**Article:**

Abadir, Karim Maher and Cornea-Madeira, Adriana orcid.org/0000-0002-0889-7145 (2019) Link of moments before and after transformations, with an application to resampling from fat-tailed distributions. Econometric Theory. pp. 630-652. ISSN 0266-4666

https://doi.org/10.1017/S026646661800021X

# Link of moments before and after transformations, with an application to resampling from fat-tailed distributions*

Karim M. Abadir[†] Adriana Cornea-Madeira[‡]

May 2, 2018

**Abstract**

Let $x$ be a transformation of $y$, whose distribution is unknown. We derive an expansion formulating the expectations of $x$ in terms of the expectations of $y$. Apart from the intrinsic interest in such a fundamental relation, our results can be applied to calculating $\mathrm{E}(x)$ by the low-order moments of a transformation which can be chosen to give a good approximation for $\mathrm{E}(x)$. To do so, we generalize the approach of bounding the terms in expansions of characteristic functions, and use our result to derive an explicit and accurate bound for the remainder when a finite number of terms is taken. We illustrate one of the implications of our method by providing

*We thank Essie Maasoumi, Natalia Markovich, Peter Phillips, Robert Taylor, Michael Wolf and three anonymous referees for their comments. This paper was invited at the 2nd conference of the International Society of Non Parametric Statistics (Cádiz), 11th International Vilnius Conference on Probability Theory and Mathematical Statistics, 23rd Midwest Econometrics Meeting (Bloomington, Indiana), 4th French Econometrics Conference (Rennes), Conference on Fat Tails (Rennes), Katholieke Universiteit Leuven, LSE, Queen Mary, University of Bologna, University of Cyprus, and University of Oxford. This research is supported by the ESRC grant RES062230790 and by the British Academy's PDF/2009/370.

[†]American University in Cairo and Imperial College London; Email: k.m.abadir@imperial.ac.uk

[‡]University of York; Email: adriana.cornea-madeira@york.ac.uk

accurate naive bootstrap confidence intervals for the mean of any fat-tailed distribution with an infinite variance, in which case currently-available bootstrap methods are asymptotically invalid or unreliable in finite samples.

*Keywords:* Expansion of functions; Remainder's bound; Complex analysis Moments; Bootstrap confidence interval; Infinite variance; Stable laws.

# 1 Introduction

Let $x \in \mathcal{X} \subseteq \mathbb{R}$ be a variate with unknown distribution, and suppose that we are interested in one of its moments, say $\mathrm{E}(x)$. We provide a methodology to calculate $\mathrm{E}(x)$ in terms of the moments of $y \in \mathcal{Y} \subseteq \mathbb{R}$, where $x = g(y)$. This fundamental relation between the moments of $x$ and $y$ has, surprisingly, not been derived before. Approximations to it have been used, typically through the leading terms of a Taylor expansion and without assessing either the goodness of such an expansion (as opposed to a more general one than Taylor's) or the precise evaluation of the remainder (as opposed to just stating its order of magnitude). In this paper, we provide an exact formula for general expansions linking these moments, in a more general context than Taylor expansions. In contrast to other types of expansions already proposed in the literature, if a finite number of terms is taken, the remainder in our expansions can be bounded explicitly and accurately, without resorting to orders of magnitude that only indicate the rate of change of the remainder rather than its size. For a striking illustration of the difference between bounds and orders, see Hallin and Seoh (1996) where it is shown how the latter can be misleading: a term of small order can be numerically huge, even in large samples. To be able to provide such accurate bounds, we generalize the approach of expanding characteristic functions and bounding their terms.

It is of interest to investigate such a fundamental relation linking the moments of $x$ and $y$. But apart from the intrinsic interest in it, it can be used in practice to approximate $\mathrm{E}(x)$ (or any other moment of $x$ that exists) by the low-order moments of a transformation that can be chosen to give a good approximation for $\mathrm{E}(x)$. Such results are useful, for example, in assessing the effect of applying transformations to data before ranking and building models; see Delaigle and Hall (2012) for the case of fat-tailed genomic data and Taylor (1986) whose result can be extended by our derivations. Another example is the case of fat-tailed data where the naive bootstrap of Efron (1979) and the moving block bootstrap of Künsch (1989) and Liu and Singh (1992) fail because of the nonexistence of higher moments (see Athreya (1987), Knight (1989), Hall (1990), Lahiri (1995)), but we can use our method to simply modify these resampling techniques to obtain valid bootstrap confidence intervals (CIs). The potential for applications is not just in statistics and econometrics, but also in economics and finance where there is interest *inter alia* in quantifying the effect of nonlinear transformations on moment conditions (such as Euler equations arising from optimization) and on asset prices which are formulated as expectations; e.g. see Yu, Yang, and Zhang (2006), Martin (2008), Backus, Chernov, and Martin (2011). The effect of higher-order terms is important and needs to be quantified, as increasingly frequent market turbulence has emphasized. Other applications include risk management, where portfolio losses (which are functions of risk factor changes) are approximated by first-order Taylor expansions called delta approximation if only the first term in the expansion is used, and delta-gamma approximation if the first two terms are used. The delta-gamma approximation is preferred because it gives a better approximation of the loss, but alternative approximations using our method can be envisaged. It is important to have a reliable approximation of the losses because it is used for backtesting, required in Basel solvency assessment; see McNeil, Frey, and Embrechts (2015).

Applications like these are important and substantial, and in this paper we illustrate how our approach can remedy the failure of the naive bootstrap for the mean of a fat-tailed distribution with infinite second moment and possibly asymmetric tails. Athreya (1987) has shown that, when the random variables are independent and identically distributed (i.i.d.) and have infinite variance, the naive bootstrap of the mean is asymptotically invalid. Lahiri (1995) arrives at the same conclusion for the moving block bootstrap when the random variables are dependent. Starting with the work of Athreya (1987), a number of important papers appeared on this topic in the last two decades; see Knight (1989), Arcones and Giné (1989), Arcones and Giné (1991), Giné and Zinn (1989), Giné and Zinn (1990), Hall (1990), Hall and LePage (1996), Athreya, Lahiri, and Wu (1998), Hall and Jing (1998), Romano and Wolf (1999), Politis, Romano, and Wolf (1999), Cavaliere, Georgiev, and Taylor (2013), Cornea-Madeira and Davidson (2015), Cavaliere, Georgiev, and Taylor (2016) and the references therein. The solutions to the failure of the naive bootstrap proposed in these papers are either based on a smaller resampling size ($m$ out of $n$ bootstrap and subsampling) or on a bootstrap sample size equal to the original sample size (parametric bootstrap, wild bootstrap, permutation bootstrap). But if the distribution of the data is allowed to be fat-tailed asymmetric, some of these bootstraps are not applicable and some, while asymptotically valid, do not perform well in finite samples as illustrated in Section 4 of this paper; see also Romano and Wolf (1999) and Cornea-Madeira and Davidson (2015) or Hall and Yao (2003) for an example in a regression context.

This paper is organised as follows. In Section 2, we introduce the expansion and the required results obtained by complex analysis. These are general results on the bounding of some functions of complex variables, and we apply them to obtain an accurate evaluation of the remainder when our expansion is truncated. In Section 3, we illustrate the expansion and the accuracy of the remainder's bound. In Section 4, we apply our expansion to obtain asymptotically-valid *naive* bootstrap CIs for the mean of a distribution with infinite variance. Section 5 concludes. The proofs are relegated to the Appendix and supplementary material provides additional tables for the illustrations of Section 3.

# 2 Expansion of $\mathrm{E}(x)$ in terms of the moments of $y$

Suppose for simplicity that we are interested in $\mathrm{E}(x)$ which is assumed to exist. We stress that the same approach will apply to the expansion of any moment of $x$, not just $\mathrm{E}(x)$. For example, for the expansion of $\mathrm{E}(x^3)$, we can replace $x$ by $z := x^3$ and apply the same method below to $\mathrm{E}(z)$. This is also true of other functions $z$ of $x$, as we shall see later in this section.

The first subsection presents two expansions, raw and centered, for the case of $g = \exp$. It also provides two propositions, derived in total generality, leading to a theorem which is then used in the second subsection for the case of general $g$. The third subsection concerns the specification of a deterministic scaling parameter $m$ in the expansion, and its implications for the choice of $g$.

## 2.1 Expansions and corresponding bounds, applied to $g = \exp$

We propose two types of expansions, raw or centered. We start by explaining the idea behind the two expansions using a simple familiar setup: (a) without recourse to complex variables; and (b) with the simple $g = \exp$, giving $x := \exp(y) \in \mathbb{R}_+$; see (1) and (2) below. Then, these two simplifications will be relaxed, respectively, in (5) of this subsection then in (14) of the next subsection.

The two types of well-known expansions are:

1. the raw (direct) expansion

$$x = \mathrm{e}^y = \sum_{j=0}^{k} \frac{y^j}{j!} + R_k, \tag{1}$$

2. the centered expansion

$$x = \mathrm{e}^{\mathrm{E}\, y}\mathrm{e}^{y-\mathrm{E}\, y} = \mathrm{e}^{\mathrm{E}\, y} \sum_{j=0}^{k} \frac{(y - \mathrm{E}\, y)^j}{j!} + R_k^{\mathrm{c}}. \tag{2}$$

Writing these expansions as $x = z + R$ generically, the expectation of $z$ always exists (since $y := \log(x)$) and so $\mathrm{E}(R)$ also exists (by the assumption that $\mathrm{E}(x)$ exists).

Before removing the first simplification used in the previous expansions (that of no complex numbers), it is easiest to explain the intuition behind our approach with the simple (1)–(2). To illustrate with the case of resampling from fat-tailed distributions, we will show in Section 4 that it is the higher-order terms that create problems:

$$x = \underbrace{\left(1 + y + \frac{y^2}{2!} + \cdots + \frac{y^k}{k!}\right)}_{\text{standard resampling applies}} + R_k, \tag{3}$$

with the difficulties arising from $R_k$ (hence for $x$ too). If, instead of using exactly $R_k$ (which inherits the problems of $x$), we are able to bound $R_k$ and replace it by a low-order term (such as a multiple of $|y|^{k+1}$), then the problem is fixed.[1] This is achievable by using complex numbers, as in the theory of characteristic functions, but not in a conventional way, as we now show.

Let $-1 \equiv \mathrm{i}^2$ and take some $m \in \mathbb{N}$ according to criteria that are application-specific and to be discussed in Subsection 2.3 and subsequently. All we need to know at this stage is that $m$ is a deterministic natural number that we now use for rewriting the centered $y$ as

$$\frac{y - \mathrm{E}\, y}{m} \equiv \zeta + 2\pi i_y, \tag{4}$$

where the left-hand side is decomposed into a multiple of $2\pi$ measured by $i_y \in \mathbb{Z}$ and a leftover $\zeta \in (-\pi, \pi]$ to be converted into an angle in the complex space. Both $i_y$ and $\zeta$ are random because $y$ is. Then, focusing on this leftover $\zeta$, write

$$x = \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left(\exp\left(\frac{\zeta}{\mathrm{i}}\right)\right)^{\mathrm{i}m} \equiv \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m}, \tag{5}$$

where $\xi_k := \sum_{j=0}^{k} \zeta^j / (\mathrm{i}^j j!)$. The standard bound

$$\left|\varrho_{x,k}\right| \equiv \left|\mathrm{e}^{\zeta/\mathrm{i}} - \sum_{j=0}^{k} \frac{(\zeta/\mathrm{i})^j}{j!}\right| \le \frac{|\zeta|^{k+1}}{(k+1)!} \tag{6}$$

may now be used for any $\zeta \in \mathbb{R}$, which is what we were seeking to do in (3). This will be helpful when we take expectations later in this section, as it controls the precision of the

[1] Of course, the bound would have to converge to $R_k$, as both bound and $R_k$ shrink to zero when the sample size expands. This will be made precise from Subsection 2.3 onwards.

remainder *and* it has a well-behaved expectation (being the finite $(k+1)$-th power of a variate with exponentially-decaying tail). This (6) is specific to $g = \exp$, unlike the general propositions and theorem that we will now derive which are valid for any $g$. We will only employ (6) after these general results.

Unlike in the standard theory of characteristic functions, the whole series expansion in (5) is additionally raised to the imaginary power $\mathrm{i}m$, a complication that we now deal with. A binomial expansion of (5) gives

$$x = \mathrm{e}^{\mathrm{E}\,y} \mathrm{e}^{2\pi m i_y} \operatorname{Re}(\xi_k^{\mathrm{i}m}) + R_{x,k}^{\mathrm{c}}, \tag{7}$$

where

$$\begin{aligned} \left|R_{x,k}^{\mathrm{c}}\right| &= \mathrm{e}^{\mathrm{E}\,y} \mathrm{e}^{2\pi m i_y} \left|\operatorname{Re}\left(\left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m} - \xi_k^{\mathrm{i}m}\right)\right| \\ &\leq \mathrm{e}^{\mathrm{E}\,y} \mathrm{e}^{2\pi m i_y} \left|\left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m} - \xi_k^{\mathrm{i}m}\right| \\ &= \mathrm{e}^{\mathrm{E}\,y} \mathrm{e}^{2\pi m i_y} \left|\xi_k^{\mathrm{i}m}\right| \left|\left(1 + \varrho_{x,k}/\xi_k\right)^{\mathrm{i}m} - 1\right|. \end{aligned} \tag{8}$$

This equation shows that, whatever goes into the expansion $\xi_k$ and the remainder $\varrho_{x,k}$, the expression $|\left(1 + \varrho_{x,k}/\xi_k\right)^{\mathrm{i}m} - 1|$ is a generic form that arises from expanding any $g$ of $x = g(y)$ into an expression raised to an imaginary power $\mathrm{i}m$. If we can bound this form accurately, rather than just give its order of magnitude, then we can evaluate remainders satisfactorily. We therefore present the following results.

**Proposition 1** Define the real-valued function $h$ of the complex $\psi$,

$$h(\psi) := \left|(1+\psi)^{\mathrm{i}m} - 1\right|, \qquad \text{with } \arg(\psi) \in [-\pi, \pi).$$

Then, the global maximum of the function is attained at $h(\psi_{\mathrm{m}}) = 1 + \mathrm{e}^{m\pi}$ by $\psi_{\mathrm{m}} = -1 - \mathrm{e}^{(2j+1)\pi/m}$ in the clockwise direction ($\arg(\psi_{\mathrm{m}}) = -\pi$) with $j \in \mathbb{Z}$.

Notice that the triangle inequality gives

$$\left|(1+\psi)^{\mathrm{i}m}-1\right| \leq \left|(1+\psi)^{\mathrm{i}m}\right|+1=\exp\left(-m\theta\right)+1 \leq 1+\exp\left(m\pi\right),$$

where the equality follows from (20) in the Appendix and $\theta := \arg(1+\psi)$. The upper bound of $1+\exp(m\pi)$ is indeed achieved and the proposition tells us which values of $\psi$ achieve it.

By choosing a large negative $j$, the solution $\psi_{\mathrm{m}} = -1-\exp\left(\left(2j+1\right)\pi/m\right)$ can be made sufficiently close to $\psi=-1$ for most practical purposes. But if the variables and expansions are such that $|\psi| \leq 1$, we need to derive another solution that takes this restriction on the size of $\psi$ into account, and we need a local bound for $h(\psi)$ when $|\psi| \leq 1$ if we want to have a precise bound on remainders of expansions. Recalling the context of the expansions, our proposition applies to $\psi := \varrho_{x,k}/\xi_k$ and the remainder $\varrho_{x,k}$ should be small relative to the leading terms $\xi_k$, so small values of $|\psi|$ are indeed relevant. The following $|\psi|$-pointwise bound is obtained for $|\psi| \leq 1$.

**Proposition 2** Define the real-valued function $h$ of the complex $\psi$,

$$h(\psi) := \left|(1+\psi)^{\mathrm{i}m}-1\right|, \qquad \text{with } \arg\left(\psi\right) \in [-\pi,\pi) \text{ and } |\psi| \leq |\psi_0| \in [0,1].$$

Then, the maximum of the function is monotonic increasing in $|\psi_0|$ and so is the bound

$$h(\psi) \leq \begin{cases} h_1(\psi_0), & |\psi_0| \in \left[0, 1-\mathrm{e}^{-\pi/m}\right), \\ 1+\mathrm{e}^{m\sin^{-1}|\psi_0|}, & |\psi_0| \in \left[1-\mathrm{e}^{-\pi/m}, 1\right], \end{cases}$$

for any given $|\psi_0|$, where

$$h_1(\psi_0) := \sqrt{1-2\mathrm{e}^{m\sin^{-1}|\psi_0|}\cos\left(m\log\left(1-|\psi_0|\right)\right)+\mathrm{e}^{2m\sin^{-1}|\psi_0|}}. \tag{9}$$

As a result of this proposition and the first one, we have the following bound.

**Theorem 1** *Define the real-valued function $h$ of the complex $\psi$,*

$$h(\psi) := \left|(1+\psi)^{\mathrm{i}m} - 1\right|, \qquad \textit{with } \arg(\psi) \in [-\pi, \pi).$$

*Then, $\left|(1+\psi)^{\mathrm{i}m} - 1\right| \le H(|\psi|)$, where*

$$H(|\psi|) := \begin{cases} h_1(\psi), & |\psi| \in \left[0, 1 - \mathrm{e}^{-\pi/m}\right), \\ 1 + \mathrm{e}^{m \sin^{-1}|\psi|}, & |\psi| \in \left[1 - \mathrm{e}^{-\pi/m}, 1\right], \\ 1 + \mathrm{e}^{m\pi}, & \textit{otherwise}, \end{cases} \tag{10}$$

*and $h_1$ is defined in (9).*

In the case of $g = \exp$ and (6), the monotonicity property in Proposition 2 and definition (10) imply that we can work out the remainder's bound in (8) as

$$\left|R^{\mathrm{c}}_{x,k}\right| \le \mathrm{e}^{\mathrm{E}\,y} \mathrm{e}^{2\pi m i_y} \left|\xi_k^{\mathrm{i}m}\right| H\left(\frac{|\zeta|^{k+1}}{(k+1)!\,|\xi_k|}\right). \tag{11}$$

For calculating expectations, the remainder from a centered expansion of $\mu := \mathrm{E}(x)$ (as opposed to expanding $x$) is denoted by $R^{\mathrm{c}}_{\mu,k}$ and is bounded by

$$\begin{aligned} \left|R^{\mathrm{c}}_{\mu,k}\right| &= \mathrm{e}^{\mathrm{E}\,y} \left|\mathrm{E}\left[\mathrm{e}^{2\pi m i_y} \operatorname{Re}\left(\left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m} - \xi_k^{\mathrm{i}m}\right)\right]\right| \\ &\le \mathrm{e}^{\mathrm{E}\,y}\, \mathrm{E}\left|\mathrm{e}^{2\pi m i_y} \operatorname{Re}\left(\left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m} - \xi_k^{\mathrm{i}m}\right)\right| \\ &\le \mathrm{e}^{\mathrm{E}\,y}\, \mathrm{E}\left[\mathrm{e}^{2\pi m i_y} \left|\xi_k^{\mathrm{i}m}\right| H\left(\frac{|\zeta|^{k+1}}{(k+1)!\,|\xi_k|}\right)\right]. \end{aligned} \tag{12}$$

We have therefore established the following result.

**Corollary 1** For $g = \exp$, the $k$-term expansions

$$x = \mathrm{e}^{\mathrm{E}\,y} \mathrm{e}^{2\pi m i_y} \operatorname{Re}(\xi_k^{\mathrm{i}m}) + R^{\mathrm{c}}_{x,k} \quad \text{and} \quad \mathrm{E}(x) = \mathrm{e}^{\mathrm{E}\,y}\, \mathrm{E}\left(\mathrm{e}^{2\pi m i_y} \operatorname{Re}(\xi_k^{\mathrm{i}m})\right) + R^{\mathrm{c}}_{\mu,k}$$

have the bounds $|R^{\mathrm{c}}_{x,k}| \leq B^{\mathrm{c}}_{x,k}$ and $|R^{\mathrm{c}}_{\mu,k}| \leq \mathrm{E}(B^{\mathrm{c}}_{x,k})$, with

$$B^{\mathrm{c}}_{x,k} := \mathrm{e}^{\mathrm{E}\, y}\mathrm{e}^{2\pi m i_y} \left|\xi_k^{\mathrm{i}m}\right| H\left(\frac{|\zeta|^{k+1}}{(k+1)!\,|\xi_k|}\right).$$

The same results apply to the raw version of our expansions, but with $\mathrm{E}(y)$ replaced by zero throughout. The alternative formulation

$$\left|\xi_k^{\mathrm{i}m}\right| = \left||\xi_k|^{\mathrm{i}m}\right| \mathrm{e}^{-m\arg(\xi_k)} = \left|\mathrm{e}^{\mathrm{i}m\log|\xi_k|}\right| \mathrm{e}^{-m\arg(\xi_k)} = \mathrm{e}^{-m\arg(\xi_k)} \tag{13}$$

does not contain imaginary quantities.

## 2.2 Expansions and corresponding bounds, applied to general $g$

We started this section with two simplifications just before (1). Having then introduced complex numbers into the expansions, hence removing the first simplification, we now relax the second one of $g = \exp$. The only point where we required $g = \exp$ and its remainder (6) was in (11)–(12) *after* the propositions and the theorem.

The series obtained above are a special case of Teixeira's expansion which expands a function ($g = \exp$ previously) in terms of another ($\log(x)$ or $\log(x) - \mathrm{E}\log(x)$ previously); e.g. see Whittaker and Watson (1997), pp.131–133 or Abadir and Talmain (1999) for an application. See also p.181 of Koenker, Machado, Skeels, and Welsh (1994). From $x = g(y)$ and (4),

$$x \equiv g\left((\mathrm{E}\, y + 2\pi m i_y) + \zeta m\right). \tag{14}$$

There are two ways to view the route we have taken earlier. First, it could be regarded as a one-term Teixeira expansion of $g$ in terms of the $m$-th power of $\xi_k^{\mathrm{i}}$ with $\xi_k := \sum_{j=0}^{k} \zeta^j/(\mathrm{i}^j j!)$. Since the nonrandom $m$ is arbitrary in $\mathbb{N}$, this one-term expansion will always exist for some $m \in \mathbb{N}$ for any analytic function $g$; see Whittaker and Watson (1997), pp.131–133 for the coefficient of the expansion. Second, we could regard our earlier setup as a $k$-term Teixeira expansion of $x^{-\mathrm{i}/m}$ in terms of $\zeta/\mathrm{i}$ as

$$x^{-\mathrm{i}/m} = g_1\left(\mathrm{E}\, y + 2\pi m i_y\right) \times \left(\xi_k + \varrho_{x,k}\right),$$

where $\xi_k$ includes the $k$ terms in powers of $\zeta/\mathrm{i}$ and the special case of $g = \exp$ gave $g_1(v) = (g(v))^{-\mathrm{i}/m}$. For both approaches, the generalization of the exponential bound (6) for the remainder $\varrho_{x,k}$ can be obtained by recourse to the inequalities of special functions. For the latter approach, we have additionally Bürmann's integral representation of the remainder of Teixeira's expansion for analytic functions (Whittaker and Watson (1997), pp.128–131) to bound $\varrho_{x,k}$. None of our propositions and the theorem (all of them listed before (11)) are affected by this generalization.

The last paragraph also implies the following. The theorem is general and can be used directly in the case of any $x = g(y)$ that is not necessarily invertible, or in the case of a composition of the type $x = g_2(g_3(y))$ where we would expand $g_2$ only. If there is such a need, the only required alteration would not be in our propositions or theorem, but in the coefficients and definition of variates in the expansion preceding Proposition 1. As an illustration of $x = g_2(g_3(y))$ with the Box-Cox transformation,

$$y := \frac{x^p - 1}{p},$$

if we are interested in the expectation of the original variate $x$, omitting the centering and scaling for ease of exposition gives

$$x = (1 + py)^{1/p} = \left(\mathrm{e}^{(\mathrm{i}p)^{-1} \log(1+py)}\right)^{\mathrm{i}} = \left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}}$$

with

$$\xi_k := \sum_{j=0}^{k} \frac{\left(\log\left(1 + py\right)\right)^j}{(\mathrm{i}p)^j\, j!}$$

and the same propositions and theorem apply as before. It is also possible to expand by something other than an exponential function, as discussed in the previous paragraph.

Finally, one should ensure that the expectation of the right-hand side of the expansion of $x$ exists, when specifying a function $g$ or $g_2$ to expand. Writing these expansions as $x = z + R$ generically, the choice of function should be such that $\mathrm{E}(z)$ exists.

### 2.3 Choice of deterministic $m$ and implications for $g$

Recall from (4) that $(y - \mathrm{E}\, y)/m \equiv \zeta + 2\pi i_y$. This means that $m$ acts like an artificial scaling parameter for the variate $y$. Choosing a large $m$ leads to the consistency-like behaviour that is needed for the remainders to converge to zero, here in the sense that the scale shrinks the variate around $\mathrm{E}(y)$ and fewer terms (smaller $k$) are needed for the expansion to be accurate; see the following sections for illustrations. This also ensures that the choice of $g = \exp$ (as opposed to some other function) is asymptotically inconsequential, although in finite samples its terms converge to zero faster than in expansions like the binomial or logarithmic that belong to the ${}_{q+1}F_q$ class of general hypergeometric ${}_pF_q$ series; see Abadir (1999) or Whittaker and Watson (1997). Other examples of ${}_pF_q$ transformations include hyperbolic functions (which are members of the ${}_0F_1$ class that converges even faster than the exponential which is a ${}_0F_0$ function) such as $\sinh(y)$ for $x \in \mathbb{R}$, $\cosh(y)$ for $x \in \mathbb{R}_+$; inverse trigonometric functions $\tan^{-1}$, $\sin^{-1}$ (see Samworth (2005) for a resampling application); and the popular Box-Cox transformation (see DiCiccio, Monti, and Young (2006) for a resampling application).

In practical applications, such as the resampling example mentioned earlier, there will be a tradeoff between reducing the magnitude of the remainder terms (requires larger $m$) and the imprecision it introduces in the evaluation of the required moments empirically; see Section 4. The optimal choice of $m$ will be application-specific, as we shall illustrate in Section 4.

# 3 Illustration of the $k$-term expansion and bound's accuracy for $\mathrm{E}(x)$

We illustrate the performance of the $k$-term expansion for $\mathrm{E}(x)$ and the corresponding bound in Corollary 1 using two distributions: the normal $y \sim \mathrm{N}(1,1)$ and the gamma $y \sim \mathrm{Gam}(\nu,\lambda)$. In the latter case, the density of the log-gamma random variable $x := \mathrm{e}^y$ is

$$f_x(u) = \frac{\lambda^\nu (\log u)^{\nu-1}}{\Gamma(\nu) u^{\lambda+1}} \qquad (\nu, \lambda > 0)\,, \tag{15}$$

for $x = u \in (1,\infty)$. For $\lambda < 2$, the log-gamma distribution is in the domain of attraction of the stable laws with infinite variance.

Tables 1–15 of the supplementary material display the $k$-term expectation, denoted $\mathrm{E}_k$, and the bound for the remainder using Monte Carlo methods with $10^5$ drawings (and same seed) from the above distributions. The precision of the $k$-term expansion is measured by the ratio $\mathrm{E}_k \,/\, \mathrm{E}_*(x)$, where $\mathrm{E}_*(x)$ is the Monte Carlo estimate of $\mathrm{E}(x)$.The $k$-term and bound's expectations can be calculated by Monte Carlo methods or numerical integration. However, numerical integration is time consuming and can be less accurate as there are many spikes in the expression of the $k$-term expansion and the remainder's bound which can be easily missed by the numerical algorithm.

Each table contains the results for the raw and the centered expansions for $k = 2,3,4$ and $m =$ 1, 10, 50, 100, 500 and 1,000. The case of the well-behaved log-normal stands in contrast to the case of the fat-tailed log-gamma whose variance does not exist. Nevertheless, the tables show that even in the fattest-tailed case of small $\lambda$ and large $\nu$, the precision of our formulae is very good. Even the 2-term expansion ($k = 2$) is accurate, especially when we choose $m$ not too small. All tables indicate that choosing a large $m$ increases the accuracy of both the expansion and bound. On the whole, centered expansions are more accurate, but we need to take $m > 1$.

Unreported calculations show that the expansions we derive using complex numbers are vastly more accurate than the ones that do not use them, like the introductory (1)–(2). To illustrate this point, take $y = 10$, then $x := e^y = 22,026.4657$. First, consider the raw expansion (1) with no complex numbers or $m$ involved. The 2-term expansion gives only $x \approx 61$ and we would need a 30-term expansion to obtain a good approximation of $x$. Second, we consider the complex (7) for $m = 1$ and different values of $k$. The 2-term expansion gives $9,447.5$ and the corresponding bound for the remainder is $83,879$, but here we only need a 12-term expansion to obtain a good approximation of $x$ with a bound of 24.83. Finally, we consider the same expansion as in (7) for fixed $k = 2$ and different values of $m$. For $m = 100$, the expansion is $22,396$ and the bound is 532. Taking $m = 10,000$, the 2-term expansion is extremely accurate and the bound for the remainder is very precise, namely 0.0519.

## 4 The transformation-based naive bootstrap, as an application of our method

The purpose of this section is to illustrate the usefulness of our expansion to solve the problem of invalidity and bad performance of the naive bootstrap when the data are fat-tailed. The traditional approach is not valid for $\bar{x}_n$ if the distribution of the $x$'s is in the domain of attraction of the stable laws and $\text{var}(x) = \infty$, as shown in Athreya (1987), Knight (1989), and Hall (1990). The $m$ out of $n$ bootstrap is similar to the naive bootstrap except that the bootstrap sample size $m$ is smaller than $n$, with $m$ satisfying the conditions $m/n \to 0$, $m \to \infty$, $n \to \infty$ which guarantee the asymptotic validity of this bootstrap. Subsampling is similar to the $m$ out of $n$ bootstrap, except that the resamplings are without replacement.

## 4.1 Our modification of the naive bootstrap and proof of its validity

We propose a simple bootstrap CI for $\mathrm{E}(x)$ where $x \in \mathbb{R}_+$ and we assume that $x$ has a distribution in the domain of attraction of the stable laws and $\mathrm{var}(x) = \infty$. The case of $x \in \mathbb{R}$ can be similarly handled by considering the negative and positive parts separately below, and the choice of $\mathrm{E}(x)$ instead of other moments is meant to keep complexity at a minimum for our application. The fact that $\mathrm{E}(x)$ is finite guarantees that all the moments of $y := \log(x)$ exist, and so do the expectations of the $k$-term sum and the remainder's bound.

Letting $z := \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \operatorname{Re}(\xi_k^{\mathrm{i}m})$, we have $x = z + R^{\mathrm{c}}_{x,k}$ and the remainder has the bound $B^{\mathrm{c}}_{x,k}$ given in Corollary 1. The same derivations will also apply to the uncentered expansions. By applying the triangle inequality twice,

$$x \in \left[ |z| - B^{\mathrm{c}}_{x,k}, |z| + B^{\mathrm{c}}_{x,k} \right] \tag{16}$$

whose endpoints can be resampled to build conservative CIs for $\mathrm{E}(x)$. To do this, consider an i.i.d. sample $x_1, \cdots, x_n$ with sample mean $\bar{x}_n$, and compute the following quantities

$$\bar{z}_n^+ := \frac{1}{n} \sum_{j=1}^{n} |z_j| \,, \quad B^{\mathrm{c}}_{\bar{x},k} := \frac{1}{n} \sum_{j=1}^{n} B^{\mathrm{c}}_{x_j,k}.$$

By the triangle inequality,

$$\varkappa_{1,n} \le \bar{x}_n \le \varkappa_{2,n}, \quad \text{with } \varkappa_{1,n} := \bar{z}_n^+ - B^{\mathrm{c}}_{\bar{x},k} \text{ and } \varkappa_{2,n} := \bar{z}_n^+ + B^{\mathrm{c}}_{\bar{x},k}. \tag{17}$$

As seen from (16), if $B^{\mathrm{c}}_{x,k}$ is too large then the resulting CI will be too conservative. While, if $k = \infty$ or $m = \infty$ for any given $n$, then $B^{\mathrm{c}}_{x,k}$ vanishes and $z$ coincides with $x$ and we are back to the original invalid naive bootstrap. Thus $k, m$ have to be finite for any finite $n$, and their value chosen depending on the thickness of the tail of $x$, as we will see how this tail affects finite-sample performance in the next subsection. As $n \to \infty$, we will allow $m$ to increase too to achieve an existence condition for the moments of $B_{x,k}^{\mathrm{c}}$, and we turn to this now.

In Proposition 3 below, we prove that the naive bootstrap of $\varkappa_{1,n}$ and $\varkappa_{2,n}$ based on resampling with replacement from the empirical distribution function (EDF) of the $y$'s (or equivalently from the joint EDF of $i_y$ and $\zeta$) is asymptotically valid. More specifically, we show that the naive bootstrap is valid for $|z|$ and $B_{x,k}^{\mathrm{c}}$ (both are in terms of powers of the well-behaved $y$), thus for their linear combination as in (17). Note that the naive bootstrap is not valid for $R_{x,k}^{\mathrm{c}}$, which is just identical to $x - z$ and thus inherits the fat tail of $x$.

**Proposition 3** Let $k$ be finite. There exists an arbitrarily-large $m$ for which

$$\Pr\left(\sup_{v\in\mathbb{R}}\left|\Pr\left(n^{1/2}\left(\varkappa_{i,n}^{\mathrm{b}}-\varkappa_{i,n}\right)\leq v \mid \boldsymbol{y}\right)-\Pr\left(n^{1/2}\left(\varkappa_{i,n}-\mathrm{E}(\varkappa_{i,n})\right)\leq v\right)\right|>\epsilon\right)\to 0, \quad (18)$$

for $i = 1, 2$ and all $0 < \epsilon < 1$ as $n \to \infty$, where $\varkappa_{i,n}^{\mathrm{b}}$ is the naive bootstrap version of $\varkappa_{i,n}$ and $\boldsymbol{y} := (y_1, \cdots, y_n)'$.

It follows from (17) and Proposition 3 that the upper limit of the CI based on (16) for $\mathrm{E}(x)$ is at most equal to the upper limit of the CI for $\mathrm{E}\,|z| + \mathrm{E}(B_{x,k}^{\mathrm{c}})$ since $\hat{q}_{\bar{x},\alpha/2} \leq \hat{q}_{\varkappa_2,\alpha/2}^{\mathrm{b}}$, where $\hat{q}_{\bar{x},\alpha/2}$ and $\hat{q}_{\varkappa_2,\alpha/2}^{\mathrm{b}}$ are the estimates of the $\alpha/2$ quantiles of the distribution of $n^{1/2}\bar{x}_n$ and $n^{1/2}\varkappa_{2,n}^{\mathrm{b}}$, respectively. Also the lower limit of the CI for $\mathrm{E}\,|z| - \mathrm{E}(B_{x,k}^{\mathrm{c}})$ is at most equal to the lower limit of the CI for $\mathrm{E}(x)$ since $\hat{q}_{\varkappa_1,1-\alpha/2}^{\mathrm{b}} \leq \hat{q}_{\bar{x},1-\alpha/2}$, where $\hat{q}_{\varkappa_1,1-\alpha/2}^{\mathrm{b}}$ is the estimate of the $1-\alpha/2$ quantile of the distribution of $n^{1/2}\varkappa_{1,n}^{\mathrm{b}}$. This gives the following result.

**Theorem 2** *Let $k$ be finite. There exists an arbitrarily-large $m$ for which the basic bootstrap confidence intervals*

$$\left[\varkappa_{1,n} - \left(\hat{q}_{\varkappa_1,1-\alpha/2}^{\mathrm{b}} - \varkappa_{1,n}\right)\ ,\ \varkappa_{2,n} - \left(\hat{q}_{\varkappa_2,\alpha/2}^{\mathrm{b}} - \varkappa_{2,n}\right)\right]$$

*provide a conservative $1-\alpha$ two-sided CI for $\mathrm{E}(x)$ in finite samples, and have asymptotically* *the correct coverage.*

The problem with the naive bootstrap of $\bar{x}_n$ (based on resampling from the EDF of the $x$'s) is that it fails to model accurately the relation among the extreme values in the sample (the higher-order terms in the expansion of $\bar{x}_n$). The probability that the largest extremes in a bootstrap sample are equal does not converge to zero, as is the case with the extremes in the original sample; see Hall and Yao (2003). In the case of our naive bootstrap for $\varkappa_{i,n}$ with $m$ finite, this probability can be made arbitrarily close to zero in $B^{\mathrm{c}}_{\bar{x},k}$: this bound is a function of the $(k+1)$-th term in the expansion, the higher-order terms being collectively bounded by a finite power of a well-behaved variate.

## 4.2 Simulation results

To complement the asymptotics of Theorem 2, we now propose a method to choose our $m$ in finite samples. The method is based on a response surface as a function of the sample size $n$, the tail exponent $\lambda$, the nominal level $\alpha$, and the number of terms $k$ in the expansion. Response surfaces have been used in various statistical and econometric applications; e.g. Mittnik, Rachev, and Kim (1998) in the context of fat tails. Response surfaces are regressions which are determined from simulated draws, here implying an optimal choice of the parameter of interest $m$. The simulated draws are from a standard Pareto Par($\lambda$) density

$$f_x(u) = \lambda u^{-\lambda-1}$$

for $x = u \in (1, \infty)$ and where $\lambda$ is the tail index (exponent) with $1 < \lambda < 2$ for $\mathrm{E}(x)$ to exist but $\mathrm{E}(x^2)$ infinite. This choice is motivated by the fact that distributions in the domain of attraction of the stable laws with infinite variance have Pareto-like tails; see Feller (1971), p.576. At the end of this section, we verify the accuracy of the performance of this response surface in selecting $m$ when distributions other than the Pareto hold, even though the surface has been optimized for Pareto data, thus confirming our large-sample invariance argument.

To obtain our response surface, we express $\log(m)$ as a transcendental function in $\alpha^{-1}$, $k^{-1}$, $\lambda^{-1}$, $\lambda^{-2}$, $\log(n)$, which is estimated by nonlinear least squares using draws from a Pareto distribution with $\lambda = 1.1, 1.2, \cdots, 2$. For each $n = 100, 200, \cdots, 1000$, $\alpha = 0.01, 0.05, 0.10$ and $k = 1, 2$ we have selected the value of $m$ that guarantees a conservative CI for $\mathrm{E}(x)$ that is closest to the nominal coverage. We have not considered larger values for $k$, as $m$ can play a similar role as $k$ gets large; see Section 3. We take $n^{\mathrm{b}} = 399$ bootstrap replications to guarantee that $\left(n^{\mathrm{b}} + 1\right)(1 - \alpha/2)$ and $\left(n^{\mathrm{b}} + 1\right)\alpha/2$ are integers; see Davison and Hinkley (2009), p.18.

After eliminating the statistically insignificant terms, we have arrived at the following response surface

$$m \approx c_1 + \exp\left(c_2 + \frac{c_3}{k} + \frac{c_4}{k}\log n + \frac{c_5}{k\lambda} + \frac{c_6 + c_7\alpha^{-1}}{k\lambda^2}\right). \tag{19}$$

The estimates of the parameters of (19) and their $t$-statistics (based on White's heteroskedasticity-consistent variance matrix estimator) are given in Table 1. The adjusted-$R^2$ of the regression is 97%. The predicted $m$ is given by the integer part of the function (19) for specific values of $\alpha$, $k$, $\lambda$, and $n$. Note that it implies $m \approx O(n^{c_4/k})$, which leads to $\operatorname{plim} i_y = 0$ as required for the existence condition (21).

We illustrate the behavior of our naive bootstrap from Theorem 2 and the performance of our response surface (19) in a simulation study. We consider $\mathrm{Par}(\lambda)$, Fréchet distribution with c.d.f. $F_x(u) = \exp(-u^{-\lambda})$, Burr distribution with c.d.f. $F_x(u) = 1-(1+u^{\beta})^{-\lambda/\beta}$ where $\beta > 0$, and the log-gamma distribution (15) with $\nu = 1$. The distributions are in the domain of attraction of the stable laws.

Table 2 gives the coverage probabilities for $\mathrm{E}(x)$ based on the naive bootstrap of $\varkappa_{i,n}$ when the data are drawn from the Pareto, using the raw version of the expansion. We see that a low $\lambda$ (fat tail) is more challenging than a high $\lambda$. But as $n \to \infty$, all our CIs have the required conservative coverage as Theorem 2 implies. We can also see that there are choices of $m$ and $k$ to fine-tune or reduce the degree of conservatism of the CIs in finite samples, which we summarize as follows:

1. For extreme quantiles, such that 99%, choose $k = 1$. For $\alpha = 0.01$ and $k = 1$, the optimal $m$ predicted by our response surface corresponds to the shaded boxes in the top panel of Table 2. As can be seen from the table, these choices of $m$ give very satisfactory CIs for $\mathrm{E}(x)$. For the case $\lambda = 1.1$ and $n = 100, 300$ and for the case $\lambda = 1.3$ and $n = 100$, the predicted $m$ is zero or $-1$ but we take $m = 1$ because $m \geq 1$ by assumption. These cases are very hard to handle by any method (in particular for small $n$), as they are very close to the non-existence of even the first moment of the distribution.

2. Considering progressively smaller quantiles, 95% and 90%, we move into the body of the distribution and choose $k = 2$. For $\alpha = 0.05, 0.10$ and $k = 2$, the optimal $m$ predicted by our response surface corresponds to the shaded boxes in the middle panel of Table 2 which shows an excellent performance again, except for only the single case of $\alpha = 0.05$, $\lambda = 1.1$ (very fat tail), $n = 100$ (small sample); a performance which improves quickly when either $\lambda$ or $n$ increase.

In the bottom panel of Table 2, we report the coverage probabilities for $\mathrm{E}(x)$ based on the $m$ out of $n$ bootstrap and subsampling with the choice of the bootstrap sample size given by the method of Bickel and Sakov (2008), p.971 where their $q$ is 0.75.[2] As can be seen from the table, these coverage probabilities are in sharp contrast to the ones of the simple naive bootstrap from Theorem 2 which performs much better. The CIs of the two bootstrap competitors are somewhat improved when $\lambda = 1.5$ (and even conservative), but sensitive to the choice of the tuning parameters used in the method of Bickel and Sakov (2008).

Using data from the Burr, Fréchet, and log-gamma distributions, respectively, Tables 3–5 repeat the exercise of Table 2 with the unchanged response surface (19) optimized under Pareto data (justified by the asymptotic invariance argument). As can be concluded from these tables, our approach using the naive bootstrap gives excellent results. Furthermore,

[2] We considered $j = 0, 1, \cdots, 9$ for $n = 100$, $j = 0, 1, \cdots, 16$ for $n = 300$, and $j = 0, 1, \cdots, 17$ for $n = 900$.

the $m$ out of $n$ bootstrap and subsampling give coverage probabilities well below the nominal ones, except for $\lambda = 1.5$ and large $n$. But, as mentioned above, these bootstraps are very sensitive to the choice of the bootstrap sample size. Similar conclusions about the behaviour of the $m$ out of $n$ bootstrap and subsampling can be drawn from Hall and Jing (1998), Romano and Wolf (1999), the supplementary material in Ibragimov and Muller (2010), Cornea-Madeira and Davidson (2015).

# 5 Concluding comments

By extending the approach of bounding terms in expansions of characteristic functions, we have provided expansions and bounds for expectations of a variate $x$ in terms of the expectations of a related variate $y$. We have then illustrated the accuracy of the expansions and bounds by simulating distributions, including ones whose higher order moments do not exist. Finally, we have shown how to apply our formulae to fix the problem of bootstrapping the mean of (asymmetric) variates that have infinite variance. Even though we have used only the naive bootstrap along with our expansions, our results are very good compared to the performance of various other bootstrap modifications that have tried to fix the problem. It shows the potential of our expansions for the bootstrap and the other applications cited in the introduction.

# Appendix: Proofs

**Proof of Proposition 1.** For any complex $\psi := a + \mathrm{i}b$, with $a, b$ real and $\theta := \arg(1+\psi) \in [-\pi, \pi)$, we have

$$(1+\psi)^{\mathrm{i}m} = |1+\psi|^{\mathrm{i}m} \exp(-m\theta), \tag{20}$$

where $|1+\psi|^{\mathrm{i}m} = \mathrm{e}^{\mathrm{i}m\log|1+\psi|}$ has modulus 1. Hence,

$$\begin{aligned} h(\psi) &= \left|(1+a+b\mathrm{i})^{\mathrm{i}m} - 1\right| = \left|\exp\left(\mathrm{i}m\log\sqrt{(1+a)^2+b^2} - m\theta\right) - 1\right| \\ &= \left|\exp\left(\mathrm{i}m\log|(1+a)\sec\theta| - m\theta\right) - 1\right| \end{aligned}$$

by $b = (1+a)\tan\theta$, and

$$\begin{aligned} h(\psi)^2 &= \left|\mathrm{e}^{-m\theta}\cos\left(m\log|(1+a)\sec\theta|\right) + \mathrm{i}\mathrm{e}^{-m\theta}\sin\left(m\log|(1+a)\sec\theta|\right) - 1\right|^2 \\ &= \left(\mathrm{e}^{-m\theta}\cos\left(m\log|(1+a)\sec\theta|\right) - 1\right)^2 \\ &+ \mathrm{e}^{-2m\theta}\sin^2\left(m\log|(1+a)\sec\theta|\right) \\ &= 1 - 2\mathrm{e}^{-m\theta}\cos\left(m\log|(1+a)\sec\theta|\right) + \mathrm{e}^{-2m\theta}. \end{aligned}$$

Optimizing $h(\psi)^2$ with respect to $a$ gives the first-order condition

$$\sin\left(m\log\left|\frac{1+a}{\cos\theta}\right|\right) = 0$$

yielding the concentrated

$$h(\psi)^2 = 1 \mp 2\mathrm{e}^{-m\theta} + \mathrm{e}^{-2m\theta}$$

which is maximized by the corner solution $\theta = -\pi$ and by

$$\cos\left(m\log\left|\frac{1+a}{\cos\theta}\right|\right) = -1.$$

Hence, with $j \in \mathbb{Z}$, the solution for $a$ can be written as

$$\log\left|\frac{1+a}{\cos(-\pi)}\right| = (2j+1)\,\pi/m$$

or $|1+a| = \exp\left((2j+1)\,\pi/m\right)$. Since $\theta = \arg(1+a+b\mathrm{i}) = -\pi$, we have that $1+a<0$ hence

$$a = -1 - \exp\left((2j+1)\,\pi/m\right)$$

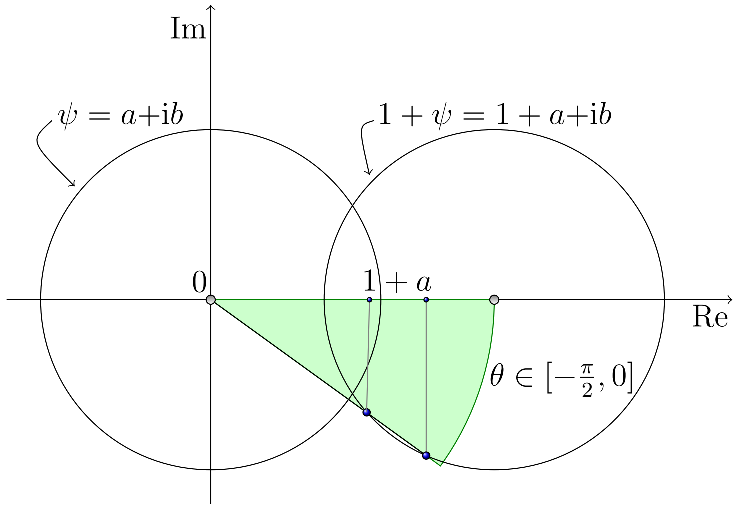


Figure 1: Illustration of the optimal solution in Proposition 2.

and the result follows. □

**Proof of Proposition 2.** We maximize $h(\psi)^2$ as in Proposition 1, but this time subject to the additional condition that $|\psi| \leq |\psi_0|$. However, now $1 + a \geq 0$ since $|a| \leq |\psi_0| \leq 1$, and the optimal solution will satisfy $\theta \in [-\pi/2, 0]$ and $a \leq 0$. Visualize the solution as the intersection point (in the lower half plane) of a ray of angle $\theta$ from the origin with a circle of radius $|\psi_0|$ centered around 1; see Figure 1. This optimization is easiest to do in terms of $|\psi|$ and $\theta$. To this end, using the definitions $|\psi|^2 = a^2 + b^2$ and $b^2 = (1+a)^2 \tan^2\theta$ gives a quadratic equation for $a$ whose solutions are

$$a = -\sin^2\theta \pm \sqrt{|\psi|^2 \cos^2\theta - \sin^2\theta\cos^2\theta}.$$

For $a \leq 0$, the top solution ($+\sqrt{\ }$) requires $|\psi|^2 \in [\sin^2\theta, \tan^2\theta]$ and the bottom ($-\sqrt{\ }$) just $|\psi|^2 \geq \sin^2\theta$. For $a \in [-|\psi|, 0]$, we need further that

$$\pm\sqrt{|\psi|^2 \cos^2\theta - \sin^2\theta\cos^2\theta} \geq \sin^2\theta - |\psi|.$$

Now $\sin^2\theta \le |\psi|^2 \le |\psi|$ since $|\psi| \le 1$, so the RHS is nonpositive: the top restriction always holds and the bottom one requires $-\sin^2\theta\,(|\psi|-1)^2 \le 0$ which always holds. As a result, $a \in [-|\psi|, 0]$ imposes no further restrictions. In either case, the objective function is

$$\begin{aligned} h(\psi)^2 &= 1 - 2\mathrm{e}^{-m\theta}\cos\left(m\log|(1+a)\sec\theta|\right) + \mathrm{e}^{-2m\theta} \\ &= 1 - 2\mathrm{e}^{-m\theta}\cos\left(m\log\left(\cos\theta \pm \sqrt{|\psi|^2 - \sin^2\theta}\right)\right) + \mathrm{e}^{-2m\theta} \end{aligned}$$

since $1+a \ge 0$ and $\cos\theta \ge 0$. This is a function of $\varphi := |\psi|^2$ and $\theta$, which is to be optimized subject to $\varphi \le |\psi_0|^2$. The augmented function is

$$1 - 2\mathrm{e}^{-m\theta}\cos\left(m\log\left(\cos\theta \pm \sqrt{\varphi - \sin^2\theta}\right)\right) + \mathrm{e}^{-2m\theta} - l\left(|\psi_0|^2 - \varphi\right),$$

leading to the Kuhn-Tucker conditions

$$l = -\mathrm{e}^{-m\theta}\frac{m\sin\left(m\log\left(\cos\theta \pm \sqrt{\varphi - \sin^2\theta}\right)\right)}{\pm\sqrt{\varphi - \sin^2\theta}\left(\cos\theta \pm \sqrt{\varphi - \sin^2\theta}\right)} \le 0, \qquad l\left(|\psi_0|^2 - \varphi\right) = 0,$$

and

$$\begin{aligned} &\cos\left(m\log\left(\cos\theta \pm \sqrt{\varphi - \sin^2\theta}\right)\right) \\ \mp\ &\frac{\sin\theta\sin\left(m\log\left(\cos\theta \pm \sqrt{\varphi - \sin^2\theta}\right)\right)}{\sqrt{\varphi - \sin^2\theta}} \\ =\ &\mathrm{e}^{-m\theta}. \end{aligned}$$

If $l = 0$, the last equation becomes $1 = \mathrm{e}^{-m\theta}$, hence $\theta = 0$ which does not lead to a maximum when substituted into $h$. Hence, $l \ne 0$ and the constraint $|\psi| = |\psi_0|$ is binding, which implies the monotonicity of $h$ in $|\psi_0|$.

Since $l \ne 0$ for the optimum, $\sin(m\log) \ne 0$ implying that $\cos(m\log) \ne 1$ unlike in Proposition 1. The objective function cannot be simplified like before, and the first-order condition on $\theta$ seems intractable. We resort instead to bounding the components of $h$. The exponentials' argument is bounded by $|\psi_0|^2 \geq \sin^2\theta$, hence $-\theta \leq \sin^{-1}|\psi_0|$ (since $\theta \in [-\pi/2, 0]$). As for the remaining component of $h$, consider the transformation $s = \pm\sqrt{\varphi - \sin^2\theta}$ hence

$$\theta = -\sin^{-1}\sqrt{\varphi - s^2} = -\cos^{-1}\sqrt{1 - \varphi + s^2}$$

and

$$\begin{aligned} h(\psi)^2 &= 1 - 2\mathrm{e}^{m\sin^{-1}\sqrt{\varphi - s^2}}\cos\left(m\log\left(\cos\sin^{-1}\sqrt{\varphi - s^2} + s\right)\right) \\ &+ \mathrm{e}^{2m\sin^{-1}\sqrt{\varphi - s^2}} \\ &= 1 - 2\mathrm{e}^{m\cos^{-1}\sqrt{1-\varphi+s^2}}\cos\left(m\log\left(\sqrt{1-\varphi+s^2} + s\right)\right) \\ &+ \mathrm{e}^{2m\cos^{-1}\sqrt{1-\varphi+s^2}}, \end{aligned}$$

where the sign of $s$ affects only the logarithmic component. Maximizing $-\cos(m\log(\sqrt{1-\varphi+s^2}+ s))$ subject to $s \in [-\sqrt{\varphi}, \sqrt{\varphi}]$ gives an interior solution of $+1$ when

$$\varphi \in \left[\left(\mathrm{e}^{(2j+1)\pi/m} - 1\right)^2, \left(\mathrm{e}^{(2j+1)\pi/m} + 1\right)^2\right],$$

where the upper bound is always bigger than 1 but the lower bound is minimized (for the interval to cover all interior solutions) by choosing $j = -1$ for any given $m$, and this latter bound is

$$\varphi = \left(\mathrm{e}^{-\pi/m} - 1\right)^2$$

or $|\psi_0| = 1 - \mathrm{e}^{-\pi/m}$, giving the solution

$$s = \frac{-1 + \varphi + \mathrm{e}^{-2\pi/m}}{2\mathrm{e}^{-\pi/m}} = \frac{\varphi}{2}\mathrm{e}^{\pi/m} + \sinh(-\pi/m)$$

and $-\cos(\log(\sqrt{1-\varphi+s^2}+s))=+1$, hence the monotonic bound

$$h(\psi)^2 \le 1+2\mathrm{e}^{m\sin^{-1}|\psi_0|}+\mathrm{e}^{2m\sin^{-1}|\psi_0|}=\left(1+\mathrm{e}^{m\sin^{-1}|\psi_0|}\right)^2.$$

Otherwise, with $-\cos(m\log\cdot)<1$, the largest corner solution is obtained when $s=-\sqrt{\varphi}<0$ (the bottom solution for $a$) and

$$-\cos\left(m\log\left(\sqrt{1-\varphi+s^2}+s\right)\right)=-\cos\left(m\log\left(1-\sqrt{\varphi}\right)\right),$$

hence

$$h(\psi)^2 \le 1-2\mathrm{e}^{m\sin^{-1}|\psi_0|}\cos\left(m\log\left(1-|\psi_0|\right)\right)+\mathrm{e}^{2m\sin^{-1}|\psi_0|}$$

and the two bounds on $h$ coincide at the switching point $|\psi_0|=1-\mathrm{e}^{-\pi/m}$. The monotonicity of this bound follows by differentiating then solving for the zeros, which shows that there are none in $(0,1-\mathrm{e}^{-\pi/m})$. □

**Proof of Proposition 3.** Since Proposition 1 bounds $H$ globally by $1+\mathrm{e}^{m\pi}$, establishing the existence of the moments of $z$ and $B^{\mathrm{c}}_{x,k}$ boils down to showing that the non-negative $\mathrm{e}^{2\pi m i_y}\left|\xi_k^{\mathrm{im}}\right|$ has finite $p$-th order moments ($p\in\mathbb{N}$ and finite) by appropriate choice of $m$. To achieve this, a sufficient condition (by the Cauchy-Schwarz inequality) is that

$$\mathrm{E}\left(\mathrm{e}^{2\pi p m i_y}\right)<\infty. \tag{21}$$

Recall that $(y-\mathrm{E}\,y)/m\equiv 2\pi i_y+\zeta$ where $\zeta\in(-\pi,\pi]$ and $i_y$ is a discrete variate. Choosing an arbitrarily large $m$ such that $\operatorname{plim} i_y=0$ as $n\to\infty$ ensures $\mathrm{E}(\mathrm{e}^{2\pi p m i_y})$ tends to 1 since $y$ is tight (it has an exponentially-decaying distribution independent of $n$). Note that the idea used here is to shift the randomness of $y$ to $\zeta$ (with probability 1) because it does not figure in the required existence condition (21) as a result of it entering $B^{\mathrm{c}}_{x,k}$ only through the bounded $H$ function.

Under no further assumptions, using the Cramér-Wold device we can then show that both $\varkappa_{1,n}$ and $\varkappa_{2,n}$ are asymptotically normal. Since these statistics are essentially usual sample means with continuous asymptotic distributions, (18) follows from Theorem 2.1 of Bickel and Freedman (1981). □

Table 1: Parameter estimates of fitted response surface.

| | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ | $c_6$ | $c_7$ |
|---|---|---|---|---|---|---|---|
| Estimate | -1.59 | 0.98 | -4.14 | 0.72 | 7.04 | -6.26 | -0.02 |
| $t$-statistic | -3.70 | 7.88 | -9.69 | 20.76 | 5.65 | -6.66 | -9.76 |

Table 2: Coverage probabilities for $E(x)$, $x \sim Par(\lambda)$.

**Transformation-based naive bootstrap, $k = 1$**

| | n = 100 | | | | n = 300 | | | | n = 900 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.98 | 0.98 | 0.99 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.58 | 0.58 | 0.60 | 2 | 0.91 | 0.91 | 0.91 | 2 | 1.00 | 1.00 | 1.00 |
| | 3 | 0.48 | 0.49 | 0.54 | 3 | 0.53 | 0.56 | 0.61 | 3 | 0.67 | 0.69 | 0.73 |
| | 4 | 0.51 | 0.53 | 0.57 | 4 | 0.60 | 0.63 | 0.67 | 4 | 0.70 | 0.73 | 0.78 |
| 1.3 | 1 | 0.99 | 0.99 | 0.99 | 1 | 1.00 | 1.00 | 1.00 | 8 | 0.98 | 0.99 | 1.00 |
| | 2 | 0.88 | 0.90 | 0.93 | 2 | 0.96 | 0.97 | 0.98 | 10 | 0.97 | 0.98 | 0.99 |
| | 3 | 0.88 | 0.90 | 0.93 | 3 | 0.97 | 0.97 | 0.99 | 12 | 0.96 | 0.97 | 0.98 |
| | 4 | 0.87 | 0.89 | 0.92 | 4 | 0.96 | 0.97 | 0.98 | 16 | 0.93 | 0.95 | 0.97 |
| 1.5 | 1 | 0.99 | 0.99 | 0.99 | 4 | 0.99 | 0.99 | 1.00 | 14 | 0.98 | 0.99 | 0.99 |
| | 2 | 0.96 | 0.97 | 0.98 | 5 | 0.99 | 0.99 | 1.00 | 18 | 0.97 | 0.98 | 0.99 |
| | 3 | 0.93 | 0.95 | 0.97 | 6 | 0.98 | 0.99 | 0.99 | 19 | 0.96 | 0.97 | 0.99 |
| | 4 | 0.86 | 0.88 | 0.91 | 7 | 0.97 | 0.98 | 0.99 | 25 | 0.94 | 0.96 | 0.98 |

**Transformation-based naive bootstrap, $k = 2$**

| | n = 100 | | | | n = 300 | | | | n = 900 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.90 | 0.90 | 0.92 | 1 | 0.97 | 0.98 | 0.99 | 3 | 0.99 | 0.99 | 1.00 |
| | 2 | 0.82 | 0.84 | 0.86 | 2 | 0.95 | 0.96 | 0.97 | 4 | 0.97 | 0.98 | 0.98 |
| | 3 | 0.77 | 0.79 | 0.81 | 3 | 0.92 | 0.93 | 0.94 | 5 | 0.93 | 0.94 | 0.95 |
| | 4 | 0.66 | 0.67 | 0.69 | 4 | 0.85 | 0.87 | 0.88 | 6 | 0.85 | 0.86 | 0.88 |
| 1.3 | 1 | 0.98 | 0.99 | 0.99 | 1 | 1.00 | 1.00 | 1.00 | 4 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.96 | 0.97 | 0.98 | 2 | 1.00 | 1.00 | 1.00 | 5 | 0.98 | 0.99 | 0.99 |
| | 3 | 0.90 | 0.91 | 0.93 | 3 | 0.99 | 0.99 | 0.99 | 6 | 0.95 | 0.96 | 0.97 |
| | 4 | 0.82 | 0.83 | 0.86 | 4 | 0.95 | 0.96 | 0.97 | 7 | 0.91 | 0.93 | 0.95 |
| 1.5 | 1 | 0.99 | 0.99 | 0.99 | 2 | 1.00 | 1.00 | 1.00 | 5 | 0.99 | 0.99 | 0.99 |
| | 2 | 0.97 | 0.98 | 0.99 | 3 | 0.99 | 0.99 | 0.99 | 6 | 0.97 | 0.97 | 0.98 |
| | 3 | 0.92 | 0.93 | 0.95 | 4 | 0.96 | 0.97 | 0.98 | 7 | 0.93 | 0.95 | 0.97 |
| | 4 | 0.86 | 0.88 | 0.91 | 5 | 0.91 | 0.93 | 0.95 | 8 | 0.90 | 0.92 | 0.95 |

**$m$ out of $n$ bootstrap**

| | n = 100 | | | n = 300 | | | n = 900 | | |
|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 1.1 | 0.24 | 0.25 | 0.26 | 0.25 | 0.26 | 0.28 | 0.27 | 0.29 | 0.30 |
| 1.3 | 0.69 | 0.73 | 0.78 | 0.76 | 0.80 | 0.86 | 0.82 | 0.86 | 0.92 |
| 1.5 | 0.88 | 0.90 | 0.94 | 0.94 | 0.96 | 0.98 | 0.97 | 0.99 | 1.00 |

**Subsampling**

| | n = 100 | | | n = 300 | | | n = 900 | | |
|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 1.1 | 0.26 | 0.27 | 0.28 | 0.27 | 0.28 | 0.30 | 0.28 | 0.30 | 0.32 |
| 1.3 | 0.71 | 0.74 | 0.78 | 0.78 | 0.81 | 0.87 | 0.84 | 0.88 | 0.93 |
| 1.5 | 0.87 | 0.90 | 0.93 | 0.94 | 0.96 | 0.98 | 0.97 | 0.99 | 1.00 |

Table 3: Coverage probabilities for $\mathrm{E}(x)$, $x \sim \mathrm{Burr}(\lambda, \beta)$, with $\beta = 2$.

Transformation-based naive bootstrap, $k = 1$

| | $n = 100$ | | | | $n = 300$ | | | | $n = 900$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.98 | 0.98 | 0.98 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.57 | 0.58 | 0.59 | 2 | 0.91 | 0.91 | 0.91 | 2 | 1.00 | 1.00 | 1.00 |
| | 3 | 0.47 | 0.49 | 0.53 | 3 | 0.54 | 0.57 | 0.62 | 3 | 0.68 | 0.70 | 0.74 |
| | 4 | 0.50 | 0.53 | 0.57 | 4 | 0.60 | 0.63 | 0.67 | 4 | 0.70 | 0.73 | 0.78 |
| 1.3 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 | 8 | 0.99 | 0.99 | 1.00 |
| | 2 | 0.87 | 0.89 | 0.92 | 2 | 0.96 | 0.97 | 0.98 | 12 | 0.96 | 0.97 | 0.98 |
| | 3 | 0.87 | 0.89 | 0.92 | 3 | 0.96 | 0.97 | 0.98 | 16 | 0.93 | 0.94 | 0.97 |
| | 4 | 0.87 | 0.89 | 0.91 | 4 | 0.96 | 0.97 | 0.98 | 20 | 0.90 | 0.92 | 0.95 |
| 1.5 | 1 | 0.99 | 0.99 | 1.00 | 3 | 0.99 | 1.00 | 1.00 | 14 | 0.98 | 0.99 | 0.99 |
| | 2 | 0.96 | 0.97 | 0.98 | 4 | 0.99 | 0.99 | 1.00 | 20 | 0.96 | 0.97 | 0.99 |
| | 3 | 0.95 | 0.96 | 0.97 | 5 | 0.99 | 0.99 | 1.00 | 25 | 0.94 | 0.95 | 0.97 |
| | 4 | 0.93 | 0.95 | 0.97 | 6 | 0.98 | 0.98 | 0.99 | 30 | 0.92 | 0.94 | 0.96 |

Transformation-based naive bootstrap, $k = 2$

| | $n = 100$ | | | | $n = 300$ | | | | $n = 900$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.89 | 0.90 | 0.92 | 1 | 0.97 | 0.98 | 0.99 | 3 | 0.99 | 0.99 | 1.00 |
| | 2 | 0.82 | 0.84 | 0.86 | 2 | 0.95 | 0.96 | 0.97 | 4 | 0.98 | 0.98 | 0.98 |
| | 3 | 0.76 | 0.77 | 0.79 | 3 | 0.93 | 0.94 | 0.95 | 5 | 0.93 | 0.93 | 0.95 |
| | 4 | 0.66 | 0.67 | 0.69 | 4 | 0.85 | 0.86 | 0.88 | 6 | 0.85 | 0.86 | 0.88 |
| 1.3 | 1 | 0.98 | 0.99 | 0.99 | 1 | 1.00 | 1.00 | 1.00 | 4 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.96 | 0.97 | 0.98 | 2 | 1.00 | 1.00 | 1.00 | 5 | 0.98 | 0.99 | 0.99 |
| | 3 | 0.89 | 0.91 | 0.92 | 3 | 0.99 | 0.99 | 0.99 | 6 | 0.95 | 0.96 | 0.97 |
| | 4 | 0.81 | 0.83 | 0.86 | 4 | 0.95 | 0.95 | 0.97 | 7 | 0.91 | 0.92 | 0.94 |
| 1.5 | 1 | 0.99 | 1.00 | 1.00 | 2 | 1.00 | 1.00 | 1.00 | 6 | 0.96 | 0.97 | 0.98 |
| | 2 | 0.97 | 0.98 | 0.98 | 3 | 0.99 | 0.99 | 1.00 | 7 | 0.94 | 0.95 | 0.97 |
| | 3 | 0.92 | 0.93 | 0.95 | 4 | 0.96 | 0.97 | 0.98 | 8 | 0.90 | 0.92 | 0.95 |
| | 4 | 0.85 | 0.87 | 0.90 | 5 | 0.92 | 0.93 | 0.95 | 9 | 0.87 | 0.90 | 0.94 |

$m$ out of $n$ bootstrap

| | $n = 100$ | | | $n = 300$ | | | $n = 900$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 1.1 | 0.25 | 0.26 | 0.27 | 0.28 | 0.27 | 0.29 | 0.27 | 0.28 | 0.30 |
| 1.3 | 0.71 | 0.75 | 0.81 | 0.77 | 0.82 | 0.88 | 0.83 | 0.87 | 0.93 |
| 1.5 | 0.91 | 0.93 | 0.97 | 0.95 | 0.97 | 0.99 | 0.97 | 0.99 | 1.00 |

Subsampling

| | $n = 100$ | | | $n = 300$ | | | $n = 900$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 1.1 | 0.26 | 0.27 | 0.29 | 0.28 | 0.29 | 0.31 | 0.29 | 0.31 | 0.33 |
| 1.3 | 0.73 | 0.76 | 0.82 | 0.79 | 0.83 | 0.89 | 0.85 | 0.87 | 0.94 |
| 1.5 | 0.91 | 0.93 | 0.96 | 0.96 | 0.98 | 0.99 | 0.98 | 0.99 | 1.00 |

Table 4: Coverage probabilities for E($x$), $x \sim$ Fréchet($\lambda$).

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Transformation-based naive bootstrap, $k=1$ | | | | | | | | | | | | |
| | $n=100$ | | | | $n=300$ | | | | $n=900$ | | | |
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.98 | 0.98 | 0.98 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.57 | 0.58 | 0.59 | 2 | 0.95 | 0.97 | 0.98 | 2 | 1.00 | 1.00 | 1.00 |
| | 3 | 0.46 | 0.48 | 0.52 | 3 | 0.96 | 0.97 | 0.98 | 3 | 0.66 | 0.68 | 0.72 |
| | 4 | 0.50 | 0.53 | 0.56 | 4 | 0.95 | 0.96 | 0.98 | 4 | 0.68 | 0.71 | 0.76 |
| 1.3 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 | 8 | 0.98 | 0.99 | 0.99 |
| | 2 | 0.86 | 0.89 | 0.92 | 2 | 0.95 | 0.97 | 0.98 | 12 | 0.96 | 0.97 | 0.98 |
| | 3 | 0.87 | 0.89 | 0.92 | 3 | 0.96 | 0.97 | 0.98 | 16 | 0.93 | 0.94 | 0.96 |
| | 4 | 0.86 | 0.88 | 0.91 | 4 | 0.96 | 0.97 | 0.98 | 20 | 0.90 | 0.92 | 0.94 |
| 1.5 | 1 | 0.99 | 0.99 | 0.99 | 2 | 0.99 | 1.00 | 1.00 | 14 | 0.98 | 0.99 | 0.99 |
| | 2 | 0.95 | 0.96 | 0.98 | 3 | 0.99 | 1.00 | 1.00 | 20 | 0.96 | 0.97 | 0.98 |
| | 3 | 0.94 | 0.96 | 0.97 | 4 | 0.99 | 0.99 | 1.00 | 25 | 0.94 | 0.95 | 0.97 |
| | 4 | 0.93 | 0.94 | 0.96 | 5 | 0.99 | 0.99 | 0.99 | 30 | 0.91 | 0.93 | 0.96 |
| Transformation-based naive bootstrap, $k=2$ | | | | | | | | | | | | |
| | $n=100$ | | | | $n=300$ | | | | $n=900$ | | | |
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.88 | 0.89 | 0.90 | 1 | 0.97 | 0.97 | 0.98 | 3 | 0.99 | 0.99 | 1.00 |
| | 2 | 0.82 | 0.83 | 0.85 | 2 | 0.95 | 0.96 | 0.97 | 4 | 0.97 | 0.98 | 0.98 |
| | 3 | 0.75 | 0.76 | 0.78 | 3 | 0.92 | 0.93 | 0.94 | 5 | 0.92 | 0.93 | 0.94 |
| | 4 | 0.64 | 0.66 | 0.68 | 4 | 0.85 | 0.86 | 0.88 | 6 | 0.85 | 0.86 | 0.88 |
| 1.3 | 1 | 0.98 | 0.99 | 0.99 | 1 | 1.00 | 1.00 | 1.00 | 4 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.96 | 0.96 | 0.97 | 2 | 1.00 | 1.00 | 1.00 | 5 | 0.98 | 0.98 | 0.99 |
| | 3 | 0.89 | 0.90 | 0.92 | 3 | 0.98 | 0.99 | 0.99 | 6 | 0.95 | 0.96 | 0.97 |
| | 4 | 0.81 | 0.83 | 0.85 | 4 | 0.95 | 0.96 | 0.97 | 7 | 0.91 | 0.92 | 0.94 |
| 1.5 | 1 | 1.00 | 1.00 | 1.00 | 2 | 1.00 | 1.00 | 1.00 | 6 | 0.97 | 0.97 | 0.99 |
| | 2 | 0.97 | 0.98 | 0.98 | 3 | 0.99 | 0.99 | 1.00 | 7 | 0.93 | 0.95 | 0.97 |
| | 3 | 0.91 | 0.92 | 0.94 | 4 | 0.96 | 0.97 | 0.98 | 8 | 0.91 | 0.93 | 0.95 |
| | 4 | 0.85 | 0.87 | 0.90 | 5 | 0.91 | 0.93 | 0.95 | 9 | 0.87 | 0.90 | 0.94 |
| $m$ out of $n$ bootstrap | | | | | | | | | | | | |
| | | $n=100$ | | | | $n=300$ | | | | $n=900$ | | |
| $\lambda$ | | 0.90 | 0.95 | 0.99 | | 0.90 | 0.95 | 0.99 | | 0.90 | 0.95 | 0.99 |
| 1.1 | | 0.24 | 0.25 | 0.26 | | 0.26 | 0.27 | 0.29 | | 0.26 | 0.27 | 0.29 |
| 1.3 | | 0.70 | 0.74 | 0.79 | | 0.78 | 0.81 | 0.88 | | 0.82 | 0.87 | 0.93 |
| 1.5 | | 0.90 | 0.92 | 0.96 | | 0.94 | 0.97 | 0.99 | | 0.97 | 0.98 | 1.00 |
| Subsampling | | | | | | | | | | | | |
| | | $n=100$ | | | | $n=300$ | | | | $n=900$ | | |
| $\lambda$ | | 0.90 | 0.95 | 0.99 | | 0.90 | 0.95 | 0.99 | | 0.90 | 0.95 | 0.99 |
| 1.1 | | 0.27 | 0.27 | 0.29 | | 0.28 | 0.30 | 0.32 | | 0.28 | 0.30 | 0.32 |
| 1.3 | | 0.72 | 0.75 | 0.80 | | 0.79 | 0.83 | 0.88 | | 0.84 | 0.88 | 0.94 |
| 1.5 | | 0.89 | 0.92 | 0.96 | | 0.95 | 0.97 | 0.99 | | 0.98 | 0.99 | 1.00 |

Table 5: Coverage probabilities for $E(x)$, $x \sim$ Log-Gamma$(\lambda, \nu)$, $\nu = 1$.

| Transformation-based naive bootstrap, $k=1$ | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n=100$ | | | | $n=300$ | | | | $n=900$ | | | |
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.97 | 0.98 | 0.98 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.57 | 0.58 | 0.60 | 2 | 0.91 | 0.91 | 0.91 | 2 | 1.00 | 1.00 | 1.00 |
| | 3 | 0.47 | 0.50 | 0.54 | 3 | 0.54 | 0.57 | 0.62 | 3 | 0.68 | 0.70 | 0.74 |
| | 4 | 0.51 | 0.53 | 0.57 | 4 | 0.61 | 0.63 | 0.68 | 4 | 0.69 | 0.72 | 0.78 |
| 1.3 | 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 | 8 | 0.98 | 0.99 | 0.99 |
| | 2 | 0.86 | 0.89 | 0.92 | 2 | 0.95 | 0.97 | 0.98 | 12 | 0.96 | 0.97 | 0.98 |
| | 3 | 0.87 | 0.89 | 0.92 | 3 | 0.96 | 0.97 | 0.98 | 16 | 0.93 | 0.94 | 0.96 |
| | 4 | 0.86 | 0.88 | 0.91 | 4 | 0.96 | 0.97 | 0.98 | 20 | 0.90 | 0.92 | 0.94 |
| 1.5 | 1 | 0.99 | 0.99 | 0.99 | 2 | 0.99 | 1.00 | 1.00 | 14 | 0.98 | 0.99 | 0.99 |
| | 2 | 0.95 | 0.96 | 0.98 | 3 | 0.99 | 1.00 | 1.00 | 20 | 0.96 | 0.97 | 0.98 |
| | 3 | 0.94 | 0.96 | 0.97 | 4 | 0.99 | 0.99 | 1.00 | 25 | 0.94 | 0.95 | 0.97 |
| | 4 | 0.93 | 0.94 | 0.96 | 5 | 0.99 | 0.99 | 0.99 | 30 | 0.91 | 0.93 | 0.96 |

| Transformation-based naive bootstrap, $k=2$ | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n=100$ | | | | $n=300$ | | | | $n=900$ | | | |
| $\lambda$ | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 | $m$ | 0.90 | 0.95 | 0.99 |
| 1.1 | 1 | 0.89 | 0.90 | 0.92 | 1 | 0.97 | 0.98 | 0.99 | 3 | 0.99 | 0.99 | 1.00 |
| | 2 | 0.82 | 0.84 | 0.86 | 2 | 0.96 | 0.96 | 0.97 | 4 | 0.97 | 0.98 | 0.99 |
| | 3 | 0.76 | 0.78 | 0.79 | 3 | 0.93 | 0.94 | 0.96 | 5 | 0.92 | 0.93 | 0.95 |
| | 4 | 0.66 | 0.67 | 0.69 | 4 | 0.86 | 0.87 | 0.88 | 6 | 0.85 | 0.86 | 0.88 |
| 1.3 | 1 | 0.99 | 0.99 | 0.99 | 1 | 1.00 | 1.00 | 1.00 | 4 | 1.00 | 1.00 | 1.00 |
| | 2 | 0.96 | 0.97 | 0.98 | 2 | 1.00 | 1.00 | 1.00 | 5 | 0.98 | 0.99 | 0.99 |
| | 3 | 0.89 | 0.90 | 0.92 | 3 | 0.99 | 0.99 | 0.99 | 6 | 0.95 | 0.96 | 0.97 |
| | 4 | 0.81 | 0.83 | 0.86 | 4 | 0.95 | 0.96 | 0.97 | 7 | 0.91 | 0.92 | 0.95 |
| 1.5 | 1 | 1.00 | 1.00 | 1.00 | 2 | 1.00 | 1.00 | 1.00 | 6 | 0.96 | 0.97 | 0.98 |
| | 2 | 0.97 | 0.98 | 0.98 | 3 | 0.99 | 0.99 | 1.00 | 7 | 0.93 | 0.95 | 0.97 |
| | 3 | 0.91 | 0.93 | 0.95 | 4 | 0.96 | 0.97 | 0.98 | 8 | 0.90 | 0.92 | 0.95 |
| | 4 | 0.85 | 0.87 | 0.90 | 5 | 0.92 | 0.93 | 0.95 | 9 | 0.88 | 0.90 | 0.94 |

| $m$ out of $n$ bootstrap | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $n=100$ | | | $n=300$ | | | $n=900$ | | |
| $\lambda$ | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 1.1 | 0.24 | 0.25 | 0.27 | 0.27 | 0.31 | 0.33 | 0.30 | 0.34 | 0.37 |
| 1.3 | 0.61 | 0.64 | 0.70 | 0.77 | 0.81 | 0.86 | 0.80 | 0.84 | 0.90 |
| 1.5 | 0.78 | 0.81 | 0.87 | 0.90 | 0.93 | 0.96 | 0.91 | 0.94 | 0.97 |

| Subsampling | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $n=100$ | | | $n=300$ | | | $n=900$ | | |
| $\lambda$ | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 1.1 | 0.26 | 0.27 | 0.28 | 0.31 | 0.34 | 0.37 | 0.33 | 0.36 | 0.40 |
| 1.3 | 0.68 | 0.72 | 0.76 | 0.86 | 0.90 | 0.93 | 0.89 | 0.93 | 0.96 |
| 1.5 | 0.85 | 0.88 | 0.92 | 0.95 | 0.98 | 0.99 | 0.97 | 0.98 | 0.99 |

# References


Abadir, K. M. (1999). An introduction to hypergeometric functions for economists. *Econometric Reviews 18*, 287–330.

Abadir, K. M. and G. Talmain (1999). Autocovariance functions of series and of their transforms. *Journal of Econometrics 124*, 227–252.

Arcones, M. A. and E. Giné (1989). The bootstrap of the mean with arbitrary bootstrap sample size. *Annales de l'Institut Henri Poincaré 25*(4), 457–481.

Arcones, M. A. and E. Giné (1991). Additions and correction to "the bootstrap of the mean with arbitrary sample size". *Annales de l'Institut Henri Poincaré 27*, 583–595.

Athreya, K. B. (1987). Bootstrap of the mean in the infinite variance case. *Annals of Statistics 15*, 724–731.

Athreya, K. B., S. N. Lahiri, and W. Wu (1998). Inference for heavy tailed distributions. *Journal of Statistical Planning and Inference 669*, 61–75.

Backus, D., M. Chernov, and I. Martin (2011). Disasters implied by equity index options. *Journal of Finance 66*, 1969–2012.

Bickel, P. J. and D. A. Freedman (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics 9*(6), 1196–1217.

Bickel, P. J. and A. Sakov (2008). On the choice of $m$ in the $m$ out of $n$ bootstrap and confidence bounds for extrema. *Statistica Sinica 18*, 967–985.

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

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

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

Davison, A. C. and D. V. Hinkley (2009). *Bootstrap Methods and Their Application.* New-York: Cambridge University Press.

Delaigle, A. and P. Hall (2012). Effect of heavy-tails on ultra high dimensional variable ranking methods. *Statistica Sinica 22*, 909–932.

DiCiccio, T. J., A. C. Monti, and A. G. Young (2006). Variance stabilization for a scalar parameter. *Journal of the Royal Statistical Society Series B 68*, 281–303.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics 7*, 1–26.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications vol. II.* New York: John Wiley and Sons.

Giné, E. and J. Zinn (1989). Necessary conditions for the bootstrap of the mean. *Annals of Statistics 17*, 684–691.

Giné, E. and J. Zinn (1990). Bootstrapping general empirical measures. *Annals of Probability 18*, 851–869.

Hall, P. (1990). Asymptotic properties of the bootstrap for heavy-tailed distributions. *Annals of Probability 18*, 1342–1360.

Hall, P. and B. Jing (1998). Comparison of the bootstrap and asymptotic approximations to the distribution of a heavy-tailed mean. *Statistica Sinica 8*, 887–906.

Hall, P. and R. LePage (1996). On bootstrap estimation of the distribution of the studentized mean. *Annals of the Institute of Statistical Mathematics 48*, 403–421.

Hall, P. and Q. Yao (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica 71(1)*, 285–317.

Hallin, M. and M. Seoh (1996). *Research developments in Probability and Statistics*, Chapter Is 131,000 a large sample size? A study on the finite-sample behavior of Edgeworth expansions. Zeist: VSP.

Ibragimov, R. and U. K. Muller (2010). t-statistic based correlation and heterogeneity robust inference. *Journal of Business and Economic Statistics 28*, 453–468.

Knight, K. (1989). On the bootstrap of the sample mean in the infinite variance case. *Annals of Statistics 17*, 1168–1175.

Koenker, R., J. A. F. Machado, C. L. Skeels, and A. H. Welsh (1994). Momentary lapses: Moment expansions and the robustness of minimum distance estimation. *Econometric Theory 10*, 172–197.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics 17*, 1217–1241.

Lahiri, S. (1995). On the asymptotic behaviour of the moving block bootstrap for normalized sums of heavy-tail random variables. *Annals of Statistics 23*, 1331–1349.

Liu, R. and K. Singh (1992). *Moving blocks jackknife and bootstrap capture weak dependence*, pp. 225–248. New-York: In: Exploring the Limits of Bootstrap, R. LePage and L. Billard eds., Wiley.

Martin, I. W. R. (2008). Disasters and the welfare cost of uncertainty. *American Economic Review 98*, 74–78.

McNeil, A. J., R. Frey, and P. Embrechts (2015). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton Series in Finance. Princeton University Press.

Mittnik, S., S. Rachev, and J.-R. Kim (1998). Chi-square-type distributions for heavy-tailed variates. *Econometric Theory 14*, 339–354.

Politis, D. N., J. P. Romano, and M. Wolf (1999). *Subsampling.* New-York: Springer.

Romano, J. P. and M. Wolf (1999). Subsampling inference for the mean in the heavy-tailed case. *Metrika 50*, 55–69.

Samworth, R. (2005). Small confidence sets for the mean of a spherically symmetric distribution. *Journal of the Royal Statistical Society Series B 67*, 343–361.

Taylor, J. M. G. (1986). The retransformed mean after a fitted power transformation. *Journal of the American Statistical Association 81*, 114–118.

Whittaker, E. T. and G. N. Watson (1997). *A course of modern analysis.* Cambridge: 4th ed. Cambridge University Press.

Yu, J., Z. Yang, and X. Zhang (2006). A class of nonlinear stochastic volatility models and its implications for pricing currency options. *Computational Statistics and Data Analysis 51*, 2218–2231.