On bootstrap sample size in extreme value theory

Jaap Geluk and Laurens de Haan
Econometric Institute, Erasmus University Rotterdam,
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

October 28, 2002

Econometric Institute Report 2002-40

Abstract

It has been known for a long time that for bootstrapping the probability distribution of the maximum of a sample consistently, the bootstrap sample size needs to be of smaller order than the original sample size. See Jun Shao and Dongsheng Tu (1995), Ex. 3.9,p. 123. We show that the same is true if we use the bootstrap for estimating an intermediate quantile.

AMS 1991 subject classification. Primary 62G30 ; secondary 60G70.

Key words. Bootstrap, regular variation.

1 Introduction

The definition, most of the properties and possible applications of regularly varying functons are due to J. Karamata. See for example Karamata (1930,1931,1933)

In particular the uniform convergence property discovered by Karamata, has been an extremely useful tool in applications. Since then the theory has been extended greatly and many new applications have been discovered. For an overview of this see the monographs by Bingham et al. (1987) or Geluk and de Haan (1987). Here we shall deal with an application in extreme value statistics.

A direct generalization of regular variation is the following property. Suppose f is a measurable function and for all $x > 0$

$$\lim_{t\to\infty} \frac{f(tx) - f(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{1.1}$$

where γ is a real-valued parameter and a a suitable positive function. For $\gamma = 0$, the right-hand side is defined by continuity. This property can be called extended regular variation. The limit function, if not identically zero, is more or less the only possible one (see e.g. Geluk et al. (1987), thm. 1.9).

In the problem we are going to consider here, the following second order relation connected with (1.1) will be used. Suppose f is a measurable function and for $x > 0$

$$\lim_{t\to\infty} \frac{\frac{f(tx)-f(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = \frac{1}{\rho}\left(\frac{x^{\rho+\gamma} - 1}{\rho+\gamma} - \frac{x^\gamma - 1}{\gamma}\right), \tag{1.2}$$

where γ is a real-valued parameter as before, ρ is a non-positive parameter, a a suitable positive function and A of constant sign and converging to zero as $t \to \infty$. As before the limit is more or less the only possible one. For $\gamma = 0$ and/or $\rho = 0$ it is defined by continuity. The function a is always regularly varying of index γ and the function $|A|$ regularly varying of index ρ. Since the function A controls the speed of convergence in (1.2), so does ρ, i.e. large values of $|\rho|$ correspond with a high speed of convergence in (1.2) (cf. de Haan et al.(1996)). Many of the properties discovered by Karamata for regularly varying functions, have analogues for the relations (1.1) and (1.2), in particular the uniform convergence property.

For the application we want to consider, we restrict ourselves to the relatively simple case where $\rho < 0$ in (1.2) and $A(t) \sim ct^\rho$ $(t \to \infty), c \neq 0$. In that situation relation (1.2) is equivalent to

$$f(x) = c_1 \frac{x^\gamma - 1}{\gamma} + c_2 x^{\gamma+\rho} + o(x^{\gamma+\rho}),\ x \to \infty. \tag{1.3}$$

See e.g. (de Haan et al. (1996), thm. 2(iii)).

Next we discuss the extreme value context. Let $X_1, X_2, \ldots X_n$ be independent and identically distributed random variables with common distribution function F. Consider the order statistics $X_{1,n} \leq \ldots \leq X_{n,n}$. The maximum $X_{n,n}$, linearly normalized, has a non-degenerate limit distribution if and only if (1.1) holds for the function

$$U := \left(\frac{1}{1-F}\right)^{\leftarrow},$$

where the arrow indicates the generalized inverse function.

Since most estimators in extreme value theory are functions of extreme and intermediate order statistics, one needs also convergence in distribution of linearly normalized intermediate (as opposed to extreme) order statistics, namely $X_{n-k,n}$ with $k = k(n) \to \infty$, $k(n)/n \to 0$ as $n \to \infty$. A convenient condition to achieve this convergence is condition (1.2) for the function U.

Let us elaborate a little on the convergence in distribution for intermediate order statistics. Empirical process theory (see e.g. de Haan, Resnick (1993,p. 295) or Drees(1998)) tells us that for $Y_1, Y_2, \ldots$ i.i.d. with distribution function $1 - 1/x,\ x \geq 1$

$$\sqrt{k}\left(\frac{k}{n}Y_{n-kx,n} - \frac{1}{x}\right) \xrightarrow{d} x^{-2}B(x)$$

for $n \to \infty, k = k(n) \to \infty, k(n)/n \to 0$ in $D(0,\infty)$, where B is Brownian motion and where $Y_{n-kx,n}$ is a simplified notation for $Y_{n-[kx],n}$.

Suppose now for simplicity that (1.3) holds for U. Since $(U(Y_{1,n}),\ldots,U(Y_{n,n})) \stackrel{d}{=} (X_{1,n},\ldots,X_{n,n})$, we can use the former instead of the latter. Note that

$$\begin{aligned} X_{n-kx,n} - U(\frac{n}{k}) &\stackrel{d}{=} U(Y_{n-kx,n}) - U(\frac{n}{k}) \\ &= c_1(\frac{n}{k})^\gamma \frac{(\frac{k}{n}Y_{n-kx,n})^\gamma - 1}{\gamma} - c_2(\frac{n}{k})^{\gamma+\rho}((\frac{k}{n}Y_{n-kx,n})^{\gamma+\rho} - 1) + o_p(\frac{n}{k})^{\gamma+\rho} \end{aligned}$$

as $n \to \infty$ locally uniformly in x. We shall need the local uniformity later, but we shall not mention it each time. It follows that

$$\begin{aligned} &\frac{X_{n-kx,n} - U(\frac{n}{k})}{c_1(\frac{n}{k})^\gamma} - \frac{x^{-\gamma}-1}{\gamma} \\ &= \frac{(\frac{k}{n}Y_{n-kx,n})^\gamma - x^{-\gamma}}{\gamma} + (\frac{n}{k})^\rho \frac{c_2}{c_1}((\frac{k}{n}Y_{n-kx,n})^{\gamma+\rho} - 1) + o_p((\frac{n}{k})^\rho) \\ &= \frac{(\frac{1}{x} + \frac{x^{-2}B(x)}{\sqrt{k}} + o_p(\frac{1}{\sqrt{k}}))^\gamma - x^{-\gamma}}{\gamma} + (\frac{n}{k})^\rho \frac{c_2}{c_1}(x^{-\gamma-\rho} - 1) + o_p((\frac{n}{k})^\rho) \\ &= \frac{x^{-1-\gamma}B(x)}{\sqrt{k}} + (\frac{n}{k})^\rho \frac{c_2}{c_1}(x^{-\gamma-\rho} - 1) + o_p((\frac{n}{k})^\rho) + o_p(\frac{1}{\sqrt{k}}). \end{aligned} \tag{1.4}$$

Note that the first term is random and decreases in size with k (variance component) and the second term is non-random and increases in size with k (bias component).

We have achieved the following. Suppose one wants to estimate a moderately high quantile of the distribution function F, i.e. suppose we want to estimate $U(r(n))$ with $r(n) \to \infty$ and $r(n) = o(n), n \to \infty$. Relation (1.4) gives the asymptotic expansion of $X_{n-k,n}$ with

$$k = k(n) = n/r(n), \tag{1.5}$$

when used as an estimator for $U(r(n))$.

In the next section we shall consider the bootstrap intermediate order statistics and we shall show that for the bootstrap version of $X_{n-kx,n}$ an expansion holds similar to (1.4) *only if* the bootstrap sample size n_1 is $o(n)$ as $n \to \infty$.

2 Bootstrap sample size

We consider the bootstrap. Suppose we have available a sample $X_1, X_2, \ldots, X_n$ from a distribution with cdf F. Let F_n be the corresponding empirical distribution function and define

$$U_n := \left(\frac{1}{1-F_n}\right)^{\leftarrow}.$$

Note that

$$U_n(\frac{n}{kx}) = X_{n-kx,n} \stackrel{d}{=} U(Y_{n-kx,n}) \tag{2.1}$$

with the Y's as before.

Let us also write (1.4) in terms of U_n:

$$\begin{aligned}
&\frac{U_n(\frac{n}{kx}) - U(\frac{n}{k})}{c_1(\frac{n}{k})^{\gamma}} = \\
&\frac{x^{-\gamma}-1}{\gamma} + \frac{x^{-1-\gamma}B(x)}{\sqrt{k}} + (\frac{n}{k})^{\rho}\frac{c_2}{c_1}(x^{-\gamma-\rho}-1) + o_p((\frac{n}{k})^{\rho}) + o_p(\frac{1}{\sqrt{k}}).
\end{aligned} \tag{2.2}$$

For the bootstrap we select independently with replacement n_1 times (with equal probability) an element from $\{X_1, X_2, \ldots, X_n\}$. This is our bootstrap sample $\{X_1^*, X_2^*, \ldots, X_{n_1}^*\}$. We form the order statistics $X_{1,n_1}^* \leq X_{2,n_1}^* \leq \ldots \leq X_{n_1,n_1}^*$.

Note the important relation

$$X_{n_1-k_1x,n_1}^* \stackrel{d}{=} U_n(Y_{n_1-k_1x,n_1}^*) \tag{2.3}$$

with $Y_1^*, Y_2^*, \ldots, Y_{n_1}^*$ i.i.d. with distribution function $1-1/x,\ x \geq 1$. Here U_n represents the randomness of the original sample and the Y^*'s the extra randomness introduced by resampling. Note that as before

$$\sqrt{k_1}\{\frac{k_1}{n_1}Y_{n_1-k_1x,n_1}^* - \frac{1}{x}\} \stackrel{d}{\to} x^{-2}B^*(x) \tag{2.4}$$

in $D(0,\infty)$ with B^* Brownian motion for $n_1 \to \infty, k_1 = k_1(n_1) \to \infty, k_1(n_1)/n_1 \to 0$.

Let us consider an expansion like (1.4) for the bootstrap intermediate order statistics. Now take

$$k_1 = n_1/r(n), \tag{2.5}$$

similar to what we did before. We use the notation $W_{n_1}^* = \frac{n_1}{k_1}/Y_{n_1-k_1x,n_1}^*$. In view of (2.3),(1.5) and (2.5)

$$\begin{aligned}
\frac{X_{n_1-k_1x,n_1}^* - U(r(n))}{c_1(\frac{n}{k})^{\gamma}} &= \frac{X_{n_1-k_1x,n_1}^* - U(\frac{n_1}{k_1})}{c_1(\frac{n}{k})^{\gamma}} \stackrel{d}{=} \frac{U_n(Y_{n_1-k_1x,n_1}^*) - U(\frac{n_1}{k_1})}{c_1(\frac{n}{k})^{\gamma}} \\
&= \frac{U_n(\frac{n}{k}(\frac{k_1}{n_1}Y_{n_1-k_1x,n_1}^*) - U(\frac{n}{k})}{c_1(\frac{n}{k})^{\gamma}} = \frac{U_n(\frac{n}{k}/W_{n_1}^*) - U(\frac{n}{k}))}{c_1(\frac{n}{k})^{\gamma}}.
\end{aligned}$$

In view of (2.2) this equals (note that $W^*_{n_1} \xrightarrow{p} x$)

$$\frac{(W^*_{n_1})^{-\gamma}-1}{\gamma}+\frac{(W^*_{n_1})^{-1-\gamma}B(W^*_{n_1})}{\sqrt{k}}+(\frac{n}{k})^{\rho}\frac{c_2}{c_1}((W^*_{n_1})^{-\gamma-\rho}-1)+o_p((\frac{n}{k})^{\rho})+o_p(\frac{1}{\sqrt{k}}).$$

Using the expansion (2.4), this becomes

$$\left(\frac{x^{-\gamma}-1}{\gamma}+\frac{x^{-\gamma-1}B^*(x)}{\sqrt{k_1}}+o_p(\frac{1}{\sqrt{k_1}})\right)+\left(\frac{x^{-1-\gamma}B(x)}{\sqrt{k}}+o_p(\frac{1}{\sqrt{k}})\right)$$
$$+\left((\frac{n_1}{k_1})^{\rho}\frac{c_2}{c_1}(x^{-\gamma-\rho}-1)+o_p(\frac{n_1}{k_1})^{\rho}\right)+o_p((\frac{n}{k})^{\rho})+o_p(\frac{1}{\sqrt{k}})=$$
$$\frac{x^{-\gamma}-1}{\gamma}+\frac{x^{-\gamma-1}B^*(x)}{\sqrt{k_1}}+(\frac{n_1}{k_1})^{\rho}\frac{c_2}{c_1}(x^{-\gamma-\rho}-1)+\frac{x^{-1-\gamma}B(x)}{\sqrt{k}}+o_p((\frac{n}{k})^{\rho})+o_p(\frac{1}{\sqrt{k}})+o_p(\frac{1}{\sqrt{k_1}})$$

Since we want the bootstrap sample to reflect the properties of the original sample (cf. (1.4)), we need to get rid of the fourth term asymptotically. This is possible only if we require $k_1 = o(k)$, i.e. $n_1 = o(n)$ (see (1.5) and (2.5)).

References

[1] N. Bingham, C. Goldie and J. Teugels (1987). Regular Variation. *Cambridge University Press.*

[2] H. Drees, On smooth statistical tail functionals, Scand. J. Statist. 25 (1998), no. 1, 187-210.

[3] J.L. Geluk, L. de Haan (1987). Regular variation, extensions and Tauberian theorems, *Mathematical Centre Tract 40, Amsterdam.*

[4] L. de Haan, S.I. Resnick,Estimating the limit distribution of multivariate extremes, Comm. Stat.-Stch. Models, 9, (1993),275-309.

[5] L. de Haan, U. Stadtmüller (1996). Generalized regular variation of second order. J. Austral. Math. Soc. Ser. A 61 (1996), no. 3, 381-395.

[6] J. Karamata, Sur un mode de croissance regulière des fonctions, Mathematica (Cluj), 4, 1930, 38-53.

[7] J. Karamata, Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjessche Transformation betreffen, Journal für die Reine und angewandte Mathematik, 164, 1931, 27-39.

[8] J. Karamata, Sur un mode de croissance regulière. Théorèmes fondamentaux, Bull. Soc. Math. France, 61, 1933, 55-62.

[9] Jun Shao, Dongsheng Tu (1995). The jackknife and the bootstrap. *Springer Verlag.*