_j}(u_j)$$
$$=\sum_{j=1}^{m}\sum_{k_j=0}^{K_j}\sum_{k_j'=0}^{K_j}\left|\frac{k_j\wedge k_j'}{K_j}-u_j\right|P_{k_j,K_j}(u_j)P_{k_j',K_j}(u_j)$$
$$=\sum_{j=1}^{m}\sum_{k_j=0}^{K_j}\sum_{k_j'=k_j}^{K_j}\left|\frac{k_j}{K_j}-u_j\right|P_{k_j,K_j}(u_j)P_{k_j',K_j}(u_j)$$
$$+\sum_{j=1}^{m}\sum_{k_j'=0}^{K_j}\sum_{k_j=k_j'+1}^{K_j}\left|\frac{k_j'}{K_j}-u_j\right|P_{k_j,K_j}(u_j)P_{k_j',K_j}(u_j)$$
$$\leq\sum_{j=1}^{m}\sum_{k_j=0}^{K_j}\left|\frac{k_j}{K_j}-u_j\right|P_{k_j,K_j}(u_j)+\sum_{j=1}^{m}\sum_{k_j'=0}^{K_j}\left|\frac{k_j'}{K_j}-u_j\right|P_{k_j',K_j}(u_j)\leq\sum_{j=1}^{m}K_j^{-1/2},$$

where the last inequality follows from part b) of the lemma.

□

Lemma 7.2. *It holds that*

$$\|B_K - C\|_\infty \leq \frac{1}{2}\sum_{j=1}^{m} K_j^{-1/2},$$

where $\|g\|_\infty := \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} |g(\mathbf{u})|$ *for* $g : [\mathbf{0},\mathbf{1}] \longrightarrow \mathbb{R}$.

Proof. We get

$$\begin{aligned}\|B_{\mathbf{K}} - C\|_\infty &\leq \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} |C(\mathbf{k}/\mathbf{K}) - C(\mathbf{u})| \prod_{j=1}^{m} P_{k_j,K_j}(u_j) \\ &\leq \sup_{\mathbf{u}\in[\mathbf{0},\mathbf{1}]} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \sum_{j=1}^{m} \left|\frac{k_j}{K_j} - u_j\right| \cdot \prod_{j=1}^{m} P_{k_j,K_j}(u_j),\end{aligned}$$

where the last inequality follows from the Lipschitz property of multivariate copula (cf., [26, Section 2]). The rest of the proof follows from part b) of Lemma 7.1. □

Lemma 7.3. *Assume bounded second order partial derivatives for* C *on* $(\mathbf{0},\mathbf{1})$ *and let* $\mathbf{u} \in (\mathbf{0},\mathbf{1})$. *Then,*

$$n^{1/2}\cdot\left(\hat{B}_{n,\mathbf{K}}(\mathbf{u}) - B_{\mathbf{K}}(\mathbf{u})\right) = n^{-1/2}\cdot\sum_{i=1}^{n} Y_{i,\mathbf{K}}(\mathbf{u}) + O\left(n^{-1/4}\cdot(\log n)^{1/2}\cdot(\log\log n)^{1/4}\right)$$

where

$$\begin{aligned}Y_{i,\mathbf{K}}(\mathbf{u}) := \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} \Bigg(&\mathbb{1}_{(-\boldsymbol{\infty},\mathbf{k}/\mathbf{K}]}(\mathbf{U}_i) - C(\mathbf{k}/\mathbf{K}) \\ &- \sum_{j=1}^{m} \partial_j C(\mathbf{k}/\mathbf{K}) \left(\mathbb{1}_{(-\infty,k_j/K_j]}(U_{i,j}) - \frac{k_j}{K_j}\right)\Bigg) \cdot \prod_{j=1}^{m} P_{k_j,K_j}(u_j)\end{aligned}$$

and $U_{i,j} = F_j(X_{i,j})$.

Proof. It holds that

$$n^{1/2}\cdot\left(\hat{B}_{n,\mathbf{K}}(\mathbf{u}) - B_{\mathbf{K}}(\mathbf{u})\right) = \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}} n^{1/2}\cdot\left(\hat{C}_n(\mathbf{k}/\mathbf{K}) - C(\mathbf{k}/\mathbf{K})\right)\prod_{j=1}^{m} P_{k_j,K_j}(u_j).$$

From [33, Section 4] we get

$$\begin{aligned}n^{1/2}\cdot\left(\hat{C}_n(\mathbf{k}/\mathbf{K}) - C(\mathbf{k}/\mathbf{K})\right) =& \bar{\alpha}_n(\mathbf{k}/\mathbf{K}) - \sum_{j=1}^{m} \partial_j C(\mathbf{k}/\mathbf{K})\, \bar{\alpha}_n(1,\ldots,k_j/K_j,\ldots,1) \\ &+ O\left(n^{-1/4}(\log n)^{1/2}\cdot(\log\log n)^{1/4}\right)\end{aligned}$$

where $\bar{\alpha}_n(\mathbf{k}/\mathbf{K}) := n^{-1/2}\cdot\sum_{i=1}^{n}\left(\mathbb{1}_{(-\boldsymbol{\infty},\mathbf{k}/\mathbf{K}]}(\mathbf{U}_i) - C(\mathbf{k}/\mathbf{K})\right)$ and $\mathbf{U}_i = \mathbf{F}(\mathbf{X}_i) \sim C$. The result of the lemma now follows directly from the definition of $Y_{i,\mathbf{K}}$. □

Lemma 7.4. *Assume bounded second order partial derivatives for C on $(\mathbf{0},\mathbf{1})$ and let $\mathbf{u}\in(\mathbf{0},\mathbf{1})$. Then, for all $i\in\{1,...,n\}$ and $\mathbf{K}\in\mathbb{N}^m$ we get that $\mathbb{E}\left[Y_{i,\mathbf{K}}\left(\mathbf{u}\right)\right]=0$ and*

$$\mathbb{V}\left[Y_{i,\mathbf{K}}\left(\mathbf{u}\right)\right]=\sigma^2\left(\mathbf{u}\right)+O\left(\sum_{j=1}^m K_j^{-1/2}\right)$$

where

$$\begin{aligned}\sigma^2\left(\mathbf{u}\right):=&C\left(\mathbf{u}\right)\cdot\left(1-C\left(\mathbf{u}\right)\right)+\sum_{j=1}^m\left(\partial_jC\left(\mathbf{u}\right)\right)^2u_j\left(1-u_j\right)\\&-2\sum_{j=1}^m\partial_jC\left(\mathbf{u}\right)C\left(\mathbf{u}\right)\left(1-u_j\right)\\&+2\sum_{j=1}^m\sum_{j'=j+1}^m\partial_jC\left(\mathbf{u}\right)\partial_{j'}C\left(\mathbf{u}\right)\left(\mathbb{P}\left[U_{i,j}\le u_j,U_{i,j'}\le u_{j'}\right]-u_ju_{j'}\right).\end{aligned}$$

Remark. The variance of $Y_{i,\mathbf{K}}\left(\mathbf{u}\right)$ only depends on n.

Proof. Since $\mathbb{E}\left[\mathbb{1}_{(-\boldsymbol{\infty},\mathbf{k}/\mathbf{K}]}(\mathbf{U}_i)\right]=C\left(\mathbf{k}/\mathbf{K}\right)$ and $\mathbb{E}\left[\mathbb{1}_{(-\infty,k_j/K_j]}(U_{i,j})\right]=\frac{k_j}{K_j}$, we get

$$\mathbb{E}\left[Y_{i,\mathbf{K}}\left(\mathbf{u}\right)\right]=0.$$

Let $f_C\left(\mathbf{k}\right):=\sum_{j=1}^m\partial_jC\left(\mathbf{k}/\mathbf{K}\right)\left(\mathbb{1}_{(-\infty,k_j/K_j]}(U_{i,j})-\frac{k_j}{K_j}\right)$. It holds that

$$\begin{aligned}\mathbb{V}\left[Y_{i,\mathbf{K}}\left(\mathbf{u}\right)\right]=&\mathbb{E}\left[Y_{i,\mathbf{K}}\left(\mathbf{u}\right)^2\right]\\=&\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}C\left(\left(\mathbf{k}\wedge\mathbf{k}'\right)/\mathbf{K}\right)\prod_{j=1}^mP_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\\&-2\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}C\left(\mathbf{k}/\mathbf{K}\right)C\left(\mathbf{k}'/\mathbf{K}\right)\prod_{j=1}^mP_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\\&-2\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^m\partial_jC\left(\mathbf{k}'/\mathbf{K}\right)\left(C\left(\frac{k_1}{K_1},...,\frac{k_j\wedge k'_j}{K_j},...,\frac{k_m}{K_m}\right)\right.\\&\left.-C\left(\mathbf{k}/\mathbf{K}\right)\frac{k'_j}{K_j}\right)\cdot\prod_{j=1}^mP_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\\&+\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}C\left(\mathbf{k}/\mathbf{K}\right)C\left(\mathbf{k}'/\mathbf{K}\right)\prod_{j=1}^mP_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\\&+2\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}C\left(\mathbf{k}/\mathbf{K}\right)\mathbb{E}\left[f_C\left(\mathbf{k}'\right)\right]\prod_{j=1}^mP_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\\&+\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\mathbb{E}\left[f_C\left(\mathbf{k}\right)f_C\left(\mathbf{k}'\right)\right]\prod_{j=1}^mP_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\end{aligned}$$

Since $\mathbb{E}\left[f_C\left(\mathbf{k}\right)\right]=0$, we have

$$
\begin{aligned}
\mathbb{V}\left[Y_{i,\mathbf{K}}\left(\mathbf{u},\cdot\right)\right]=&\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}C\left(\left(\mathbf{k}\wedge\mathbf{k}'\right)/\mathbf{K}\right)\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right) && \text{(I)}\\
&-\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}C\left(\mathbf{k}/\mathbf{K}\right)C\left(\mathbf{k}'/\mathbf{K}\right)\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right) && \text{(II)}\\
&-2\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\partial_jC\left(\mathbf{k}'/\mathbf{K}\right)\left(C\left(\frac{k_1}{K_1},\dots,\frac{k_j\wedge k'_j}{K_j},\dots,\frac{k_m}{K_m}\right)\right.\\
&\left.-C\left(\mathbf{k}/\mathbf{K}\right)\frac{k'_j}{K_j}\right)\cdot\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right) && \text{(III)}\\
&+\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\sum_{j'=1}^{m}\partial_jC\left(\mathbf{k}/\mathbf{K}\right)\partial_{j'}C\left(\mathbf{k}'/\mathbf{K}\right)\\
&\left(\mathbb{P}\left[U_{i,j}\leq\frac{k_j}{K_j},U_{i,j'}\leq\frac{k'_{j'}}{K_{j'}}\right]-\frac{k_j}{K_j}\cdot\frac{k'_{j'}}{K_{j'}}\right)\cdot\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right) && \text{(IV)}
\end{aligned}
$$

By the Lipschitz property of multivariate copula (cf., [26]) and Lemma 7.1 we conclude for term (I)

$$
\begin{aligned}
\text{(I)}=&C\left(\mathbf{u}\right)+\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}O\left(\sum_{j=1}^{m}\left|\frac{k_j\wedge k'_j}{K_j}-u_j\right|\right)\cdot\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k'_j,K_j}\left(u_j\right)\\
=&C\left(\mathbf{u}\right)+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right).
\end{aligned}
$$

By Lemma 7.2 we get for term (II)

$$
\begin{aligned}
\text{(II)}&=-B_{\mathbf{K}}\left(\mathbf{u}\right)^2\\
&=-\left(C\left(\mathbf{u}\right)+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right)\right)^2\\
&=-C\left(\mathbf{u}\right)^2+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right).
\end{aligned}
$$

For term (III), we get by Taylor series expansion that

$$
\begin{aligned}
\text{(III)} =& -2\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\left(\left[\partial_j C\left(\mathbf{u}\right)+O\left(\sum_{j'=1}^{m}K_{j'}^{-1/2}\right)\right]\right.\\
&\cdot\left[C\left(\mathbf{u}\right)+O\left(\sum_{j'=1,j'\neq j}^{m}\left|\frac{k_{j'}}{K_{j'}}-u_{j'}\right|\right)+O\left(\left|\frac{k_j\wedge k_j'}{K_j}-u_j\right|\right)\right.\\
&\left.\left.-\left(C\left(\mathbf{u}\right)+O\left(\sum_{j'=1}^{m}K_{j'}^{-1/2}\right)\right)\cdot\frac{k_j'}{K_j}\right]\right)\cdot\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k_j',K_j}\left(u_j\right)\\
=& -2\sum_{j=1}^{m}\left(\partial_j C\left(\mathbf{u}\right)\left[C\left(\mathbf{u}\right)-C\left(\mathbf{u}\right)u_j\right]\right)+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right)
\end{aligned}
$$

And, finally, for term (IV) we have

$$
\begin{aligned}
\text{(IV)} =& \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{K}}\sum_{\mathbf{k}'=\mathbf{0}}^{\mathbf{K}}\sum_{j=1}^{m}\sum_{j'=1}^{m}\left(\partial_j C\left(\mathbf{u}\right)+O\left(\sum_{j''=1}^{m}K_{j''}^{-1/2}\right)\right)\left(\partial_{j'} C\left(\mathbf{u}\right)+O\left(\sum_{j''=1}^{m}K_{j''}^{-1/2}\right)\right)\\
&\cdot\left(\mathbb{P}\left[U_{i,j}\leq u_j,U_{i,j'}\leq u_{j'}\right]+O\left(\left|\frac{k_j}{K_j}-u_j\right|\right)+O\left(\left|\frac{k_{j'}'}{K_{j'}}-u_{j'}\right|\right)-\frac{k_j}{K_j}\cdot\frac{k_{j'}'}{K_{j'}}\right)\\
&\cdot\prod_{j=1}^{m}P_{k_j,K_j}\left(u_j\right)P_{k_j',K_j}\left(u_j\right)\\
=&\sum_{j=1}^{m}\sum_{j'=1}^{m}\partial_j C\left(\mathbf{u}\right)\partial_{j'} C\left(\mathbf{u}\right)\left(\mathbb{P}\left[U_{i,j}\leq u_j,U_{i,j'}\leq u_{j'}\right]+2O\left(\sum_{j''=1}^{m}K_{j''}^{-1/2}\right)-u_ju_{j'}\right)\\
&+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right)\\
=&\sum_{j=1}^{m}\sum_{j'=1}^{m}\partial_j C\left(\mathbf{u}\right)\partial_{j'} C\left(\mathbf{u}\right)\left(\mathbb{P}\left[U_{i,j}\leq u_j,U_{i,j'}\leq u_{j'}\right]-u_ju_{j'}\right)+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathbb{V}\left[Y_{i,\mathbf{K}}\left(\mathbf{u}\right)\right] =& C\left(\mathbf{u}\right)\cdot\left(1-C\left(\mathbf{u}\right)\right)+\sum_{j=1}^{m}\left(\partial_j C\left(\mathbf{u}\right)\right)^2u_j\left(1-u_j\right)-2\sum_{j=1}^{m}\partial_j C\left(\mathbf{u}\right)C\left(\mathbf{u}\right)\left(1-u_j\right)\\
&+2\sum_{j=1}^{m}\sum_{j'=j+1}^{m}\partial_j C\left(\mathbf{u}\right)\partial_{j'} C\left(\mathbf{u}\right)\left(\mathbb{P}\left[U_{i,j}\leq u_j,U_{i,j'}\leq u_{j'}\right]-u_ju_{j'}\right)\\
&+O\left(\sum_{j=1}^{m}K_j^{-1/2}\right).
\end{aligned}
$$

The first term comes from (I) and (II), the second from (IV), the third from (III), the forth from (IV). $\square$

Lemma 7.5. *Let* $\sum_{j=1}^{m} K_j^{-1/2} \to 0$, $n \to \infty$. *Then for all* $\mathbf{u} \in (\mathbf{0}, \mathbf{1})$ *and all sequences* $(\mathbf{u}_n)_{n\in\mathbb{N}}$ *in* $(\mathbf{0}, \mathbf{1})$ *with* $\lim_{n\to\infty} \mathbf{u}_n = \mathbf{u}$

$$n^{-1/2} \sum_{i=1}^{n} Y_{i,\mathbf{K}}(\mathbf{u}_n) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2(\mathbf{u})\right).$$

Proof. We want to use the Lindeberg Central Limit Theorem (see [30, Section 1.9.3] or [20, Theorem 11.2.5]). Let $X_{i,n} := n^{-1/2} Y_{i,\mathbf{K}}(\mathbf{u}_n), 1 \le i \le n$. Because of Lemma 7.1, the remainder term in Lemma 7.4 does not depend on $\mathbf{u}$. Therefore, we have $\mathbb{E}[X_{i,n}] = 0$ and

$$\begin{aligned}
\sigma_n^2 &:= \mathbb{V}\left[\sum_{i=1}^{n} X_{i,n}\right] \\
&= \mathbb{V}\left[Y_{1,\mathbf{K}}(\mathbf{u}_n)\right] \\
&= \sigma^2(\mathbf{u}_n) + O\left(\sum_{j=1}^{m} K_j^{-1/2}\right) \\
&\xrightarrow{n\to\infty} \sigma^2(\mathbf{u}),
\end{aligned}$$

because σ^2 is continuous. From the convergence $|X_{1,n}| \to 0$ $a.s.$ and the dominated convergence theorem, we get the Lindeberg Condition, i.e.

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}\left[X_{i,n}^2 \mathbb{1}_{\{|X_{i,n}|>\epsilon\sigma_n\}}\right] \sigma_n^{-2} &= \mathbb{E}\left[Y_{1,n}^2 \mathbb{1}_{\{|X_{1,n}|>\epsilon\sigma_n\}}\right] \sigma_n^{-2} \\
&\xrightarrow{n\to\infty} 0
\end{aligned}$$

for all $\epsilon > 0$. The result follows from [30, Section 1.9.3]. $\square$

Lemma 7.6. *Let* $p \in (0,1)$. *Suppose that* $\tilde{C}'(C^{\leftarrow}(p)) > 0$ *exists, then*

$$\sqrt{n}\left(\hat{B}_{n,\mathbf{K}}^{\leftarrow}(p) - C^{\leftarrow}(p)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma\left(C^{\leftarrow}(p), \ldots, C^{\leftarrow}(p)\right)}{\tilde{C}'\left(C^{\leftarrow}(p)\right)}\right),$$

where $C^{\leftarrow}(p) := \inf\{u \in [0,1] \,|\, C(u,\ldots,u) \ge p\}$ *and* $\tilde{C}(u) := C(u,\ldots,u)$.

Proof. We argue similarly to the proof of Theorem A in [30, Section 2.3.3]. Fix $p \in (0,1)$ and let

$$G_n(t) := \mathbb{P}\left[\frac{n^{1/2}\left(\hat{B}_{n,\mathbf{K}}^{\leftarrow}(p) - C^{\leftarrow}(p)\right)}{\tilde{\sigma}} \le t\right],$$

where $\tilde{\sigma}^2 := \frac{\sigma(C^{\leftarrow}(p),\dots,C^{\leftarrow}(p))}{\tilde{C}'(C^{\leftarrow}(p))}$. Let $u_n := t\tilde{\sigma}n^{-1/2} + C^{\leftarrow}(p)$. We have

$$\begin{aligned} G_n(t) &= \mathbb{P}\left[\hat{B}^{\leftarrow}_{n,\mathbf{K}}(p) \le u_n\right] \\ &= \mathbb{P}\left[p \le \hat{B}_{n,\mathbf{K}}(u_n,\dots,u_n)\right] \end{aligned}$$

Put $c_{nt} := \frac{n^{1/2}(C(u_n,\dots,u_n)-p)}{\sigma(u_n)}$. Then it holds that

$$G_n(t) = \mathbb{P}\left[-c_{nt} \le Z_n\right],$$

where $Z_n := \frac{n^{1/2}\left(B_{n,\mathbf{K}}(u_n,\dots,u_n)-C(u_n,\dots,u_n)\right)}{\sigma(u_n)}$. Furthermore, we get

$$\begin{aligned} \Phi(t) - G_n(t) &= \mathbb{P}\left[Z_n < -c_{nt}\right] - (1-\Phi(t)) \\ &= \mathbb{P}\left[Z_n < -c_{nt}\right] - \Phi(-c_{nt}) + \Phi(t) - \Phi(c_{nt}) \end{aligned} \tag{1}$$

Since C and $\partial_j C, 1 \le j \le m$ are continuous, we have

$$\begin{aligned} \lim_{n\to\infty} c_{nt} &= \lim_{n\to\infty}\left(t\cdot\frac{\tilde{\sigma}}{\sigma(u_n,\dots,u_n)}\cdot\frac{\tilde{C}(u_n)-\tilde{C}(C^{\leftarrow}(p))}{t\tilde{\sigma}n^{-1/2}}\right) \\ &= t\cdot\frac{\tilde{\sigma}}{\sigma(C^{\leftarrow}(p),\dots,C^{\leftarrow}(p))}\cdot\tilde{C}'(C^{\leftarrow}(p)) \\ &= t. \end{aligned}$$

Further, by Theorem 2.2 and Polya's Theorem (see [30, Section 1.5.3]), since Φ is continuous, we have

$$\lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[Z_n \le x\right] - \Phi(x)\right| = 0.$$

Using these two properties, (1) results in

$$\begin{aligned} \lim_{n\to\infty}\left|\Phi(t) - G_n(t)\right| &\le \lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[Z_n < x\right] - \Phi(x)\right| + \lim_{n\to\infty}\left|\Phi(t) - \Phi(c_{nt})\right| \\ &= 0. \end{aligned}$$

□

References

[1] Aitchison, J. (1964). Confidence-region tests. *J. R. Stat. Soc., Ser. B*, 26:462–476.

[2] Belalia, M. (2016). On the asymptotic properties of the Bernstein estimator of the multivariate distribution function. *Stat. Probab. Lett.*, in press.

[3] Bodnar, T. and Dickhaus, T. (2014). False discovery rate control under Archimedean copula. *Electron. J. Stat.*, 8(2):2207–2241.

[4] Bouzebda, S. and Zari, T. (2013). Strong approximation of empirical copula processes by Gaussian processes. *Statistics*, 47(5):1047–1063.

[5] Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Stat. Probab. Lett.*, 80(23-24):1925–1932.

[6] Cerqueti, R., Costantini, M., and Lupi, C. (2012). A copula-based analysis of false discovery rate control under dependence assumptions. Economics & Statistics Discussion Paper 065/12, Università degli Studi del Molise, Dipartimento di Scienze Economiche, Gestionali e Sociali (SEGeS).

[7] Cottin, C. and Pfeifer, D. (2014). From Bernstein polynomials to Bernstein copulas. *J. Appl. Funct. Anal.*, 9(3-4):277–288.

[8] Deheuvels, P. (1979). La fonction de dependance empirique et ses propriétés. Un test non paramètrique d'independance. *Bull. Cl. Sci., V. Sér., Acad. R. Belg.*, 65:274–292.

[9] Dickhaus, T. (2014). *Simultaneous Statistical Inference with Applications in the Life Sciences*. Springer-Verlag Berlin Heidelberg.

[10] Dickhaus, T. and Gierl, J. (2013). Simultaneous test procedures in terms of p-value copulae. Proceedings on the 2nd Annual International Conference on Computational Mathematics, Computational Geometry & Statistics (CMCGS 2013), pages 75–80. Global Science and Technology Forum (GSTF).

[11] Diers, D., Eling, M., and Marek, S. D. (2012). Dependence modeling in non-life insurance using the Bernstein copula. *Insur. Math. Econ.*, 50(3):430–436.

[12] Embrechts, P., Lindskog, F., and McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In Rachev, S., editor, *Handbook of Heavy Tailed Distributions in Finance*, pages 329–384. Elsevier Science B.V.

[13] Gabriel, K. R. (1969). Simultaneous test procedures - some theory of multiple comparisons. *Ann. Math. Stat.*, 40:224–250.

[14] Härdle, W. K. and Okhrin, O. (2010). De copulis non est disputandum - Copulae: an overview. *AStA Adv. Stat. Anal.*, 94(1):1–31.

[15] Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom. J.*, 50(3):346–363.

[16] Janssen, P., Swanepoel, J., and Veraverbeke, N. (2012). Large sample behavior of the Bernstein copula estimator. *J. Statist. Plann. Inference*, 142(5):1189–1197.

[17] Joe, H. (2014). *Dependence modeling with copulas.* Boca Raton, FL: CRC Press.

[18] Kiefer, J. (1961). On large deviations of the empiric D. F. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.*, 11:649–660.

[19] Lambert, P. (2007). Archimedean copula estimation using Bayesian splines smoothing techniques. *Comput. Stat. Data Anal.*, 51(12):6307–6320.

[20] Lehmann, E. L. and Romano, J. P. (2005). Testing statistical hypotheses (springer texts in statistics).

[21] Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24:1097–1130.

[22] McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative risk management. Concepts, techniques, and tools.* Princeton, NJ: Princeton University Press.

[23] Nelsen, R. B. (2006). *An introduction to copulas. 2nd ed.* Springer Series in Statistics. New York, NY: Springer.

[24] Nikoloulopoulos, A. K., Joe, H., and Li, H. (2009). Extreme value properties of multivariate t copulas. *Extremes*, 12(2):129–148.

[25] Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Stat.*, 4:912–923.

[26] Sancetta, A. and Satchell, S. (2004). The bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, 20(03):535–562.

[27] Schmidt, R., Faldum, A., and Gerß, J. (2015). Adaptive designs with arbitrary dependence structure based on Fisher's combination test. *Stat. Methods Appl.*, 24(3):427–447.

[28] Schmidt, R., Faldum, A., Witt, O., and Gerß, J. (2014). Adaptive designs with arbitrary dependence structure. *Biom. J.*, 56(1):86–106.

[29] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, 18(3):764–782.

[30] Serfling, R. J. (1980). *Approximation theorems of mathematical statistics.* New York: Wiley.

[31] Stange, J., Bodnar, T., and Dickhaus, T. (2015). Uncertainty quantification for the family-wise error rate in multivariate copula models. *AStA Adv. Stat. Anal.*, 99(3):281–310.

[32] Stange, J., Dickhaus, T., Navarro, A., and Schunk, D. (2016). Multiplicity- and dependency-adjusted p-values for control of the family-wise error rate. *Stat. Probab. Lett.*, 111:32–40.

[33] Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab.*, 12(2):361–379.

[34] Swanepoel, J. W. H. (1986). A note on proving that the (modified) bootstrap works. *Comm. Statist. A—Theory Methods*, 15(11):3193–3203.

[35] Westfall, P. H. and Young, S. S. (1993). *Resampling-based multiple testing: examples and methods for p-value adjustment.* Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics. Wiley, New York.