MONASH University

MONASH BUSINESS SCHOOL

ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

Consistency of full-sample bootstrap for estimating high-quantile, tail probability, and tail index

Svetlana Litvinova and Mervyn J. Silvapulle

April 2020

Working Paper 15/20

AACSB ACCREDITED EFMD EQUIS ACCREDITED ASSOCIATION OF MBAS AMBA ACCREDITED

Consistency of full-sample bootstrap for estimating high-quantile, tail probability, and tail index

Svetlana Litvinova and Mervyn J. Silvapulle

Department of Econometrics and Business Statistics, Monash University, Australia

svetlana.litvinova@monash.edu, mervyn.silvapulle@monash.edu

Abstract

We show that the full-sample bootstrap is asymptotically valid for constructing confidence intervals for high-quantiles, tail probabilities, and other tail parameters of a univariate distribution. This resolves the doubts that have been raised about the validity of such bootstrap methods. In our extensive simulation study, the overall performance of the bootstrap method was better than that of the standard asymptotic method, indicating that the bootstrap method is at least as good, if not better than, the asymptotic method for inference. This paper also lays the foundation for developing bootstrap methods for inference about tail events in multivariate statistics; this is particularly important because some of the non-bootstrap methods are complex.

JEL Classifications: C13, C15, C18.

Keywords: Full-sample bootstrap; Intermediate order statistic; Extreme value index; Hill estimator; Tail probability; Tail quantile.

Corresponding author: Svetlana Litvinova, Department of Econometrics and Business Statistics, Monash University, Australia. email: svetlana.litvinova@monash.edu.

1 Introduction

In risk management, the quantities of interest are often related to rare and costly extreme events. Some examples of statistical inference questions that arise in such context are: (1) What is the probability that an insurance claim would exceed a given large threshold? (2) What is the level of financial loss that an investment portfolio exceeds with only a very small probability, say, 0.001? The main statistical method used for answering such questions is confidence intervals for tail parameters. For example, it may be desired to construct a confidence interval for the tail probability $p(x) = 1 - F(x)$, where F is the unknown population distribution function and x is a given large value; here large x means that it may be larger than the largest order statistic in the sample and hence the sample proportion of observations that exceeds x, which is zero, is not a useful estimator of $p(x)$. Doubts have been raised about the validity of full-sample bootstrap for constructing confidence intervals for such tail parameters. The objective of this paper is to show that the full-sample bootstrap is valid.

There is a large literature on estimation of confidence intervals for high-quantiles and tail probabilities. For excellent accounts of the topic, see Chapter 4 in Beirlant *et al.* (2004), section 4.3 in de Haan and Ferreira (2006), and sections 6.3 and 6.4 in Embrechts *et al.* (1997). Coles (2001) provides an excellent introductory account to the topic. A topic that frequently arises in this area is construction of confidence intervals for tail parameters (for eg., Tajvidi 2003). Although m out of n and subsample bootstrap are in general asymptotically valid in these cases they do not appear to perform well in finite samples (for eg., Cornea-Madeira and Davidson 2015). Geluk and de Haan (2002) examined the validity of the bootstrap method for intermediate order statistics; see also Gomes *et al.* (2016). The validity of bootstrap for inference in heavy-tailed distributions is a delicate issue (Davidson 2012). The question of whether or not full-sample nonparametric bootstrap in the iid setting is consistent for inference on high quantiles and tail probabilities in univariate distributions, has been an open question for sometime. This paper provides a rigorous answer to this question.

The main contribution of this paper is that it is shown that the full-sample bootstrap methods for constructing confidence intervals for the tail parameters, high quantiles, tail probabilities, and the extreme value tail index are valid. In our simulation study, the overall performance of the proposed bootstrap method was better than that of the asymptotic method, but none of them performed uniformly the best. This corroborates the validity of the main result of this paper, namely that the bootstrap is valid. The results in this paper also provide the foundation for developing full-sample bootstrap methods for tails of multivariate distributions and for time series. This is particularly important since methods based on the asymptotic distribution of estimators of tail parameters in the multivariate setting are typically rather complex.

The method of estimation is based on tail empirical process (Einmahl 1990). Our technique for establishing bootstrap validity draws from the probability theory literature on bootstrapping tail empirical processes (Csörgő and Mason 1989). Therefore, in this short paper we state the main results as five theorems, and relegate the mathematical details to the Supplementary Materials. The validity of bootstrap for inference on some parameters of a heavy-tailed distribution has been established using different machinery, namely wild bootstrap and permutation bootstrap (see, Cavaliere *et al.* 2016, 2013); it would be interest to know whether or not their machinery could be applied to the setting of this paper. Lahiri (2003) provides a chapter on bootstrapping heavy-tailed data and extremes for time-series and stationary processes.

The paper is structured as follows. Section 2 introduces some notation, assumptions, and preliminaries. Section 3 provides a statement of the main results. The proofs of these main results are provided in the Supplementary Materials. Numerical studies are provided in the next two Sections.

2 Notation, assumptions, and preliminaries

A discussion of the results in this section may be found in Beirlant *et al.* (2004), de Haan and Ferreira (2006), Embrechts *et al.* (1997), and Coles (2001), among others. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed with common distribution function F, and let $X_{1,,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ denote the corresponding order statistics. If there exists a sequence of constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that $P((X_{n,n} - b_n)/a_n \leq x) \longrightarrow G(x)$ as $n \to \infty$, and G is nondegenerate, then (a) G is called an *extreme value distribution* function, (b) F is said to be in the *maximum domain of attraction* of G, denoted by $F \in D(G)$, and (c) the class of extreme value distributions is $G_\gamma(ax+b)$, where $G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}$ $(x > -\gamma^{-1}, b \in \mathbb{R}, a > 0)$. The parameter γ, called the *extreme value index* or the *tail index*, characterizes the tail behaviour of F; in what follows, the term 'tail' refers to 'right tail'. If $\gamma < 0$, $\gamma = 0$, or $\gamma > 0$ then F has a finite end point, *light tail*, or *heavy tail*, respectively. *In what follows, we assume that* $F \in D(G_\gamma)$ *for some* $\gamma \in \mathbb{R}$.

Let $U(y) = F^{\leftarrow}(1 - y^{-1})$, where $F^{\leftarrow}$ denotes the inverse function of F. There is a one-to-one correspondence between F and U. Often, regularity conditions on the tails of F are expressed in terms of U. The distribution function $F \in D(G_\gamma)$ if and only if there exists a positive function $a(t)$ ($t \in \mathbb{R}$) such that (for example, Theorem 1.1.6, de Haan and Ferreira 2006).

$$\lim_{t\to\infty} \{U(tx) - U(t)\}/a(t) = \{x^\gamma - 1\}/\gamma, \quad (x > 0). \tag{2.1}$$

If (2.1) holds, then U is said to satisfy the *first-order condition* of regular variation. In what follows, we assume that this is satisfied.

The asymptotic methods of statistical inference on the tail parameters are based on the asymptotic normality of estimators of such parameters. In order to establish the asymptotic normality of such estimators, a second-order refinement of (2.1) is also assumed. The function U is said to satisfy the *second-order condition of regular variation* if there exists a positive or negative function $A(\cdot)$ with $\lim_{t\to\infty} A(t) = 0$, such that

$$\lim_{t\to\infty} [\{U(tx) - U(t)\}/a(t) - \{x^\gamma - 1\}/\gamma]/A(t) = H_{\gamma,\rho}(x) \quad (x > 0), \tag{2.2}$$

where $H_{\gamma,\rho}(x) = \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} duds$, $\rho \le 0$ is called the second order parameter of regular variation, and $A(t)$ is regularly varying with index ρ. Let $\gamma_- = \min\{0, \gamma\}$ and $\gamma_+ = \max\{0, \gamma\}$. *In what follows we assume that (2.2) is satisfied,* $\gamma \neq \rho$, *and that* $\rho < 0$ *if* $\gamma > 0$. Under the additional assumption (2.2), we have [page 103 and Lemma B.3.16 in de Haan and Ferreira 2006]

$$\lim_{t\to\infty} \left[\{\log U(tx) - \log U(t)\}/q(t) - (x^{\gamma_-} - 1)/\gamma_-\right]/Q(t) = H_{\gamma_-,\rho'}(x) \quad (x > 0), \tag{2.3}$$

and $Q(\cdot)$ does not change sign eventually with $Q(t) \to 0$ as $n \to \infty$. When $\gamma > 0$ and $\rho = 0$ the limit (2.3) vanishes.

Let $k(n)$ denote a sequence of positive integers satisfying $k(n) \to \infty$ and $\{k(n)/n\} \to 0$ as $n \to \infty$; in what follows we write k for $k(n)$. Then $X_{n-k,n}$ is called an *intermediate order statistic*, and the sequence $\{k\}$ is called an *intermediate order sequence*. The estimators of tail quantities presented in this paper are all based on the largest $(k+1)$ order statistics $\{X_{n-k,n}, \ldots, X_{n,n}\}$. For asymptotic normality to hold, k must satisfy some conditions. These are stated below in which (A.4) implies (A.3).

Condition A: (A.1). $k \to \infty$ and $(k/n) \to 0$ as $n \to \infty$. (A.2). The function U satisfies (2.1) and (2.2). (A.3). $k^{1/2} A(n/k) \to 0$ as $n \to \infty$. (A.4). $k^{1/2} Q(n/k) \to 0$ as $n \to \infty$. (A.5). If $\gamma < -1/2$ then $k^{|\gamma|} Q_0(n/k) = O(1)$.

Next we introduce a condition on the distribution function F, and assume that this holds in what follows.

Condition B. The function U corresponding to F satisfies the first and second order conditions of regular variation, (2.1) and (2.2) respectively. Further, F and its probability density function f satisfy the following conditions (these are the same as those in Lemma 6.1.1 of Csörgő and Horváth 1993 on page 369, and also in Proposition 2.4.9 of de Haan and Ferreira 2006): (i) F is differentiable on (x_*, x^*) where $x_* = \sup\{x : F(x) = 0\}$, $x^* = \inf\{x : F(x) = 1\}$, and $-\infty \leq x_* < x^* \leq \infty$. (ii) $f(x) > 0$, $(a < x < b)$. (iii) $\sup_{0<t<1} t(1-t)f'\{Q(t)\}/f^2\{Q(t)\} < C$ for some $C > 0$, where Q is the quantile function of F defined by $Q(t) = F^{\leftarrow}(t) = \inf\{x : F(x) \geq t\}, (0 < t < 1), Q(0) = Q(0+)$.

The estimators of tail parameters studied in this paper are based on the method of moments, which is a widely used general method of inference in this area (Einmahl *et al.* 2008, Dekkers *et al.* 1989). Derivations of these estimators may be found in section 3.5 of de Haan and Ferreira (2006); see (3.2.2), (3.5.2), and (3.5.9) in de Haan and Ferreira (2006). Let us introduce the following statistics and estimators: $H_n = k^{-1}\sum_{i=0}^{k-1}[\log X_{n-i,n} - \log X_{n-k,n}]$, $M_n = k^{-1}\sum_{i=0}^{k-1}(\log X_{n-i,n} - \log X_{n-k,n})^2$, $\hat{\gamma}_+ = H_n$, $\hat{\gamma}_- = 1 - 2^{-1}\{1 - (H_n^2/M_n)\}^{-1}$, $\hat{\gamma} = \hat{\gamma}_+ + \hat{\gamma}_-$, $\hat{a}(n/k) = X_{n-k,n}H_n(1 - \hat{\gamma}_-)$, and $\hat{b}(n/k) = X_{n-k,n}$.

3 Asymptotic validity of bootstrap

Let $X_1^*, \dots, X_n^*$ denote a simple random sample from the empirical distribution function F_n of $\{X_1, \dots, X_n\}$, and let $X_{1,n}^*, \dots, X_{1,n}^*$ denote the corresponding order statistics. Let $\{\hat{b}^*(n/k), \hat{\gamma}^*, \hat{a}^*(n/k), \hat{x}_n^*, \hat{p}_n^*\}$ denote the bootstrap statistics corresponding to $\{\hat{b}(n/k), \hat{\gamma}, \hat{a}(n/k), \hat{x}_n, \hat{p}_n\}$. Let P^* denote the bootstrap probability conditional on the sample $\{X_1, \dots, X_n\}$. The following results establish the consistency of the full-sample bootstrap to estimate asymptotic distributions of $\hat{b}(n/k), \hat{\gamma}, \hat{a}(n/k), \hat{x}_n$, and $\hat{p}_n$. In each of these theorems, the convergence in distribution of the non-bootstrap statistic is already known; references to the corresponding non-bootstrap result is provided in the Supplementary materials at the beginning of the proof of each theorem.

Theorem 3.1. *Suppose that (A.1), (A.2), (A.3), and (2.2) are satisfied for some $\gamma \in \mathbb{R}$, $\rho \leq 0$, and $\rho \neq \gamma$. Then, $P^*[Y_n^* \leq y \mid X_1, \dots, X_n] - P[Y_n \leq y] \to 0$, in probability, as $n \to \infty$, where $Y_n^* = k^{1/2}\{\hat{b}^*(n/k) - \hat{b}(n/k)\}/\hat{a}(n/k)$ and $Y_n = k^{1/2}\{\hat{b}(n/k) - U(n/k)\}/\hat{a}(n/k)$, $(y \in \mathbb{R})$.*

Since $\hat{b}(n/k) = X_{n-k}$, it follows from the above theorem that bootstrap method consistently estimates the asymptotic distribution of the intermediate order statistic X_{n-k}.

Theorem 3.2. *Suppose that (A.1), (A.2), (A.4), and (2.2) are satisfied for some* $\gamma \in \mathbb{R}$, $\rho \leq 0$, *and* $\rho \neq \gamma$. *Then,* $P^*[Y_n^* \leq y \mid X_1, \ldots, X_n] - P[Y_n \leq y] \to 0$, *in probability, as* $n \to \infty \quad (y \in \mathbb{R})$, *where* $Y_n^* = k^{1/2}(\hat{\gamma}^* - \hat{\gamma})$ *and* $Y_n = k^{1/2}(\hat{\gamma} - \gamma)$.

It follows from the foregoing theorem that asymptotically valid confidence intervals for the tail index γ may be constructed by bootstrap.

Theorem 3.3. *Suppose that (A.1), (A.2), (A.4), and (2.2) are satisfied for some* $\gamma \in \mathbb{R}$, $\rho \leq 0$, *and* $\rho \neq \gamma$. *Then,* $P^*[Y_n^* \leq y \mid X_1, \ldots, X_n] - P[Y_n \leq y] \to 0$, *in probability, as* $n \to \infty \quad (y \in \mathbb{R})$, *where* $Y_n^* = k^{1/2}\left[\hat{a}^*(n/k)/\hat{a}(n/k)) - 1\right]$ *and* $Y_n = k^{1/2}\left[\hat{a}(n/k)/a(n/k)) - 1\right]$.

Next, we consider estimation of a high quantile. Let p_n be a given small number in the range $0 < p_n < 1$; exactly what is meant by 'small' is made more precise later. For now, we may think of p_n being close to or even smaller than $(1/n)$. Let $x(p_n) = F^{\leftarrow}(1 - p_n)$ denote the upper p_nth quantile of F. We refer to $x(p_n)$ as a high-quantile since p_n is small. Let $d_n = k/(np_n)$. An estimator of $x(p_n)$ is $\hat{x}(p_n) = \hat{b}\,(n/k) + \hat{a}\,(n/k)\,\{d_n^{\hat{\gamma}} - 1\}/\hat{\gamma}$; see (4.3.3 in de Haan and Ferreira 2006 for arguments leading to this estimator. If $\hat{\gamma} = 0$ then $(d_n^{\hat{\gamma}} - 1)/\hat{\gamma}$ is interpreted as $\log d_n$. Let $q_\gamma(t) = \int_1^t s^{\gamma-1} \log s \; ds$, $(t > 0)$. The asymptotic distribution of $\hat{x}(p_n)$ is provided in the next theorem; it can be used for constructing a confidence interval for the quantile x_n.

Theorem 3.4. *Suppose that (i) (A.1), (A.2), and (A.4) are satisfied, (ii) (2.2) is satisfied for some* $\gamma \in \mathbb{R}$ *and* $\rho < 0$ *or* $\rho = 0$ *with* $\gamma < 0$, *and (iii)* $np_n = o(k)$ *and* $\log(np_n) = o(k^{1/2})$. *Let* $x_n = x(p_n), \hat{x}_n = \hat{x}(p_n)$, *and* $\hat{x}_n^* = \hat{x}^*(p_n)$ *Then, as* $n \to \infty$, $P^*[Y_n^* \leq y \mid X_1, \ldots, X_n] - P[Y_n \leq y] \to 0$, *in probability, as* $n \to \infty \quad (y \in \mathbb{R})$, *where* $Y_n^* = k^{1/2}(\hat{x}_n^* - \hat{x}_n)/\{(a(n/k)q_{\hat{\gamma}}(d_n)\}$ *and* $Y_n = k^{1/2}(\hat{x}_n - x_n)/\{a(n/k)q_{\hat{\gamma}}(d_n)\}$.

Next, we consider estimation of tail probability. Let x_n be a given large number, and consider estimation of the tail probability $p(x_n) = 1 - F(x_n)$. Let

$$\hat{p}(x_n) = (k/n)\left(\max\left\{0, 1 + \hat{\gamma}\{x_n - \hat{b}\,(n/k)\}/\hat{a}\,(n/k)\right\}\right)^{-1/\hat{\gamma}}; \tag{3.1}$$

see section 4.4 in de Haan and Ferreira (2006) for some details leading to this estimator. Let $w_\gamma(t) = t^{-\gamma}\int_1^t s^{\gamma-1}\log s\; ds, \quad (t > 0)$. The asymptotic distribution of the estimator $\hat{p}(x_n)$ is given in the next result; it can be used for constructing a confidence interval for $p(x_n)$.

Theorem 3.5. *Suppose that (i) (A.1), (A.2), and (A.4) are satisfied, (ii) (2.2) is satisfied for some* $\gamma > -(1/2)$, *(iii)* $\rho < 0$, *or* $\rho = 0$ *with* $\gamma < 0$, *and (iv)* $d_n \to \infty$ *and* $w_\gamma(d_n) = o(k^{1/2})$.

Let $p_n = p(x_n), \hat{p}_n = \hat{p}(x_n)$*,* $\hat{p}_n^* = \hat{p}^*(x_n)$*, and* $\hat{d}_n = k/(n\hat{p}_n)$*. Then,* $P^*[Y_n^* \leq y \mid X_1, \ldots, X_n] - P[Y_n \leq y] \to 0$*, in probability, as* $n \to \infty \quad (y \in \mathbb{R})$*, where* $Y_n^* = k^{1/2}\left[(\hat{p}_n^*/\hat{p}_n) - 1\right]/w_{\hat{\gamma}^*}(\hat{d}_n^*)$ *and* $Y_n = k^{1/2}\left[(\hat{p}_n/p_n) - 1\right]/w_{\hat{\gamma}}(\hat{d}_n)$*.*

Based on Theorems 3.1 – 3.5, asymptotically valid confidence interval for $b(n/k), \gamma, a(n/k), x(p_n)$ for a given small p_n, and $p(x_n)$ for a given large x_n, can be constructed by full-sample bootstrap. In this paper we study performance of the percentile, basic, and t-bootstrap methods for p_n by a simulation study.

To construct a confidence interval for the tail probability $p(x_n)$, where x_n is large number, suppose that the conditions of Theorem 3.5 are satisfied. Let $\hat{p}_{n,\alpha}^*$ denote the α-quantile of the distribution $P^*\{\hat{p}_n^* \leq x|X_1, \ldots, X_n\}$. Then, the *Efron's percentile* $(1-\alpha)$-confidence interval for $p(x_n)$ is $I_{Ep}(\alpha) = \left(\hat{p}_{n,\alpha/2}^*, \hat{p}_{n,1-\alpha/2}^*\right)$. Let $\zeta_{n,\alpha}$ denote the α-quantile of the distribution $P^*\left(k^{1/2}\log(\hat{p}_n^*/\hat{p}_n)/w_{\hat{\gamma}}(d_n) \leq x|X_1, \ldots, X_n\right)$, then the *percentile* $(1-\alpha)$-confidence interval for p_n is

$$I_p(\alpha) = \left(\hat{p}_n \exp\left\{-\zeta_{n,1-\alpha/2} w_{\hat{\gamma}}(d_n) k^{-1/2}\right\}, \hat{p}_n \exp\left\{-\zeta_{n,\alpha/2} w_{\hat{\gamma}}(d_n) k^{-1/2}\right\}\right). \quad (3.2)$$

To define the t-bootstrap confidence interval, first note that the limiting distribution in Theorem 3.5 is in fact normal with mean zero and variance

$$\sigma^2(\gamma) = (\gamma^2+1)I(\gamma \geq 0) + \frac{(1-\gamma)^2(1-3\gamma+4\gamma^2)}{(1-2\gamma)(1-3\gamma)(1-4\gamma)} I(\gamma < 0)$$

where $I(\cdot)$ denotes the indicator function (page 141, de Haan and Ferreira 2006). Let $\hat{\sigma}^2 = \sigma^2(\hat{\gamma})$, and let $\xi_{n,\alpha}$ denote the α-quantile of the distribution $P^*\left(k^{1/2}\log(\hat{p}_n^*/\hat{p}_n)/(w_{\hat{\gamma}}(d_n)\hat{\sigma}) \leq x|X_1, \ldots, X_n\right)$. Then a $(1-\alpha)$-level t-confidence interval for p_n is

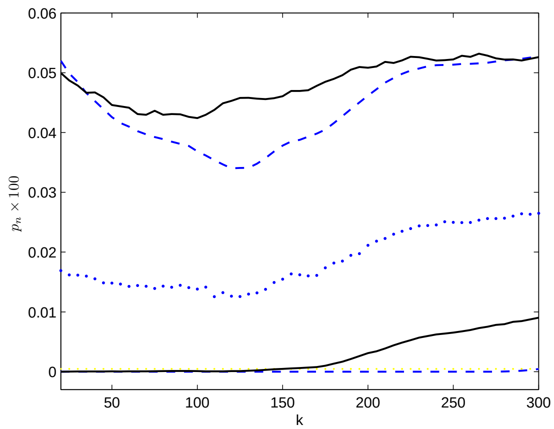
$$I_t(\alpha) = \left(\hat{p}_n \exp\left\{-\xi_{n,1-\alpha/2} w_{\hat{\gamma}}(d_n)\hat{\sigma} k^{-1/2}\right\}, \hat{p}_n \exp\left\{-\xi_{n,\alpha/2} w_{\hat{\gamma}}(d_n)\hat{\sigma} k^{-1/2}\right\}\right). \quad (3.3)$$

4 Application to Danish Fire Insurance Data

To illustrate the bootstrap method for constructing confidence intervals for a tail probability we use Danish fire insurance data. The data consists of insurance claims exceeding one million Danish Krone (DKK); each claim corresponds to total losses due to damages to buildings and contents, and loss of profits. There are $n = 2156$ observations in the data set. The data set is well known and studied, see McNeil (1997), Resnick (1997), and Lee and Qi (2019), for example.

For illustrative purposes, suppose that we wish to construct a 95% confidence interval for the probability of a loss over 300 million DKK. There are no observations exceeding this level and

Figure 1: The 95% confidence interval for $p_n = P(X > 300)$. The confidence intervals obtained by the percentile bootstrap and by the normal approximation methods are given as functions of k ($k = 20, 25, \ldots, 300$). The dotted line is the estimate of the tail probability p_n. The solid lines are upper and lower bounds of the percentile bootstrap confidence interval. The dashed lines are upper and lower bounds of the asymptotic confidence intervals.

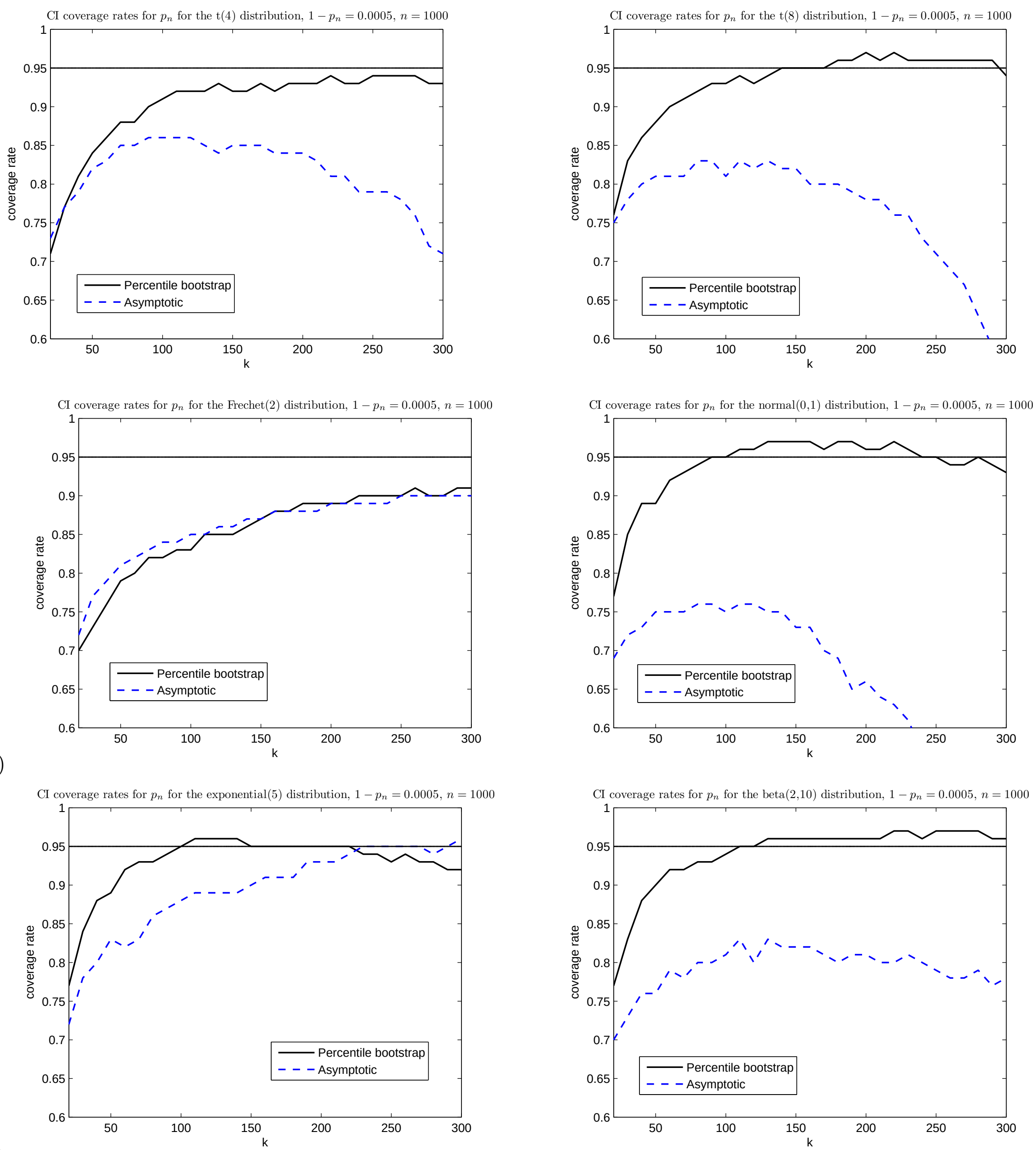


hence a method based on extreme value theory is essential; we have also performed the calculations for other thresholds. Figure 1 provides pointwise confidence intervals obtained by the percentile bootstrap and asymptotic methods. The asymptotic confidence interval is based on $\{k^{1/2}/w_{\hat{\gamma}}(\hat{d}_n)\}(p_n\hat{p}_n^{-1} - 1)$ being asymptotically normal. The tail probability estimates for the asymptotic method are stable for k ranging from 45 to 105. Confidence intervals obtained by the bootstrap and the asymptotic methods are close. The bootstrap confidence intervals are slightly more stable across different values of k. The asymptotic confidence intervals are slightly shorter on average; but this does not mean that they are better since one needs to take coverage probability into account. Since, in the simulation study, the bootstrap method provided better coverage than the asymptotic method, the indications are that the former is more reliable than the latter. We also carried out the calculations for $pr(X > 100)$. Out of the 2156 observations, there are only 3 that exceed this level; therefore, again methods based on extreme value theory are essential. For this case also, the relative performance of the two methods is essentially the same.

5 Simulation Study

We carried out a simulation study to evaluate the reliability of the bootstrap method. First, we outline the design of the study and then present a summary of the results; detailed results of the

Figure 2: Coverage rates of the asymptotic and bootstrap 95% confidence intervals for the tail probability $1 - F(x_n)$, where x_n was chosen such that $p_n := 1 - F(x_n) = 0.0005$ and $n = 1000$.



b)

c)

study, mostly in the form of graphs, are presented in the Supplementary Materials.

Design

For the population distribution function F, we considered the following cases. (I) Heavy tail ($\gamma > 0$): $t(2), t(4), t(8)$, Frechet(1), and Frechet(2). [For $t(\alpha)$ and Frechet(α), the tail index $\gamma = \alpha^{-1}$.] (II) Light tail ($\gamma = 0$): exp(5) and $N(0,1)$. (III) Bounded support ($\gamma < 0$): beta(2, 10), beta(1, 2).

For each distribution, we considered sample sizes $n = 200, 1000$. First, we choose a small tail probability, denoted p_n, and compute $x_n = F^{-1}(1 - p_n)$. The unknown tail quantity to be estimated is p_n that corresponds to the known high-quantile, x_n. We considered $np_n = 1, 0.5, 0.25$ when $n = 1000$, and $np_n = 5, 1, 0.5$ when $n = 200$. To this end, we apply Theorem 3.5. The main simulation steps are: (1) Choose a population distribution function F, sample size n, and a small tail probability, p_n. (2) Generate a simple random sample $X_1, \ldots, X_n$ from F. For each given k in a range of values, estimate p_n corresponding to x_n, and a confidence interval based on the asymptotic normality of the estimator. (3) Draw a simple random sample $\{X_1^*, \ldots, X_n^*\}$ from $\{X_1, \ldots, X_n\}$. (4) For each k, compute the estimate $\hat{p}_n^*$. (5) Repeat the previous two steps 1000 times and compute the percentile, basic, and t bootstrap confidence intervals, for each k.

Results

The coverage rates for $n = 1000$ and $1 - p = 0.0005$ are presented in Figure 2; the results for $1 - p = 0.0001, 0.00025$ and $n = 200$ are presented in a working paper of the authors. Here, we present a summary of the main observations. We use the Efron's percentile bootstrap method for comparison; the other ones mentioned in Section 3 did not perform as well. The two main observations of the simulation study:

(1) For every model considered, the bootstrap method performed reasonably well in terms of coverage rate. More specifically, the observed coverage rate was reasonably close to the nominal level compared to other simulation results published in this area.

(2) The bootstrap method performed better than the asymptotic method in most cases, and in the remaining small number of cases the difference was small. More specifically, the bootstrap method had better coverage rates than the asymptotic method for all the distributions, except for some cases of Frechet(1) and Frechet(2) for which the two coverage rates were close. In these cases, the bootstrap method was slightly better than the asymptotic method for smaller tail probabilities, and the asymptotic method was slightly better for less extreme tail probabilities. In general, the lighter

the tail of the population distribution, the more advantage the bootstrap method had over the asymptotic method in terms of coverage rate. In terms of length of the confidence interval, the two methods were comparable. Overall, the bootstrap method performed better than the asymptotic method.

Acknowledgement

We wish to thank Professor John Einmahl for helpful comments and suggestions. Svetlana Litvinova would like to thank Monash University for support through a postgradauate scholarship; this paper is based on her doctoral dissertation. Mervyn Silvapulle gratefully acknowledges support from the Australian Government through the Australian Research Council. We acknowledge that, in 2018, Professor Chen Zhou of Erasmus University (joint work with Laurens de Haan and Alex Koning) presented a seminar at Monash University which contained some results that overlap with those presented in this paper.

References

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley & Sons (Chichester).

Cavaliere, G., Georgiev, I., and Taylor, A. M. R. (2013). Wild bootstrap of the sample mean in the infinite variance case. *Econometric Reviews*, **32**(2), 204–219.

Cavaliere, G., Georgiev, I., and Taylor, A. M. R. (2016). Sieve-based inference for infinite-variance linear processes. *The Annals of Statistics*, **44**(4), 1467–1494.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag Inc.

Cornea-Madeira, A. and Davidson, R. (2015). A parametric bootstrap for heavy-tailed distributions. *Econometric Theory*, **31**(3), 449–470.

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. Wiley, New York.

Csörgő, S. and Mason, D. M. (1989). Bootstrapping empirical functions. *The Annals of statistics*, **17**(4), 1447–1471.

Davidson, R. (2012). Statistical inference in the presence of heavy tails. *Econometrics Journal*, **15**, C31–C53.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer.

Dekkers, A. L. M., Einmahl, J. H. J., and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, **17**(4), 1833–1855.

Einmahl, J. (1990). The empirical distribution function as a tail estimator. *Statistica Neerlandica*, **44**(2), 79–82.

Einmahl, J. H. J., Krajina, A., and Segers, J. (2008). A method of moments estimator of tail dependence. *Bernoulli*, **14**, 1003–1026.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Springer-Verlag Inc.

Geluk, J. and de Haan, L. (2002). On bootstrap sample size in extreme value theory. *Publ. Inst. Math. (N.S.)*, **71**(85), 21–25.

Gomes, M. I., Caeiro, F., Henriques-Rodrigues, L., and Manjunath, B. G. (2016). Bootstrap methods in statistics of extremes. In F. Longin, editor, *Extreme Events in Finance: A Handbook of Extreme Value Theory and its Applications*, chapter 6, pages 117–138. Wiley, New York.

Lahiri, S. N. (2003). Bootstrapping heavy-tailed data and extremes. In *Resampling Methods for Dependent Data*, chapter 11, pages 261–279. Springer, New York, NY.

Lee, Y. and Qi, Y. (2019). Adjusted empirical likelihood method for the tail index of a heavy-tailed distribution. *Statistics and Probability Letters*, **152**, 50–58.

McNeil, A. (1997). Estimating the tails of loss severity distribution using extreme value theory. *Astin Bulletin: A Journal of the International Actuarial Association*, **27**, 117–137.

Resnick, S. I. (1997). Discussion of the danish data on large fire insurance losses. *Astin Bulletin: A Journal of the International Actuarial Association*, **27**, 139–151.

Tajvidi, N. (2003). Confidence intervals and accuracy estimation for heavytailed generalized pareto distributions. *Extremes*, **6**(2), 111–123.

Supplementary materials for "Consistency of full-sample bootstrap tail estimators: high-quantile, tail probability, and tail index"

Svetlana Litvinova and Mervyn J. Silvapulle

Department of Econometrics and Business Statistics, Monash University, Australia

svetlana.litvinova@monash.edu, mervyn.silvapulle@monash.edu

Abstract

The proofs of the results in the main paper are provided in this Supplement. Initially, some known results are stated. Then, several lemmas are established, and then the proofs for the main results are provided. Simulation results on coverage rates and lengths of confidence intervals are also provided.

Keywords: Full-sample bootstrap; Intermediate order statistic; Extreme value index; Hill estimator; Tail probability; Tail quantile .

1 Some known results on tail quantile processes.

Let $V_1, \ldots, V_n$ be independent and identically distributed random variables with the Uniform(0,1) distribution. Let $V_{1,n} \leq \ldots, V_{n,n}$ denote the order statistics of $V_1, \ldots, V_n$. Let $[x]$ denote the smallest integer larger than or equal to x. Let $V_n(s) = V_{[ns],n} \quad (0 < s \leq 1)$ with $V_n(0) = V_{1,n}$ denote the *empirical quantile function*. Define the *empirical quantile process* of $\{V_1, \ldots, V_n\}$ by $\beta_n(s) = n^{1/2}\{s - V_n(s)\} \quad (0 \leq s \leq 1)$. A result on β_n that we use is the following: [Csörgő *et al.* 1986, Theorem 2.1] There exists a probability space $(\Omega, \mathcal{A}, P)$ which carries a sequence of independent uniform (0,1) random variables $\{V_i\}_{i\in\mathbb{N}}$ and a sequence of Brownian bridges $\{B_i\}_{i\in\mathbb{N}}$, where $\mathbb{N}$ denotes the set of positive integers, such that

$$\sup_{\lambda/n \leq s \leq 1-\lambda/n} |\beta_n(s) - B_n(s)|/\{s(1-s)\}^{1/2-\nu} = O_P(n^{-\nu}), \ (0 < \lambda < \infty; 0 \leq \nu < 1/2). \tag{1.1}$$

Csörgő and Mason (1989) established the following result similar to (1.1) for the bootstrap version of the uniform empirical quantile process: There exists a probability space, denoted $(\Omega', \mathcal{A}', P')$ that

Corresponding author: Svetlana Litvinova, Department of Econometrics and Business Statistics, Monash University, Australia. email: svetlana.litvinova@monash.edu.

carries the aforementioned sequences $\{V_i\}_{i\in\mathbb{N}}$ and $\{B_i\}_{i\in\mathbb{N}}$, together with a sequence of independent uniform (0,1) random variables $\{V_i'\}_{i\in\mathbb{N}}$ and a sequence of Brownian bridges $\{B_i'\}_{i\in\mathbb{N}}$ such that the two sets of random elements $\{V_i\}_{i\in\mathbb{N}}\cup\{B_i(s) : 0 \le s \le 1\}_{i\in\mathbb{N}}$ and $\{V_i'\}_{i\in\mathbb{N}}\cup\{B_i'(s) : 0 \le s \le 1\}_{i\in\mathbb{N}}$ are independent. Let $\{V_{i,n}', \beta_n'(s), V_n'(s)\}$ be defined in the same way as $\{V_{i,n}, \beta_n(s), V_n(s)\}$ except that $\{V_i\}_{i\in\mathbb{N}}$ is replaced by $\{V_i'\}_{i\in\mathbb{N}}$ $(i = 1,\dots,n; 0 \le s \le 1)$.

Let $V_i^* = V_n(V_i')$ and $V_n^*(s) = V_n\{V_n'(s)\}$ $(i = 1,\dots,n;\ 0 \le s \le 1)$. Then, $\{V_1^*,\dots,V_n^*\}$ is a simple random sample from $\{V_1,\dots,V_n\}$; in what follows, we treat $\{V_1^*,\dots,V_n^*\}$ as a bootstrap sample from $\{V_1,\dots,V_n\}$, and $V_n^*(s)$ as the corresponding *bootstrap empirical quantile function.* Let $\{V_{1,n}^*,\dots V_{n,n}^*\}$ denote the order statistics of $\{V_1^*,\dots V_n^*\}$. Then, we have $V_{[ks],n}^* = V_n(V_{[ks],n}') = V_n(V_{[nks/n],n}') = V_n\{V_n'(ks/n)\} = V_n^*(ks/n)$.

Define the *bootstrap uniform quantile process* as $\beta_n^*(s) = n^{1/2}\{V_n(s) - V_n^*(s)\}$ $(0 \le s \le 1)$. Csörgő and Mason (1989) (see Theorem 2.1) showed that

$$\sup_{\lambda/n\le s\le 1-\lambda/n} \left|\beta_n^*(s) - B_n'(s)\right|/\{s(1-s)\}^{1/2-\nu} = O_{P'}(n^{-\nu}) \quad (0 < \lambda < \infty; 0 \le \nu < 1/4) \tag{1.2}$$

where the P' in $O_{P'}(n^{-\nu})$ refers to the probability in the space that carries both $\{V_i, B_i(\cdot)\}_{i\in\mathbb{N}}$ and $\{V_i', B_i'(\cdot)\}_{i\in\mathbb{N}}$.

Let k denote an *intermediate order sequence*; therefore $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. Define the *uniform tail quantile process* as $\beta_{k,n}(s) = (n/k)^{1/2}\beta_n(ks/n)$ and the *uniform tail quantile bootstrap processes* as $\beta_{k,n}^*(s) = (n/k)^{1/2}\beta_n^*(ks/n)$ $(0 < s \le 1)$. Then, by simple substitution, we obtain $\beta_{k,n}(s) = k^{1/2}\{s - (n/k)V_{[ks],n}\}$ and $\beta_{k,n}^*(s) = k^{1/2}\{(n/k)V_{[ks],n} - (n/k)V_{[ks],n}^*\}$ $(0 \le s \le 1)$.

Let $B_{k,n}(s) = (n/k)^{1/2}B_n(ks/n)$ and $B_{k,n}'(s) = (n/k)^{1/2}B_n'(ks/n)$ $(0 < s \le 1)$ where the Brownian bridges B_n and B_n' are the ones introduced earlier to approximate the quantile processes β_n and β_n' $(n = 1, 2,\dots)$ respectively. Then, we have (Theorem 2.3 in Peng and Qi 2017)

$$\sup_{1\le sk\le n-1} s^{-\nu}\left|\beta_{k,n}(s) - B_{k,n}(s)\right| = o_P(1) \quad (0 \le \nu < 1/2). \tag{1.3}$$

Let W denote a standard Wiener process on $[0, 1]$. The following result about the order of magnitude of W near the origin, is important (see Remark 1 in Csörgő and Mason 1989): for $0 < \delta \le 1/2$, we have

$$a^{-\delta}\sup_{0\le s\le a} s^{-1/2+\delta}|B_n(s)| \to \sup_{0\le s\le 1} s^{-1/2+\delta}|W(s)| \tag{1.4}$$

in distribution, as $a \downarrow 0$. In addition, we make use of the following results on the empirical quantile function: For $0 < \rho < \infty$, we have (Csörgő and Mason 1989, Theorem 2.3 and page 1454)

$$\sup_{0<s<1} s/V_n(s), \quad \sup_{\rho/n\le s<1} V_n(s)/s, \quad \sup_{0<s<1} s/V_n^*(s), \quad \sup_{\rho/n\le s<1} V_n^*(s)/s \; = O_{P'}(1), \tag{1.5}$$

$$\sup_{0<s\le 1} \frac{ks}{(nV_{[ks],n})}, \quad \sup_{1/k\le s\le 1} \frac{nV_{[ks],n}}{ks}, \quad \sup_{0<s\le 1} \frac{ks}{(nV^*_{[ks],n})}, \quad \sup_{1/k\le s\le 1} \frac{nV^*_{[ks],n}}{ks} \; = O_{P'}(1). \tag{1.6}$$

2 Preliminary lemmas to prove the theorems in the paper

Lemma 2.1. *For $0 \le \nu < 1/2$ and $\epsilon > 0$, we have*

$$\sup_{0\le s\le 1} s^{-\nu}|B_{k,n}(s)| = O_{P'}(1), \quad \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{\sup_{0\le s\le\delta} s^{-\nu}|B_{k,n}(s)| < \epsilon\right\} = 1. \tag{2.1}$$

Proof. Let $\delta > 0$, $\mu = (1/2) - \nu$, and $a = (k\delta/n)$. Then using the definition of $B_{k,n}$, we obtain, $\sup_{0\le s\le\delta} s^{-\nu}|B_{k,n}(s)| = \delta^\mu a^{-\mu} \sup_{0\le t\le a} t^{-(1/2)+\mu}|B_n(t)|$. Now, the proof follows from (1.4). □

Lemma 2.2. *Let $0 \le \nu < 1/2$, $\mu = (1/2) - \nu$, $0 < \lambda \le 1$, and $0 < \delta < \min\{\mu, 1/4\}$. Then*

$$\sup_{\lambda/k\le s\le 1} \left|\beta^*_{k,n}(s) - B'_{k,n}(s)\right|/s^\nu = O_{P'}(k^{-\delta}) = o_{P'}(1). \tag{2.2}$$

Proof. For any $0 < t \le k/n$, we have $(n/k)^\mu t^{\mu-1/2} \le (n/k)^\delta t^{\delta-1/2}$. Then,

$$\begin{aligned} &\sup_{\lambda/k\le s\le 1} \left|\beta^*_{k,n}(s) - B'_{k,n}(s)\right|/s^\nu = \sup_{\lambda/n\le t\le k/n} (n/k)^\mu \left|\beta^*_n(t) - B'_n(t)\right|/t^{1/2-\mu} \\ \le\; &\sup_{\lambda/n\le t\le k/n} n^\delta\left|\beta^*_n(t) - B'_n(t)\right|/\{t(1-t)\}^{1/2-\delta} k^{-\delta} \sup_{1/n\le t\le k/n} (1-t)^{1/2-\delta} = O_{P'}(k^{-\delta}) = o_{P'}(1). \end{aligned}$$

where the ineqequality uses (1.2). □

Lemma 2.3. *Let $\epsilon > 0$ be given. Then,*

$$\sup_{0<s\le 1} s^{1/2+\epsilon}\left|k^{1/2}\big[\log\{nk^{-1}V_{[ks],n}\} - \log\{nk^{-1}V^*_{[ks],n}\}\big] - s^{-1}B'_{k,n}(s)\right| = o_{P'}(1). \tag{2.3}$$

Proof. We consider the supremum over $[1/k, 1]$ and that over $(0, 1/k]$ separately. First, we consider the former; therefore, let $s \in [1/k, 1]$. Let $\xi_l(s)$ and $\xi_u(s)$ denote the minimum and maximum of $\{(n/k)V^*_{[ks],n}, (n/k)V_{[ks],n}\}$. By the mean value theorem we have

$$k^{1/2}\big[\log\{nk^{-1}V_{[ks],n}\} - \log\{nk^{-1}V^*_{[ks],n}\}\big] = -{\xi_n}^{-1}k^{1/2}\left(nk^{-1}V_{[ks],n} - nk^{-1}V^*_{[ks],n}\right) = \xi_n^{-1}(s)\beta^*_{k,n}(s),$$

where $\xi_n(s) \in [\xi_l(s), \xi_u(s)]$. Then,

$$\sup_{1/k\le s\le 1} s^{1/2+\epsilon}\left|k^{1/2}\big[\log\{nk^{-1}V_{[ks],n}\} - \log\{nk^{-1}V^*_{[ks],n}\}\big] - s^{-1}B'_{k,n}(s)\right| \le A_1 + A_2,$$

$$A_1 = \sup_{1/k \le s \le 1} \frac{s}{\xi_n(s)} \sup_{1/k \le s \le 1} \frac{\left|\beta^*_{k,n}(s) - B'_{k,n}(s)\right|}{s^{1/2-\epsilon}}, \quad A_2 = \sup_{1/k \le s \le 1} \frac{B'_{k,n}(s)}{s^{1/2-\epsilon}} \frac{|\xi_n(s) - s|}{\xi_n(s)}. \tag{2.4}$$

It may be verified using (1.4), (1.5), (1.6), and Lemmas 2.1 and 2.2, that $A_1 = o_{P'}(1)$ and $A_2 = o_{P'}(1)$. Next consider the proof for the supremum over $0 < s < k^{-1}$. First, substituting $s = 1/k$ in the foregoing result for the supremum over $k^{-1} \le s \le 1$ we have

$$(1/k)^{1/2+\epsilon} \left| k^{1/2} \left[\log\{nk^{-1} V_{[ks],n}\} - \log\{nk^{-1} V^*_{[ks],n}\} \right] - k B'_{k,n}(1/k) \right| = o_{P'}(1). \tag{2.5}$$

Then, the proof of the Lemma for the interval $0 < s < 1/k$ follows by (2.2). □

Lemma 2.4. *Let $\epsilon > 0$ be given. Then, we have* $\sup_{\lambda/k \le s \le 1} s^{\gamma+(1/2)+\epsilon} \left| L_n(s) - B'_{k,n}(s)/s^{\gamma+1} \right| = o_{P'}(1)$, *for $\gamma \in \mathbb{R}$ and $0 < \lambda \le 1$, where*

$$L_n(s) = k^{1/2} \gamma^{-1} \left\{ \left(nk^{-1} V^*_{[ks],n} \right)^{-\gamma} - 1 \right\} - k^{1/2} \gamma^{-1} \left\{ \left(nk^{-1} V_{[ks],n} \right)^{-\gamma} - 1 \right\}$$

and the function $\gamma^{-1}(s^{-\gamma} - 1)$ is interpreted to be equal to $-\log s$ when $\gamma = 0$.

Proof. For $\gamma = 0$ the result is established in Lemma 2.3. For $\gamma \ne 0$, we consider the suprema over $\lambda/k \le s \le 1/k$ and $1/k \le s \le 1$ separately. First, let us consider $1/k \le s \le 1$ with $\gamma \ne 0$. Let $\xi_l(s)$ and $\xi_u(s)$ denote the minimum and maximum of $\{(n/k) V^*_{[ks],n}, (n/k) V_{[ks],n}\}$ $(1/k \le s \le 1)$. Let $\varphi(t) = \gamma^{-1}(t^{-\gamma} - 1)$. Applying the mean value theorem to $\varphi(t)$ and noting that $\dot{\varphi}(t) = -t^{-\gamma-1}$, we obtain $L_n(s) = \xi_n^{-\gamma-1}(s) \beta^*_{k,n}(s)$, where $\xi_n(s) \in [\xi_l(s), \xi_u(s)]$. Then,

$$\sup_{1/k \le s \le 1} s^{\gamma+(1/2)+\epsilon} \left| L_n(s) - B'_{k,n}(s)/s^{\gamma+1} \right| \le A_1 + A_2$$

$$A_1 = \sup_{1/k \le s \le 1} \left(\frac{s}{\xi_n(s)} \right)^{\gamma+1} \sup_{1/k \le s \le 1} \frac{\left| \beta^*_{k,n}(s) - B'_{k,n}(s) \right|}{s^{1/2-\epsilon}},$$

$$A_2 = \sup_{1/k \le s \le 1} \frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}} \left| 1 - \{s/\xi_n(s)\}^{\gamma+1} \right|.$$

By (1.6) and (2.2), $A_1 = o_{P'}(1)$. Next, let $\delta \in (0,1)$. Choose $k > \delta^{-1}$. Then,

$$\begin{aligned} A_2 &\le \sup_{0 < s \le \delta} \frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}} \sup_{1/k \le s \le 1} \left| 1 - \{\frac{s}{\xi_n(s)}\}^{\gamma+1} \right| + \sup_{0 < s \le 1} \frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}} \sup_{\delta < s \le 1} \left| 1 - \{\frac{s}{\xi_n(s)}\}^{\gamma+1} \right| \\ &\le B_1(n,k,\delta) B_2(n,k) + B_3(n,k) B_4(n,k,\delta), \qquad \text{say.} \end{aligned}$$

By using arguments similar to those for Lemma 2.3, it may be verified that $B_1(n,k,\delta)$ can be made arbitrarily small for large $\{n,k\}$ and small δ. Further, $B_2(n,k) = O_{p'}(1)$, and $B_4(n,k,\delta) = o_{P'}(1)$. Therefore, $A_2 = o_{P'}(1)$ and hence the proof follows for the range $k^{-1} \le s \le 1$. The proof for the supremum over $s \in [\lambda/k, 1/k]$ follows as in the proof of the previous Lemma. This completes the proof of the Lemma. □

3 Proofs of Theorems

Let $V_1, \dots, V_n$ be independent and identically distributed as Uniform(0,1) random variables, and $W_i = 1 - V_i$, $(i = 1, \dots n)$. Let $X_i = F^{-1}(W_i)$, $(i = 1, \dots n)$. Then, $X_1, \dots, X_n$ are independent and identically distributed random variables from the distribution function F, and $X_i = F^{-1}(1-V_i)$ $(i = 1, \dots n)$. Let $\{V_i'\}_{i=1}^n$ and $\{V_i^*\}_{i=1}^n$ be defined as at the beginning of Section 1, more specifically as in Csörgő and Mason (1989). Let $X_i^* = F^{\leftarrow}(1 - V_i^*)$ $(i = 1, \dots n)$. Since $\{V_1^*, \dots, V_n^*\} \subseteq \{V_1, \dots, V_n\}$, it follows that $\{X_1^*, \dots, X_n^*\} \subseteq \{X_1, \dots, X_n\}$. Let $\{X_{1,n} \leq \dots \leq X_{n,n}\}$ and $\{X_{1,n}^* \leq \dots \leq X_{n,n}^*\}$ denote order statistics of $\{X_1, \dots, X_n\}$ and $\{X_1^*, \dots, X_n^*\}$, respectively. Then, since $U(x) = F^{\leftarrow}(1 - x^{-1})$, we have $X_{n-[ks]+1,n} = F^{\leftarrow}(1 - V_{[ks],n}) = U(1/V_{[ks],n})$ and $X_{n-[ks]+1,n}^* = F^{\leftarrow}(1 - V_{[ks],n}^*) = U(1/V_{[ks],n}^*)$ $(0 < ks \leq n)$. Let $\mathcal{V}_s = k/(nV_{[ks],n})$ and $\mathcal{V}_s^* = k/(nV_{[ks],n}^*)$; $\mathcal{V}_s$ and $\mathcal{V}_s^*$ also depend on (k, n) but not shown explicitly.

3.1 Proofs of main results in terms of joint probability P'

Let $0 < \tau < 1$ be given, and let $\mathcal{T}$ denote the interval $\tau/k < s < 1+(\tau/k)$. Throughout this section, suprema of functions of the real variable s over the interval $\mathcal{T}(k, \tau)$ arises frequently. To simplify notation we write $\sup_{s\in\mathcal{T}} G(s)$ for $\sup_{\tau/k<s<1+(\tau/k)} G(s)$ for any function $G(s)$. Typically, the main results are proved for $\tau/k < s \leq 1$; in general, the proof could be extended to $\tau/k < s < 1 + (\tau/k)$ without much difficulty. The next result is a bootstrap version of a result from Drees (1998) (see also de Haan and Ferreira 2006a, Theorem 2.4.2).

Lemma 3.1. *Suppose that the conditions of Theorem 3.1 are satisfied. Let γ denote the tail index of F ($\gamma \in \mathbb{R}$). Let $\epsilon > 0$ and $0 < \tau < 1$ be given. Then, for suitably chosen function $a_0(\cdot)$, we have*

$$\sup_{s\in\mathcal{T}} s^{\gamma+(1/2)+\epsilon} \left| k^{1/2} \left\{ X_{n-[ks]+1,n}^* - X_{n-[ks]+1,n} \right\} / a_0(n/k) - s^{-\gamma-1} B_{k,n}'(s) \right| = o_{P'}(1). \tag{3.1}$$

Proof. Let $\Lambda_n(s; \gamma, \rho, \delta) = \max\left\{\mathcal{V}_s^{*(\gamma+\rho+\delta)}, \mathcal{V}_s^{*(\gamma+\rho-\delta)}\right\} + \max\left\{\mathcal{V}_s^{\gamma+\rho+\delta}, \mathcal{V}_s^{\gamma+\rho-\delta}\right\}$. Let $\epsilon > 0$ and $\delta > 0$ be given. We apply inequality (2.3.17) in Theorem 2.3.6 of de Haan and Ferreira (2006a) with $t := t_n = n/k$ and $x := x_n(s) = k/(nV_{[ks],n})$ $(s \in \mathcal{T})$. Since $\sup_{\mathcal{T}} V_{[ks],n} \leq V_{k+1,n} = o_{P'}(1)$, we have that $tx \geq \inf_{\mathcal{T}} t_n x_n(s) \to \infty$, in probability$[P']$. Hence, for any fixed t_0, $t_n x_n(s) > t_0$ with probability approaching 1, uniformly on $s \in \mathcal{T}$. For the rest of the proof of this theorem, the inequalities are implicitly assumed to hold uniformly over $s \in \mathcal{T}$, with probability approaching 1 as $n \to \infty$, without further comment.

Let $a_0(t), A_0(t), A_*(t)$, and $a_*(t)$ be as in Corollary 2.3.5 and Theorem 2.3.6 in de Haan and Ferreira (2006a). Invoking the latter theorem, we we obtain

$$\left|\frac{U(1/V_{[ks],n}) - U(n/k)}{a_0(n/k)} - \frac{\mathcal{V}_s^{\gamma} - 1}{\gamma} - A_0\left(n/k\right)\Psi_{\gamma,\rho}(\mathcal{V}_s)\right| \le \varepsilon \left|A_0\left(n/k\right)\right| \Lambda_n(s;\gamma,\rho,\delta). \tag{3.2}$$

Similarly, for the bootstrap sample, with $x_n(s) = k/(nV^*_{[ks],n})$, we have

$$\left|\frac{U(1/V^*_{[ks],n}) - U(n/k)}{a_0(n/k)} - \frac{(\mathcal{V}_s^*)^{\gamma} - 1}{\gamma} - A_0\left(n/k\right)\Psi_{\gamma,\rho}\left(\mathcal{V}_s^*\right)\right| \le \varepsilon \left|A_0\left(n/k\right)\right| \Lambda_n(s;\gamma,\rho,\delta). \tag{3.3}$$

Let

$$\begin{aligned}
D_0(s) &= s^{\gamma+1/2+\epsilon}\frac{B'_{k,n}(s)}{s^{\gamma}+1}, \quad D_1(s) = s^{\gamma+1/2+\epsilon}k^{1/2}\frac{X^*_{n-[ks]+1,n} - X_{n-[ks]+1,n}}{a_0(n/k)}\\
D_2(s) &= s^{\gamma+1/2+\epsilon}k^{1/2}\left\{(\mathcal{V}_s^{*\gamma}-1)/\gamma - (\mathcal{V}_s^{\gamma}-1)/\gamma\right\}\\
D_3(s) &= s^{\gamma+1/2+\epsilon}k^{1/2}A_0\left(n/k\right)\left\{\Psi_{\gamma,\rho}\left(\mathcal{V}_s^*\right) - \Psi_{\gamma,\rho}\left(\mathcal{V}_s\right)\right\}\\
D_4(s) &= \varepsilon k^{1/2}\left|A_0\left(n/k\right)\right| s^{\gamma+(1/2)+\epsilon}\Lambda_n(s;\gamma,\rho,\delta).
\end{aligned}$$

Subtracting (3.2) from (3.3), and multiplying by $k^{1/2}s^{\gamma+(1/2)+\epsilon}$, we obtain,

$$|D_1(s) - D_2(s) - D_3(s)| \le D_4(s) \quad (s \in \mathcal{T}).$$

By Lemma 2.4, for any $\gamma \in \mathbb{R}$, we have $\sup_{s\in\mathcal{T}} |D_2(s) - D_0(s)| = o_{P'}(1)$. Substituting $\gamma = 1$ in this inequality, we obtain $\sup_{s\in\mathcal{T}} s^{(3/2)+\epsilon}k^{1/2}\left|\mathcal{V}_s^* - \mathcal{V}_s\right| = O_{P'}(1)$. By the mean value theorem, $\Psi_{\gamma,\rho}(\mathcal{V}_s^*) - \Psi_{\gamma,\rho}(\mathcal{V}_s) = \dot{\Psi}_{\gamma,\rho}(\zeta_n(s))(\mathcal{V}_s^* - \mathcal{V}_s)$ where $\zeta_n(s)$ lies between $\mathcal{V}_s^*$ and $\mathcal{V}_s$. By (1.6), we have

$$\sup_{s\in\mathcal{T}} s\zeta_n(s) = O_{P'}(1), \quad \sup_{s\in\mathcal{T}} \{s\zeta_n(s)\}^{-1} = O_{P'}(1), \quad \sup_{s\in\mathcal{T}} \log\{s\zeta_n(s)\} = O_{P'}(1). \tag{3.4}$$

Now, using the functional form of $\Psi_{\gamma,\rho}(t)$ in (2.3.16) of de Haan and Ferreira (2006a) on page 46, it may be verified that $\sup_{s\in\mathcal{T}} |D_3(s)| = o_{P'}(1)$. Choose $\delta > 0$ such that $\delta < 1/2 + \epsilon - \rho$. By (1.6), $\sup_{s\in\mathcal{T}}(ks/(nV_{[ks],n}))^{\gamma+\rho\pm\delta} = O_{P'}(1)$. Then,

$$\sup_{s\in\mathcal{T}} s^{\gamma+1/2+\epsilon}\mathcal{V}_s^{\gamma+\rho\pm\delta} = O_{P'}(1), \quad \sup_{s\in\mathcal{T}} s^{\gamma+1/2+\epsilon}\mathcal{V}_s^{*(\gamma+\rho\pm\delta)} = O_{P'}(1). \tag{3.5}$$

By (3.5) and $k^{1/2}A(n/k) = o(1)$, we obtain $\sup_{s\in\mathcal{T}} |D_4(s)| = o_{P'}(1)$. We have shown that $|D_1(s) - D_2(s) - D_3(s)| \le D_4(s)$, $\sup_{s\in\mathcal{T}} |D_2(s) - D_0(s)| = o_{P'}(1)$, $\sup_{s\in\mathcal{T}} |D_3(s) = o_{P'}(1)$, and $\sup_{s\in\mathcal{T}} |D_4(s)| = o_{P'}(1)$. Therefore, $\sup_{s\in\mathcal{T}} |D_1(s) - D_0(s)| = o_{P'}(1)$. □

The proof of the next lemma follows from the previous lemma, (3.1), and (2.1); therefore the proof is omitted.

Lemma 3.2. *Suppose that the conditions of Lemma 3.1 are satisfied, and let* $b_0(n) = (X^*_{n,n} - X_{n,n})I(\gamma < -1/2)$, *where* I *denotes the indicator function. Then, for* $\epsilon > 0$ *and* $0 < \tau < 1$ *we have*

$$\sup_{0<s\leq 1+\tau/k} s^{\gamma+(1/2)+\epsilon}\left| k^{1/2}\frac{X^*_{n-[ks]+1,n} - X_{n-[ks]+1,n} - b_0(n)}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1). \tag{3.6}$$

Lemma 3.3. *Suppose that the conditions of Theorem 3.1 are satisfied. Then,*

$$k^{1/2}\{\hat{b}^*(n/k) - \hat{b}(n/k)\}/\hat{a}(n/k) = B'_{k,n}(1) + o_{P'}(1). \tag{3.7}$$

Proof. Let a_0 be as in Theorem 2.3.6 in de Haan and Ferreira (2006a). By the definitions of $\hat{b}^*(n/k)$ and $\hat{b}(n/k)$, we have

$$k^{1/2}\frac{\hat{b}^*(n/k) - \hat{b}(n/k)}{\hat{a}(n/k)} = k^{1/2}\frac{X^*_{n-k,n} - X_{n-k,n}}{a_0(n/k)}\ \frac{a_0(n/k)}{a(n/k)}\ \frac{a(n/k)}{\hat{a}(n/k)}.$$

By Lemma 3.1 with $s = 1$, we have $k^{1/2}\{X^*_{n-k,n} - X_{n-k,n}\}/a_0(n/k) = B'_{k,n}(1) + o_{P'}(1)$. By Lemma 4.7 in de Haan and Resnick (1993), we have $\hat{a}(n/k)/a(n/k) = 1 + o_{P'}(1)$. By Theorem 2.3.6 and Corollary 2.3.5 in de Haan and Ferreira (2006a), we have $a_0(n/k)/a(n/k) = 1 + o(1)$. Therefore, $k^{1/2}\{\hat{b}^*(n/k) - \hat{b}(n/k)\}/\hat{a}(n/k) = [B'_{k,n}(1) + o_{P'}(1)][1 + o_{P'}(1)][1 + o(1)] = B'_{k,n}(1) + o_{P'}(1)$. □

The next result assumes Condition B stated in the main paper.

Proposition 3.1. *Let* $X_1, X_2, \ldots$ *be a given sequence of iid random variables with distribution function* F*, and suppose that* F *satisfies Condition B. Let* $0 \leq \epsilon \leq 1/2$*. Then, for the sequence of Brownian bridges* B_n *in (1.1), we have*

$$\sup_{1/(n+1)\leq t\leq n/(n+1)} n^{\epsilon}\{t(1-t)\}^{\epsilon-1/2}|n^{1/2}f\{Q(t)\}\big(Q(t) - X_{[nt],n}\big) - B_n(t)| = O_p(1) \tag{3.8}$$

as $n \to \infty$. *[Note: de Haan and Ferreira (2006a) use the notation* $[x]$ *for* $\lfloor x \rfloor$.*]*

Proof. The fact that there exists a sequence of Brownian bridges B_n that satisfy (3.8) is the same as Theorem 6.2.1 in Csörgő and Horváth (1993). To show that the B_n can be chosen to be the same as the one in (1.1) follows from the proof of this theorem given in Csörgő and Horváth (1993) on page 381. The proof therein uses Theorem 4.2.1 which is the same as (1.1). □

Let $\tilde{W}_{k,n}(s) = (n/k)^{1/2}B_n(1 - ks/n)$, where B_n is the sequence of Brownian bridges that appear in (1.1). Also, let $C_n = \{s : n/(n+1) < ks < (k+1)\}$. The first part of the following lemma is the same as Lemma 2.4.10 of de Haan and Ferreira (2006a), and the second part follows from the first.

Proposition 3.2. *Let $\epsilon > 0$ and $\gamma \in \mathbb{R}$ be given. Let V be a uniform(0,1) random variable, $\{V_1, V_2, \ldots\}$ be a given sequence of iid uniform (0,1) random variables, $Y = V^{-1}$, and $Y_i = V_i^{-1}$. Then we have the following:*

$$\sup_{s\in C_n} s^{\gamma+(1/2)+\epsilon}\left|k^{1/2}\frac{\{(k/n)Y_{n-\lfloor ks\rfloor,n)}\}^{\gamma}-1}{\gamma}-k^{1/2}\frac{s^{-\gamma}-1}{\gamma}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1) \qquad (3.9)$$

$$\sup_{s\in C_n} s^{\gamma+(1/2)+\epsilon}\left|k^{1/2}\frac{\{(k/(nV_{[ks],n})\}^{\gamma}-1}{\gamma}-k^{1/2}\frac{s^{-\gamma}-1}{\gamma}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1). \qquad (3.10)$$

Proof. The proof of (3.9) appears on page 53 of de Haan and Ferreira (2006a) in their initial part of the proof of Lemma 2.4.10 therein, and it is based on (3.8). To prove (3.10), let

$$\Delta_{k,n}(s,\gamma) = k^{1/2}\left(\left[\{k/(nV_{\lfloor ks\rfloor+1,n})\}^{\gamma}-1\right]\gamma^{-1}-\left[\{k/(nV_{[ks],n})\}^{\gamma}-1\right]\gamma^{-1}\right) \quad (0 < s \leq 1). \quad (3.11)$$

Then $\Delta_{k,n}(s,\gamma) = 0$ if $s \notin \{1/k, 2/k, \ldots\}$. Let

$$A_{k,n}(s,\gamma) = s^{\gamma+(1/2)+\epsilon}\left[k^{1/2}\frac{\{(k/(nV_{[ks],n})\}^{\gamma}-1}{\gamma}-k^{1/2}\frac{s^{-\gamma}-1}{\gamma}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma+1}}\right].$$

Now, substitute $Y_{n-\lfloor ks\rfloor,n} = \{V_{\lfloor ks\rfloor+1,n}\}^{-1}$ in (3.9), and add and subtract $k^{1/2}\{\{(k/(nV_{[ks],n})\}^{\gamma}-1\}/\gamma$ to the expression within the absolute sign in (3.9) to obtain $\sup_{s\in C_n}|A_{k,n}(s,\gamma)+\Delta_{k,n}(s,\gamma)| = o_{P'}(1)$. Since $\Delta_{k,n}(s,\gamma) = 0$ for $i/k < s < (i+1)/k$ $(i = 1, \ldots, k)$, it follows from the continuity properties of $A_{k,n}$ that $\sup_{i/k<s\leq(i+1)/k}|A_{k,n}(s,\gamma)| = \sup_{i/k<s<(i+1)/k}|A_{k,n}(s,\gamma)+\Delta_{k,n}(s,\gamma)| = o_{P'}(1)$. By a similar argument, $\sup_{\{n/(n+1)\}/k\leq s\leq 1/k}|A_{k,n}(s,\gamma)| = o_{P'}(1)$. Therefore, $\sup_{s\in C_n}|A_{k,n}(s,\gamma)| \leq \sup_{s\in C_n}|A_{k,n}(s,\gamma)+\Delta_{k,n}(s,\gamma)| = o_{P'}(1)$, which is (3.10). ☐

Remark: It can also be verified, using (3.10) that, for $0 < \lambda \leq 1$

$$\sup_{\lambda<ks<k+1} s^{\gamma+(1/2)+\epsilon}\left|k^{1/2}\frac{\{(k/(nV_{[ks],n})\}^{\gamma}-1}{\gamma}-k^{1/2}\frac{s^{-\gamma}-1}{\gamma}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1). \qquad (3.12)$$

Let $\gamma_- = \min\{0,\gamma\}$. As in the rest of this paper, the function $(s^{-\gamma}-1)/\gamma$ is interpreted as $-\log s$ when $\gamma = 0$.

Lemma 3.4. *Suppose that the conditions of Theorem 3.2 are satisfied. Then, for $\epsilon > 0$, we have*

$$\sup_{s\in C_n} s^{\gamma_-+1/2+\epsilon}\left|k^{1/2}\frac{\log X_{n-[ks]+1,n}-\log U(\frac{n}{k})}{q_0(\frac{n}{k})}-k^{1/2}\frac{s^{-\gamma}-1}{\gamma}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1) \qquad (3.13)$$

$$\sup_{s\in C_n} s^{\gamma_-+1/2+\epsilon}\left|k^{1/2}\frac{\log X^*_{n-[ks]+1,n}-\log U(\frac{n}{k})}{q_0(\frac{n}{k})}-k^{1/2}\frac{s^{-\gamma}-1}{\gamma}-\frac{\tilde{W}_{k,n}(s)+B'_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1) \quad (3.14)$$

$$\sup_{s\in C_n} s^{\gamma_-+1/2+\epsilon}\left|k^{1/2}\frac{\log X^*_{n-[ks]+1,n}-\log X_{n-[ks]+1,n}}{q_0(\frac{n}{k})}-\frac{B'_{k,n}(s)}{s^{\gamma_-+1}}\right| = o_{P'}(1). \qquad (3.15)$$

Proof. Proof of (3.13): Let $\delta > 0$. Recall $\mathcal{V}_s = k/(nV_{[ks],n})$ and $\mathcal{V}_s^* = k/(nV_{[ks],n}^*)$; $\mathcal{V}_s$ and $\mathcal{V}_s^*$ also depend on (k, n) but not shown explicitly. We apply inequality at the top of page 104 in de Haan and Ferreira (2006a) with $t := t_n = n/k$ and $x := x_n(s) = k/(nV_{[ks],n})$ $(s \in C_n)$. Since $\sup_{s \in C_n} V_{[ks],n} \le V_{k+1,n} = o_{P'}(1)$, $tx \ge \inf_{s \in C_n} t_n x_n(s) \ge \{V_{k+1,n}\}^{-1} \to \infty$, in probability$[P']$. Hence, for any fixed t_0, $t_n x_n(s) > t_0$ with probability approaching 1, uniformly on $s \in C_n$. For the rest of the proof of this theorem, the inequalities are implicitly assumed to hold uniformly over $s \in C_n$, with probability approaching 1 as $n \to \infty$, without further comment. Let $A_2(s) = (\mathcal{V}_s^{\gamma_-} - 1)/\gamma_-$, $A_3(s) = Q_0(n/k)\, \Psi_{\gamma_-, \rho'}(\mathcal{V}_s)$, and $A_4(s) = \varepsilon\, |Q_0(n/k)|\, \mathcal{V}_s^{\gamma_- + \rho'} \max\left\{\mathcal{V}_s^{\delta}, \mathcal{V}_s^{-\delta}\right\}$. Applying the inequality at the top of page 104 in de Haan and Ferreira (2006a), with the foregoing choices for t and x, and then multiplying the result by $k^{1/2} s^{\gamma_- + (1/2) + \epsilon}$, we obtain,

$$s^{\gamma_- + 1/2 + \epsilon} k^{1/2} \left| \frac{\log X_{n-[ks]+1,n} - \log U(\frac{n}{k})}{q_0(\frac{n}{k})} - A_2(s) - A_3(s) \right| \le s^{\gamma_- + 1/2 + \epsilon} k^{1/2} |A_4(s)| \tag{3.16}$$

Let $D_{k,n}(s, \gamma) = \left[k^{1/2}\{s^{-\gamma} - 1\}/\gamma + \tilde{W}_{k,n}(s)/s^{\gamma+1}\right]$. Since (3.10) holds or any $\gamma \in \mathbb{R}$, we have $\sup_{s \in C_n} s^{\gamma_- + (1/2) + \epsilon} \left| k^{1/2} A_2(s) - D_{k,n}(s, \gamma_-) \right| = o_{P'}(1)$. Using the form of $\Psi_{\gamma,\rho}$ in Corollary 2.3.5 of de Haan and Ferreira (2006a) on page 46, it may be verified that $\sup_{s \in C_n} s^{\gamma_- + (1/2) + \varepsilon} k^{1/2} |A_3(s)| = o_{P'}(1)$. Using(3.5) and Assumption (A.4), it may be verified that $\sup_{s \in C_n} s^{\gamma_- + (1/2) + \varepsilon} k^{1/2} |A_4(s)| = o_{P'}(1)$. The result follows by substituting these in (3.16). The proof of the next part is similar to that of the previous one, and hence omitted. □

For $\gamma \ge -(1/2)$, let $b_0(n) = b_0^*(n) = \log U(n/k)$; for $\gamma < -(1/2)$, let

$$b_0(n) = \log U(1/V_{1,n}) + \gamma_-^{-1} q_0(n/k), \;\; b_0^*(n) = \log U(1/V_{1,n}^*) + \gamma_-^{-1} q_0(n/k).$$

The statement of the two results in the next lemma, and their proofs are similar to those of the previous Lemma 3.4. Therefore, the proof is omitted.

Lemma 3.5. *Suppose that conditions of Theorem 3.2 are satisfied. Then, for $\epsilon > 0$, we have*

$$\sup_{0 < ks < k+1} s^{\gamma_- + 1/2 + \epsilon} \left| k^{1/2} \frac{\log X_{n-[ks]+1,n} - b_0(n)}{q_0(\frac{n}{k})} - k^{1/2} \frac{s^{-\gamma_-} - 1}{\gamma_-} - \frac{\tilde{W}_{k,n}(s)}{s^{\gamma_- + 1}} \right| = o_{P'}(1), \tag{3.17}$$

$$\sup_{0 < ks < k+1} s^{\gamma_- + 1/2 + \epsilon} \left| k^{1/2} \frac{\log X_{n-[ks],n}^* - b_0^*(n)}{q_0(\frac{n}{k})} + k^{1/2} \frac{s^{-\gamma_-} - 1}{\gamma_-} - \frac{\tilde{W}_{k,n}(s) + B_{k,n}'(s)}{s^{\gamma_- + 1}} \right| = o_{P'}(1). \tag{3.18}$$

Lemma 3.6. *Suppose that the conditions of Theorem 3.2 are satisfied. Then, for $\epsilon > 0$, we have*

$$k^{1/2}\frac{\log X_{n-[ks]+1,n}-\log X_{n-k,n}}{q_0(\frac{n}{k})}-k^{1/2}\frac{s^{-\gamma_-}-1}{\gamma_-}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma_-+1}}+\tilde{W}_{k,n}(1) = (1+s^{-(\gamma_-+1/2+\epsilon)})o_{P'}(1), \tag{3.19}$$

$$k^{1/2}\frac{\log X^*_{n-[ks]+1,n}-\log X^*_{n-k,n}}{q_0(\frac{n}{k})}-k^{1/2}\frac{s^{-\gamma_-}-1}{\gamma_-}-\frac{\tilde{W}_{k,n}(s)+B'_{k,n}(s)}{s^{\gamma_-+1}} + \tilde{W}_{k,n}(1)+B'_{k,n}(1) = (1+s^{-(\gamma_-+1/2+\epsilon)})o_{P'}(1), \tag{3.20}$$

$$k^{1/2}\frac{\log X^*_{n-[ks]+1,n}-\log X^*_{n-k,n}-(\log X_{n-[ks]+1,n}-\log X_{n-k,n})}{q_0(\frac{n}{k})}-\frac{B'_{k,n}(s)}{s^{\gamma_-+1}} +B'_{k,n}(1) = (1+s^{-(\gamma_-+1/2+\epsilon)})o_{P'}(1) \tag{3.21}$$

where $o_{P'}(1)$ in (3.19), (3.20), and (3.21) do not depend on s ($0 < ks < k+1$).

Proof. It follows from (3.17) that

$$\sup_{k<ks<k+1} s^{\gamma_-+1/2+\epsilon}\left|k^{1/2}\frac{\log X_{n-k,n}-b_0(n)}{q_0(\frac{n}{k})}-k^{1/2}\frac{s^{-\gamma_-}-1}{\gamma_-}-\frac{\tilde{W}_{k,n}(s)}{s^{\gamma_-+1}}\right| = o_{P'}(1), \tag{3.22}$$

By a continuity argument similar to that for (3.2), (3.22) also holds with $\sup_{k<ks<k+1}$ replaced by $\sup_{k\le ks<k+1}$ and hence at $s=1$, from which we obtain

$$k^{1/2}b_0(n)/q_0(\frac{n}{k}) = k^{1/2}\log X_{n-k,n}/q_0(\frac{n}{k})-\tilde{W}_{k,n}(1)+o_{P'}(1).$$

Substituting this in (3.17) leads to (3.19). By a similar argument applied to (3.18), we have

$$k^{1/2}\{\log X^*_{n-k,n}-b^*_0(n)\}/q_0(\frac{n}{k})-\{\tilde{W}_{k,n}(1)+B'_{k,n}(1)\} = o_{P'}(1). \tag{3.23}$$

Substitution of this in (3.18) leads to (3.20). The third part follows from the first two. □

The asymptotic distributions of various tail statistics turn out to be functionals of a Wiener process. Let $W(s)$, $0 \le s \le 1$, denote a standard Wiener process, and let us define the following random variables that are functionals of the same W:

$$Q = 2\int_0^1 \frac{s^{-\gamma}-1}{\gamma_-}\left(\frac{W(s)}{s^{\gamma_-+1}}-W(1)\right)ds, \qquad P = \int_0^1 \frac{W(s)}{s^{\gamma_-+1}}-W(1)ds, \tag{3.24}$$

$$R = (1-\gamma_-)^2(1-2\gamma_-)\left(2^{-1}(1-2\gamma_-)Q-2P\right), \qquad \Gamma = \gamma_+P+R \tag{3.25}$$

$$A = \gamma_+W(1)+(1-\gamma_-)(3-4\gamma_-)P-2^{-1}(1-\gamma_-)(1-2\gamma_-)^2Q \tag{3.26}$$

These random variables will appear in the limiting distributions of various statistics. In finite samples, the following random variables that correspond these will appear: Let $(P'_n, Q'_n, A'_n, \Gamma'_n)$

be defined in the same way except that the standard Wiener process W is replaced by $B'_{k,n}$. Similarly, let $(\tilde{P}_n, \tilde{Q}_n, \tilde{A}_n, \tilde{W}_n)$ be defined in the same way except that W is replaced by $\tilde{W}_{k,n}$, where $\tilde{W}_{k,n}(s) = (n/k)^{1/2} B_n(1 - ks/n)$.

Since $\tilde{W}_{k,n}(s)$ and $B'_n(ks/n)$ are equal in distribution, it follows that $\tilde{P}_n \overset{d}{=} P'_n$, but they are not the same random variables. Similar comments apply to the other pairs of random variables of the form $(\tilde{Y}_{k,n}, Y'_{k,n})$, with different symbols for Y, that will appear in the derivations.

Lemma 3.7.

$$k^{1/2}\left\{H_n/q_0(n/k) - (1-\gamma_-)^{-1}\right\} = \tilde{P}_n + o_{P'}(1), \tag{3.27}$$

$$k^{1/2}\left\{H^*_n/q_0(n/k) - (1-\gamma_-)^{-1}\right\} = (P'_n + \tilde{P}_n) + o_{P'}(1), \tag{3.28}$$

$$k^{1/2}\left\{(H^*_n - H_n)/q_0(n/k)\right\} = P'_n + o_{P'}(1). \tag{3.29}$$

Proof. The proof of (3.27) follows from (3.19); (3.28) follows by similar arguments from (3.20). Finally, (3.29) follows from the previous two parts. □

Lemma 3.8.

$$k^{1/2}(H_n - \gamma_+) = \gamma_+\tilde{P}_n + o_{P'}(1), \qquad k^{1/2}(H^*_n - H_n) = \gamma_+ P'_n + o_{P'}(1). \tag{3.30}$$

Proof. A proof of the first part of (3.30) is given on page 109 in de Haan and Ferreira (2006a). To prove the second part, it follows from (3.29) that $k^{1/2}(H^*_n - H_n) = q_0(n/k)[P'_n + o_{P'}(1)]$. Since $\{q_0(t) - \gamma_+\}/Q(t) = O(1)$ as $t \to \infty$ [see de Haan and Ferreira 2006a, page 109], it follows from $k^{1/2}Q(t) = o(1)$ (see Assumption (A.4)) that $\{q_0(n/k) - \gamma_+\} = \big[\{q_0(n/k) - \gamma_+\}/Q(n/k)\big]Q(n/k) = O(1)O(k^{-1/2}) = o(1)$. Now, the proof follows easily. □

Lemma 3.9.

$$k^{1/2}\left(\left(\frac{H_n}{q_0(n/k)}\right)^2 - \left(\frac{1}{1-\gamma_-}\right)^2\right) = \frac{2}{1-\gamma_-}\tilde{P}_n + o_{P'}(1), \tag{3.31}$$

$$k^{1/2}\left(\frac{(H^*_n)^2 - H_n^2}{q_0^2(n/k)}\right) = \frac{2}{1-\gamma_-}P'_n + o_{P'}(1). \tag{3.32}$$

Proof. Proof of (3.31) follows from (3.27) by an application of mean value theorem; similarly, (3.29) follows from (3.29). By similar arguments, we may also show the following, but it is not directly required:

$$k^{1/2}\left(\left(\frac{H^*_n}{q_0(n/k)}\right)^2 - \left(\frac{1}{1-\gamma_-}\right)^2\right) = \frac{2}{1-\gamma_-}(\tilde{P}_n + P'_n) + o_{P'}(1), \tag{3.33}$$

□

Let

$$\lambda_1 = \frac{1}{1-\gamma_-}, \qquad \lambda_2 = \frac{2}{(1-\gamma_-)(1-2\gamma_-)}. \tag{3.34}$$

Lemma 3.10.

$$k^{1/2}\left(\{q_0(n/k\}^{-2} M_n - \lambda_2\right) - \tilde{Q}_n = o_{P'}(1), \tag{3.35}$$

$$k^{1/2}\left(\{q_0(n/k\}^{-2} M_n^* - \lambda_2\right) - (\tilde{Q}_n + Q_n') = o_{P'}(1), \tag{3.36}$$

$$k^{1/2}\{q_0(n/k\}^{-2}(M_n^* - M_n) - Q_n' = o_{P'}(1). \tag{3.37}$$

Proof. Let $0 < \epsilon < 1/4$. By (3.19), for $0 < s < 1$,

$$\begin{aligned} &\frac{\log X_{n-[ks]+1,n} - \log X_{n-k,n}}{q_o(n/k)} \\ = &\frac{s^{-\gamma_-}-1}{\gamma_-} + k^{-1/2}\left(\frac{\tilde{W}_{k,n}(s)}{s^{\gamma_-+1}} - \tilde{W}_{k,n}(1)\right) + k^{-1/2}(1+s^{-1/2-\epsilon-\gamma_-})o_{P'}(1) \\ = &f_{1n}(s) + f_{2n}(s) + f_{3n}(s), \quad \text{say.} \end{aligned} \tag{3.38}$$

Then

$$k^{1/2}\frac{M_n}{(q_o(n/k))^2} = \int_0^1 k^{1/2}\{f_{1n}(s) + f_{2n}(s) + f_{3n}(s)\}^2\, ds. \tag{3.39}$$

Note that if $2\{(1/2)+\epsilon+\gamma_-\} \geq 1$ then $\int_0^1 f_{3n}^2(s)\, ds$ is not finite. Therefore, though the integral in (3.39) is finite, it is not possible to evaluate it by expanding the squared expression under the integral sign in (3.39) and evaluating the integral of each resulting term. Therefore, we express $\int_0^1$ as $\int_0^{1/k} + \int_{1/k}^1$, and evaluate the two integrals separately. This is equivalent to evaluating the first term of M_n and the rest separately. Let $k^{1/2}M_n/q_0^2(n/k) = A_1(k,s) + A_2(k,n)$ where

$$A_1(k,n) = \left[k^{-1/4}\{\log X_{n,n} - \log X_{n-k,n}\}/q_0(n/k)\right]^2 \tag{3.40}$$

$$A_2(k,n) = \sum_{i=1}^{k-1}\left[k^{-1/4}\{\log X_{n-i,n} - \log X_{n-k,n}\}/q_0(n/k)\right]^2. \tag{3.41}$$

It follows from (3.38) with $s = 1/k$ that

$$A_1(k,n) = \left[k^{-1/4}f_{1n}(k^{-1}) + k^{-1/4}f_{2n}(k^{-1}) + k^{-1/4}f_{3n}(k^{-1})\right]^2 = o_{P'}(1).$$

Next, note that $A_2(k,n) = \int_0^{1/k} k^{1/2}\{f_{1n}(s)+f_{2n}(s)+f_{3n}(s)\}^2\, ds = o_{P'}(1)$. Expanding the integrand and evaluating each resulting integral, it may be verified that

$$k^{1/2}\frac{M_n}{(q_o(n/k))^2} = \frac{2k^{1/2}}{(1-\gamma_-)(1-2\gamma_-)} + 2\int_0^1 \frac{s^{-\gamma}-1}{\gamma_-}\left(\frac{\tilde{W}_{k,n}(s)}{s^{\gamma_-+1}} - \tilde{W}_{k,n}(1)\right) ds + o_{P'}(1).$$

This completes the proof of (3.35). The proof of (3.36) is similar, and hence omitted. The proof of (3.37) follows by subtracting (3.35) from (3.36). □

Lemma 3.11.

$$k^{1/2}\left(\frac{H_n^2}{M_n}-\frac{\lambda_1^2}{\lambda_2}\right)=(1-2\gamma_-)\tilde{P}_n-\frac{(1-2\gamma_-)^2}{4}\tilde{Q}_n+o_{P'}(1). \tag{3.42}$$

$$k^{1/2}\left(\frac{(H_n^*)^2}{M_n^*}-\frac{H_n^2}{M_n}\right)=(1-2\gamma_-)P_n'-\frac{(1-2\gamma_-)^2}{4}Q_n'+o_{P'}(1). \tag{3.43}$$

Proof. Consider the left hand side of (3.42):

$$k^{1/2}\left(\frac{H_n^2}{M_n}-\frac{\lambda_1^2}{\lambda_2}\right)=\frac{\lambda_1^2}{\lambda_2}\frac{q_0^2(n/k)}{M_n}k^{1/2}\left(\lambda_2-\frac{M_n}{q_0^2(n/k)}\right)+\frac{q_0^2(n/k)}{M_n}k^{1/2}\left(\frac{H_n^2}{q_0^2(n/k)}-\lambda_1^2\right) \tag{3.44}$$

The first part follows from (3.35) and (3.31). The proof of the second part is similar. □

Lemma 3.12.

$$k^{1/2}(\hat{\gamma}_--\gamma_-)=\tilde{R}_n+o_{P'}(1),\qquad k^{1/2}(\hat{\gamma}_-^*-\hat{\gamma}_-)=R_n'+o_{P'}(1),\quad k^{1/2}(\hat{\gamma}^*-\hat{\gamma})=\Gamma_n'+o_{P'}(1). \tag{3.45}$$

Proof. It follows from (3.34) that $\gamma_-=1-2^{-1}\{1-\lambda_1^2/\lambda_2\}$. Then, by the definition of $\hat{\gamma}_-$,

$$k^{1/2}(\hat{\gamma}_--\gamma_-)\;=\;2^{-1}k^{1/2}\left(\lambda_1^2/\lambda_2-H_n^2/M_n\right)\left(1-\lambda_1^2/\lambda_2\right)^{-1}\left(1-H_n^2/M_n\right)^{-1}. \tag{3.46}$$

Now, the first part of the lemma may be verified using (3.42). The proof of the second part is similar and hence omitted. Using the definitions of $\hat{\gamma}^*$ and $\hat{\gamma}$, and Lemmas 3.12 and 3.8, we obtain

$$k^{1/2}(\hat{\gamma}^*-\hat{\gamma})\;=\;k^{1/2}(\hat{\gamma}_-^*-\hat{\gamma}_-)+k^{1/2}(\hat{\gamma}_+^*-\hat{\gamma}_+)=R_n'+\gamma_+P_n'+o_{P'}(1)=\Gamma_n'+o_{P'}(1).\blacksquare$$

Lemma 3.13.

$$k^{1/2}\left\{\hat{a}^*(n/k)/\hat{a}(n/k)-1\right\}=A_n'+o_{P'}(1). \tag{3.47}$$

Proof. Using the definitions, we have

$$k^{1/2}\left(\frac{\hat{a}^*(n/k)}{\hat{a}(n/k)}-1\right)\;=\;k^{1/2}\left(\frac{X_{n-k,n}^*}{X_{n-k,n}}\,\frac{H_n^*}{H_n}\,\frac{(1-\hat{\gamma}_-^*)}{(1-\hat{\gamma}_-)}-1\right). \tag{3.48}$$

To prove the lemma, we substitute asymptotic representations for the three terms $X_{n-k,n}^*/X_{n-k,n}$, H_n^*/H_n, and $(1-\hat{\gamma}_-^*)/(1-\hat{\gamma}_-)$ in (3.48). By (3.15) and an application of mean value theorem, we obtain $k^{1/2}\{q_0(n/k)\}^{-1}\log\{X_{n-k,n}^*/X_{n-k,n}\}=B_{k,n}'(1)+o_{P'}(1)$. Since $q_0(n/k)=\gamma_++o(1)$, it may be verified that $X_{n-k,n}^*/X_{n-k,n}=1+k^{-1/2}\{\gamma_+B_{k,n}'(1)+o_{P'}(1)\}$. Hence,

$$\frac{X_{n-k,n}^*}{X_{n-k,n}}=1+k^{-1/2}\{\gamma_+B_{k,n}'(1)+o_{P'}(1)\}. \tag{3.49}$$

We adopt a similar procedure for (H_n^*/H_n). First, we write

$$k^{1/2}\left(\frac{H_n^*}{H_n}-1\right)=(1-\gamma_-)k^{1/2}\frac{H_n^*-H_n}{q_0(n/k)}\frac{q_0(n/k)}{(1-\gamma_-)H_n}. \tag{3.50}$$

Then, it may be verified, using Lemma 3.7, that $(H_n^*/H_n)=1+k^{-1/2}\{(1-\gamma_-)P_n'+o_{P'}(1)\}$. Using Lemma 3.12, we may verify that $(1-\hat{\gamma}_-^*)/(1-\hat{\gamma}_-)=1+k^{-1/2}(-R_n'(1-\gamma_-)^{-1}+o_{P'}(1))$. Now the lemma follows by substituting the foregoing asymptotic representations in (3.48). □

3.2 Estimation of high quantile

In this subsection we establish the main results for constructing a confidence interval for a high quantile. Let p_n be a given small number in the range $0 < p_n < 1$, and let $x(p_n) = F^{\leftarrow}(1-p_n)$ denote the unknown upper p_nth quantile of F. Recall that $d_n = k/(np_n)$. An estimator of $x(p_n)$ is (see, de Haan and Ferreira 2006a, Theorem 4.3.1)

$$\hat{x}(p_n) = \hat{b}\,(n/k) + \hat{a}\,(n/k)\,\{d_n^{\hat{\gamma}} - 1\}/\hat{\gamma};$$

the bootstrap counterpart of $\hat{x}(p_n)$ is $\hat{x}^*(p_n) = \hat{b}^*\,(n/k) + \hat{a}^*\,(n/k)\,\{d_n^{\hat{\gamma}^*} - 1\}/\hat{\gamma}^*$. Let $q_\gamma(t) = \int_1^t s^{\gamma-1}\log s\; ds$, $(t > 0)$. The asymptotic distribution of the bootstrap quantity $\hat{x}^*(p_n)$ is provided in the next lemma; it can be used for constructing a confidence interval for the high-quantile $x(p_n)$.

Lemma 3.14. *Suppose that the conditions of Theorem 3.4 are satisfied. Then,*

$$\frac{k^{1/2}}{q_{\hat{\gamma}}(d_n)}\frac{\hat{x}^*(p_n) - \hat{x}(p_n)}{\hat{a}\,(n/k)} = \Gamma'_n + (\gamma_-)^2 B'_{k,n}(1) - (\gamma_-)A'_n + o_{P'}(1). \tag{3.51}$$

$$\frac{k^{1/2}}{q_{\hat{\gamma}^*}(d_n)}\frac{\hat{x}^*(p_n) - \hat{x}(p_n)}{\hat{a}\,(n/k)} = \Gamma'_n + (\gamma_-)^2 B'_{k,n}(1) - (\gamma_-)A'_n + o_{P'}(1). \tag{3.52}$$

Proof. To simplify notation, we write $\hat{x}_n$ and $\hat{x}_n^*$ for $\hat{x}^*(p_n)$ and $\hat{x}(p_n)$ respectively; this simplified notation is applicable only to this proof. Starting from the definitions of these quantities, we have

$$\begin{aligned} k^{1/2}\frac{\hat{x}_n^* - \hat{x}_n}{q_{\hat{\gamma}}(d_n)\hat{a}\,(n/k)} &= k^{1/2}\frac{\hat{b}^*\,(n/k) - \hat{b}\,(n/k)}{q_{\hat{\gamma}}(d_n)\hat{a}\,(n/k)} + k^{1/2}\left(\frac{\hat{a}^*\,(n/k)}{\hat{a}\,(n/k)} - 1\right)\frac{d_n^{\hat{\gamma}} - 1}{\hat{\gamma}q_{\hat{\gamma}}(d_n)} \\ &+ \frac{k^{1/2}}{q_{\hat{\gamma}}(d_n)}\frac{\hat{a}^*\,(n/k)}{\hat{a}\,(n/k)}\left\{(d_n^{\hat{\gamma}^*} - 1)/\hat{\gamma}^* - (d_n^{\hat{\gamma}} - 1)/\hat{\gamma}\right\} := A1 + A2 + A3, \end{aligned} \tag{3.53}$$

say. The technical details when $\gamma_- = 0$ and when $\gamma_- < 0$ are different, as mentioned earlier. For example, $(x^\gamma - 1)/\gamma$ needs to be interpreted as $\log x$ at $\gamma = 0$. In this proof, we provide the details for the case $\gamma_- = 0$.

It follows from Remark 4.3.3 and (4.3.5) in de Haan and Ferreira (2006a), that

$$q_\gamma(t) \sim \begin{cases} \gamma^{-1}t^\gamma \log t \\ (\log t)^2/2, \\ \gamma^{-2}, \end{cases} \qquad \frac{t^\gamma - 1}{\gamma} \sim \begin{cases} \gamma^{-1}t^\gamma \\ 2\log t, \\ -\gamma^{-1}, \end{cases} \qquad \frac{t^\gamma - 1}{\gamma q_\gamma(t)} \sim \begin{cases} (\log t)^{-1} & \gamma > 0 \\ 2(\log t)^{-1}, & \gamma = 0 \\ -\gamma, & \gamma < 0 \end{cases} \tag{3.54}$$

as $t \to \infty$, Then,

$$\frac{d_n^{\hat{\gamma}} - 1}{\hat{\gamma}q_{\hat{\gamma}}(d_n)} \sim -\hat{\gamma}_- = -\gamma_- + o_{P'}(1), \qquad \frac{1}{q_{\hat{\gamma}}(d_n)} \sim (\hat{\gamma}_-)^2 = (\gamma_-)^2 + o_{P'}(1).$$

Substituting these into A1 and A2 in (3.53), we obtain $A1 = (\gamma_-)^2 B'_{k,n}(1) + o_{P'}(1)$ by Lemma 3.3, and $A2 = -(\gamma_-)A'_n + o_{P'}(1)$ by Lemma 3.13.

To evaluate $A3$, first notice that $(d_n^\gamma - 1)\gamma^{-1} = \int_1^{d_n} s^{\gamma-1} ds$. Since $k^{1/2}\hat{a}^*(n/k)/\hat{a}(n/k) = 1 + O_{P'}(1)$ by Lemma 3.13, we have

$$A3 = \frac{k^{1/2}}{q_{\hat{\gamma}}(d_n)} \frac{\hat{a}^*\left(\frac{n}{k}\right)}{\hat{a}\left(\frac{n}{k}\right)} \left(\frac{d_n^{\hat{\gamma}^*} - 1}{\hat{\gamma}^*} - \frac{d_n^{\hat{\gamma}} - 1}{\hat{\gamma}} \right) = \frac{k^{1/2}}{q_{\hat{\gamma}}(d_n)} \int_1^{d_n} s^{\hat{\gamma}-1}(s^{\hat{\gamma}^*-\hat{\gamma}} - 1) ds \ + o_{P'}(1).$$

Let us ignore the $o_{P'}(1)$ for now and consider the other term. By the mean value theorem, $\log s^{\hat{\gamma}^*-\hat{\gamma}} - \log 1 = \nu(s)^{-1}\{s^{\hat{\gamma}^*-\hat{\gamma}} - 1\}$, where $\nu(s) = s_1^{\hat{\gamma}^*-\hat{\gamma}}$ for some $s_1 \in [1, s]$. Therefore,

$$k^{1/2}\{q_{\hat{\gamma}}(d_n)\}^{-1} \int_1^{d_n} s^{\hat{\gamma}-1}(s^{\hat{\gamma}^*-\hat{\gamma}} - 1) ds = k^{1/2}(\hat{\gamma}^* - \hat{\gamma}) + B_2$$

where $B_2 = k^{1/2}\{q_{\hat{\gamma}}(d_n)\}^{-1} \int_1^{d_n} s^{\hat{\gamma}-1} \log s^{\hat{\gamma}^*-\hat{\gamma}}(s_1^{\hat{\gamma}^*-\hat{\gamma}} - 1) ds$.

Since $k^{-1/2} \log np_n = o(1)$ by assumption, we have

$$|\log d_n^{\hat{\gamma}^*-\hat{\gamma}}| = |(\hat{\gamma}^* - \hat{\gamma}) \log d_n| = (\hat{\gamma}^* - \hat{\gamma})[\log k - \log np_n] = o_{P'}(1). \qquad (3.55)$$

Choose n_0 large enough such that $d_n > 1$ for $n \geq n_0$. Let $n > n_0$. Then

$$\begin{aligned} |B_2| &\leq |d_n^{\hat{\gamma}^*-\hat{\gamma}} - 1| \frac{k^{1/2}}{q_{\hat{\gamma}}(d_n)} \int_1^{d_n} |s^{\hat{\gamma}-1} \log s^{\hat{\gamma}^*-\hat{\gamma}}| ds = |d_n^{\hat{\gamma}^*-\hat{\gamma}} - 1| O_{P'}(1) \\ |\log d_n^{\hat{\gamma}^*-\hat{\gamma}}| &= |(\hat{\gamma}^* - \hat{\gamma}) \log d_n| = (\hat{\gamma}^* - \hat{\gamma})[\log k - \log np_n] = o_{P'}(1) \end{aligned}$$

since $k^{-1/2} \log np_n = o(1)$ by assumption. Therefore, $B_2 = o_{P'}(1)$. Then, $A_3 = k^{1/2}(\hat{\gamma}^* - \hat{\gamma}) + o_{P'}(1) = \Gamma_n + o_{P'}(1)$ by (3.45).

Since $q_{\hat{\gamma}^*}(d_n)/q_{\hat{\gamma}}(d_n) = 1 + o_{P'}(1)$ (see, de Haan and Ferreira 2006a, Corollary 4.3.2, page135), and $\hat{a}^*(n/k)\{\hat{a}(n/k)\}^{-1} = 1 + o_{P'}(1)$ by Lemma 3.13, using (3.53), we have

$$k^{1/2} \frac{\hat{x}_n^* - \hat{x}_n}{q_{\hat{\gamma}}(d_n)\hat{a}(n/k)} = A1 + A2 + A3 = \Gamma'_n + (\gamma_-)^2 B'_{k,n}(1) - (\gamma_-)A'_n + o_{P'}(1).$$

□

3.3 Estimation of tail probability

In this subsection, we prove the main results for constructing confidence intervals for an unknown tail probability. Let x_n be a given sequence of real numbers. The results in this subsection are for constructing a bootstrap confidence interval for the unknown tail probability $p_n := pr(X \geq x_n)$. In the proof of Lemma 3.14, we used the two quantities

$$\hat{x}_{p_n} = \hat{b}(n/k) + \hat{a}(n/k)\{d_n^{\hat{\gamma}} - 1\}/\hat{\gamma}, \quad \hat{x}^*_{p_n} = \hat{b}^*(n/k) + \hat{a}^*(n/k)\{d_n^{\hat{\gamma}^*} - 1\}/\hat{\gamma}^*.$$

In the technical arguments provided below, we use these two quantities and results about their properties obtained in the proof of Lemma 3.14. One difference worthy of noting is that, in the context of this subsection, $\hat{x}_{p_n}$ is an estimator of the known quantity x_n. Therefore, while we use $\hat{x}_{p_n}$ in the following technical details, it is used only to facilitate using the previously obtained results in the current proofs, not as an empirically useful statistic.

Recall that

$$\hat{p}_n = \frac{k}{n}\left(\max\left\{0, 1+\hat{\gamma}\frac{x_n - \hat{b}\left(\frac{n}{k}\right)}{\hat{a}\left(\frac{n}{k}\right)}\right\}\right)^{-1/\hat{\gamma}}. \tag{3.56}$$

and $w_\gamma(t) = t^{-\gamma}\int_1^t s^{\gamma-1}\log s\, ds$, $t>0$.

Lemma 3.15. *Suppose that the conditions of Theorem 3.5 are satisfied. Then,*

$$\frac{k^{1/2}}{w_{\hat{\gamma}}(\hat{d}_n)}\left(\frac{\hat{p}_n^*}{\hat{p}_n} - 1\right) = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1). \tag{3.57}$$

$$\frac{k^{1/2}}{w_{\hat{\gamma}^*}(\hat{d}_n^*)}\left(\frac{\hat{p}_n^*}{\hat{p}_n} - 1\right) = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1). \tag{3.58}$$

Proof. Let

$$Y_n = \left(\frac{\hat{\gamma}}{d_n^{\hat{\gamma}}}\frac{x_n - \hat{x}_n}{\hat{a}\left(\frac{n}{k}\right)}\right), \qquad Y_n^* = \left(\frac{\hat{\gamma}^*}{d_n^{\hat{\gamma}^*}}\frac{x_n - \hat{x}_n^*}{\hat{a}^*\left(\frac{n}{k}\right)}\right).$$

Using the definitions of $\hat{p}_n, \hat{x}_n, \hat{p}_n^*$, and $\hat{x}_n^*$, and simplifying the expressions, we obtain

$$\hat{p}_n = \frac{k}{nd_n}(1+Y_n)^{-1/\hat{\gamma}}, \quad \hat{p}_n^* = \tfrac{k}{nd_n}(1+Y_n^*)^{-1/\hat{\gamma}^*}, \quad \frac{\hat{p}_n^*}{\hat{p}_n} = (1+Y_n^*)^{-1/\hat{\gamma}^*}(1+Y_n)^{1/\hat{\gamma}}. \tag{3.59}$$

By the mean value theorem, $\log(1+Y_n) - \log 1 = \zeta_n^{-1}Y_n$, where ζ_n lies between 1 and $1+Y_n$. Using the previously established results and assumptions about the rates of convergence of d_n, and (3.54), it may be verified that $Y_n = o_{P'}(1)$. Therefore, $\zeta_n = 1 + o_{P'}(1)$ and $\log(1+Y_n) = (1+o_{P'}(1))Y_n$.

Similarly, by the mean value theorem, $\log(1+Y_n^*) - \log 1 = \xi_n^{-1}Y_n$, where ξ_n lies between 1 and Y_n^*. As in the previous proof for $Y_n = o_{P'}(1)$, it may be verified that $Y_n^* = o_{P'}(1)$, $\xi_n = 1 + o_{P'}(1)$, and $\log(1+Y_n^*) = (1+o_{P'}(1))Y_n^*$.

Now,

$$\frac{k^{1/2}}{w_{\hat{\gamma}}(d_n)}\log\frac{\hat{p}_n^*}{\hat{p}_n} = \frac{k^{1/2}}{w_{\hat{\gamma}}(d_n)}\Big[-\{\hat{\gamma}^*\}^{-1}\log(1+Y_n^*) + \{\hat{\gamma}\}^{-1}\log(1+Y_n)\Big]$$

$$= k^{1/2}\frac{\hat{x}_n^* - \hat{x}_n}{q_{\hat{\gamma}}(d_n)\hat{a}\left(\frac{n}{k}\right)}\frac{\hat{a}\left(\frac{n}{k}\right)}{\hat{a}^*\left(\frac{n}{k}\right)}d_n^{\hat{\gamma}-\hat{\gamma}^*}(1+o_{P'}(1)) + k^{1/2}\frac{x_n - \hat{x}_n}{q_{\hat{\gamma}}(d_n)\hat{a}\left(\frac{n}{k}\right)}\left(1 - \frac{\hat{a}\left(\frac{n}{k}\right)}{\hat{a}^*\left(\frac{n}{k}\right)}d_n^{\hat{\gamma}-\hat{\gamma}^*}\right)(1+o_{P'}(1))$$

By (3.55) $|d_n^{\hat{\gamma}^*-\hat{\gamma}} - 1| = o_{P'}(1)$. By (3.47), $\hat{a}^*(n/k)/\hat{a}(n/k) = 1 + O_P'(k^{-1/2}) = 1 + o_{P'}(1)$. By Theorem 3.4, $k^{1/2}q_{\hat{\gamma}}^{-1}(d_n)\hat{a}^{-1}(n/k)(x_n - \hat{x}_n) = O_{P'}(1)$. By Lemma 3.15,

$$k^{1/2}\frac{\hat{x}_n^* - \hat{x}_n}{q_{\hat{\gamma}}(d_n)\hat{a}\left(\frac{n}{k}\right)} = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1).$$

Therefore,

$$\begin{aligned}\frac{k^{1/2}}{w_{\hat\gamma}(d_n)}\log\frac{\hat p_n^*}{\hat p_n} &= \left(\Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1)\right)(1+o_{P'}(1)) + o_{P'}(1)\\ &= \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1).\end{aligned} \tag{3.60}$$

It follows from (3.60) that $\log(\hat p_n^*/\hat p_n) = o_{P'}(1)$, and hence $|\hat p_n^*/\hat p_n - 1| = o_{P'}(1)$. Therefore, by the delta method, we have (3.57). Let $\hat d_n = k/(n\hat p_n)$ and $\hat d_n^* = k/(n\hat p_n^*)$. Then (see Corollary 4.4.4, de Haan and Ferreira 2006a) since $\hat\gamma^* - \hat\gamma = O_{P'}(k^{-1/2})$ and $\hat p_n^*/\hat p_n = 1 + o_{P'}(1)$, we have $\{w_{\hat\gamma}(\hat d_n)/w_\gamma(d_n)\}$ and $\{w_{\hat\gamma^*}(\hat d_n^*)/w_{\hat\gamma}(\hat d_n)\}$ converge to 1, in probability. Therefore, we also have

$$\frac{k^{1/2}}{w_{\hat\gamma^*}(\hat d_n^*)}\log\frac{\hat p_n^*}{\hat p_n} = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1). \tag{3.61}$$

□

Lemma 3.16. *Let $\boldsymbol{V} = (V_1, V_2, \ldots)$ and $\boldsymbol{V}' = (V_1', V_2', \ldots)$ be two sequences of random variables as in (1.2), $S_n(\boldsymbol{V}, \boldsymbol{V}')$ be a function of $V_1, \ldots, V_n$ and $V_1', \ldots, V_n'$, and $S_n(\boldsymbol{V}, \boldsymbol{V}') = o_{P'}(1)$. Then $S_n(\boldsymbol{V}, \boldsymbol{V}') = o_{P^*}(1)$, in probability, where P^* denotes the probability conditional on V'.*

Proof. Let $\epsilon > 0$ and $\delta > 0$ be given. Suppose that $S_n(\boldsymbol{V}, \boldsymbol{V}') = o_{P'}(1)$. Let $Z_n(\boldsymbol{V};\delta) = P^*\left[S_n(\boldsymbol{V}, \boldsymbol{V}') \geq \delta|\boldsymbol{V}\right]$, where P^* denotes the conditional probability given $\boldsymbol{V}$. Then $E[Z_n^2(\boldsymbol{V};\delta)] \leq E[Z_n(\boldsymbol{V};\delta)] = P'(S_n(\boldsymbol{V}, \boldsymbol{V}' \geq \delta) \overset{P'}{\to} 0$ as $n \to \infty$. and, by Chebyshev inequality, $P_{\boldsymbol{V}}[Z_n(\boldsymbol{V};\delta) > \epsilon] = o_{P'}(1)$. Therefore, $P_{\boldsymbol{V}}\left(P_{\boldsymbol{V}'|\boldsymbol{V}}\left[|S_n(\boldsymbol{V}, \boldsymbol{V}')| \geq \delta|\boldsymbol{V}\right] > \epsilon\right) \to 0$ as $n \to \infty$, and hence $S_n(\boldsymbol{V}, \boldsymbol{V}') = o_{P^*}(1)$ in probability $[P]$.

□

Proposition 3.3.

$$(P_n', Q_n', A_n', \Gamma_n') \overset{d}{\to} (P, Q, A, \Gamma), \quad (\tilde P_n, \tilde Q_n, \tilde A_n, \tilde\Gamma_n) \overset{d}{\to} (P, Q, A, \Gamma), \text{ as } n \to \infty.$$

Proof. Since$B_{k,n}'(s) = W(s) - (ks/n)W(n/k)$, $(0 \leq s \leq 1)$, in distribution, we have

$$\frac{B_{k,n}'(s)}{s^{\gamma_-+1}} \overset{d}{=} \frac{W(s)}{s^{\gamma_-+1}} - \frac{k}{n}W\left(\frac{n}{k}\right)s^{-\gamma_-} = \frac{W(s)}{s^{\gamma_-+1}} - \left(\frac{k}{n}\right)^{1/2}\left(\frac{k}{n}\right)^{1/2} W\left(\frac{n}{k}\right)s^{-\gamma_-}. \tag{3.62}$$

Since $(k/n)^{1/2}W(n/k)$ is bounded in probability, and $s^{-\gamma_-}$ is uniformly continuous on $[0,1]$, we have $\sup_{0\leq s\leq 1}|\,(k/n)^{1/2}(k/n)^{1/2}W(n/k)\,s^{-\gamma_-} = o_{P'}(1)$. Therefore, the right hand side of (3.62) is tight and hence it converges weakly to $W(s)s^{-(\gamma_-+1)}$, $(0 < s < 1)$. □

3.4 Proofs of the theorems

In this subsection, we provide the proofs of the theorems in the main paper. These results are in terms of the conditional probability P^* given the observations $X_1, \ldots, X_n$.

Prof of Theorem 3.1. For a statement and proof of the corresponding non-bootstrap version of this theorem, see Lemma 4.3 in de Haan and Resnick (1993); according to this Lemma, under the conditions of this theorem, $k^{1/2}\{\hat{b}(n/k) - b(n/k)\}/a(n/k)$ converges in distribution to B as $n \to \infty$. By Lemma 3.3, $[k^{1/2}/\hat{a}(n/k)]\{\hat{b}^*(n/k) - \hat{b}(n/k)\} = B'_{k,n}(1) + o_{P'}(1)$. Then, by Lemma 3.16, $[k^{1/2}/\hat{a}(n/k)]\{\hat{b}^*(n/k) - \hat{b}(n/k)\} = B'_{k,n}(1) + o_{P^*}(1)$, in probability $[P]$. Now, the result follows by Proposition 3.3. Since B_n is independent of $\boldsymbol{V}$, and hence of $\boldsymbol{X}$, it follows from Proposition 3.3 that $[k^{1/2}/\hat{a}(n/k)]\{\hat{b}^*(n/k) - \hat{b}(n/k)\} \xrightarrow{d^*} B$, in probability, as $n \to \infty$. ■

Prof of Theorem 3.2. For a statement and proof of the corresponding non-bootstrap version of this theorem, see Lemma 4.6 in de Haan and Resnick (1993); according to this Lemma, under the conditions of this theorem, $k^{1/2}\{\hat{\gamma} - \gamma\} \xrightarrow{d} \Gamma$ as $n \to \infty$. By (3.45), $k^{1/2}(\hat{\gamma}^* - \hat{\gamma}) = \Gamma'_n + o_{P'}(1)$. Then, by Lemma 3.16, $k^{1/2}(\hat{\gamma}^* - \hat{\gamma}) = \Gamma'_n + o_{P^*}(1)$, in probability $[P]$; note that the remainder term is $o_{P^*}(1)$, not $o_{P'}(1)$. Since Γ'_n is independent of $\boldsymbol{V}$, and hence of $\boldsymbol{X}$, the result follows by Proposition 3.3. ■

Proof of Theorem 3.3. For this proof, p_n is given and x_n is the unknown quantile corresponding to p_n. For a statement and proof of the corresponding non-bootstrap version of this theorem, see Lemma 4.7 in de Haan and Resnick (1993); according to this Lemma, under the conditions of this theorem, $k^{1/2}\left(\{\hat{a}(n/k)/a(n/k)\} - 1\right) \xrightarrow{d} A$ as $n \to \infty$. By Lemma 3.13, $k^{1/2}\left(\{\hat{a}^*(n/k)/\hat{a}(n/k)\} - 1\right) = A'_n + o_{P'}(1)$. Then, by Lemma 3.16, $k^{1/2}\left(\{\hat{a}^*(n/k)/\hat{a}(n/k)\} - 1\right) = A'_n + o_{P^*}(1)$ in probability $[P]$. Now, the result follows by Proposition 3.3. ■

Proof of Theorem 3.4. For this proof, x_n is known, and $p_n = P(X > p_n)$ is the unknown tail probability. For a statement and proof of the corresponding non-bootstrap version of this theorem, see Theorem 4.3.1 in de Haan and Ferreira (2006b); according to this result,

$$\frac{k^{1/2}}{q_{\hat{\gamma}}(d_n)} \frac{\hat{x}_n - x_n}{\hat{a}\left(\frac{n}{k}\right)} \xrightarrow{d} \Gamma + (\gamma_-)^2 B - (\gamma_-)A \quad \text{as } n \to \infty. \tag{3.63}$$

under the conditions of Theorem 3.4. By Lemma 3.14,

$$[k^{1/2}/q_{\hat{\gamma}}(d_n)][\{\hat{x}^*_n - \hat{x}_n\}/\hat{a}\,(n/k)] = \Gamma'_n + (\gamma_-)^2 B'_{k,n}(1) - (\gamma_-)A'_n + o_{P'}(1).$$

The Lemma also shows that the result holds with the scaling factor, $q_{\hat{\gamma}}(d_n)\hat{a}\,(n/k)$ replaced by its bootstrap counterpart $q_{\hat{\gamma}^*}(d_n)\hat{a}^*(n/k)$. Then, by Lemma 3.16,

$$\{k^{1/2}/q_{\hat{\gamma}^*}(d_n)\}\{(\hat{x}_n^* - \hat{x}_n)/\hat{a}^*(n/k)\} = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P^*}(1)$$

in probability $[P]$. Therefore, by Proposition 3.3,

$$\{k^{1/2}/q_{\hat{\gamma}^*}(d_n)\}\{(\hat{x}_n^* - \hat{x}_n)/\hat{a}^*(n/k)\} \xrightarrow{d^*} \Gamma + (\gamma_-)^2 B - (\gamma_-)A. \blacksquare$$

Proof of Theorem 3.5. For a statement and proof of the corresponding non-bootstrap version of this theorem, see Theorem 4.4.1 in de Haan and Ferreira (2006b); it is shown there that

$$\frac{k^{1/2}}{w_{\hat{\gamma}}(d_n)}\left(\frac{\hat{p}_n}{p_n} - 1\right) \xrightarrow{d} \Gamma + (\gamma_-)^2 B - (\gamma_-)A. \tag{3.64}$$

By Lemma 3.15,

$$\{k^{1/2}/w_{\hat{\gamma}}(\hat{d}_n)\}\{\hat{p}_n^*/\hat{p}_n - 1\} = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P'}(1).$$

Then, by Lemma 3.16,

$$\{k^{1/2}/w_{\hat{\gamma}}(\hat{d}_n)\}\left(\hat{p}_n^*/\hat{p}_n - 1\right) = \Gamma_n' + (\gamma_-)^2 B_{k,n}'(1) - (\gamma_-)A_n' + o_{P^*}(1)$$

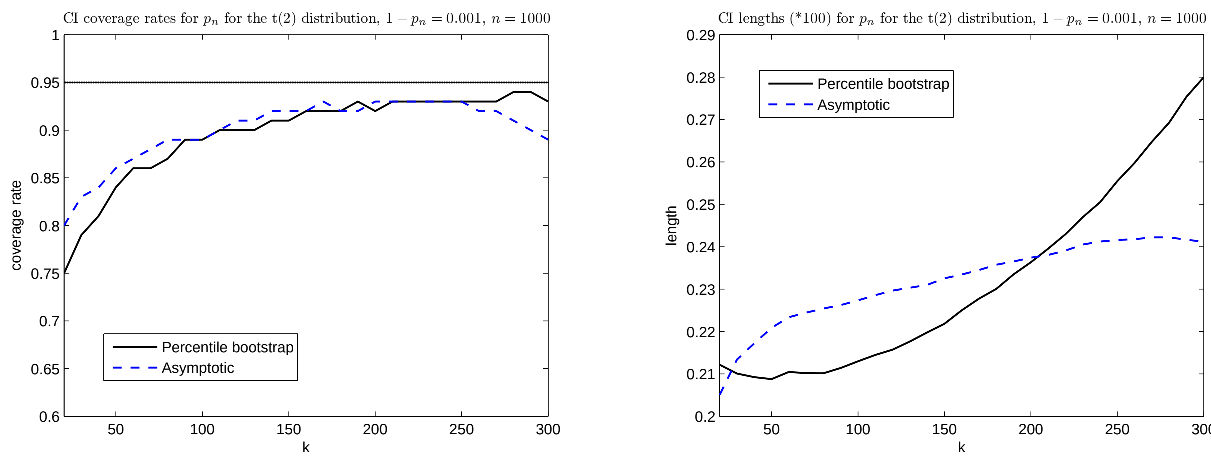
in probability $[P]$. Therefore, $\{k^{1/2}/w_{\hat{\gamma}}(\hat{d}_n)\}\left(\hat{p}_n^*/\hat{p}_n - 1\right) \xrightarrow{d^*} \Gamma + (\gamma_-)^2 B - (\gamma_-)A$, in probability, by Proposition 3.3. Again, the result holds with the scaling factor $w_{\hat{\gamma}}(\hat{d}_n)$ replaced by its bootstrap counterpart $w_{\hat{\gamma}^*}(\hat{d}_n^*)$ where $d_n^* = k/(n\hat{p}_n^*)$. $\blacksquare$

Appendix 2: Simulation Results

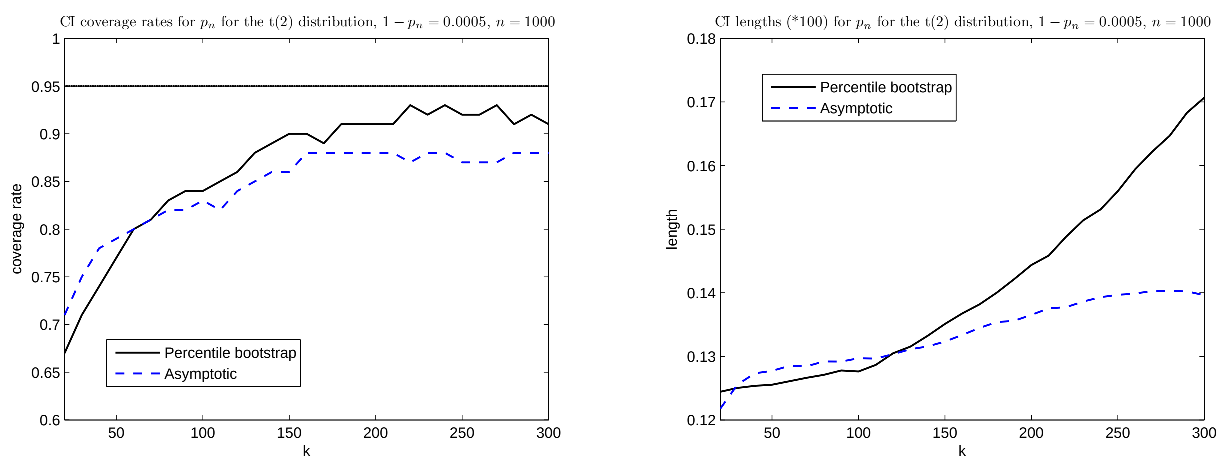
The figure on each page that appears in this section provides the coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the tail probability $1-F(x_n)$, where x_n was chosen such that $p_n := 1-F(x_n)$ takes the values $0.001, 0.0005$, or 0.00025. Each figure has three rows indicated (a), (b), and (c), and two columns. The column on the left provides the coverage rates of the confidence intervals for the unknown tail probability $1-F(x_n)$, and the column on the right provides the mean length of the confidence intervals. The three rows provide the results for $p_n = 0.001, 0.0005$, and 0.00025 respectively. The asymptotic method for constructing the confidence interval refers to the asymptotic distribution of the estimator of the tail probability (see Theorem 4.4.1 in de Haan and Ferreira 2006b). Since these details are the same for every figure in this section, they are mentioned only briefly in the legend of each figure. The coverage rates and length of the confidence intervals are plotted against k, as is typical of the literature on tail estimation; these are pointwise confidence intervals in the sense that the confidence interval is constructed separately for each fixed k.

Figure 1: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the t(2) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

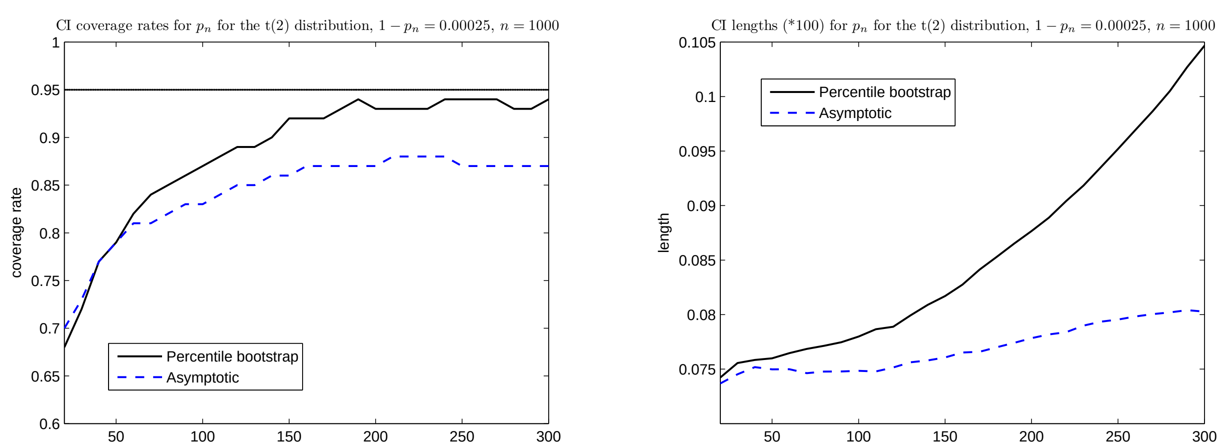
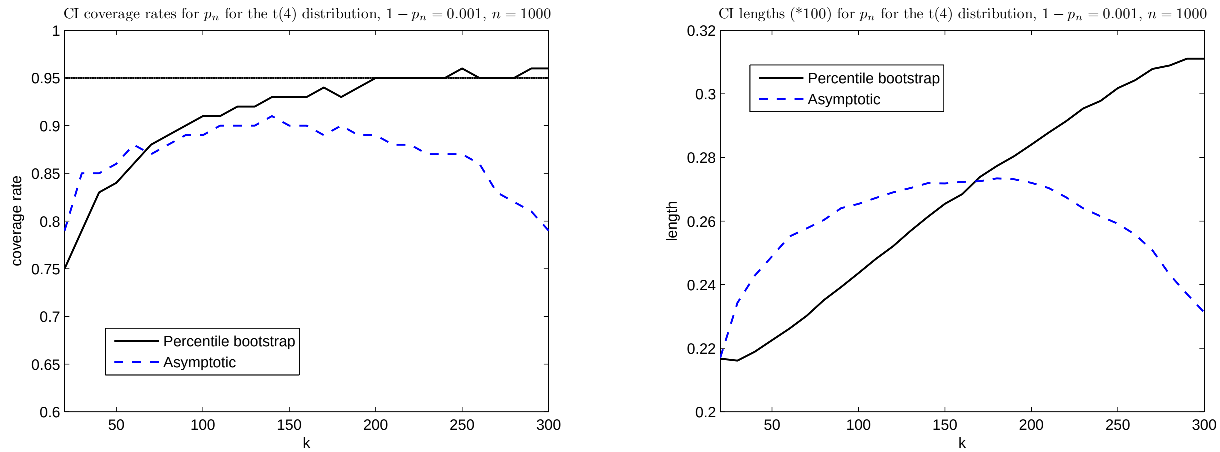
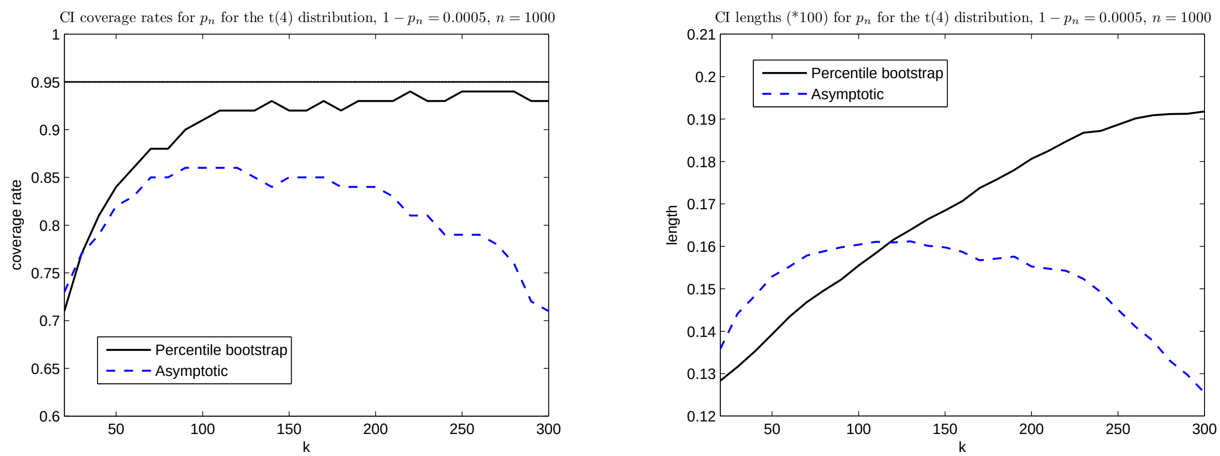


Figure 2: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the t(4) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

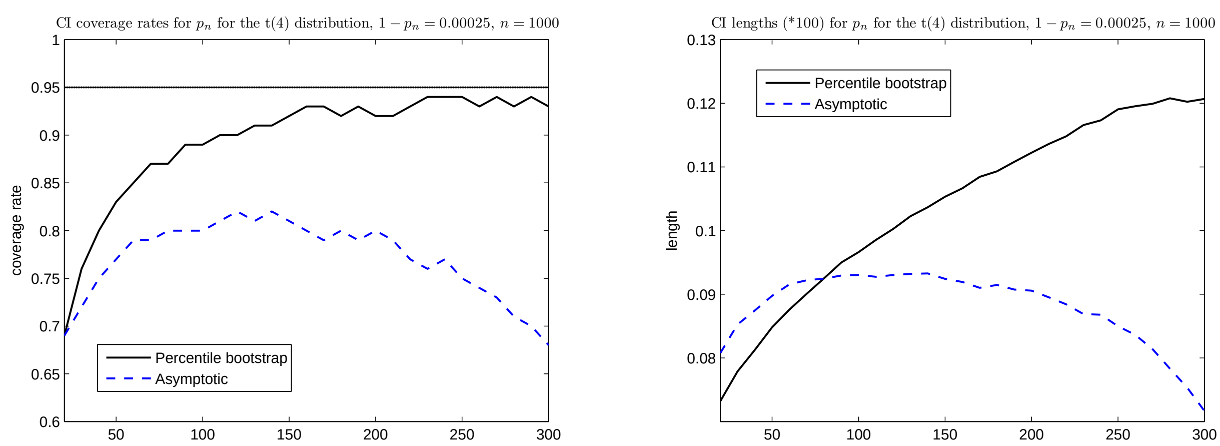
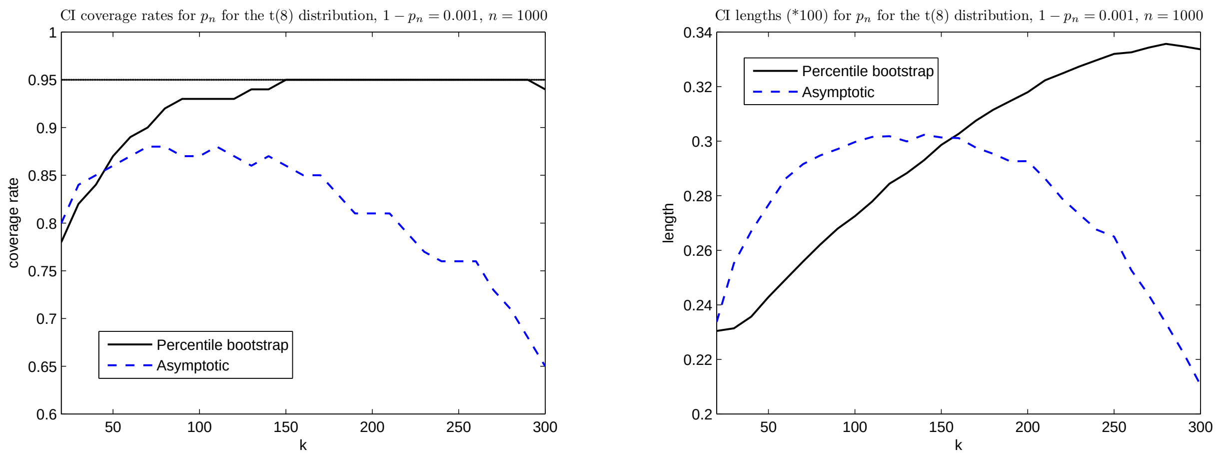
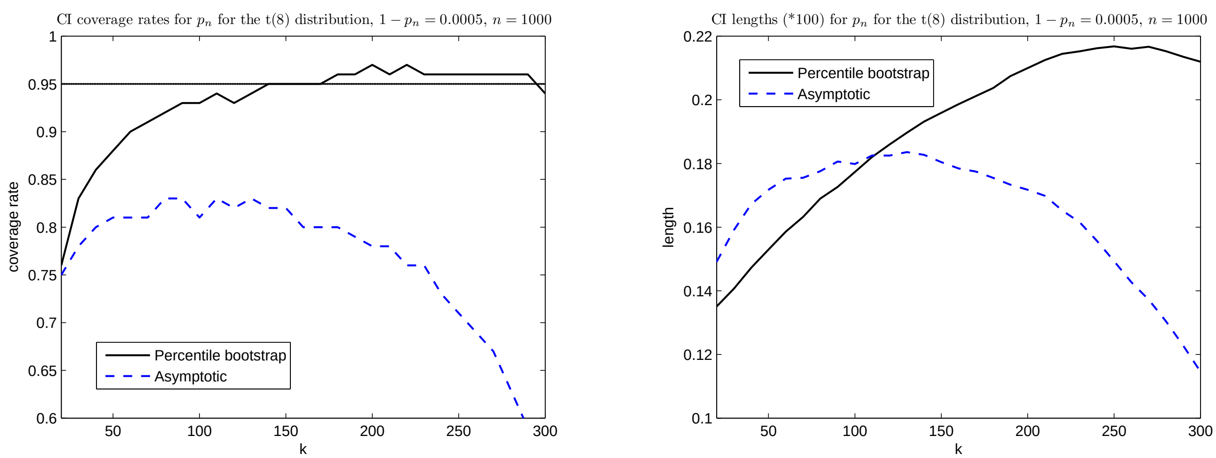


Figure 3: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the t(8) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

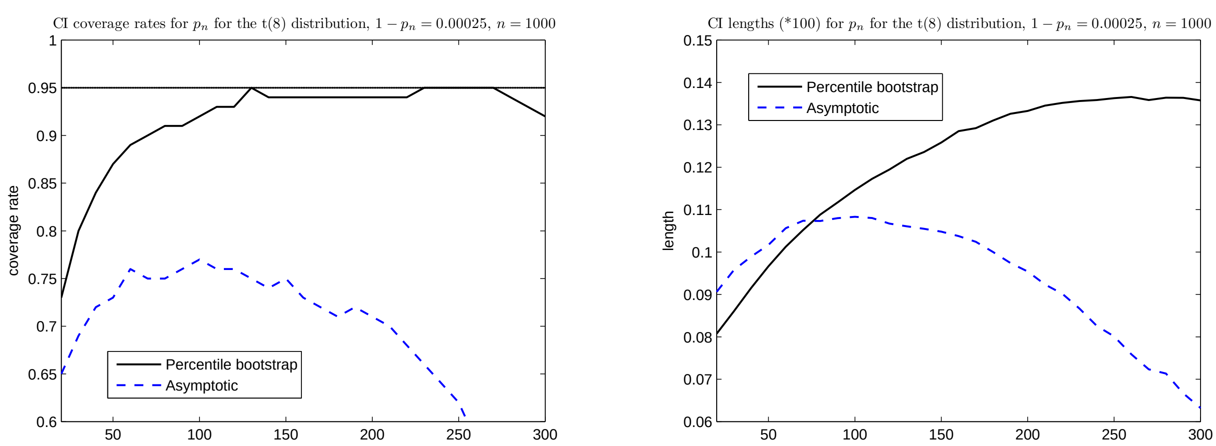
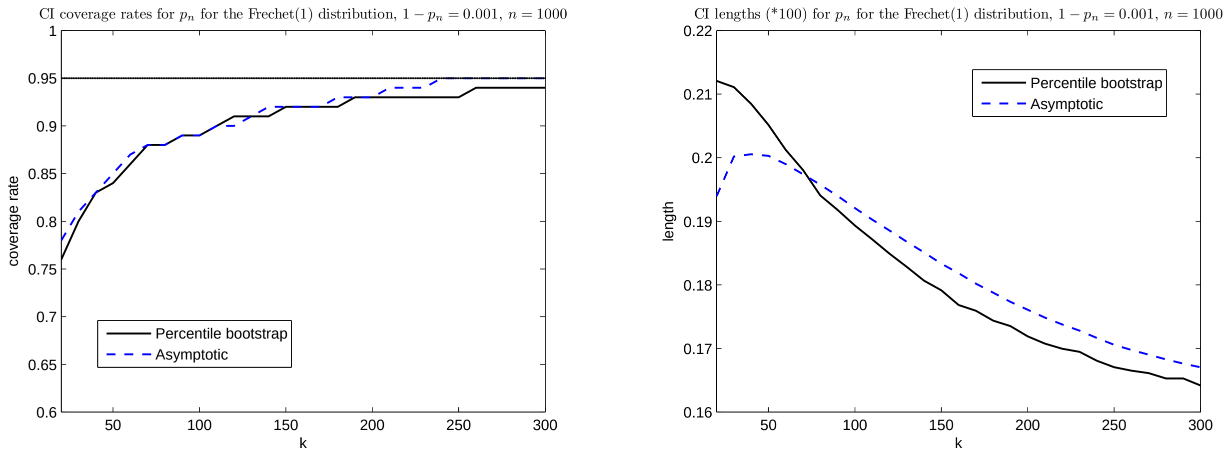
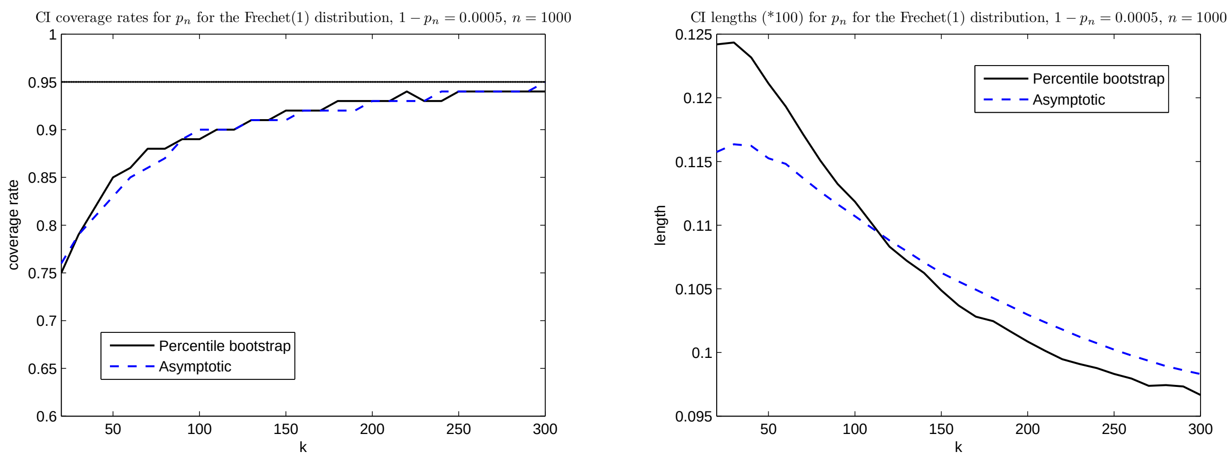


Figure 4: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the Frechet(1) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

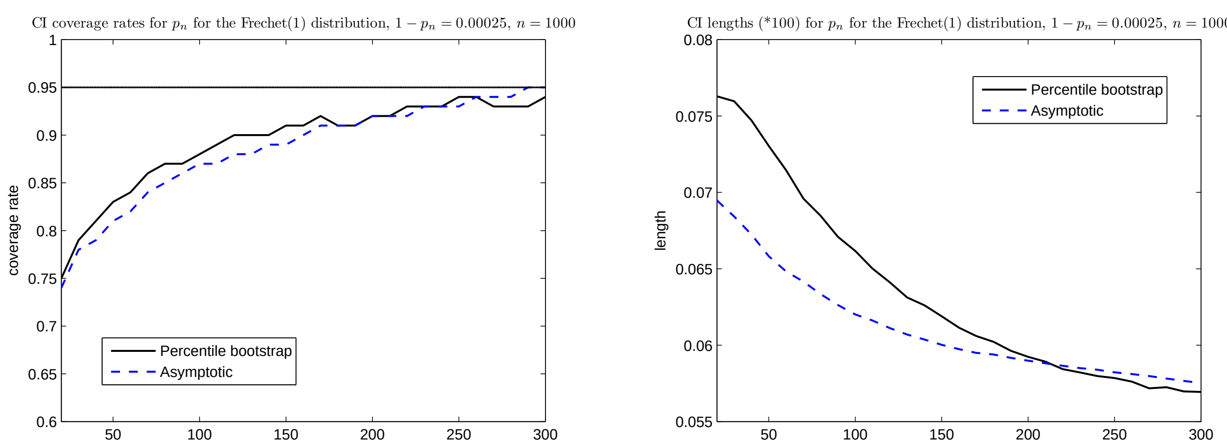
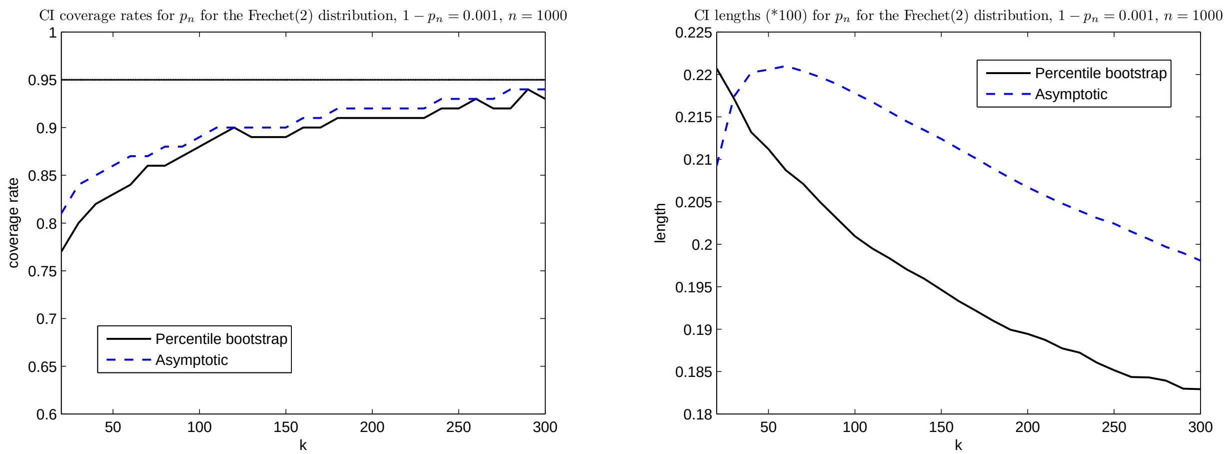
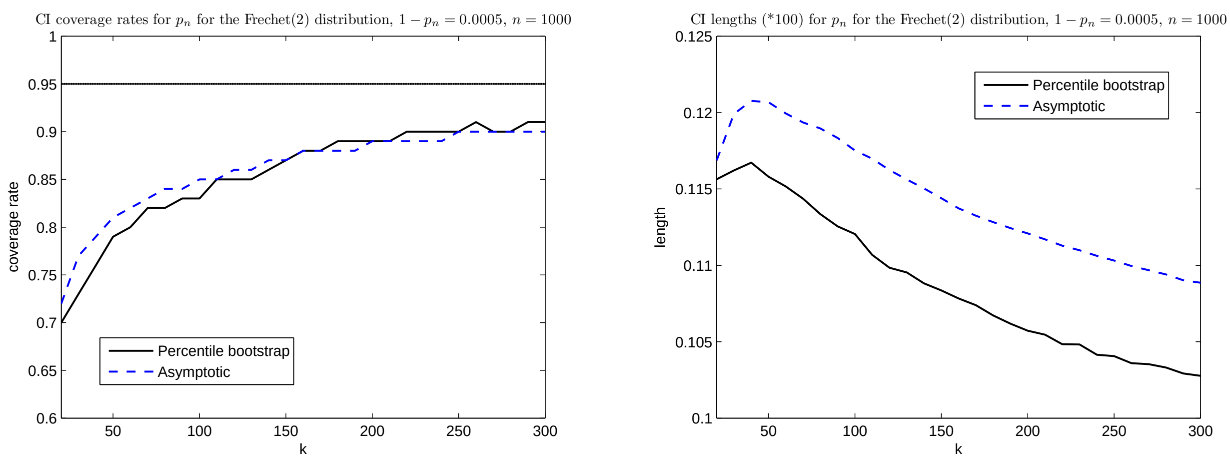


Figure 5: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the Frechet(2) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

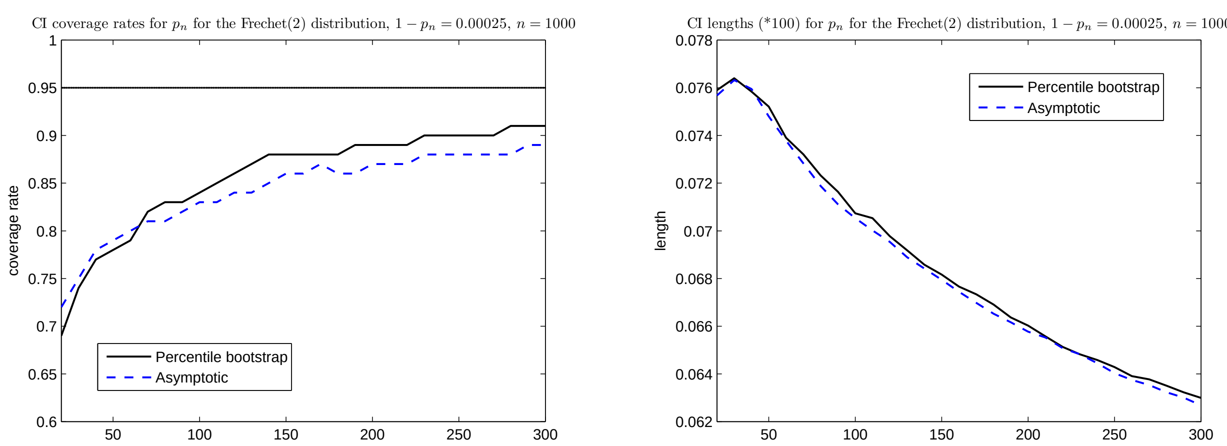
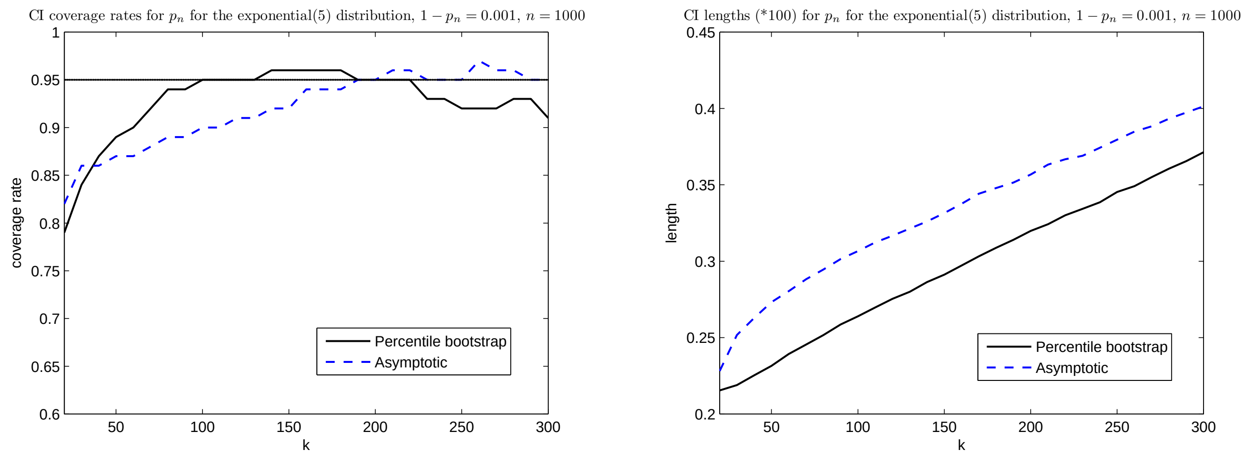
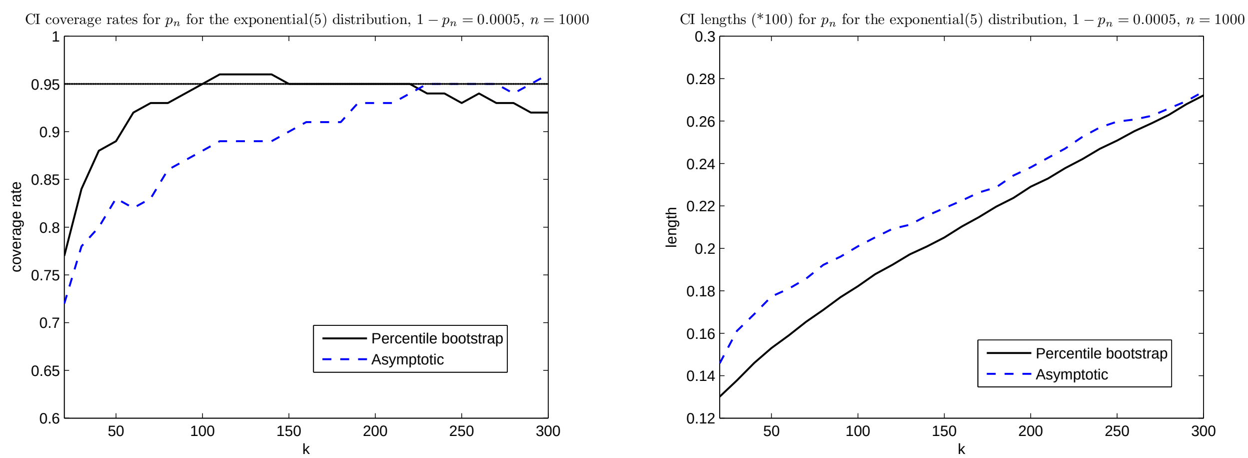


Figure 6: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the exponential(5) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

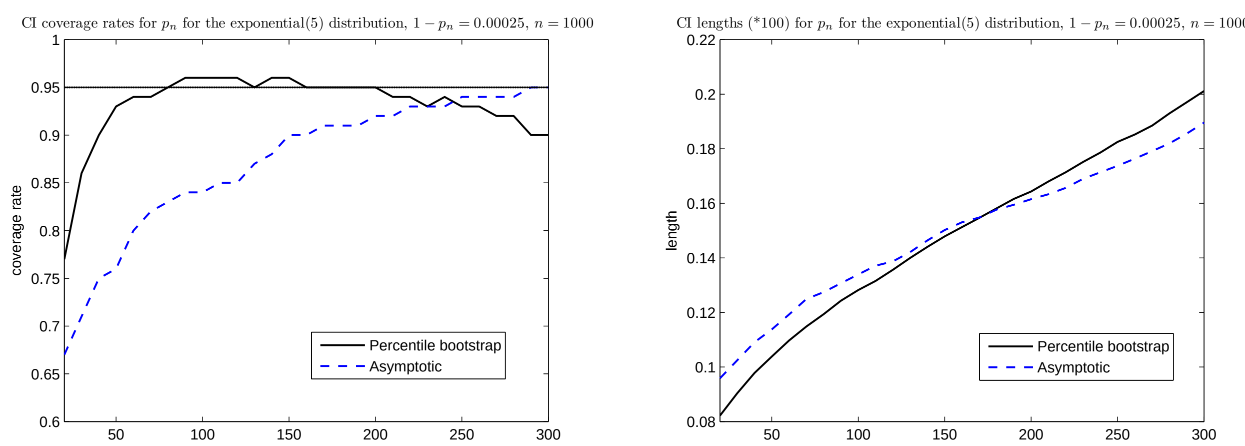
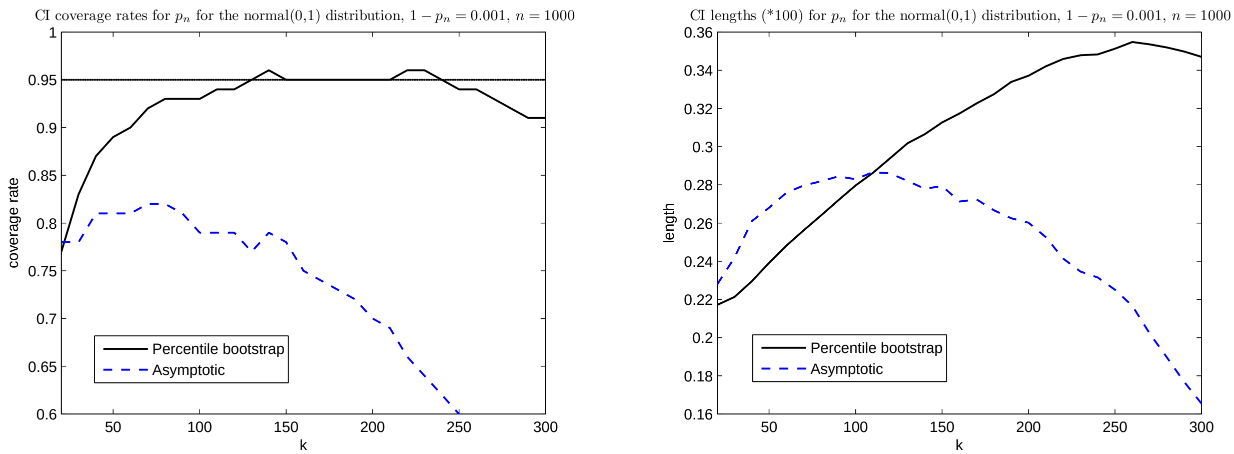
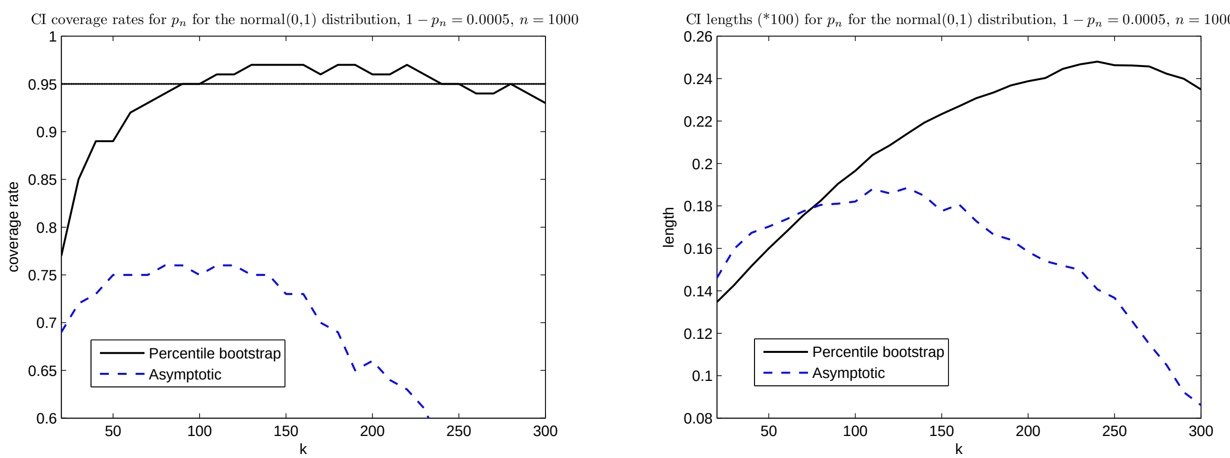


Figure 7: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the normal(0,1) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

a)



b)



c)

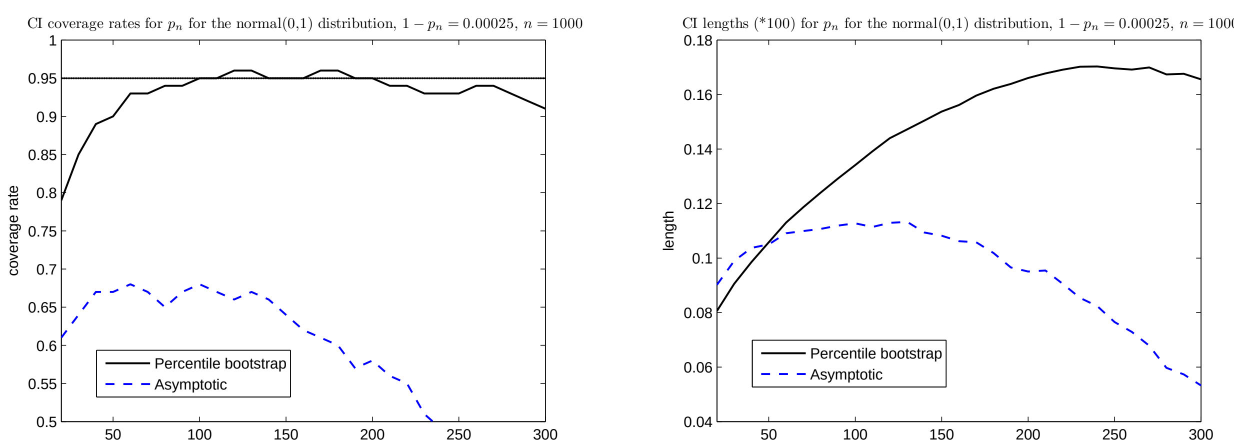
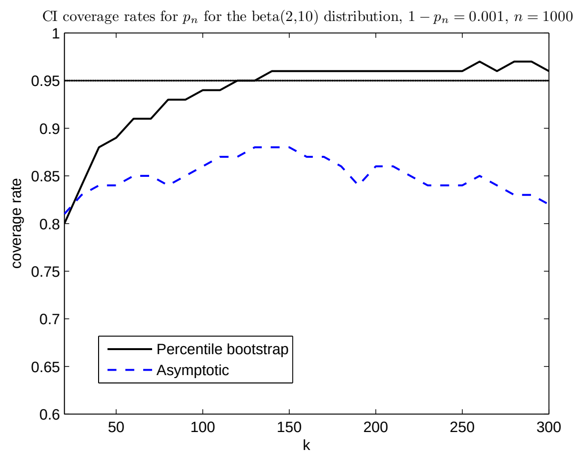
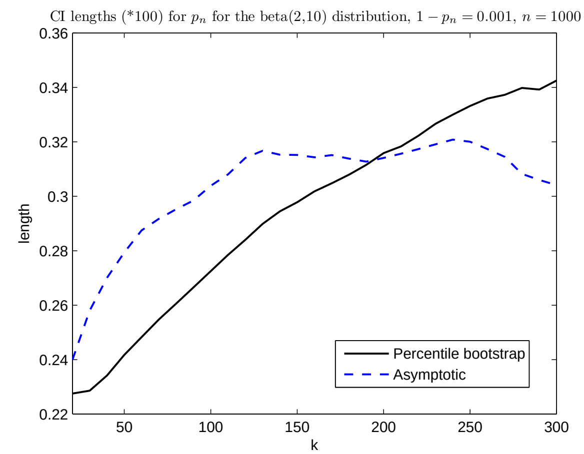


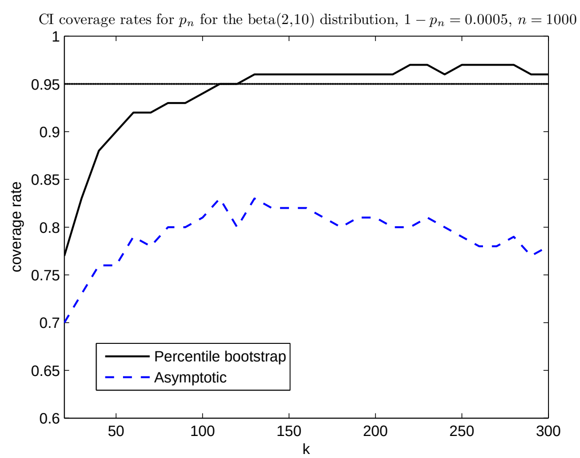
Figure 8: Coverage rates and average lengths of the asymptotic and bootstrap 95% confidence intervals for the beta(2,10) distribution; $n = 1000$. The coverage rates and average lengths of the confidence intervals for the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0005$ and c) $p_n = 0.00025$ are plotted as a function of k.

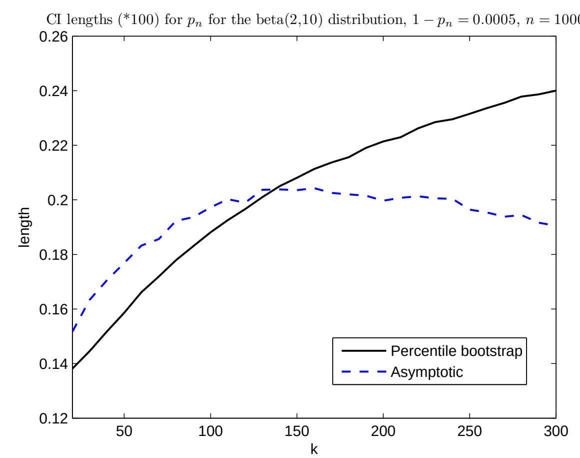
a)



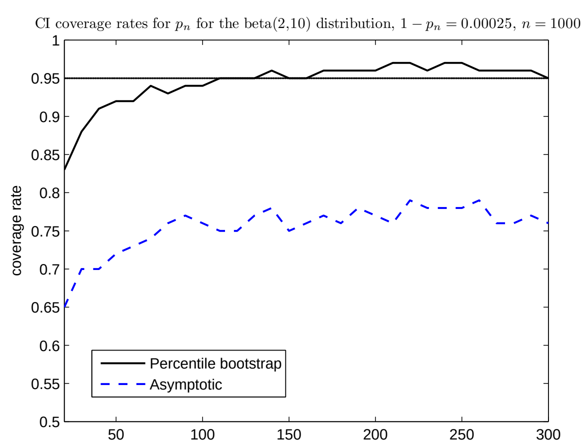


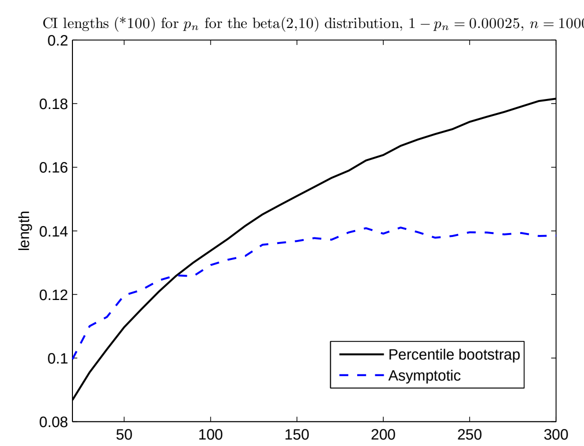
b)





c)





References

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. Wiley, New York.

Csörgő, M., Csörgő, S., Horváth, L., and Mason, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.*, **14**(1), 31–85.

Csörgő, S. and Mason, D. M. (1989). Bootstrapping empirical functions. *The Annals of statistics*, **17**(4), 1447–1471.

de Haan, L. and Ferreira, A. (2006a). *Extreme Value Theory: An Introduction*. Springer.

de Haan, L. and Ferreira, A. (2006b). *Extreme Value Theory: An Introduction*. Springer.

de Haan, L. and Resnick, S. (1993). Estimating the limit distribution of multivariate extremes. *Communications in Statistics Stochastic Models*, **2**, 275–309.

Drees, H. (1998). On smooth statistical tail functionals. *Scandinavian Journal of Statistics*, **25**(1), 187–210.

Peng, L. and Qi, Y. (2017). *Inference for Heavy-Tailed Data: Applications in Insurance and Finance*. Academic Press.