Approximating moments by nonlinear transformations, with an application to resampling from fat-tailed distributions*

Karim M. Abadir — Imperial College London

Adriana Cornea — University of Exeter

Abstract

We provide a methodology to calculate the expectation of a variate x in terms of the moments of a transformation of x. Apart from the intrinsic interest in such a fundamental relation that relates the moments of a variate and its nonlinear transformations, our results can be used in practice to approximate $\mathrm{E}(x)$ by the low-order moments of a transformation which can be chosen to give a good approximation for $\mathrm{E}(x)$. To obtain an accurate evaluation of the remainder, we derive results for the bounding of functions of complex variables. Our results are useful, for example, in resampling applications like bootstrap confidence intervals for fat-tailed data. They are also useful in economics and finance in quantifying the effect of taking nonlinear transformations on moment conditions and on asset prices which are formulated as expectations. We illustrate with an application to solving the problem of bootstrapping from a density that has a fat tail, in which case currently-available bootstrap methods are invalid.

*We thank Essie Maasoumi and Michael Wolf for their comments. This research is supported by the ESRC grant RES062230790 and by the British Academy's PDF/2009/370.

1 Introduction

Let $x \in \mathcal{X} \subseteq \mathbb{R}$ be a variate with unknown density, 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 an invertible transformation $y := g^{-1}(x) \in \mathcal{Y} \subseteq \mathbb{R}$. This fundamental relation between the moments of x and y has, surprisingly, not been derived anywhere before. Approximations to it have been used on an ad-hoc basis, 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) or the precise evaluation of the remainder. 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, the remainder term in our expansions can be bounded explicitly and accurately, without restoring to orders of magnitude that indicate the rate of change of the remainder rather than its size.

It is of interest to investigate such a fundamental relation. 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 which can be chosen to give a good approximation for $\mathrm{E}(x)$. Such results are useful, for example, in assessing the effect of the Box-Cox transformation on the mean (Taylor (1986)) or in resampling applications like bootstrap confidence intervals (CIs) for fat-tailed data where the standard bootstrap fails because of the nonexistence of higher moments (Athreya (1987), Knight (1989), Politis, Romano and Wolf (1999)). Applications like these are important and substantial, and in this paper we illustrate how the failure of the standard bootstrap for the mean with infinite second moment can be alleviated by our approach.

The potential for applications is not just statistical and econometric, but also in economics and finance where there is interest *inter alia* in quantifying the effect of taking nonlinear transformations on moment conditions (such as Euler's, arising from optimization) and on asset prices which are formulated as expectations; e.g. see Yu et al. (2006), Martin (2008), Backus, Chernov and Martin (2011). The effect of higher-order terms is important and needs to be quantified, as recent market turbulence has emphasized.

In Section 2, we expand on the bootstrap motivating example that we have just stated. In Section 3, we introduce the expansion and the required results obtained by complex analysis. These are needed for the bounding of functions of complex variables to obtain an accurate evaluation of the remainder of the approximation. In Section 4, we illustrate the expansion and the accuracy of the remainder's bound. In Section 5, we apply our expansions to obtain asymptotically-valid bootstrap CIs for a mean from a distribution with infinite variance, a case where exisiting methods are invalid. Section 6 concludes. The proofs are relegated to the Appendix. We follow the notation conventions proposed in Abadir and Magnus (2002).

2 A motivating example

2.1 Invalid bootstrap without transformation

Let $y \sim \text{Expo}(\lambda)$ with $\lambda > 0$. Then, $x := \mathrm{e}^y \in (1, \infty)$ has the fat-tailed standard Pareto density

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

and $\mathrm{E}(x) = \lambda/(\lambda - 1)$. Table 1 shows the coverage probabilities for $\mathrm{E}(x)$ when both the ordinary bootstrap and the m out of n bootstrap are used. The results are based on 10,000 samples of size 100, with 399 bootstrap replications for each sample. We build basic bootstrap confidence intervals (see Davison and Hinkley (1997), p.28). The lower and upper limits of the CIs are

$$\bar{x}_n - (\bar{x}^b_{(B+1)(1-\alpha)} - \bar{x}_n), \qquad \bar{x}_n - (\bar{x}^b_{(B+1)\alpha} - \bar{x}_n),$$

where $x^b_{(B+1)(1-\alpha)}$ and $\bar{x}^b_{(B+1)\alpha}$ are the $1 - \alpha$ and α quantiles of the m out of n bootstrap distribution or of the ordinary bootstrap distribution depending on the bootstrap used. B is the number of bootstrap replications. The ordinary bootstrap, introduced by Efron (1979), is based on resamplings with replacement; see Godfrey (2009) for a comprehensive survey. It is not valid for $\lambda < 2$ because the variance of x does not exist; see Athreya (1989), Knight (1989). The m out of n bootstrap is similar to the ordinary bootstrap except that the bootstrap sample 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 the bootstrap. As Table 1 shows, the ordinary bootstrap works very badly, as expected. For the m out of n bootstrap, the best coverage probabilites are achieved when m is set at 10% of n but they are far from the nominal coverage probabilities except for $\lambda = 1.5$. Unreported simulations (see also Cornea and Davidson (2011)) indicate that the coverage probability is very sensitive to the choice of m for which there is little theoretical guidance. Moreover, for a regular statistic like the mean studied here, the bootstrap CI approach should convey the same conclusion as the bootstrap hypothesis testing approach (see van Giersbergen and Kiviet (2002)). However, this is not the case if we consider the same setup as in Table 1 with $\lambda = 1.5$ and compute the error in rejection probability (ERP) of the m out of n bootstrap. The ERP is as high as 30% at the 1%, 5% and 10% when $m = 10$ and improves slighly if we increase m. Also considering $n = 2,000$ does not alter by much our results. Essentially the same conclusion holds for subsampling which is asymptotically equivalent to the m out of n bootstrap in an i.i.d. setting; see Politis, Romano and Wolf (1999, p.48).

2.2 Valid bootstrap with transformation

On the other hand, the tail of $y := \log(x)$ declines exponentially (this holds more generally – by the transformation theorem– for the log-transform of any variate with integrable density having fat tails), hence making the ordinary bootstrap of the mean based on y valid. This is illustrated in Table 2. Here, we have $\mathrm{E}(y) = \lambda^{-1}$ and

$$\mathrm{E}(x) = (1 - \lambda^{-1})^{-1} = (1 - \mathrm{E}(y))^{-1},$$

so we can obtain indirectly a CI for $\mathrm{E}(x)$ by bootstrapping the mean of y and applying the inverse transformation to the endpoints of the CI. More exactly the $1 - 2\alpha$ bootstrap CI has limits

$$2\bar{x}_n - 1/(1-\bar{y}^b_{(B+1)(1-\alpha)}), \quad 2\bar{x}_n - 1/(1-\bar{y}^b_{(B+1)\alpha})$$

where $\bar{y}^b_{(B+1)(1-\alpha)}$ and $\bar{y}^b_{(B+1)\alpha}$ are the $1-\alpha$ and α quantiles of the ordinary bootstrap distribution. However, we do not know the link between the expectations in general, and we have to resort to approximating $\mathrm{E}(x)$ by the moments of another variate like y which satisfies the bootstrap validity conditions. This link, in a more general setup, is the purpose of the paper.

3 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 any function z of x.

We propose two types of expansions, raw or centered. We start by explaining the idea behind the two expansions: (a) without recourse to complex variables; and (b) with the simple $g = \exp$, giving $x := \exp(y) \in \mathbb{R}_+$. The two types of expansions are:

1. the raw (direct) expansion

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

2. the centered expansion

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

The moments on the right-hand side exist because of the existence of $\mathrm{E}(x)$. The series obtained above is a special case of Teixeira's expansion which expands a function in terms of another; e.g. see Whittaker and Watson (1997, pp.131-133) or Abadir and Talmain (2005) for an application. In this illustration, g was the exponential function but the derivations to follow would apply to other functions. For more elaborate cases from the hypergeometric ${}_pF_q$ family, see Abadir (1999) or Whittaker and Watson (1997), but one should bear in mind the convergence radius of the expansion used. For example, it would be more problematic to expand the logarithmic function, due to its slower (and conditional) convergence: $\log(\cdot)$ is of the hypergeometric ${}_2F_1$ class, whereas $\exp(\cdot)$ is ${}_0F_0$. Other examples include hyperbolic functions (which are members of the ${}_0F_1$ class that converges even faster than ${}_0F_0$) such as

$$x := \cosh(y) = \sum_{j=0}^{k} \frac{y^{2j}}{(2j)!} + R_k, \qquad (\text{for } x \in \mathbb{R}_+), \tag{3}$$
$$x := \sinh(y) = \sum_{j=0}^{k} \frac{y^{2j+1}}{(2j+1)!} + R_k, \qquad (\text{for } x \in \mathbb{R}),$$

or trigonometric functions on restricted domains.

We can control the precision of R_k and R_k^{c},[1] but a better way to proceed with this will be given in (4) below. The intuition behind using expansions like (2) for approximations (in applications like resampling) is that it is the higher-order terms in the remainder that create problems with fat-tailed distributions, whereas bounding R_k^{c} by a low-power term will get around such difficulties. Furthermore, the conditions that lead to the sample mean being a consistent estimator of the population mean will typically also lead to the asymptotically correct coverage by the centered expansions. We shall not dwell on these asymptotic issues in this section, since our focus is on the expansions themselves. Nevertheless, we will touch briefly on this topic at the end of this section, once the derivations are made and provide an application in Section 5.

We have centered the expansion around the mean of y, but there is a more attractive expansion if we are willing to introduce complex numbers and use bounding results from the theory of characteristic functions. Let $-1 \equiv \mathrm{i}^2$ and $(y - \mathrm{E}\, y)/m \equiv \zeta + 2\pi i_y$, where $i_y \in \mathbb{Z}$, $m \in \mathbb{N}$, and $\zeta \in (-\pi, \pi]$, then write

$$x = \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left(\exp\left(\frac{\zeta}{\mathrm{i}} \right) \right)^{\mathrm{i}m} \equiv \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left(\xi_k + \varrho_{x,k} \right)^{\mathrm{i}m}, \tag{4}$$

[1]Of course, R_k and R_k^{c} can be calculated exactly in theory, but their moments could be empirically problematic (e.g. with fat-tailed densities), hence our approach to bound them instead.

where $\xi_k := \sum_{j=0}^{k} \zeta^j/(\mathrm{i}^j j!)$. Notice that i_y is random, but m is deterministic and to be chosen later.[2] We can now state the bound

$$(5) \qquad \left|\varrho_{x,k}\right| \equiv \left|\mathrm{e}^{\mathrm{i}t} - \sum_{j=0}^{k} \frac{(\mathrm{i}t)^j}{j!}\right| \leq \frac{|t|^{k+1}}{(k+1)!}$$

for any $t \in \mathbb{R}$. This is helpful when taking expectations, as it controls the precision by calculating the moment of the next power. A binomial expansion of (4) gives

$$(6) \qquad x = \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \operatorname{Re}(\xi_k^{\mathrm{i}m}) + R^{\mathrm{c}}_{x,k}, \qquad R^{\mathrm{c}}_{x,k} = O_p\left(\zeta^{k+1}\right).$$

More precisely,

$$(7) \quad \begin{aligned} \left|R^{\mathrm{c}}_{x,k}\right| &= \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left|\operatorname{Re}\left(\left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m} - \xi_k^{\mathrm{i}m}\right)\right| \\ &\leq \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left|\left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}m} - \xi_k^{\mathrm{i}m}\right| = \mathrm{e}^{\mathrm{E}\, y} \mathrm{e}^{2\pi m i_y} \left|\xi_k^{\mathrm{i}m}\right| \left|\left(1 + \varrho_{x,k}/\xi_k\right)^{\mathrm{i}m} - 1\right|. \end{aligned}$$

Following from the representation of x in terms of a power series in y, the form $|\left(1 + \varrho_{x,k}/\xi_k\right)^{\mathrm{i}m} - 1|$ is common to all the expansion mentioned in the previous paragraphs. If we can bound it accurately, then we can evaluate these remainders satisfactorily. We therefore present the following results.

Proposition 1 *Define the real-valued function h of the complex ψ,*

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

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

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

where the equality follows from (15) in the Appendix. The upper bound of $1 + \exp(m\pi)$ is indeed achieved and the proposition tells us which values of ψ achieve it.

[2] This m differs from the one referred to in the previous section, in the context of "m out of n bootstrap".

By choosing a large negative j, the solution $\psi_{\mathrm{m}} = -1 - \exp\left((2j+1)\,\pi/m\right)$ can be made sufficiently close to -1 for most practical purposes. For $|\psi| \leq 1$, the next proposition provides a pointwise logarithmic bound for $h(\psi)$.

Proposition 2 *Define the real-valued function h_1 of the complex ψ,*

$$h_1(\psi) := \left|(1+\psi)^{\mathrm{i}m} - 1\right| - c_m \log\left(1+|\psi|\right), \qquad \textit{with } \arg\left(\psi\right) \in [-\pi, \pi) \textit{ and } |\psi| \leq 1.$$

Then, $h_1(\psi) \leq 0$ when $c_m = (1+\mathrm{e}^{m\pi})/\log 2$, with $h_1(\psi) = 0$ at $\psi = 0$.

As a result of the previous two propositions, the simple bound

$$\left|(1+\psi)^{\mathrm{i}m} - 1\right| \leq (1+\mathrm{e}^{m\pi}) \min\left(\frac{\log\left(1+|\psi|\right)}{\log 2}, 1\right)$$

applies. The same idea can be applied to any monotonic increasing function of $|\psi|$ (the choice of log was due to its slow variation), however a sharper $|\psi|$-pointwise bound is obtained by the following.

Proposition 3 *Define the real-valued function h of the complex ψ,*

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

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

$$h(\psi) \leq \begin{cases} \sqrt{1 - 2\mathrm{e}^{m\sin^{-1}|\psi_0|}\cos\left(m\log\left(1-|\psi_0|\right)\right) + \mathrm{e}^{2m\sin^{-1}|\psi_0|}}, & |\psi_0| \in \left[0, 1-\mathrm{e}^{-\pi/m}\right), \\ 1 + \mathrm{e}^{m\sin^{-1}|\psi_0|}, & |\psi_0| \in \left[1-\mathrm{e}^{-\pi/m}, 1\right], \end{cases}$$

for any given $|\psi_0|$.

As a result of this proposition and the first one, when $\arg\left(\psi\right) \in [-\pi, \pi)$ we have the bound

$$\left|(1+\psi)^{\mathrm{i}m} - 1\right| \leq H(|\psi|),$$

where

(8)

$$H(|\psi|) := \begin{cases} \sqrt{1 - 2\mathrm{e}^{m\sin^{-1}|\psi|}\cos\left(m\log\left(1-|\psi|\right)\right) + \mathrm{e}^{2m\sin^{-1}|\psi|}}, & |\psi| \in \left[0, 1-\mathrm{e}^{-\pi/m}\right), \\ 1 + \mathrm{e}^{m\sin^{-1}|\psi|}, & |\psi| \in \left[1-\mathrm{e}^{-\pi/m}, 1\right], \\ 1 + \mathrm{e}^{m\pi}, & \text{otherwise.} \end{cases}$$

Given the monotonicity property in Proposition 3 and (8), we can work out the bound in (7) as

$$\left|R^{\mathrm{c}}_{x,k}\right| \leq \mathrm{e}^{\mathrm{E}\,y}\mathrm{e}^{2\pi m i_y}\left|\xi_k^{\mathrm{i}m}\right| H\left(\frac{|\zeta|^{k+1}}{(k+1)!\,|\xi_k|}\right) =: B^{\mathrm{c}}_{x,k} \tag{9}$$

by (5). For calculating expectations, the error for $\mu := \mathrm{E}(x)$ is denoted by $R^{\mathrm{c}}_{\mu,k}$ and is bounded by

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

Note the alternative formulation

$$\left|\xi_k^{\mathrm{i}m}\right| = \left|\left|\xi_k\right|^{\mathrm{i}m}\right|\mathrm{e}^{-m\arg(\xi_k)} = \left|\mathrm{e}^{\mathrm{i}m\log|\xi_k|}\right|\mathrm{e}^{-m\arg(\xi_k)} = \mathrm{e}^{-m\arg(\xi_k)}, \tag{10}$$

which does not contain imaginary quantities.

We conclude this section by analyzing the choice of the deterministic m. Recall that $(y-\mathrm{E}\,y)/m \equiv \zeta + 2\pi i_y$, which was introduced before (4). This means that m acts like a scaling integer. Choosing a large m leads to the consistency-like behaviour we mentioned at the outset, 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. 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 5.

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

We illustrate the performance of the k-term expansion and bound given in (6) for $\mathrm{E}(x)$ 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 $x := \mathrm{e}^y \in (1, \infty)$ is

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

hence being in the domain of attraction of stable laws with exponent (index) λ when $\lambda < 2$: its variance does not exist for $1 < \lambda < 2$ but its mean does.

Tables 3-17 display the k-term expectation, denoted 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 $\mathrm{E}(x)$.[3] 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 tables show that even in the fattest-tailed case of small λ and large ν, the performance is very good. Even the 2-term expansion ($k = 2$) is accurate, especially when we choose m not too small. All the tables indicate that choosing a large m increases the accuracy of both the expansion and bound. Centered expansion are, on the whole, more accurate when we take $m > 1$. 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.

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, consider $y = 10$, then $x := \mathrm{e}^y = 22,026.4657$. First, consider the raw expansion (1) with no complex numbers or m involved. The 2-term expansion gives only $x \approx 61$. We need a 30-term expansion to obtain a good approximation of x.

Second, we consider (6) for $m = 1$ and different values of k. The 2-term expansion gives $9,447.5$ and the corresponding bound for the remainder (9) is $83,879$. However, 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 (6) for fixed $k = 2$ and different values of

[3] 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.

m. For $m = 100$ the expansion is $22,396$ while 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.

5 Application: transformation-based bootstrap

The purpose of this section is to illustrate the usefulness of our expansion to solve a problem with the validity of the bootstrap. We propose a simple bootstrap CI for $\mathrm{E}(x)$ where $x \in \mathbb{R}_+$ and, as in Section 2, we assume that $\mathrm{var}(x) = \infty$. The case of $x \in \mathbb{R}$ can be similarly handled by considering the negative and positive parts separately, and the choice of $\mathrm{E}(x)$ instead of other moments is meant to keep complexity at a minimum for our application. The same idea can be applied to statistics other than the mean for which the second moment does not exist. The fact that $\mathrm{E}(x)$ is finite guarantees that all the moments of y exist and that the k-term and bound's expectations are finite.[4]

Letting $y := \log(x)$ and $z := \mathrm{e}^{\mathrm{E}\, y}\mathrm{e}^{2\pi m i_y}\,\mathrm{Re}(\xi_k^{\mathrm{i}m})$, we have $x = z + R^{\mathrm{c}}_{x,k}$ and the remainder has the bound $B^{\mathrm{c}}_{x,k}$ given in (9). By applying the triangle inequality twice,

$$|z| - B^{\mathrm{c}}_{x,k} \leq x \leq |z| + B^{\mathrm{c}}_{x,k}, \tag{11}$$

which can be used to build conservative CIs for $\mathrm{E}(x)$. To do this, consider an i.i.d. sample $x_1, \cdots, x_n$ and compute the following quantities

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

By the triangle inequality,

$$t_{1,n} \leq \bar{x}_n \leq t_{2,n}, \qquad t_{1,n} := \bar{z}_n^+ - B^{\mathrm{c}}_{\bar{x}_n,k}, \qquad t_{2,n} := \bar{z}_n^+ + B^{\mathrm{c}}_{\bar{x}_n,k}. \tag{12}$$

In Proposition 4, we prove that the ordinary bootstrap of $t_{1,n}$ and $t_{2,n}$ based on resamplings from the EDF of the y's (or equivalently from the joint EDF of i_y and ζ) is asymptotically valid. Denote by $t^{\mathrm{b}}_{1,n}$ and $t^{\mathrm{b}}_{2,n}$ the bootstrap versions of $t_{1,n}$ and $t_{2,n}$.

[4] This is in contrast, for instance, with the Nargar-type expansion of moments (Nagar (1959), Sargan (1974), Maasoumi (1977)) which does not hold if higher-order moments do not exist.

As seen from (11), if $B^{\mathrm{c}}_{x,k}$ is too large then the CI is too conservative. While, if $k \to \infty$ or $m \to \infty$, then $B^{\mathrm{c}}_{x,k}$ vanishes and z coincides with x and we are back to the original invalid bootstrap. Thus k, m have to be finite 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 practice, k, m can be chosen based on Hill's (1975) estimator.

We now show that the standard bootstrap is valid for $|z|$ and $B^{\mathrm{c}}_{x,k}$, then their linear combination as in (12). Note that the bootstrap is not valid for $R^{\mathrm{c}}_{x,k}$, which is just $x - z$ and inherits the fat tail of x.

Proposition 4 *Let k, m be finite. Then, for $i = 1, 2$ and all $0 < \epsilon < 1$,*

$$\Pr\left(\sup_{v\in\mathbb{R}} \left|\Pr\left(n^{1/2} t^{\mathrm{b}}_{i,n} \le v \mid \boldsymbol{y}\right) - \Phi(v)\right| > \epsilon\right) \to 0. \tag{13}$$

It follows from (12) and Proposition 4 that the upper limit of the CI for $\mathrm{E}(x)$ is at most equal to the upper limit of the CI for $\mathrm{E}\,|z| + \mathrm{E}(B^{\mathrm{c}}_{x,k})$ since $\bar{x}_{n,\alpha} \le t^{\mathrm{b}}_{2,n,\alpha}$, where $\bar{x}_{n,\alpha}$ and $t^{\mathrm{b}}_{2,n,\alpha}$ are the estimates of the α-quantiles of the distribution of $n^{1/2}\bar{x}_n$ and $n^{1/2}t^{\mathrm{b}}_{2,n}$, respectively. Also the lower limit of the CI for $\mathrm{E}\,|z| - \mathrm{E}(B^{\mathrm{c}}_{x,k})$ is at most as large as the lower limit of the CI for $\mathrm{E}(x)$ since $t^{\mathrm{b}}_{1,n,\alpha} \le \bar{x}_{n,\alpha}$, where $t_{1,n,\alpha}$ is the estimate of the α-quantile of the distribution of $n^{1/2}t^{\mathrm{b}}_{1,n}$. Thus, $[2t_{1,n} - t^{\mathrm{b}}_{1,n,1-\alpha}, 2t_{2,n} - t^{\mathrm{b}}_{2,n,\alpha}]$ is a conservative $1-2\alpha$ two-sided CI for $\mathrm{E}(x)$.

To illustrate these results with a simulation, we take the same setup as in Tables 1 and 2. Table 18 gives the bootstrap coverage probabilities for $\mathrm{E}(x)$ based on the approach just described, using the raw version of the expansion. We see that, as in Table 1, a low λ (fat tail) is more challenging that a high λ, but as $n \to \infty$, all CIs have the required conservative coverage as Proposition 4 implies. We can also see that there are choices for k, m 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 as the 99%, choose $k = 1$. The optimal m can be predicted by a response surface based on a count regression for m with regressors $n^{-1/2}$, λ^{-1},

λ^{-2}. The data considered are those values of m for which the CP is closest to the nominal one for each $n = 100, 200, \cdots, 1000$ and $\lambda = 1.1, 1.2, \cdots, 1.9, 1.95, 2$. The response surface with the t-statistics in the brackets is given by

$$\begin{array}{lllll} \log \hat{\mu}_i = & 1.44 & -37.90 n_i^{-1/2} & +14.41 \lambda_i^{-1} & -15.42 \lambda_i^{-2}, \\ & (1.31) & (-16.70) & (4.12) & (-5.66) \end{array}$$

$i = 1, \cdots, 110$. The residual deviance is 31. The predicted m is given in the shaded boxes in the top panel of Table 18. For the case $n = 100$, $\lambda = 1.1$, the predicted m is 0 but we take $m = 1$ as $m \geq 1$ by assumption. This case is very hard to handle as it is very close to the non-existence of the first moment of the distribution.

2. Considering progressively smaller quantiles, as α declines, we move into the body of the distribution and:

   (a) for small samples: choose $k = 1$ and $m = 1$ if the tail is fat, while a thinner tail admits $k = 2$ but with a higher m;

   (b) for large samples and fat or thin tails: choose $k = 2$ and progressively increase m.

   For $k = 2$, the optimal m can be predicted by a response surface based on a count regression for m with regressors $n^{-1/2}$, λ^{-2}, $n^{-1}\lambda^{-2}$ and $n^{-1/2}\alpha^{-1}$. The data is selected as above for $\alpha = 0.90, 0.95$, but this time ties are possible in the form of conservative and liberal CPs that are equally distant from the nominal one, in which case we choose the former (conservative). The estimation results are

$$\begin{array}{llllll} \log \hat{\mu}_i & = 3.22 & -66.36 n_i^{-1/2} & -1.16 \lambda_i^{-2} & +102.58 n_i^{-1} \lambda_i^{-2} & +40.36 n_i^{-1/2} \alpha_i^{-1}, \\ & (14.10) & (-4.75) & (-4.48) & (1.09) & (3.31) \end{array}$$

   $i = 1, \cdots, 216$. The residual deviance is 6.32. The m predicted by this regression is given in the shaded boxes in the second panel of Table 18 and show an excellent performance again.

The optimal m given in the shaded boxes of Table 18 indicate that the count regressions are useful instruments for choosing m in practice and in general. We fitted response surfaces for Pareto *tail* quantiles, which also apply in large samples to fat-tailed distributions more generally. Unreported simulation results indicate that the optimal m predicted by the count regressions above give conservative CIs for the mean of distributions other than Pareto, which have Pareto-like tails, such as the log-gamma distribution and the Burr distribution (or Singh-Maddala distribution) used to model household incomes (see Davidson and Flachaire (2007)).

6 Concluding comments

Using complex analysis, we provided expansions and bounds for the expectation of a variate x in terms of the moments of a transformation y of x. We then illustrated the accuracy of the expansions and bounds by simulating distributions, including ones whose higher order moments do not exist. Finally, we show how to use our formulae to fix the problem of bootstrapping from a density that has a fat tail: even though we use the simple bootstrap along with our expansions, our results are very good compared to the bad performance of various sophisticated bootstrap modifications that tried to fix the problem. It shows the potential of our expansions for this and the other applications cited in the introduction.

In this paper, we assumed g to be invertible because we dealt with $x \equiv g(g^{-1}(x))$ and expanded a chosen g that converges fast, the motivating reason being that we wanted an accurate expansion for x and its expectation. This invertibility of g is not required for our propositions to apply. They are general and can be used directly in the case of any function $x = g(z)$ that is not necessarily invertible (e.g. by a Taylor expansion of $g(z)^{-\mathrm{i}}$), or even the composition $x = g_1(g_2(z))$ where we would expand g_1 only. If so, the only required alteration would not be in our propositions, but in the coefficients of the expansion preceding Proposition 1. To illustrate with the Box-Cox transformation

$$z := \frac{x^p - 1}{p} \tag{14}$$

if we are interested in the expectation of the original variate x, omitting the centring for ease of exposition gives

$$x = (1+pz)^{1/p} = \left(\mathrm{e}^{(\mathrm{i}p)^{-1}\log(1+pz)}\right)^{\mathrm{i}} = \left(\xi_k + \varrho_{x,k}\right)^{\mathrm{i}}, \qquad \text{with } \xi_k := \sum_{j=0}^{k} \frac{(\log(1+pz))^j}{(\mathrm{i}p)^j \, j!}$$

and the same propositions apply as before. It is also possible to expand by something other than an exponential function, as discussed in Section 3, or simply expand $(1+pz)^{-\mathrm{i}/p}$ by the binomial (relevant special case of a Taylor series) but the latter's convergence would be conditional.

APPENDIX

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

$$(1+\psi)^{\mathrm{i}m} = |1+\psi|^{\mathrm{i}m} \exp(-m\theta), \tag{15}$$

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

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

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

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

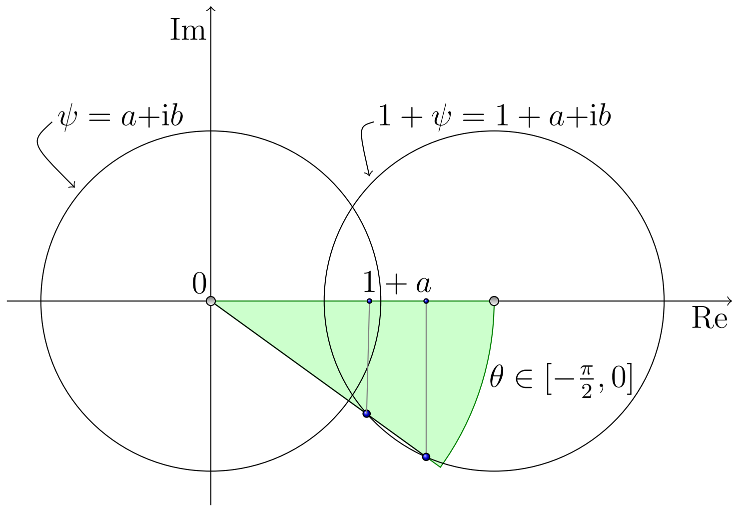

Figure 1: Illustration of the solution of Proposition 3.

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

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

and the result follows. □

Proof of Proposition 2. This follows from the previous proposition and the fact that $\log(1+|\psi|)$ is monotonic increasing in $|\psi|$. □

Proof of Proposition 3. 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 θ 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 θ. 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 $\sin^2\theta \leq |\psi|^2 \leq |\psi|$ since $|\psi| \leq 1$, so the RHS is nonpositive: the top restriction always holds and the bottom one requires $-\sin^2\theta\,(|\psi| - 1)^2 \leq 0$ 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 $1+a \geq 0$ and $\cos\theta \geq 0$. This is a function of $\varphi := |\psi|^2$ and θ, which is to be optimized subject to $\varphi \leq |\psi_0|^2$. 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

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

and

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

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

Since $l \neq 0$ for the optimum, $\sin(m\log) \neq 0$ implying that $\cos(m\log) \neq 1$ unlike in Proposition 1. The objective function cannot be simplified like before, and the first-order condition on θ 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 $-\theta \leq \sin^{-1}|\psi_0|$. 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

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

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

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

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 shows that there are none in $(0, 1-\mathrm{e}^{-\pi/m})$. □

Proof of Proposition 4. Since $y := \log(x)$, the transformation theorem ensures that the tails of y decay exponentially and all finite-order (hence $k < \infty$) moments exist, which is sufficient for the standard bootstrap to be valid; see Theorem 2.1 of Bickel and Freedman (1981). The condition $k < \infty$ amounts to requiring $R^{\mathrm{c}}_{x,k}$ is nondegenerate. Recalling that $(y - \mathrm{E}\,y)/m \equiv \zeta + 2\pi i_y$, where the random $i_y \in \mathbb{Z}$ is discrete, it follows that both ζ, i_y possess the required moments when m is finite, which establishes the validity of the bootstrap for $|z|$. As for $B^{\mathrm{c}}_{x,k}$, using (10) with $|\arg(\xi_m)| \leq \pi$ then (8),

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

where the moments of the right-hand side exist as in the case of $|z|$. Under no further assumptions, using the Cramér-Wold device we can then show that both $n^{1/2}t_{1,n}$ and $n^{1/2}t_{2,n}$ have a normal distribution asymptotically. Since these statistics are essentially sample means, (13) follows from Theorem 2.1 of Bickel and Freedman (1981). □

REFERENCES

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

Abadir, K. M. and J. R. Magnus (2002). Notation in econometrics: a proposal for a standard. *Econometrics Journal*, 5, 76–90.

Abadir, K. M. and G. Talmain (2005). Autocovariance functions of series and of their transforms. *Journal of Econometrics*, 124, 227–252.

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

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

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

Cornea, A. and R. Davidson (2011). A parametric bootstrap for heavy-tailed distributions. Mimeo. (R&R *Econometric Theory*).

Davidson, R. and E. Flachaire (2007). Asymptotic and Bootstrap Inference for Inequality and Poverty Measures. *Journal of Econometrics*, 141, 141–166.

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

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

Godfrey, L. (2009). *Bootstrap tests for regression models*. Palgrave MacMillan, Basingstoke.

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

Hendry, D.E. (1984). Monte Carlo experimentation in Econometrics. In Griliches, Z. and Intriligator, M. D. (eds), *Handbook of Econometrics*, vol., 2, North-Holland, Amsterdam.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3, 1163–1174.

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

Maasoumi, E. (1977). A study of improved methods of estimating the reduced form coefficients based upon 3SLS. Ph.D. Thesis, London School of Economics and Political Science.

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

Nagar, A. L. (1959). The bias and the moment matrix of the general k-class estimator of the parameters in simultaneous equations, *Econometrica*, 27, 575–595.

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

Sargan, J. D (1974). The validity of Nagar's expansion for the moments of econometric estimators. *Econometrica*, 42, 169–176.

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

van Giersbergen, N. P. A. and J. F. Kiviet (2002). How to implement the bootstrap in static or stable dynamic regression models: test statistic versus confidence region approach. *Journal of Econometrics*, 108, 133–156.

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

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

Table 1: Bootstrap coverage probabilities for E(x) without transformation.

	ordinary bootstrap			m out of n bootstrap		
	0.90	0.95	0.99	0.90	0.95	0.99
$\lambda = 1.1$	0.20	0.21	0.23	0.27	0.28	0.28
$\lambda = 1.3$	0.48	0.52	0.56	0.73	0.76	0.79
$\lambda = 1.5$	0.63	0.66	0.63	0.89	0.91	0.95

Table 2: Ordinary bootstrap coverage probabilities for E(x) after transformation.

	0.90	0.95	0.99
$\lambda = 1.1$	0.89	0.93	0.98
$\lambda = 1.3$	0.89	0.94	0.98
$\lambda = 1.5$	0.88	0.93	0.98

Table 3: $y \sim \text{Gam}(2, 1.1)$, $k = 2$, $\text{E}_*(x) = 48.113837381$

	raw			centered		
m	E_k	Bound	$\text{E}_k \,/\, \text{E}_*(x)$	E_k	Bound	$\text{E}_k \,/\, \text{E}_*(x)$
1	36.02031518	79.52389775	0.748647731	31.85063757	112.8084388	0.661984978
10	22.75139093	739.1814387	0.472865857	59.81391499	232.8180046	1.243174900
50	50.59838790	3.901968792	1.051639001	49.48870901	2.079265927	1.028575389
100	48.72227366	0.882972177	1.012645765	48.45340538	0.488301544	1.007057595
500	48.13800244	0.034209684	1.000502248	48.12736420	0.019144562	1.000281142
1000	48.11987728	0.008543904	1.000125534	48.11721864	0.004783143	1.000070276

Table 4: $y \sim \text{Gam}(2, 1.1)$, $k = 3$, $\text{E}_*(x) = 48.113837381$

	raw			centered		
m	E_k	Bound	$\text{E}_k \,/\, \text{E}_*(x)$	E_k	Bound	$\text{E}_k \,/\, \text{E}_*(x)$
1	686.7873982	4271.121409	14.27421789	1295.665613	9251.947117	26.92916806
10	62.26001429	66.30685719	1.29401473	53.4084961	18.29835898	1.110044407
50	48.13614219	0.183205259	1.000463584	48.12306332	0.087691534	1.000191752
100	48.11525298	0.022820961	1.000029422	48.11441994	0.010941749	1.000012108
500	48.11383966	0.000182483	1.000000047	48.11383832	8.74994E-05	1.000000019

Table 5: $y \sim \text{Gam}(2, 1.1)$, $k = 4$, $\text{E}_*(x) = 48.113837381$

	raw			centered		
m	E_k	Bound	$\text{E}_k \,/\, \text{E}_*(x)$	E_k	Bound	$\text{E}_k \,/\, \text{E}_*(x)$
1	656.6211587	17041.23391	13.64724151	1280.201915	34238.65768	26.60776992
10	45.93426141	5.088503652	0.954699602	47.10456151	2.057393128	0.979023168
50	48.10824978	0.008049875	0.999883867	48.11153053	0.003306909	0.999952054
100	48.11348322	0.000503168	0.999992639	48.11369172	0.000206678	0.999996973

Table 6: $y \sim \text{Gam}(2, 1.3)$, $k = 2$, $\text{E}_*(x) = 16.2588462282$

	raw			centered		
m	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$
1	10.90027129	39.49266418	0.67042096	12.04682277	28.36346446	0.740939585
10	20.05461030	42.98010744	1.233458391	19.51817111	14.75765738	1.200464709
50	16.56024987	0.450947284	1.018537825	16.41722038	0.232551018	1.009740799
100	16.33345748	0.106994115	1.004588963	16.29821230	0.056215388	1.002421210
500	16.26182072	0.004208903	1.000182946	16.26041783	0.002224563	1.000096661
1000	16.25958977	0.001051677	1.000045732	16.25923910	0.000555951	1.000024164

Table 7: $y \sim \text{Gam}(2, 1.3)$, $k = 3$, $\text{E}_*(x) = 16.2588462282$

	raw			centered		
m	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$
1	357.3868597	2437.040583	21.98107139	179.4132451	1207.692768	11.03480792
10	17.27127102	3.435483172	1.062269166	16.65696979	1.263222364	1.024486582
50	16.26061542	0.017805411	1.000108814	16.25954838	0.008141218	1.000043186
100	16.25895783	0.002222322	1.000006864	16.25889039	0.001016834	1.000002716
500	16.25884641	1.77665E-05	1.000000011	16.2588463	8.12314E-06	1.000000004

Table 8: $y \sim \text{Gam}(2, 1.3)$, $k = 4$, $\text{E}_*(x) = 16.2588462282$

	raw			centered		
m	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$
1	347.0041084	9138.275446	21.34248049	175.6001415	4476.162379	10.80028306
10	16.05949307	0.393932431	0.987738788	16.17271909	0.155668905	0.994702752
50	16.25840397	0.000633313	0.999972799	16.25867084	0.000250341	0.999989213
100	16.25881833	3.95755E-05	0.999998284	16.25883519	1.56374E-05	0.999999321

Table 9: $y \sim \text{Gam}(1, 1.1)$, $k = 2$, $\text{E}_*(x) = 7.64132294$

	raw			centered		
m	E_k	Bound	$\text{E}_k / \text{E}_*(x)$	E_k	Bound	$\text{E}_k / \text{E}_*(x)$
1	5.348890454	18.11267799	0.69999534	5.547758808	13.9716725	0.726020723
10	9.070487448	21.76358109	1.18703103	9.250433075	11.70554432	1.210580045
50	7.779876421	0.208442261	1.018132132	7.738604908	0.144474011	1.012731037
100	7.675580602	0.049193328	1.00448321	7.665445087	0.034536134	1.003156802
500	7.642688123	0.001931828	1.000178657	7.642285167	0.001361787	1.000125924
1000	7.641664198	0.000482679	1.000044659	7.641563478	0.000340293	1.000031478

Table 10: $y \sim \text{Gam}(1, 1.1)$, $k = 3$, $\text{E}_*(x) = 7.64132294$

	raw			centered		
m	E_k	Bound	$\text{E}_k / \text{E}_*(x)$	E_k	Bound	$\text{E}_k / \text{E}_*(x)$
1	132.5229086	1132.385351	17.34292734	120.6061654	847.8742229	15.78341424
10	8.141174023	1.725595451	1.065414207	7.936898676	0.954853041	1.038681225
50	7.642193478	0.008490408	1.000113925	7.641844525	0.005513082	1.000068258
100	7.641377893	0.001059513	1.000007191	7.641355802	0.000688324	1.0000043
500	7.641323032	8.47005E-06	1.000000012	7.641322996	5.50162E-06	1.000000007

Table 11: $y \sim \text{Gam}(1, 1.1)$, $k = 4$, $\text{E}_*(x) = 7.64132294$

	raw			centered		
m	E_k	Bound	$\text{E}_k / \text{E}_*(x)$	E_k	Bound	$\text{E}_k / \text{E}_*(x)$
1	137.4851024	3506.24111	17.99231669	116.763729	2896.920893	15.28056462
10	7.544974283	0.194066642	0.987391102	7.580423302	0.115830347	0.992030223
50	7.641105291	0.000311884	0.999971516	7.641192612	0.00018637	0.999982944
100	7.641309205	1.94899E-05	0.999998202	7.64131473	1.1645E-05	0.999998925

Table 12: $y \sim \text{Gam}(1, 1.3)$, $k = 2$, $\text{E}_*(x) = 4.1276252974$

	raw			centered		
m	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$
1	3.222377874	7.159466827	0.780685659	3.459858499	5.498965123	0.838220102
10	4.610126678	1.932521013	1.116895635	4.478563404	1.101408133	1.085021794
50	4.150268872	0.033086837	1.005485860	4.142876889	0.022165928	1.003695004
100	4.133258332	0.008031035	1.001364716	4.131424273	0.005412930	1.000920378
500	4.127850246	0.000318227	1.000054498	4.127777070	0.000214898	1.000036770
1000	4.127681532	7.95323E-05	1.000013624	4.127663239	5.37101E-05	1.000009192

Table 13: $y \sim \text{Gam}(1, 1.3)$, $k = 3$, $\text{E}_*(x) = 4.1276252974$

	raw			centered		
m	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$
1	41.10209787	298.1583280	9.957807433	35.42614118	246.8631630	8.582693104
10	4.179104175	0.166211312	1.012471790	4.157808888	0.098258707	1.007312580
50	4.127715779	0.001092510	1.000021921	4.127678059	0.000686532	1.000012783
100	4.127630986	0.000136464	1.000001378	4.127628611	8.57695E-05	1.000000803
500	4.127625307	1.08802E-06	1.000000002	4.127625303	6.83128E-07	1.000000001

Table 14: $y \sim \text{Gam}(1, 1.3)$, $k = 4$, $\text{E}_*(x) = 4.1276252974$

	raw			centered		
m	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$	E_k	Bound	$\text{E}_k\,/\,\text{E}_*(x)$
1	41.16945050	1034.516608	9.974124958	35.33930121	867.5052936	8.561654382
10	4.116393816	0.020057047	0.997278948	4.120839821	0.011688021	0.998356082
50	4.127602697	3.22434E-05	0.999994525	4.127612121	1.87732E-05	0.999996808
100	4.127623875	2.01233E-06	0.999999655	4.127624469	1.17092E-06	0.999999799

Table 15: $y \sim \mathrm{N}(1,1)$, $k=2$, $\mathrm{E}_*(x) = 4.4703660027$

	raw			centered		
m	E_k	Bound	$\mathrm{E}_k\,/\,\mathrm{E}_*(x)$	E_k	Bound	$\mathrm{E}_k\,/\,\mathrm{E}_*(x)$
1	2.98772013	11.18380332	0.668339042	4.098873968	3.711456770	0.916898967
10	4.573313241	0.160009011	1.023028816	4.499708029	0.045612240	1.006563674
50	4.474510726	0.005887077	1.000927155	4.471546512	0.001749706	1.000264074
100	4.471402220	0.001467883	1.000231797	4.470661165	0.000436851	1.000066026
500	4.470407452	5.86654E-05	1.000009272	4.470377810	1.74653E-05	1.000002641
1000	4.470376365	1.46655E-05	1.000002318	4.470368954	4.36445E-06	1.000000660

Table 16: $y \sim \mathrm{N}(1,1)$, $k=3$, $\mathrm{E}_*(x) = 4.4703660027$

	raw			centered		
m	E_k	Bound	$\mathrm{E}_k\,/\,\mathrm{E}_*(x)$	E_k	Bound	$\mathrm{E}_k\,/\,\mathrm{E}_*(x)$
1	49.77001850	409.4277092	11.13332073	11.14246140	66.45933835	2.492516585
10	4.472402982	0.011294884	1.000455663	4.470738257	0.002604166	1.000083272
50	4.470369340	8.97957E-05	1.000000747	4.470366608	2.07783E-05	1.000000135
100	4.470366211	1.12231E-05	1.000000047	4.470366041	2.59425E-06	1.000000008
500	4.470366003	8.13632E-08	1.000000000	4.470366003	1.52491E-08	1.000000000

Table 17: $y \sim \mathrm{N}(1,1)$, $k=4$, $\mathrm{E}_*(x) = 4.4703660027$

	raw			centered		
m	E_k	Bound	$\mathrm{E}_k\,/\,\mathrm{E}_*(x)$	E_k	Bound	$\mathrm{E}_k\,/\,\mathrm{E}_*(x)$
1	51.04836227	1304.891883	11.41928026	11.12050304	190.6877168	2.487604601
10	4.469867668	0.000738275	0.999888525	4.470274316	0.000136150	0.999979490
50	4.470365169	1.17793E-06	0.999999813	4.470365852	2.13008E-07	0.999999966
100	4.470365951	6.65426E-08	0.999999988	4.470365993	9.62026E-09	0.999999998

Table 18: Transformation-based bootstrap coverage probabilities for $\mathrm{E}(x)$

				$k = 1$				
		$n = 100$				$n = 1000$		
	m	0.90	0.95	0.99	m	0.90	0.95	0.99
$\lambda = 1.1$	1	0.98	0.98	0.99	1	1	1	1
	2	0.58	0.58	0.60	2	0.99	0.99	0.99
	3	0.48	0.49	0.54	3	0.67	0.71	0.75
	4	0.51	0.53	0.57	4	0.71	0.74	0.79
$\lambda = 1.3$	1	0.99	0.99	0.99	9	0.99	0.99	0.99
	2	0.88	0.90	0.93	12	0.97	0.98	0.99
	3	0.88	0.90	0.93	16	0.95	0.96	0.98
	4	0.87	0.89	0.92	20	0.92	0.93	0.96
$\lambda = 1.5$	1	0.99	0.99	0.99	16	0.98	0.99	0.99
	2	0.96	0.97	0.98	20	0.97	0.98	0.99
	3	0.93	0.95	0.97	25	0.95	0.96	0.98
	4	0.86	0.88	0.91	30	0.93	0.94	0.97
				$k = 2$				
		$n = 100$				$n = 1000$		
	m	0.90	0.95	0.99	m	0.90	0.95	0.99
$\lambda = 1.1$	1	0.90	0.90	0.92	3	0.99	0.99	0.99
	2	0.82	0.84	0.86	4	0.98	0.98	0.99
	3	0.77	0.79	0.81	5	0.93	0.94	0.95
	4	0.66	0.67	0.69	6	0.86	0.87	0.89
$\lambda = 1.3$	1	0.98	0.99	0.99	4	0.99	0.99	0.99
	2	0.96	0.97	0.98	5	0.99	0.99	0.99
	3	0.90	0.91	0.93	6	0.96	0.96	0.98
	4	0.82	0.83	0.86	7	0.92	0.94	0.95
$\lambda = 1.5$	1	0.99	0.99	0.99	6	0.97	0.97	0.98
	2	0.97	0.98	0.99	7	0.94	0.95	0.97
	3	0.92	0.93	0.95	8	0.91	0.93	0.96
	4	0.86	0.88	0.91	9	0.89	0.91	0.94
				$k = 3$				
		$n = 100$				$n = 1000$		
	m	0.90	0.95	0.99	m	0.90	0.95	0.99
$\lambda = 1.1$	1	0.88	0.88	0.90	1	1	1	1
	2	0.81	0.82	0.83	2	0.99	1	1
	3	0.59	0.60	0.61	3	0.98	0.98	0.99
	4	0.40	0.42	0.43	4	0.84	0.84	0.86
$\lambda = 1.3$	1	0.97	0.98	0.99	1	1	1	1
	2	0.88	0.89	0.90	2	0.99	1	1
	3	0.69	0.71	0.75	3	0.98	0.99	0.99
	4	0.58	0.60	0.65	4	0.87	0.89	0.91
$\lambda = 1.5$	1	0.98	0.98	0.99	1	1	1	1
	2	0.86	0.88	0.91	2	0.99	0.99	1
	3	0.74	0.76	0.82	3	0.97	0.98	0.99
	4	0.68	0.72	0.78	4	0.89	0.91	0.94