



HAL
open science

Asymptotically Unbiased Estimation of the Coefficient of Tail Dependence

Yuri Goegebeur, Armelle Guillou

► **To cite this version:**

Yuri Goegebeur, Armelle Guillou. Asymptotically Unbiased Estimation of the Coefficient of Tail Dependence. *Scandinavian Journal of Statistics*, 2013, 40 (1), 10.1111/j.1467-9469.2012.00800.x . hal-01311656

HAL Id: hal-01311656

<https://hal.science/hal-01311656>

Submitted on 11 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotically unbiased estimation of the coefficient of tail dependence

YURI GOEGBEUR

Department of Mathematics and Computer Science, University of Southern Denmark

ARMELLE GUILLOU

Institut Recherche Mathématique Avancée, Université de Strasbourg et CNRS

Abstract

We introduce and study a class of weighted functional estimators for the coefficient of tail dependence in bivariate extreme value statistics. Asymptotic normality of these estimators is established under a second order condition on the joint tail behavior, some conditions on the weight function and for appropriately chosen sequences of intermediate order statistics. Asymptotically unbiased estimators are constructed by judiciously chosen linear combinations of weighted functional estimators, and variance optimality within this class of asymptotically unbiased estimators is discussed. The finite sample performance of some specific examples from our class of estimators and some alternatives from the recent literature are evaluated with a small simulation experiment.

Keywords: bias-correction, coefficient of tail dependence, multivariate extremes, second order condition.

Running title: Unbiased estimation of tail dependence

1 Introduction

In recent years, the area of multivariate extreme value theory has received a lot of attention, and many directions have been explored to infer on the characteristics of multivariate extremes, including e.g. the estimation of indices or functions describing tail dependence and the estimation of the probability of extreme failure sets. We refer to Ledford & Tawn (1996, 1997), Beirlant & Vandewalle (2002), Heffernan & Tawn (2004), Draisma *et al.* (2004), Peng (1999, 2010) to name but a few.

Focusing on the bivariate case where we have at our disposal pairs (X_i, Y_i) , $i = 1, \dots, n$, being independent copies of the random vector (X, Y) with joint distribution function F , it is mathematically convenient to assume that the marginals, F_X and F_Y , are known and that they are unit Fréchet distributions, i.e. $F_X(u) = F_Y(u) = \exp(-1/u)$ for $u > 0$. In this case, denoting by $M_{X,n} = \max_{1 \leq i \leq n} X_i$ and by $M_{Y,n} = \max_{1 \leq i \leq n} Y_i$, it is well-known that, if there exists appropriate sequences a_n and $b_n > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{M_{X,n} - a_n}{b_n} \leq x, \frac{M_{Y,n} - a_n}{b_n} \leq y \right) = G(x, y) \quad (1)$$

where G is a non-degenerate distribution function, then G is called a bivariate extreme value distribution.

We say that $M_{X,n}$ and $M_{Y,n}$ are *asymptotically independent* if $G(x, y) = G(x, \infty)G(\infty, y)$, for all x and y . This case is highly relevant in practice as for instance the bivariate normal distribution with correlation coefficient $|\rho| < 1$ (see our examples in Section 4.1) has asymptotically independent margins. Unfortunately, if we want to estimate the probability that both X and Y exceed some high thresholds, convergence (1) is not helpful. Indeed, noting that (1) can be reformulated in terms of the random pair (X, Y) as

$$\mathbb{P} \left(\frac{M_{X,n} - a_n}{b_n} \leq x, \frac{M_{Y,n} - a_n}{b_n} \leq y \right) = F^n(a_n + b_n x, a_n + b_n y),$$

and taking the logarithm, yields after some rearrangements that

$$\lim_{n \rightarrow \infty} n \mathbb{P} \left(\frac{X - a_n}{b_n} > x, \frac{Y - a_n}{b_n} > y \right) = \log G(x, y) - \log G(x, \infty) - \log G(\infty, y),$$

which is exactly 0 in the case of asymptotic independence. To solve this issue, Ledford & Tawn (1996, 1997) introduced a submodel assuming that the function $t \mapsto q(t) := \mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)$ is regularly varying at zero with index $1/\eta$. This means that $q(t) = t^{1/\eta}\ell(t)$, where ℓ is a function slowly varying at zero, i.e. an ultimately positive function satisfying $\ell(tx)/\ell(t) \rightarrow 1$ as $t \downarrow 0$ for all $x > 0$ (see Bingham *et al.*, 1987, p. 15-18). The parameter η is called the coefficient of tail dependence, and satisfies $\eta \in (0, 1]$. Ledford & Tawn (1996, 1997) motivate their model by showing that many important and commonly used bivariate distribution functions can be written in this form; we also refer to Heffernan (2000) for an extensive list of examples. The case $\eta = 1$ and $\lim_{t \rightarrow 0} \ell(t) = c$ for some $0 < c \leq 1$, corresponds to the asymptotic dependence case with degree c , whereas $\eta \in (0, 1)$ or $\eta = 1$ with $\lim_{t \rightarrow 0} \ell(t) = 0$ implies asymptotic independence with degree $2\eta - 1$. Ledford & Tawn (1996) identify three types of asymptotic independence:

- if $\eta \in (0, 1/2)$, the pairs (X, Y) which exceed a same high threshold occur less frequently than if X and Y are independent;
- $\eta = 1/2$: extremes of X and Y are close to independence and are exactly independent if $\ell(\cdot) = 1$;
- $\eta \in (1/2, 1)$ or $\eta = 1$ and $\ell(t) \rightarrow 0$, the pairs (X, Y) which exceed a same high threshold occur more frequently than under exact independence.

Thus the degree of asymptotic dependence depends on both η and ℓ . Larger values of η indicate a stronger association between extreme values of the two components X and Y .

As we can imagine, different methods have been proposed in the literature in order to estimate η from observed data. In particular, as q is regularly varying with index $1/\eta$, setting $Z_i = \min\{X_i, Y_i\}$, we have

$$\mathbb{P}(Z_i > z) = \mathbb{P}(X_i > z, Y_i > z) = z^{-1/\eta}\mathcal{L}(z) \tag{2}$$

with $\mathcal{L}(\cdot)$ a slowly varying function at infinity. Thus the parameter η can be viewed as the extreme value index of the minimum of the two components, and hence all the classical type of estimators can be used, for instance the Hill (1975), moment (Dekkers *et al.*, 1989) or maximum

likelihood (Smith, 1987) estimators. Unfortunately, as usual in the extreme value context, the bias of all these estimators is a big challenge. In the univariate case, bias-reduced estimators have been introduced in e.g. Feuerverger & Hall (1999), Beirlant *et al.* (1999), and more recently in Gomes *et al.* (2008), to name but a few, and such procedures have recently been considered in the bivariate case in Beirlant & Vandewalle (2002) using an exponential regression model and in Beirlant *et al.* (2011) under a Hall-type condition on the slowly varying function \mathcal{L} . The procedure by Beirlant & Vandewalle (2002) performs quite well with respect to bias, though it should be noted that it is not necessarily asymptotically unbiased. The Beirlant *et al.* (2011) procedure on the other hand is asymptotically unbiased, though it does not take the uncertainty arising from the marginal transformation by means of the empirical distribution function into account. Both estimation methods are maximum likelihood methods and no explicit expressions for the estimators are available.

In this paper we reconsider the bias-issue when estimating the coefficient of tail dependence η . In Section 2 we introduce a weighted functional estimator for η and establish its asymptotic normality under a second order condition on the joint tail behavior of the underlying distribution function, some conditions on the weight function and for appropriately chosen sequences of intermediate order statistics. A possible way to eliminate the asymptotic bias of such estimators, consisting in taking appropriate linear combinations, is discussed in Section 3, where we also derive the minimum variance asymptotically unbiased estimator. Unlike the above mentioned maximum likelihood approaches to bias-corrected estimation of η , our estimators have explicit expressions, making them computationally inexpensive. Some simulations are discussed in Section 4, in which we compare the finite sample efficiency of some examples of weighted estimators with the likelihood based methods of Beirlant & Vandewalle (2002) and Beirlant *et al.* (2011). The proofs of the results are postponed to the appendix, which is online available as a supporting information.

2 A functional estimator for η

In this section we introduce a class of weighted functional estimators for the coefficient of tail dependence. Given a sample of independent and identically distributed (i.i.d.) random vectors $(X_1, Y_1), \dots, (X_n, Y_n)$, Ledford & Tawn (1996) proposed to standardize the marginal distributions to unit Fréchet margins, for instance by their empirical distribution function combined with the inverse probability integral transform, after which η can be estimated on the basis of the minima of the components by a classical estimator for the extreme value index, e.g. the Hill estimator (Hill, 1975) or the moment estimator (Dekkers *et al.*, 1989). This idea was also used in Beirlant & Vandewalle (2002), Draisma *et al.* (2004) and Beirlant *et al.* (2011), and will also form the basis for our methodology.

Formally, let $R(X_i)$ denote the rank of X_i among (X_1, \dots, X_n) and $R(Y_i)$ that of Y_i among (Y_1, \dots, Y_n) , define

$$Z_i = \min \left\{ -\frac{1}{\log(R(X_i)/(n+1))}, -\frac{1}{\log(R(Y_i)/(n+1))} \right\}, \quad i = 1, \dots, n, \quad (3)$$

the minimum of unit Fréchet transformed margins, and denote by $Z_{1,n} \leq \dots \leq Z_{n,n}$ the corresponding ascending order statistics. Instead of assuming a transformation to unit Fréchet margins, we can alternatively work with unit Pareto margins, with distribution function given by $F(u) = 1 - 1/u$, $u > 1$. In that case (3) has to be replaced by

$$Z_i = \min \left\{ \frac{1}{1 - R(X_i)/(n+1)}, \frac{1}{1 - R(Y_i)/(n+1)} \right\}, \quad i = 1, \dots, n.$$

Now, consider the functional

$$T_K(z) := \int_0^1 \log \frac{z(t)}{z(1)} d(tK(t)),$$

for any measurable function $z : [0, 1] \rightarrow \mathbb{R}$ (provided the right-hand side is defined and finite; otherwise $T_K(z) = 0$), leading to the following class of estimators for η :

$$\begin{aligned} \hat{\eta}_m(K) &:= T_K(Q_n) \\ &= \int_0^1 \log \frac{Q_n(t)}{Q_n(1)} d(tK(t)) \end{aligned} \quad (4)$$

where $Q_n(t) := Z_{n-\lfloor mt \rfloor, n}$, $0 < t < n/m$, is the empirical quantile function, and K is a weight or kernel function with support on $(0, 1)$.

In order to establish the asymptotic distribution of (4) we have to introduce some conditions on the kernel function K as well as a second order condition on the tail of the joint distribution function F . Concerning the weight function K we assume the following.

Condition \mathcal{K} : *Let $K(\cdot)$ be a kernel such that*

- (i) $K(\cdot)$ is continuously differentiable on $(0, 1)$,
- (ii) $\int_0^1 (-\log t) d(tK(t)) = 1$,
- (iii) there exists $M > 0$, $0 \leq r < 1/2$ and $p < 1$ such that $|K(u)| \leq Mu^{-r}$ and $|K'(u)| \leq Mu^{-p-r}$ on $(0, 1)$.

The conditions on K are not too restrictive and besides they are easy to verify for a given kernel function. For instance, our introduced class (4) includes

- the log weight-type estimator if $K(u) = (-\log u)^\nu / \Gamma(\nu + 1)$, $u \in (0, 1)$, $\nu \geq 1$,
- estimators based on the weight functions proposed in Gomes *et al.* (2007) in the framework of the estimation of the positive extreme value index γ :

$$K(u) = (1 + \nu)u^\nu \quad \text{and} \quad K(u) = (1 + \nu)^2 u^\nu (-\log u),$$

with $u \in (0, 1)$ and $\nu \geq 0$.

Similar conditions can be found in e.g. Gardes & Girard (2008a) and Goegebeur & Guillou (2011) in the framework of the estimation of Weibull-type tails. We refer also to Mason (1981) for general results on asymptotic normality for linear combinations of order statistics.

For what concerns the second order tail behavior of F we use the following slightly modified version of the condition of Draisma *et al.* (2004) (see also de Haan & Stadtmüller, 1996). This condition is not too restrictive and commonly used in estimation problems involving bivariate

extreme values.

Condition \mathcal{SO} : Let (X, Y) be a random vector with joint distribution function F and continuous marginal distribution functions F_X and F_Y such that

$$\lim_{t \downarrow 0} q_1(t)^{-1} \left(\frac{\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{q(t)} - c(x, y) \right) =: c_1(x, y) \quad (5)$$

exists for all $x \geq 0, y \geq 0$ with $x + y > 0$, a positive function q and a function q_1 both tending to zero as $t \downarrow 0$, and c_1 a function neither constant nor a multiple of c . Moreover, we assume that the convergence is uniform on $\{(x, y) \in [0, \infty)^2 | x^2 + y^2 = 1\}$, that c_1 is continuous and $c_1(x, x) = x^{1/\eta}(x^\tau - 1)/\tau$.

Remember that $q(t) := \mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)$. It can be shown that (5) implies that q and $|q_1|$ are regularly varying at zero with index $1/\eta$ and $\tau \geq 0$, respectively. The function c is homogeneous of order $1/\eta$, that is $c(tx, ty) = t^{1/\eta}c(x, y)$. In this paper, we only consider the case where $\eta < 1$ and $\tau > 0$ (similar restrictions were used in Beirlant & Vandewalle, 2002, and Beirlant *et al.*, 2011). We refer to Section 4 for some examples of distributions satisfying \mathcal{SO} .

Our main result is stated in Theorem 1 below. It is essentially based on a modified version of Lemma 6.2 in Draisma *et al.* (2004) by including a bias term, combined with the functional delta method. We refer to the supporting information for further details.

Theorem 1 Assume Condition \mathcal{K} and the second order condition \mathcal{SO} with a function c that is continuously differentiable. If $m \rightarrow \infty$ such that $m/n \rightarrow 0$ and $\sqrt{m}q_1(q^{-1}(m/n)) \rightarrow \lambda$, finite, there exists a standard Brownian motion \overline{W} , such that

$$\sqrt{m}(T_K(Q_n) - \eta) \xrightarrow{d} \eta \int_0^1 t^{-1} \overline{W}(t) d(tK(t)) - \eta \overline{W}(1)K(1) + \lambda \frac{\eta}{\tau} \int_0^1 (t^{\eta\tau} - 1) d(tK(t)).$$

In particular $\sqrt{m}(T_K(Q_n) - \eta)$ is asymptotically normal $\mathcal{N}(\lambda \mathcal{AB}(K), \mathcal{AV}(K))$ where

$$\begin{aligned} \mathcal{AB}(K) &:= \frac{\eta}{\tau} \int_0^1 (t^{\eta\tau} - 1) d(tK(t)), \\ \mathcal{AV}(K) &:= \eta^2 \int_0^1 \int_0^1 \frac{\min\{s, t\}}{st} d(tK(t)) d(sK(s)) - \eta^2 K^2(1). \end{aligned}$$

We can now illustrate this theorem in the special cases of the log weight-type kernels and the two weights proposed in Gomes *et al.* (2007). This leads to the following corollaries.

Corollary 1 (Log weight). *Under the conditions of Theorem 1, if $K(u) = (-\log u)^\nu / \Gamma(\nu + 1)$, $\nu \geq 1$, then*

$$\sqrt{m}(T_K(Q_n) - \eta) \xrightarrow{d} \mathcal{N}\left(-\lambda \frac{\eta^2}{(1 + \eta\tau)^{\nu+1}}, \eta^2 \frac{\Gamma(1 + 2\nu)}{\Gamma^2(1 + \nu)}\right).$$

Corollary 2 *Under the conditions of Theorem 1, if $K(u) = (1 + \nu)u^\nu$, $\nu \geq 0$, then*

$$\sqrt{m}(T_K(Q_n) - \eta) \xrightarrow{d} \mathcal{N}\left(-\lambda\eta^2 \frac{1 + \nu}{1 + \nu + \eta\tau}, \eta^2 \frac{(1 + \nu)^2}{1 + 2\nu}\right).$$

The Hill weight function $K(u) = 1$ is included in this family ($\nu = 0$).

Corollary 3 *Under the conditions of Theorem 1, if $K(u) = (1 + \nu)^2 u^\nu (-\log u)$, $\nu \geq 0$, then*

$$\sqrt{m}(T_K(Q_n) - \eta) \xrightarrow{d} \mathcal{N}\left(-\lambda\eta^2 \left(\frac{1 + \nu}{1 + \nu + \eta\tau}\right)^2, 2\eta^2 \frac{(1 + \nu)^4}{(1 + 2\nu)^3}\right).$$

3 Bias correction and estimation of the second order parameter

3.1 An asymptotically unbiased estimator with minimum variance

In this section we introduce a class of bias-corrected estimators for the parameter η , and discuss variance optimality in the considered class. In particular, we propose to combine two kernel-type estimators for η in order to cancel the asymptotic bias appearing in Theorem 1. The elimination of bias by the construction of an appropriately weighted sum of two estimators is also referred to as the generalized Jackknife methodology; see Gray & Schucany (1972) for further details. More precisely, let $K_1(\cdot)$ and $K_2(\cdot)$ be two different kernels, both satisfying Condition \mathcal{K} and let α be a real constant. Clearly

$$K_\alpha(t) := \alpha K_1(t) + (1 - \alpha)K_2(t) \tag{6}$$

satisfies also Condition \mathcal{K} , and hence Theorem 1 yields the asymptotic bias of this new estimator T_{K_α} , namely

$$\mathcal{AB}(K_\alpha) = \frac{\eta}{\tau} \int_0^1 (t^{\eta\tau} - 1) d(tK_\alpha(t)) = \alpha \mathcal{AB}(K_1) + (1 - \alpha) \mathcal{AB}(K_2).$$

Equating the right-hand side of the latter equation to 0 leads to the value of α that eliminates the asymptotic bias:

$$\alpha^* = \frac{\mathcal{AB}(K_2)}{\mathcal{AB}(K_2) - \mathcal{AB}(K_1)}, \quad (7)$$

provided $\mathcal{AB}(K_1) \neq \mathcal{AB}(K_2)$. This result is formalized in the following proposition.

Proposition 1 *Under the second order condition \mathcal{SO} with a function c that is continuously differentiable, and assuming that K_1 and K_2 satisfy Condition \mathcal{K} with $\mathcal{AB}(K_1) \neq \mathcal{AB}(K_2)$, we have that if $m \rightarrow \infty$ such that $m/n \rightarrow 0$ and $\sqrt{m}q_1(q^{-1}(m/n)) \rightarrow \lambda$, finite, then*

$$\sqrt{m} (T_{K_{\alpha^*}}(Q_n) - \eta) \xrightarrow{d} \mathcal{N}(0, \mathcal{AV}(K_{\alpha^*})).$$

The bias-correcting weight depends on η and τ , i.e. $\alpha^* = \alpha^*(\eta, \tau)$, which are unknown and hence need to be estimated from the data. The following proposition states that replacing η and τ by initial consistent estimators $\tilde{\eta}$ and $\tilde{\tau}$, respectively, which possibly depend on sequences of upper order statistics different from the one used for the bias-corrected estimator, does not change the limiting distribution of the normalized bias-corrected estimator. Denote $\hat{\alpha}^* := \alpha^*(\tilde{\eta}, \tilde{\tau})$.

Proposition 2 *Assume (i) the second order condition \mathcal{SO} with a function c that is continuously differentiable, (ii) kernel functions K_1 and K_2 that satisfy Condition \mathcal{K} with $\mathcal{AB}(K_1) \neq \mathcal{AB}(K_2)$ and that are such that α^* is continuously differentiable with respect to η and τ , and (iii) initial consistent estimators $\tilde{\eta}$ and $\tilde{\tau}$ for η and τ , respectively. Then, if $m \rightarrow \infty$ such that $m/n \rightarrow 0$ and $\sqrt{m}q_1(q^{-1}(m/n)) \rightarrow \lambda$, finite, we have that*

$$\sqrt{m} (T_{K_{\hat{\alpha}^*}}(Q_n) - \eta) \xrightarrow{d} \mathcal{N}(0, \mathcal{AV}(K_{\alpha^*})).$$

We refer to the next subsection for an example of a consistent estimator for the second order parameter τ .

Now, inspired by Drees (1998a) and Proposition 3 in Gardes & Girard (2008b), we construct an asymptotically unbiased functional estimator with minimum variance.

Theorem 2 *Let $\alpha_{opt} := (1 + \eta\tau)^2/(\eta^2\tau^2)$, $K_1(t) = 1$ and $K_2(t) = (1 + \eta\tau)t^{\eta\tau}$. Then $K_{\alpha_{opt}}(\cdot)$ defined as in (6) is the asymptotically unbiased weight function with minimum variance among unbiased weight functions satisfying Condition \mathcal{K} .*

We refer to the appendix for a proof of Theorem 2. Also, direct computations yield that the asymptotic variance of this optimal asymptotically unbiased estimator is given by $\eta^2(1 + \eta\tau)^2/(\eta\tau)^2$.

Corollary 4 *Under the assumptions of Theorem 1 and Theorem 2,*

$$\sqrt{m} \left(T_{K_{\alpha_{opt}}}(Q_n) - \eta \right) \xrightarrow{d} \mathcal{N} \left(0, \eta^2 \frac{(1 + \eta\tau)^2}{(\eta\tau)^2} \right).$$

From Theorem 1 we also obtain immediately the limiting distribution of the estimator for η that is obtained with $K_{\alpha_{opt}}$ in case one mis-specifies the parameter $\eta\tau$ at some value. Let K_a denote the kernel function $K_{\alpha_{opt}}$ with $\eta\tau$ fixed at the value a , i.e. $K_a(t) := (1 + a)^2/a^2 - (1 + a)(1 + 2a)t^a/a^2$, $a > 0$.

Corollary 5 *Under the second order condition \mathcal{SO} with a function c that is continuously differentiable, we have that if $m \rightarrow \infty$ such that $m/n \rightarrow 0$ and $\sqrt{m}q_1(q^{-1}(m/n)) \rightarrow \lambda$, finite, then*

$$\sqrt{m} (T_{K_a}(Q_n) - \eta) \xrightarrow{d} \mathcal{N} \left(\lambda d(a, \eta\tau), \eta^2 \frac{(1 + a)^2}{a^2} \right),$$

where

$$d(a, \eta\tau) := -\eta^2 \frac{(1 + a)(a - \eta\tau)}{a(1 + \eta\tau)(1 + a + \eta\tau)}.$$

3.2 Estimation of the second order parameter τ

In this section we introduce a simple estimator for the second order parameter τ , the coefficient of regular variation of the function $|q_1|$. In the univariate Pareto-type framework, the estimation of second order parameters has been well studied and several good working estimators are available, we refer to Goegebeur *et al.* (2010) and the references therein for a recent account on the topic. In the multivariate setting little work has been done in this respect. Peng (2010) introduced an estimator for a rate parameter appearing in a second order condition that is slightly different from condition \mathcal{SO} , but apart from this we are not aware of other attempts.

As the basic building block for our estimator we use the statistic

$$S(x, y) := \sum_{i=1}^n \mathbf{1}\{X_i \geq x, Y_i \geq y\}.$$

Inspired by Goegebeur *et al.* (2010) and Peng (2010) we propose the following estimator for τ

$$\hat{\tau}_k(x) := -\frac{1}{\log 2} \log \left| \frac{H(x) - H(2x)}{H(2x) - H(4x)} \right|,$$

where

$$H(x) := \frac{S(X_{n-\lfloor kx \rfloor + 1, n}, Y_{n-\lfloor kx \rfloor + 1, n})}{S(X_{n-\lfloor k2x \rfloor + 1, n}, Y_{n-\lfloor k2x \rfloor + 1, n})}.$$

The consistency of this estimator is established in the following proposition.

Proposition 3 *Assume that the second order condition \mathcal{SO} is satisfied with a function c that has continuous first order partial derivatives. If $k, n \rightarrow \infty$, such that $k/n \rightarrow 0$ and $\sqrt{nq(k/n)}q_1(k/n) \rightarrow \infty$, then $\hat{\tau}_k(x) \xrightarrow{\mathbb{P}} \tau$.*

4 Simulation results

The aim of this section is to illustrate the finite sample performance of our bias-reduced estimator compared to alternative methodologies. To this aim, we first study some bivariate models. Here we consider three distributions which satisfy our framework and one, the bivariate normal case, which does not. We also shortly describe some alternative estimation methods from the recent literature, and compare these with the methodology presented in this paper in a small simulation experiment.

4.1 Some bivariate distributions

First note that, because of the assumed continuous marginal distribution functions in condition \mathcal{SO} we have that $F_X(X)$ and $F_Y(Y)$ are uniform (0,1) random variables, and hence

$$\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty) = tx + ty - 1 + C(1 - tx, 1 - ty),$$

where C is the copula function of the joint distribution function F . Verification of condition \mathcal{SO} for a specific distribution can therefore be based completely on the copula function C .

- **The Farlie Gumbel Morgenstern distribution**

The Farlie Gumbel Morgenstern copula function is given by

$$C(u, v) = uv [1 + \xi(1 - u)(1 - v)], \quad (u, v) \in [0, 1]^2,$$

with $\xi \in [-1, 1]$. Straightforward calculations lead to

$$\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty) = t^2 xy [1 + \xi - \xi t(x + y) + \xi t^2 xy].$$

In the case where $\xi \in (-1, 1]$,

$$\frac{\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{\mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[1 - \frac{\xi t}{1 + \xi} (x + y - 2) + O(t^2) \right],$$

from which one easily verifies that condition \mathcal{SO} is satisfied with $\eta = 0.5$, $c(x, y) = xy$, $c_1(x, y) = xy(x + y - 2)/2$, $q_1(t) = -2\xi t/(1 + \xi)$, so $\tau = 1$, while for the case $\xi = -1$

$$\frac{\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{\mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[\frac{x + y}{2} + \frac{t}{4} (x + y - 2xy) + O(t^2) \right],$$

and hence condition \mathcal{SO} is satisfied with $\eta = 1/3$, $c(x, y) = xy(x + y)/2$, $c_1(x, y) = xy(2xy - x - y)/2$, $q_1(t) = -t/2$, so $\tau = 1$.

For the simulation we consider $\xi = -1$ and -0.25 .

- **The Frank distribution**

The copula function for the bivariate Frank distribution is given by

$$C(u, v) = -\frac{1}{\xi} \log \left[1 - \frac{(1 - e^{-\xi u})(1 - e^{-\xi v})}{1 - e^{-\xi}} \right], \quad (u, v) \in [0, 1]^2,$$

where $\xi > 0$. Tedious computations based on expansions of the above copula function lead to the following approximation

$$\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty) = \frac{\xi}{1 - e^{-\xi}} t^2 xy \left[1 - \frac{\xi t}{2} (x + y) + O(t^2) \right],$$

from which we deduce that

$$\frac{\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{\mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[1 - \frac{\xi t}{2} (x + y - 2) + O(t^2) \right].$$

Hence condition \mathcal{SO} is satisfied with $\eta = 0.5$, $c(x, y) = xy$, $c_1(x, y) = xy(x + y - 2)/2$, $q_1(t) = -\xi t$ and $\tau = 1$.

In the simulation we use $\xi = 2$ and 5 .

• **The Ali-Mikhail-Haq distribution**

For this distribution the copula function is given by

$$C(u, v) = \frac{uv}{1 - \xi(1 - u)(1 - v)}, \quad (u, v) \in [0, 1]^2,$$

with $\xi \in [-1, 1]$. From the copula we easily establish that

$$\begin{aligned} \mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty) = \\ t^2 xy [1 + \xi - \xi t(x + y) + \xi(1 + \xi)t^2 xy - \xi^2 t^3 xy(x + y) + O(t^4)]. \end{aligned}$$

First we consider the case where $\xi \in (-1, 1]$. Using Taylor's theorem we obtain

$$\frac{\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{\mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[1 - \frac{\xi t}{1 + \xi} (x + y - 2) + O(t^2) \right],$$

and hence condition \mathcal{SO} is satisfied with $\eta = 0.5$, $c(x, y) = xy$, $c_1(x, y) = xy(x + y - 2)/2$, $q_1(t) = -2\xi t/(1 + \xi)$, so $\tau = 1$. In case $\xi = -1$ we obtain

$$\frac{\mathbb{P}(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{\mathbb{P}(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[\frac{x + y}{2} - \frac{t^2}{2} (x + y)(xy - 1) + O(t^3) \right],$$

leading to $\eta = 1/3$, $c(x, y) = xy(x + y)/2$, $c_1(x, y) = xy(x + y)(xy - 1)/4$, $q_1(t) = -2t^2$, and $\tau = 2$.

We will consider $\xi = -1$.

• **The bivariate normal distribution**

The bivariate normal distribution with mean 0, variance 1 and correlation coefficient $\rho \in (-1, 1)$ satisfies condition \mathcal{SO} with $\eta = (1 + \rho)/2$. We refer to Draisma *et al.* (2004) and Ledford & Tawn (1997) for further details. Unfortunately, it falls outside the scope of the present paper since $\tau = 0$. However, we add this distribution in our simulation study in order to examine the robustness of our approach. We consider the cases $\rho = -0.5$ and $\rho = 0.5$, for which $\eta = 0.25$ and $\eta = 0.75$, respectively.

4.2 Estimators

In the simulation experiment we compare the optimal (in the sense of minimal asymptotic variance) bias-corrected estimator with the Hill estimator and two estimators from the recent extreme value literature, described below. Note that the optimal weight in the bias-corrected estimator involves the parameter $\eta\tau$. In the simulation experiment, for all distributions, except the bivariate normal, we will explore two strategies namely fixing $\eta\tau$ at its true value and a mis-specification of this parameter at $\eta\tau = 1$. For the bivariate normal distribution, which has $\tau = 0$, we consider fixing $\eta\tau$ at 0.5 and 1. Fixing second (or higher) order distributional parameters at a canonical choice, e.g. taking $\eta\tau = 1$, is a common approach in extreme value statistics. For instance, in the univariate extreme value framework, the second order parameter ρ is often fixed at -1 in practical implementations of bias-corrected estimators for the extreme value index; see e.g. Feuerverger & Hall (1999) and Gomes & Martins (2004). Alternatively, the second order parameter τ could also be estimated with the estimator proposed in Section 3.2. Though this estimator is proven to be consistent, its practical use needs further investigation, and this will be pursued in future research on the estimation of the second order parameter in the multivariate extreme value framework.

Beirlant & Vandewalle (2002) proposed to estimate η by applying the maximum likelihood method to the following approximate exponential regression model for scaled log-ratios

$$j \log \left(\frac{Z_{n-j+1,n} - Z_{n-m,n}}{Z_{n-j,n} - Z_{n-m,n}} \right) \approx \frac{\eta}{1 - (j/m)^\eta} f_j, \quad j = 1, \dots, m-1,$$

where f_1, \dots, f_{m-1} are i.i.d. standard exponential random variables. The resulting estimator will here be denoted as $\hat{\eta}_{BV}$. Beirlant & Vandewalle (2002) proved that $\hat{\eta}_{BV}$, when appropriately

normalized, is asymptotically normally distributed, though it should be mentioned that it is not asymptotically unbiased. Nevertheless, in the simulations reported in their paper, the estimator performs quite well with respect to bias.

Beirlant *et al.* (2011) introduced the extended Pareto distribution as approximate model for relative excesses over a threshold

$$\frac{\mathbb{P}(Y_w > tz)}{\mathbb{P}(Y_w > t)} \approx [z(1 + \delta_w(t) - \delta_w(t)z^{-\tau})]^{-1/\eta},$$

where $Y_w := \min(Z_1, Z_2 w / (1 - w))$, Z_1 and Z_2 unit Fréchet random variables, w a tuning parameter, and obtained an estimator for η , denoted here as $\hat{\eta}_B$, from linearized score functions. We refer to Beirlant *et al.* (2011) for more details about w and $\delta_w(t)$. The estimator was proven to be asymptotically unbiased with minimal asymptotic variance $\eta^2(1 + \eta\tau)^2 / (\eta\tau)^2$, though the uncertainty arising from the marginal transformations by means of the empirical distribution functions was not explicitly taken into account. The parameter w was introduced to estimate probabilities in joint tail regions, but has little practical relevance for the estimation of η , and therefore we fix it at 0.5. Concerning the second order parameter τ , Beirlant *et al.* (2011) showed that replacing it by a consistent estimator does not change the limiting distribution of $\hat{\eta}_B$, though no estimator for that parameter was proposed. As suggested by these authors, in the simulation experiment we implement their estimator with τ fixed at the canonical choice 1.

In Figure 1 we show the asymptotic standard deviations of the estimators under study as a function of η . Similar to the univariate case, the Hill estimator, which will be denoted by $\hat{\eta}_H$ has the smallest possible variance, but it is not unbiased. The estimators $T_{K_{\alpha_{opt}}}$ and $\hat{\eta}_B$ have the same asymptotic variance, which is here shown for three values of τ namely $\tau = 0.5, 1, 2$. From Corollary 4, the increase in variance relative to the Hill estimator is immediately clear, and is given by $(1 + \eta\tau)^2 / (\eta\tau)^2$. The Beirlant & Vandewalle (2002) estimator $\hat{\eta}_{BV}$ has approximately the same variance as $T_{K_{\alpha_{opt}}}$ and $\hat{\eta}_B$ in case $\tau = 1$. For values of τ below 1, $\hat{\eta}_{BV}$ will typically have a smaller asymptotic variance than $T_{K_{\alpha_{opt}}}$ and $\hat{\eta}_B$, though it is not asymptotically unbiased.

Insert Figure 1 here

4.3 Results

The simulation experiment considers the distributions listed in Section 4.1, with both unit Pareto and unit Fréchet margins. For each of the distributions we generated 1000 samples of size $n = 500$, and computed all the above mentioned estimators for $m = 5, \dots, 499$, where m denotes the number of upper order statistics of the Z_i observations used in the estimation of η . In Figures 2 till 6 we show the sample mean (left) and the empirical mean squared error (MSE) (right) as a function of m for the estimators $T_{K_{\alpha_{opt}}}$ with the true value of $\eta\tau$ (black solid line), $T_{K_{\alpha_{opt}}}$ with $\eta\tau = 1$ (black dashed line), $\hat{\eta}_H$ (black dotted line), $\hat{\eta}_B$ (grey solid line) and $\hat{\eta}_{BV}$ (grey dashed line). For the bivariate normal copula function the results are shown in Figures 7 and 8. Here $T_{K_{\alpha_{opt}}}$ is computed with $\eta\tau = 0.5$ and $\eta\tau = 1$. From the simulation results we can draw the following conclusions:

(i) The Hill estimator is generally biased, though the bias seems to be a more severe problem for unit Fréchet marginal distributions than for unit Pareto marginal distributions. This observation is in agreement with the theoretical considerations in Drees (1998a, b). For the distributions where the bias is a problem, like e.g. the Farlie Gumbel Morgenstern with unit Fréchet margins and the distributions based on Frank's copula, the estimators $T_{K_{\alpha_{opt}}}$, $\hat{\eta}_B$ and $\hat{\eta}_{BV}$ typically outperform $\hat{\eta}_H$ in terms of minimal MSE. Also, the sample paths of these estimators show a longer stable part compared to $\hat{\eta}_H$.

(ii) The estimators $\hat{\eta}_B$ and $\hat{\eta}_{BV}$ exhibit a very similar behavior, both in terms of bias and MSE, but $\hat{\eta}_B$ behaves somewhat less erratic than $\hat{\eta}_{BV}$, as can be seen from its more stable sample paths.

(iii) As expected, no estimator performs uniformly best, but $T_{K_{\alpha_{opt}}}$ is highly competitive compared to $\hat{\eta}_B$ and $\hat{\eta}_{BV}$, especially if one takes into account that, unlike the maximum likelihood based estimators, it can be computed directly from the data without needing an iterative optimization scheme. Also, for $T_{K_{\alpha_{opt}}}$, mis-specifying $\eta\tau$ at the canonical choice 1, often turns out to work better in practice than using the true value of that parameter. This can be explained from Corollary 5, at least for the cases with $\tau > 0$. Indeed, from Corollary 5 the asymptotic

variance is decreasing in the parameter a , and hence the estimator with $\eta\tau$ fixed at the value 1 will have a smaller variance compared to the estimator obtained with the true value of $\eta\tau$ (for the distributions under consideration in the simulation $\eta\tau < 1$). This difference in variance can be considerable, e.g. in case $\eta = 1/3$ and $\tau = 1$ we have an asymptotic variance of 1.78 in case one uses the true value of $\eta\tau$, compared to an asymptotic variance of only 0.44 when one fixes $\eta\tau$ at 1. Using the true value of $\eta\tau$ of course eliminates the bias, at least theoretically. However, in the simulations fixing $\eta\tau$ at 1 performs often as well as using the true value of $\eta\tau$ in terms of bias, and in some cases even better, which then lead to the better performance in terms of minimum MSE.

Insert Figures 2-8 here

5 Conclusion

In this paper we considered the estimation of the coefficient of tail dependence in bivariate extreme value statistics, and introduced a class of weighted functional estimators for this parameter. By using the functional delta method the estimators were shown to be asymptotically normal under a second order condition on the joint tail behavior, some conditions on the kernel function and for an intermediate sequence of upper order statistics. By taking appropriately chosen weighted sums of two (biased) estimators we obtained a class of asymptotically unbiased estimators, and we established variance optimality within this class. The simulations indicated that the estimator is highly competitive with recent alternatives from the extreme value literature, especially if one takes their low computational demands into account. In future work we will focus on the further development of estimators for the second order parameter τ , as well as their practical implementation.

Acknowledgement

The authors are very grateful to the associate editor and the two referees for their careful reading of the paper and their comments that lead to significant improvements of the initial

draft. Yuri Goegebeur's research was supported by a grant from the Danish Natural Science Research Council.

Supporting information

Additional information for this article is available online including: a supplementary appendix containing detailed proofs of Theorem 1, Proposition 2, Theorem 2, Proposition 3.

References

- [1] Beirlant, J., Dierckx, G., Goegebeur, Y. & Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177–200.
- [2] Beirlant, J., Dierckx, G. & Guillou, A. (2011). Bias-reduced estimators for bivariate tail modelling. *Insurance Math. Econom.* **49**, 18–26.
- [3] Beirlant, J. & Vandewalle, B. (2002). Some comments on the estimation of a dependence index in bivariate extreme value in statistics. *Statist. Probab. Lett.* **60**, 265–278.
- [4] Bingham, N.H., Goldie, C.M. & Teugels, J.L. (1987). *Regular variation*. Cambridge University Press, Cambridge.
- [5] Dekkers, A.L.M., Einmahl, J.H.J. & de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17**, 1833–1855.
- [6] Draisma, G., Drees, H., Ferreira, A. & de Haan, L. (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli* **10**, 251–280.
- [7] Drees, H. (1998a). A general class of estimators of the extreme value index. *J. Statist. Plann. Inference* **66**, 95–112.
- [8] Drees, H. (1998b). On smooth statistical tail functionals. *Scand. J. Stat.* **25**, 187–210.
- [9] Feuerverger, A. & Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760–781.

- [10] Gardes, L. & Girard, S. (2008a). Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *J. Statist. Plann. Inference* **138**, 1416–1427.
- [11] Gardes, L. & Girard, S. (2008b). A moving window approach for non parametric estimation of the conditional tail index. *J. Multivariate Anal.* **99**, 2368–2388.
- [12] Goegebeur, Y., Beirlant, J. & de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Plann. Inference* **140**, 2632–2652.
- [13] Goegebeur, Y. & Guillou, A. (2011). A weighted mean excess function approach to the estimation of Weibull-type tails. *TEST* **20**, 138–162.
- [14] Gomes, M.I., de Haan, L. & Rodrigues, L.H. (2008). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **70**, 31–52.
- [15] Gomes, M.I. & Martins, M.J. (2004). Bias reduction and explicit semi-parametric estimation of the tail index. *J. Statist. Plann. Inference* **124**, 361–378.
- [16] Gomes, M.I., Miranda, C. & Viseu, C. (2007). Reduced bias tail index estimation and the Jackknife methodology. *Stat. Neerl.* **61**, 243–270.
- [17] Gray, H.L. & Schucany, W.R. (1972). *The generalized Jackknife statistic*. Marcel Dekker.
- [18] de Haan, L. & Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Aust. Math. Soc.* **61**, 381–395.
- [19] Heffernan, J. E. (2000). A directory of coefficients of tail dependence. *Extremes* **3**, 279–290.
- [20] Heffernan, J. E. & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66**, 497–546.
- [21] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163–1174.
- [22] Ledford, A.W. & Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.

- [23] Ledford, A.W. & Tawn, J.A. (1997). Modelling dependence within joint tail regions. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **59**, 475–499.
- [24] Mason, D.M. (1981). Asymptotic normality of linear combinations of order statistics with a smooth score function. *Ann. Statist.* **9**, 899–908.
- [25] Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.* **43**, 399–409.
- [26] Peng, L. (2010). A practical way for estimating tail dependence functions. *Statist. Sinica* **20**, 365–378.
- [27] Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15**, 1174–1207.

Corresponding author: Yuri Goegebeur, Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark.

E-mail: yuri.goegebeur@stat.sdu.dk.

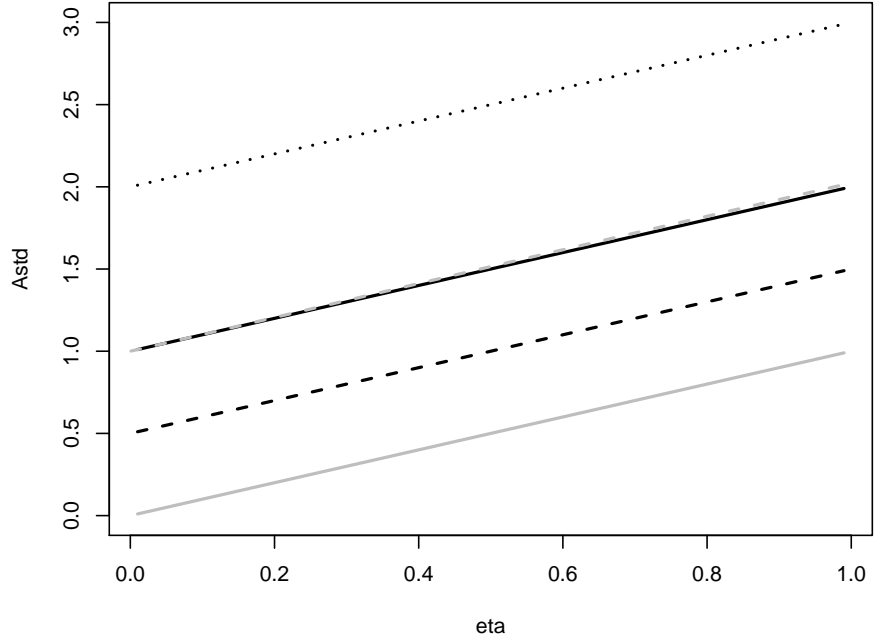


Figure 1: Asymptotic standard deviations as a function of η : $T_{K_{\alpha_{opt}}}$ and $\hat{\eta}_B$ when $\tau = 1$ (black solid line), $\tau = 2$ (black dashed line) and $\tau = 0.5$ (black dotted line), $\hat{\eta}_H$ (grey solid line), and $\hat{\eta}_{BV}$ (grey dashed line) estimators.

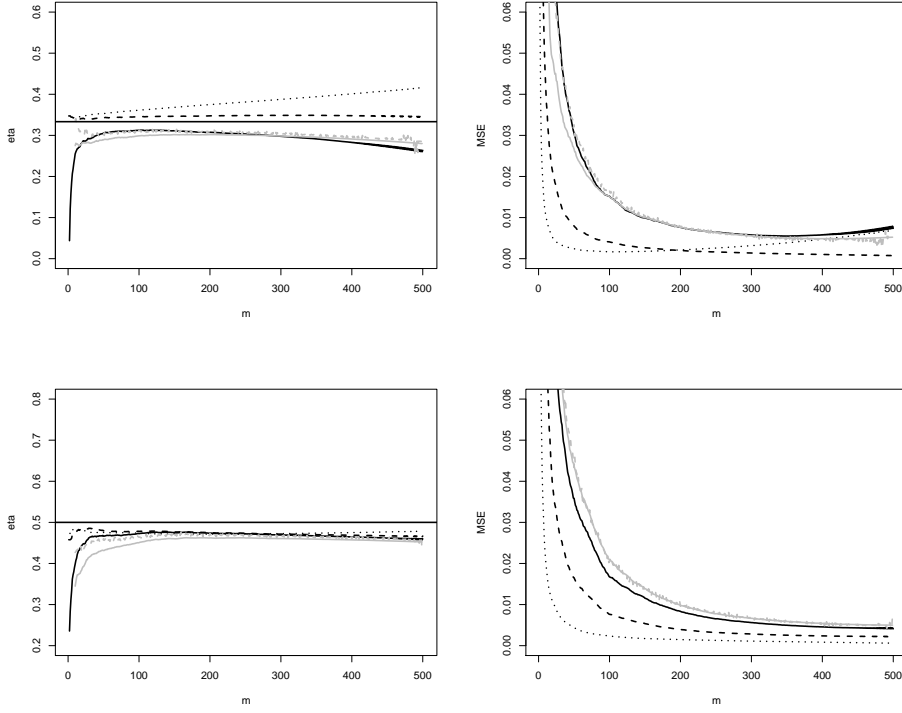


Figure 2: Bivariate Farlie Gumbel Morgenstern copula with $\xi = -1$ (row 1) and $\xi = -0.25$ (row 2), sample mean (left) and MSE (right), transformation to unit Pareto margins: Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and the true value of $\eta\tau$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .

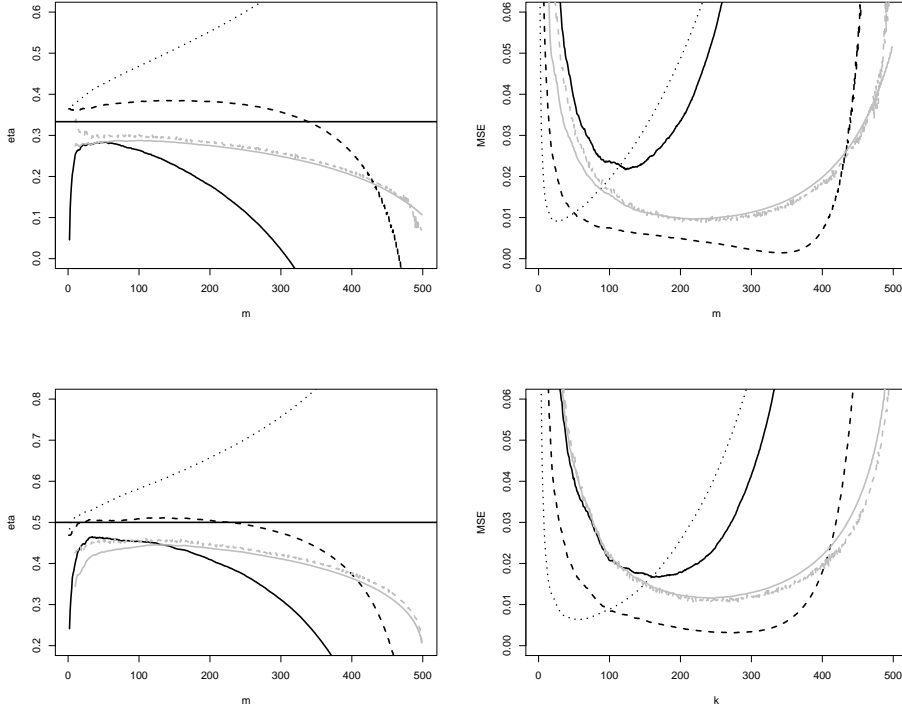


Figure 3: Bivariate Farlie Gumbel Morgenstern copula with $\xi = -1$ (row 1) and $\xi = -0.25$ (row 2), sample mean (left) and MSE (right), transformation to unit Fréchet margins: Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and the true value of $\eta\tau$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey, solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .

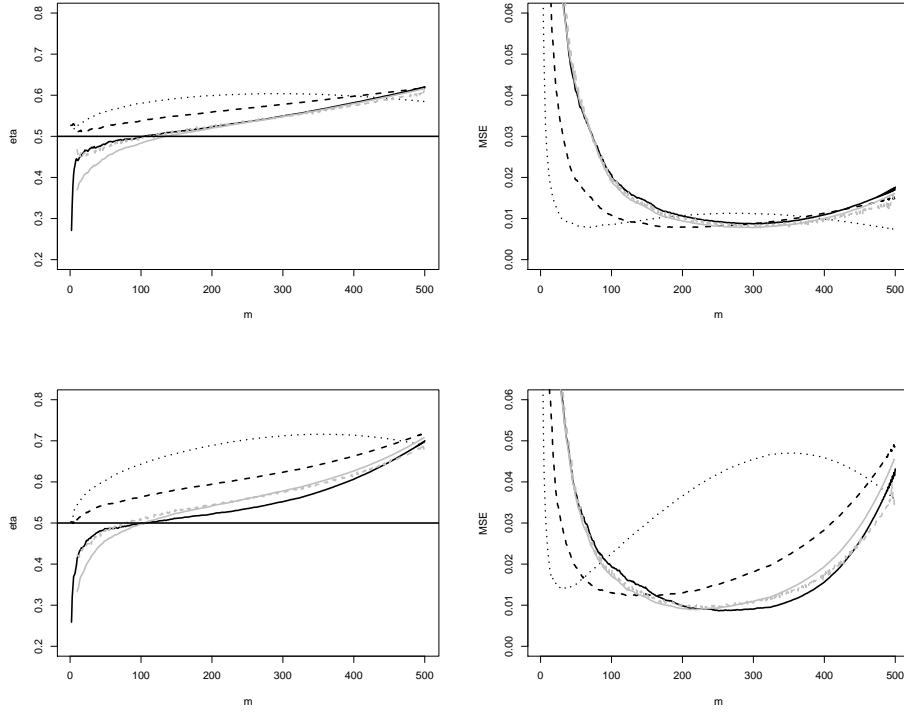


Figure 4: Bivariate Frank copula with $\xi = 2$ (row 1) and $\xi = 5$ (row 2), sample mean (left) and MSE (right), transformation to unit Pareto margins: Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 0.5$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey, solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .

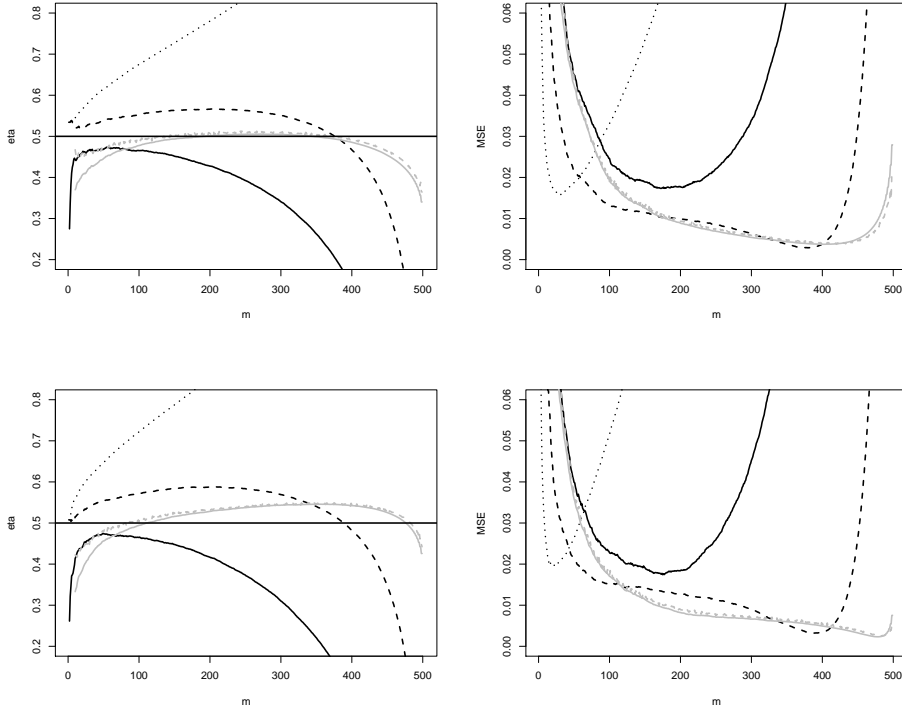


Figure 5: Bivariate Frank copula with $\xi = 2$ (row 1) and $\xi = 5$ (row 2), sample mean (left) and MSE (right), transformation to unit Fréchet margins: Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 0.5$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey, solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .

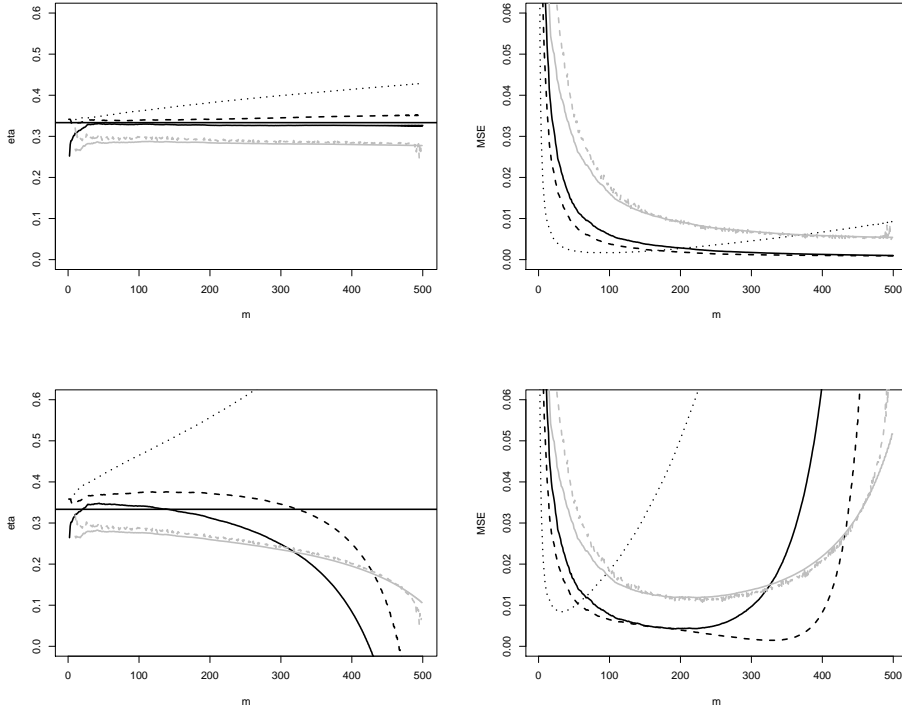


Figure 6: Bivariate Ali-Mikhail-Haq copula with $\xi = -1$, unit Pareto margins (row 1) and unit Fréchet margins (row 2), sample mean (left) and MSE (right): Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 2/3$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey, solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .

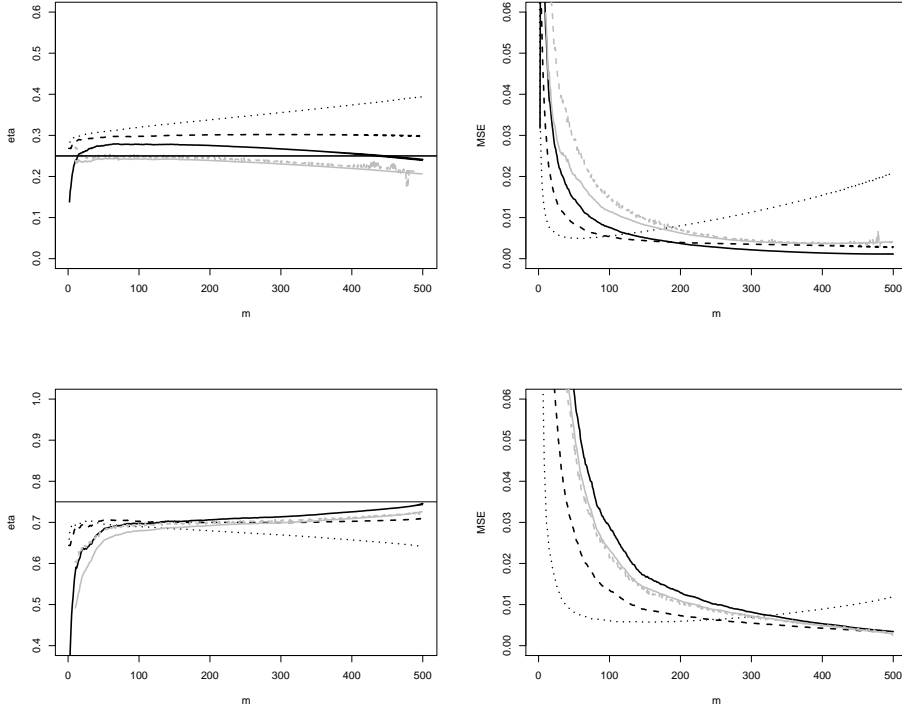


Figure 7: Bivariate normal copula with $\mu_1 = \mu_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$, $\rho = -0.5$ (row 1) and $\rho = 0.5$ (row 2), sample mean (left) and MSE (right), transformation to unit Pareto margins: Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 0.5$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey, solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .

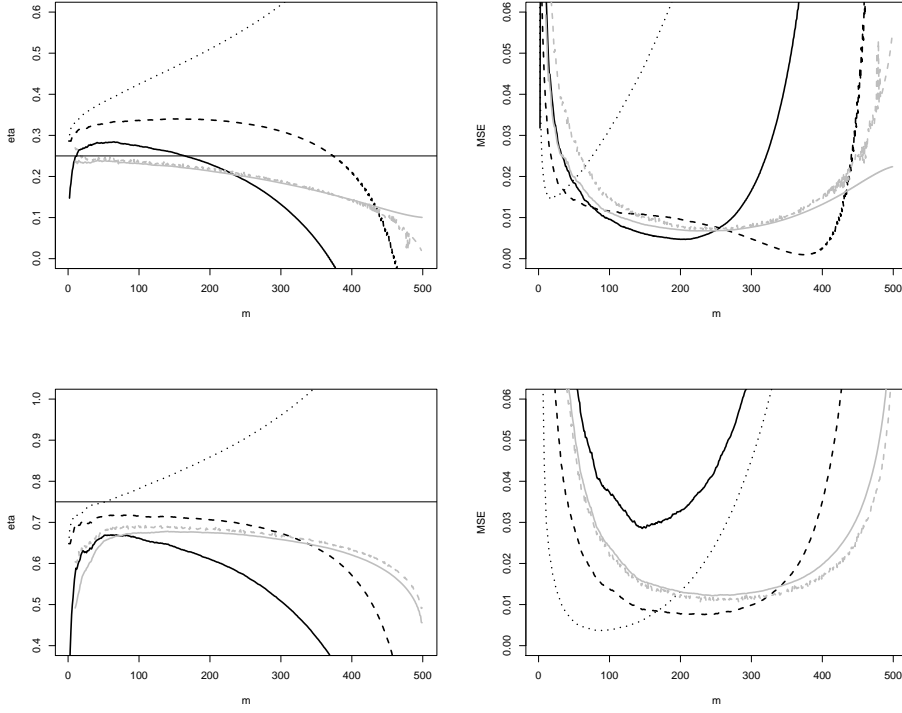


Figure 8: Bivariate normal copula with $\mu_1 = \mu_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$, $\rho = -0.5$ (row 1) and $\rho = 0.5$ (row 2), sample mean (left) and MSE (right), transformation to unit Fréchet margins: Hill (black, dotted), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 0.5$ (black, solid), $T_K(Q_n)$ with optimal bias-correcting weight and $\eta\tau = 1$ (black, dashed), $\hat{\eta}_B$ (grey, solid) and $\hat{\eta}_{BV}$ (grey, dashed) as a function of m . The horizontal line is the true value of η .