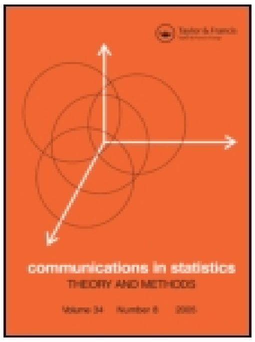

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Fat tail distributions and local thin tail alternatives

Geert Gielens , Stefan Straetmans & Casper G. De vries

To cite this article: Geert Gielens , Stefan Straetmans & Casper G. De vries (1996) Fat tail distributions and local thin tail alternatives, Communications in Statistics - Theory and Methods, 25:4, 705-710, DOI: 10.1080/03610929608831726

To link to this article: https://doi.org/10.1080/03610929608831726

Published online: 27 Jun 2007.

COMMUN. STATIST.—THEORY METH., 25(4), 705-710 (1996)

FAT TAIL DISTRIBUTIONS AND LOCAL THIN TAIL ALTERNATIVES

Geert GIELENS

Center for Economic Studies, K.U.Leuven,
Naamsestraat 69, 3000 Leuven, Belgium;

Stefan STRAETMANS
Casper G. de VRIES

Tinbergen Instituut and Erasmus Universiteit Rotterdam,
Oostmaaslaan 950, 3063 DM Rotterdam, The Netherlands;

Key Words and Phrases: Hill estimator; tail index; bias; stable distribution

ABSTRACT

The behaviour of the Hill estimator for the tail index of fat tailed distributions in the presence of local alternatives which have a thin tail is investigated. The converse problem is also briefly addressed. A local thin tail alternative can severely bias the Hill statistic. The relevance of this issue for the class of stable distributions is discussed. We conduct a small simulation study to support the analysis. In the conclusion it is argued that for moderate out of sample quantile analysis the problem of local alternatives may be less pressing.

1. INTRODUCTION

It is well known that the Hill estimator for the tail index of fat tailed distributions (dfs) is a biased estimator except in the case of the Pareto law; see e.g. Hall (1982) and Goldie and Smith (1987). The source of the bias studied in these articles stems from the second order term in the expansion of the df at infinity. In particular, this term is mostly taken to be hyperbolic. In a recent paper De Haan and Stadtmüller (1992) show that this specific expansion applies to a wide class of distributions, like e.g. the Student-t and the type II extreme value dfs. Nevertheless, other sources for the bias exist as well. In this note we investigate the bias that is due to a local thin tailed alternative. To introduce the topic, consider the Student-t dfs versus the stable laws. The interesting difference between the Student model and the stable model is that while the Student model has all moments bounded only as the degrees of freedom tend to infinity, with the stable model there is a discrete jump in the number of moments once the characteristic exponent equals two. We will say that in the latter case there is a local thin tailed alternative.

In the next section we generalise the problem of stable laws to the class of heavy tailed dfs. It is shown that for any fat tailed df with tail index α, one can always construct a local alternative which has thin tails. Conversely, a local fat tailed alternative to a thin tailed df is straightforward to obtain as well. Two test functions in which local alternatives from the other class are nested are presented. In section 3, we obtain explicit expressions for the bias in the Hill estimator in the case of the stable dfs and for the two test functions. The fourth section presents a small Monte Carlo study that complements the theoretical results. We conclude by arguing that the problem of local alternatives may have little bearing on moderate out of sample quantile analysis.

2. FAT TAILS AND LOCAL ALTERNATIVES

First consider the class of symmetric stable distributions with characteristic function:

$$\varphi_{\phi}(t) = exp\,(-\left|t\right|^{\phi}) \quad , 0 < \phi \leq 2\,, \tag{1}$$

where ϕ is the characteristic exponent. For $\phi < 2$, $\alpha = \phi$ determines the maximal bounded moment. But when $\phi = 2$, the case of the normal, all moments are finite. Note the continuity of φ_0 in ϕ, but the jump in the number of bounded moments at $\phi = 2$. Also note that the densities are continuous in ϕ, as can be directly inferred from the Bergström-Feller series expansion for the densities.

The discontinuity in the number of finite moments, through the presence of a local thin tailed alternative, is not particular to the stable laws. Let $H_\alpha(x)$ be any heavy tailed distribution, with tail index α, in the sense that $H_\alpha(x)$ varies regularly at infinity. Let the distribution of the maximum of L(x) be in the domain of attraction of $\exp(e^{-x})$. Then the mixture $F(x) = \lambda H_\alpha(x) + (1-\lambda) L(x)$, with $0 \leq \lambda \leq 1$, varies regularly at infinity and has bounded moments $m < \alpha$ as long as $\lambda > 0$. But for $\lambda = 0$, F(x) is thin tailed. A specific example is the mixture of a Pareto and an exponential df:

$$F(x) = 1 - \lambda x^{-\alpha} - (1-\lambda)\, e^{-(x-1)} \qquad x \geq 1 \quad , \quad 0 \leq \lambda \leq 1 \;. \tag{2}$$

Equation (2) will be our first test function. To increase the analogy between the stable laws and F(x), note that the mixing parameter λ can be made functionally dependent on the tail index α.

A similar procedure can be used to construct a fat tailed local alternative to a thin tailed df. Consider e.g. our second test function:

$$G(x) = 1 - (x^{-\alpha})^{\lambda}\ (e^{1-x})^{1-\lambda}. \tag{4}$$

For $\lambda < 1$, G(x) is thin tailed, and becomes the fat tailed Pareto at $\lambda = 1$. In this case the fat tailed Pareto is a local alternative to the thin tailed distribution.

3. THE BIAS IN THE HILL ESTIMATOR

In this section we give explicit expressions for the bias in the Hill estimator. The Hill statistic is defined as follows (see Goldie and Smith (1987)):

$$\widehat{\frac{1}{\alpha}} = \hat{\gamma} = \frac{1}{M}\sum_{i=1}^{M} \log\left(X_i \middle| X_i \geq s\right) - \log s, \tag{5}$$

where X_i is the i-th descending order statistic of a sample with lenght n and γ is the inverse of the tail index. Here, M denotes the random number of observations that exceed s, where s is some high threshold level. To compute the bias we need to evaluate the following integral:

$$E[\hat{\gamma}\mid s]=\frac{1}{1-F(s)}\int_{s}^{+\infty}\log x\, f(x)dx-\log s. \tag{6}$$

We start with the stable dfs. By using the asymptotic expansion for symmetric stable dfs with $1<\phi<2$ as x tends to infinity (Ibragimov and Linnik (1971, ch. 2)), we obtain the following result:

Proposition 1. For symmetric stable dfs, $1<\phi<2$,

$$E[\hat{\gamma}\mid s]-\gamma=-\frac{1}{2\phi}\cdot\frac{b}{b+s^{\phi}}+o(s^{-\phi}),\ \text{where } b=-\frac{1}{2}\frac{\Gamma(2\phi)\sin(\phi\pi)}{\Gamma(\phi)\sin\left(\frac{\phi\pi}{2}\right)}. \tag{7}$$

Proposition 2. For the test function F(x), as defined in equation (2) :

$$E[\hat{\gamma}\mid s]=\frac{(1-\lambda)\, e E_1(s)+\lambda\,\alpha^{-1}\, s^{-\alpha}}{(1-\lambda)\, e^{1-s}+\lambda\, s^{-\alpha}} \quad . \tag{8}$$

Proof. Straightforward calculus on the analogue of equation (6) yields (8). Here $E_1(x)$ is the exponential integral (see e.g. Abramowitz and Stegun, 1970). (Q.E.D.)

Proposition 3. For the test function G(x) as defined in equation (4):

$$E[\hat{\gamma}\mid s]=s^{\alpha\lambda}e^{s(1-\lambda)}(1-\lambda)^{\alpha\lambda}\,\Gamma(-\alpha\lambda,(1-\lambda)s),\ \text{if } \lambda<1. \tag{9}$$

If $\lambda=1$, G(x) reduces to the Pareto law and the Hill statistic is unbiased.

4. MONTE CARLO EVIDENCE

This section presents some Monte Carlo evidence on how the presence of a local alternative affects the performance of the Hill statistic. The exercise is limited to the class of stable dfs and the first test function. For both cases the inverse of the average estimate produced by the Hill statistic is reported. This average is in each case based on a sample of 3000 observations and 1000

TABLE I: The Hill statistic and stable laws

α	M	average α	M^*	average α
1.1	50	1.14	785	0.93
1.5	50	1.64	141	1.73
1.95	50	4.76	15	4.31

TABLE II: The Hill statistic and F(x)

		α					
λ	s	1	2	3	4	5	10
0.001	8.80	3.21	7.04	7.48	7.48	7.48	7.48
0.01	7.64	1.63	5.11	6.71	6.88	6.89	6.89
0.1	6.40	1.21	3.55	6.04	6.59	6.64	6.65
0.5	5.00	1.08	2.52	4.48	5.54	5.60	5.60
0.9	3.01	1.03	2.12	3.21	3.91	4.01	3.81

replications. These numbers were chosen for reason of comparability with a recent preprint by McCulloch (1994) that studies the bias of the Hill statistic in case of the stable laws. McCulloch fixes the number of order statistics at M=50. A summary of his results is replicated and reported in the third column of Table I. Instead of fixing the M level across different characteristic exponents, one might want to evaluate the bias at the M level that minimises the asymptotic MSE (see Goldie and Smith (1987)). This MSE minimising M level is given in the fourth column as M^*, and the fifth column reports the average α estimates. Using M^* somewhat helps to reduce the bias when $\phi = \alpha = 1.95$, but it remains considerable nevertheless. Due to the fact that for $\phi = 1.95$ the normal is a local alternative, only increasing the sample could help to alleviate the problem.

Table II reports the average estimate produced by the Hill statistic when the DGP is given by equation (2). The bias problem is investigated when λ tends to zero. The s levels were chosen such that, given some α level, the closer the thin tail local alternative is, the further one needs to go into the tail to capture the fat tail property.

It is evident from this table that the presence of a local thin tailed alternative pushes the estimates upwards. This causes a serious bias problem for the lower α and λ values. On the other hand the presence of a local alternative lessens the bias problem for the highest α and lowest λ combinations.

5. CONCLUSIONS

In this paper, we study the behavior of the Hill estimator conditional upon a DGP that lies close to a local alternative which is thin tailed. Calculation and simulation of the bias in the Hill estimator shows that it has limited value to discern between fat or thin tailed dfs, when the DGP lies close to the local alternative. Put differently, given a fixed sample size n, it is always possible to formulate a fat tailed df which is indistinguishable from the thin tailed DGP. An interesting extension of this study would be to consider the extreme value index estimators, like Pickand's estimator, which are not restricted to the domain of fat tailed dfs. Lastly, we plan to investigate the issue for moderate out of sample quantile estimation. From the second test function (4) it can be seen that for λ close to 1 the erroneous presumption that $G(x)$ is fat tailed has little bearing on the stated quantile problem.

BIBLIOGRAPHY

Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, No 55.

Goldie, C. M. and Smith, R. L. (1987). "Slow Variation with Remainder: Theory and Applications", *Quarterly Journal of Mathematics, Oxford 2nd series*, 38, 1987, 45-71.

Haan, L. de and Stadtmüller, U. (1992). "Generalized Regular Variation of Second Order", report 9268/A Econometric Institute Erasmus University Rotterdam.

Hall, P. (1982). " On Some Simple Estimates of an Exponent of Regular Variation", *Journal of the Royal Statistical Society, Series B*, 44, 37-42.

Ibragimov, I. A. and Linnik, Y. V. (1971). "*Independent and Stationary Sequences of Random Variables*", Wolters-Noordhoff, Groningen.

McCulloch, J. H. (1994). "Measuring Tail Thickness in Order to Estimate the Stable Index α : A Critique", unpublished manuscript.