Adaptive Estimation of Heavy Tail Distributions with application to Hall model

Dimitris N. Politis, Vyacheslav A. Vasiliev and Sergey E. Vorobeychikov

Abstract The problem of tail index estimation of Hall distribution is considered. We propose the estimators of tail index using the truncated estimation method developed for ratio type functionals. It is shown that the truncated estimator constructed on the sample of fixed size has a guaranteed accuracy in the sense of the L_{2m}-norm, $m \geq 1$. The asymptotic properties of estimators are although investigated. These properties make it possible to find the rates of decreasing of the χ^2 divergence in the almost surely sense between distribution and its adaptive estimator. Simulations confirm theoretical results.

1 Introduction

The models with heavy tail distributions are of interest in many applications connected with financial mathematics, insurance theory [1, 4, 15], telecommunication [16] and physics [2]. Usually it is assumed that the distribution function contains as an unknown multiplier a slowly varying function. The problem of tail index estimation was studied by Hill [9] who proposed the estimators based on the order statistics. The estimator is optimal in mean square sense on the class of distribution functions with heavy tails in presence of unknown slowly varying function. It should be noted that Hill's estimators are unstable and can diverge essentially from the estimated parameter for large sample sizes [4, 17].

Dimitris N. Politis
Department of Mathematics,University of California, San Diego, La Jolla, CA 92093-0112, USA
Tel.: (858) 534-5861Fax: (858) 534-5273 e-mail: dpolitis@ucsd.edu

Vyacheslav A. Vasiliev and Sergey E. Vorobeychikov
Department of Applied Mathematics and Cybernetics Tomsk State University Lenin 36, Tomsk, 634050, Russia Tel.: +7(3822)529-578 Fax: +7(3822)529-895 e-mail: vas@mail.tsu.ru and e-mail: sev@mail.tsu.ru

Later another approaches to estimation problem were proposed (see, e.g., [6, 10] and references therein). In [18] a least squares estimator for tail index was proposed which is based on the estimation of parameters in linear regression. The geometric-type estimators of the tail index are proposed and investigated in [2].

Some estimators have the form of ratio statistics, see e.g. Embrechts et al. [4]. For example, formula (1.7) of Markovich [10] describes a well–known class of ratio estimators which are generalization of Hill's estimator in the sense that an arbitrary threshold level instead of an order statistic is used – see e.g. Novak [11]-[14], Resnick and Starica [17], or Goldie and Smith [5].

In this paper, the truncated estimation method of ratio type functionals, proposed by Vasiliev [19], is used to obtain estimators with guaranteed accuracy in the sense of the L_{2m}-norm, $m \geq 1$. The estimators are constructed on the basis of empirical functionals without usage of non-parametric approach in an effort to obtain (or get close to) the parametric optimal rate of convergence. These estimators can be used to construct the adaptive estimators of distribution functions. It allows one to find the rates of decreasing $\varphi_\varepsilon^{-1}(n),\ \varepsilon > 0$ of the χ^2 divergence in the almost surely sense between distributions and their adaptive estimators.

As an example we have found the rate of decreasing for Hall distribution [7, 8] with unknown tail index. Similar results for convergence of the χ^2 divergence in probability are presented, e.g., in [6].

2 Adaptive distribution estimation

Let $\mathscr{F} = \{F_\Delta(x),\ x \in G \subseteq \mathscr{R}^1,\ \Delta \in \mathscr{D} \subseteq \mathscr{R}^q\}$ be the parametric family of heavy tail distributions. Here $\mathscr{D}$ is an admissible set of the unknown parameter Δ. Denote Δ_n an estimator of Δ.

Suppose that for every $\Delta \in \mathscr{D}$ the density $f_\Delta(x) = dF_\Delta(x)/dx$ exists. It is easy to verify that the χ^2 divergence between F_Δ and F_{Δ_n} has the form

$$\chi^2(F_\Delta, F_{\Delta_n}) = \int\limits_G \frac{dF_{\Delta_n}(x)}{dF_\Delta(x)} dF_{\Delta_n}(x) - 1 = \int\limits_G \left(\frac{f_{\Delta_n}(x)}{f_\Delta(x)} - 1\right)^2 f_\Delta(x) dx.$$

The problem is to construct estimators F_{Δ_n} of concrete well-known distributions F_Δ on the basis of a special type parameter estimators Δ_n with known rates of decreasing $\varphi_\varepsilon^{-1}(n),\ \varepsilon > 0$ of the χ^2 divergence in the following sense

$$\lim_{n\to\infty} \varphi_\varepsilon(n) \chi^2(F_\Delta, F_{\Delta_n}) = 0 \ \text{ a.s.} \tag{1}$$

Suppose the following

Assumption (A). Assume there exists the number $\delta_0 > 0$ such that for true value Δ the set $\mathscr{D}_0 = \{\delta :\ \Delta + \delta \subseteq \mathscr{D},\ ||\delta|| \leq \delta_0\}$ is not empty and

$$\sup_{\delta \in \mathscr{D}_0} \int_G ||\nabla_\Delta f_{\Delta+\delta}(x)||^2 f_\Delta^{-1}(x) dx < \infty.$$

Then, using the Taylor expansion for the function $f_{\Delta_n}(x)$ on the set $\Omega_n = \{\omega : ||\Delta_n - \Delta|| \le \delta_0\}$ we have

$$\chi^2(F_\Delta, F_{\Delta_n}) = \int_G (f_{\Delta_n}(x) - f_\Delta(x))^2 f_\Delta^{-1}(x) dx$$

$$= \int_G \left[\sum_{i=1}^q \frac{\partial f_{\Delta+\alpha(\Delta_n-\Delta)}(x)}{\partial \Delta_i} (\Delta_n - \Delta)_i \right]^2 \cdot f_\Delta^{-1}(x) dx$$

$$\le ||\Delta_n - \Delta||^2 \cdot \int_G ||\nabla_\Delta f_{\Delta+\alpha(\Delta_n-\Delta)}(x)||^2 f_\Delta^{-1}(x) dx \le C_0 ||\Delta_n - \Delta||^2,$$

where $\alpha \in (0,1)$.

Thus to prove (1) it is enough to find the functions $\varphi_\varepsilon(n)$ and estimators Δ_n such that

$$\lim_{n\to\infty} \varphi_\varepsilon(n) ||\Delta_n - \Delta||^2 = 0 \ \text{ a.s.} \tag{2}$$

The general truncated estimation method presented in [19] makes possible to obtain estimators of tail indexes of various type distributions with the properties

$$E||\Delta_n - \Delta||^{2p} \le r^{-1}(n,p), \quad n \ge 1, \tag{3}$$

which are fulfilled for every $p \ge 1$ and some functions $r(n,p) \to \infty$ as $n \to \infty$ and/or $p \to \infty$.

Define $\overline{\Omega}_n$ a complement of the set Ω_n. Suppose that there exists a number p_0 such that the series

$$\sum_{n\ge 1} r^{-1}(n, p_0) < \infty.$$

Then using inequalities

$$P(\chi^2(F_\Delta, F_{\Delta_n}) > C_0 ||\Delta_n - \Delta||^2) \le P(\overline{\Omega}_n) = P(||\Delta_n - \Delta|| > \delta_0)$$

$$\le \delta_0^{-2p_0} E||\Delta_n - \Delta||^{2p_0} \le \delta_0^{-2p_0} r^{-1}(n, p_0),$$

and the Borel-Cantelli lemma we have

$$||\Delta_n - \Delta||^{-2} \chi^2(F_\Delta, F_{\Delta_n}) \to 0 \quad \text{a.s.} \tag{4}$$

Define $\varphi(n,p) = (n^{-2} r(n,p))^{1/p}$. By making use the Borel-Cantelli lemma for every $p \ge 1$ in particular it follows

$$\lim_{n\to\infty} \varphi(n,p) ||\Delta_n - \Delta||^2 = 0 \ \text{ a.s.} \tag{5}$$

From (4) and (5) we get

$$\varphi(n,p)\chi^2(F_\Delta,F_{\Delta_n}) \to 0 \quad \text{a.s.}$$

and the function $\varphi_\varepsilon(n)$ can be defined as $\varphi_\varepsilon(n) = \varphi(n,p_\varepsilon)$ with an appropriate chosen $p_\varepsilon > 0$.

We will apply this approach to the adaptive estimation problem of the Hall model for distribution function $F_\Delta(x) = 1 - C_1 x^{-\gamma} - C_2 x^{-1/\alpha}$, $\gamma^{-1} = \alpha + \theta$.

In the next section the estimator $\Delta_n = \hat{\gamma}_n$ of γ with needed properties will be constructed and investigated.

Define $\rho = \theta/\alpha - 1$.

The following theorem presents the main result of the paper.

Theorem 1 *For every $\varepsilon > 0$ there exist numbers p_ε such that the property (1) for the Hall model is fulfilled with*

$$\varphi_\varepsilon(n) = n^{\frac{\rho+1}{\rho+2}-\varepsilon}.$$

3 Estimation of heavy tail index of the Hall model

The problem is to estimate by i.i.d. observations $X_1,\ldots,X_n$ the parameter $\gamma = 1/\beta$ of the Hall distribution function [7]

$$F_\Delta(x) = 1 - C_1 x^{-1/\beta} - C_2 x^{-1/\alpha}, \quad x \geq c,$$

where $\beta > 0$, $\alpha > 0$; $\beta = \alpha + \theta \geq \beta_0 > 0$.

Then the tail distribution function

$$P(x) = C_1 x^{-\gamma}(1 + C_3 x^{-\gamma(\rho+1)}),$$

where $C_3 = C_2/C_1$, $\gamma = 1/\beta$, $\rho = \theta/\alpha - 1$.

The density function has the form

$$f(x) = C_1\gamma x^{-(\gamma+1)} + (C_2/\alpha)x^{-(1/\alpha+1)}$$

and Assumption (A) is fulfilled for $\mathscr{D} = \{\Delta = \gamma,\ \gamma > 0\}$, $\mathscr{D}_0 = \{\delta :\ |\delta| \leq \gamma/2\}$ and $\delta_0 = \gamma/2$.

To construct the estimator for γ we find its appropriate representation.

For some $X = (x_1,x_2)$, $x_1 > x_2 > c$ by the definition of $P(x)$ we have

$$\log P(x_1) = \log C_1 - \gamma \log x_1 + \log(1 + C_3 x_1^{-\gamma(\rho+1)}),$$

$$\log P(x_2) = \log C_1 - \gamma \log x_2 + \log(1 + C_3 x_2^{-\gamma(\rho+1)}).$$

Thus we can find γ as a solution of this system

$$\gamma = \frac{\log(P(x_2)/P(x_1))}{\log(x_1/x_2)} - \log\left[1 + \frac{C_3(x_2^{-\gamma(\rho+1)} - x_1^{-\gamma(\rho+1)})}{1 + C_3 x_1^{-\gamma(\rho+1)}}\right]$$

and it is natural to define the estimators γ_n of γ as follows

$$\gamma_n(X) = \frac{\log(P_n(x_2)/P_n(x_1))}{\log(x_1/x_2)} \cdot \chi(P_n(x_1) \geq \log^{-1} n), \quad n > 1.$$

Here $P_n(x)$ is the empirical tail distribution function

$$P_n(x) = \frac{1}{n}\sum_{i=1}^{n} \chi(X_i \geq x).$$

To get the estimator γ_n with the optimal rate of convergence (in the sense of L_2-norm see [3, 6, 8]), we put for $p \geq 1$ the sequence $X(n) = (x_1(n), x_2(n))$, where

$$x_1(n) = \mathrm{e} \cdot x_2(n), \qquad x_2(n) = n^{\frac{p}{\gamma[2p(\rho+2)-1]}} \tag{6}$$

The deviation of this estimator has the form

$$\gamma_n(X) - \gamma = \Big\{ \log(P_n(x_2(n))/P(x_2(n))) - \log(P_n(x_1(n))/P(x_1(n)))$$

$$-\log\left[1 + \frac{C_3(1-\mathrm{e}^{-1})}{1 + C_3 x_1^{-\gamma(\rho+1)}(n)} x_2^{-\gamma(\rho+1)}(n)\right]\Big\} \cdot \chi(P_n(x_1(n)) \geq \log^{-1} n)$$

$$-\gamma \cdot \chi(P_n(x_1(n)) < \log^{-1} n). \tag{7}$$

For any $m \geq 1$ and $x \geq c$ it follows

$$E(P_n(x) - P(x))^{2m} \leq \frac{2B_m P(x)}{n^m}, \quad n \geq 1, \tag{8}$$

where B_m is a constant from the Burkholder inequality.

We will use the following inequality

$$|\log(P_n(x)/P(x))| = |\log\left(1 + \frac{P_n(x) - P(x)}{P(x)}\right) \cdot \chi(P_n(x) - P(x) > 0)$$

$$+\log\left(1 + \frac{|P_n(x) - P(x)|}{P_n(x)}\right) \cdot \chi(P_n(x) - P(x) \leq 0)|$$

$$\leq |P_n(x) - P(x)| \cdot \left[\frac{1}{P(x)} + \frac{1}{P_n(x)}\right] = |P_n(x) - P(x)| \cdot \left[\frac{2}{P(x)} + \left(\frac{1}{P_n(x)} - \frac{1}{P(x)}\right)\right]$$

$$\leq \frac{2|P_n(x)-P(x)|}{P(x)} + \frac{(P_n(x)-P(x))^2}{P(x)P_n(x)}.$$

Then using the c_r-inequality and (8) for $i=1,2$ we estimate

$$E\log^{2p}(P_n(x_i)/P(x_i))\cdot\chi\Big(P_n(x_1)\geq\log^{-1}n\Big)$$

$$\leq \frac{C}{n^pP^{2p-1}(x_1)} + \frac{C\log^{2p}n}{n^{2p}P^{2(p-1)}(x_1)}. \tag{9}$$

In what follows, C will denote a generic non-negative constant whose value is not critical (and not always the same).

Further, by the Chebyshev inequality and (8) we have

$$P(P_n(x)<\log^{-1}n) = P(P(x)-P_n(x)>P(x)-\log^{-1}n)$$

$$\leq \frac{E[P_n(x)-P(x)]^{4p}}{[P(x)-\log^{-1}n]^{4p}} \leq \frac{C}{n^{2p}P^{4p-1}(x)} \leq C\frac{x^{(4p-1)\gamma}}{n^{2p}}. \tag{10}$$

From (7), (9) and (10) it follows

$$E(\gamma_n(X(n))-\gamma)^{2p} \leq Cr^{-1}(n,p), \tag{11}$$

$$r(n,p) = n^{\frac{2p(\rho+1)}{2p(\rho+1)+2p-1}p}$$

and we can put according to the definition of $\varphi(n,p)=(n^{-2}r(n,p))^{1/p}$ with the $r(n,p)$ defined in (11)

$$p_\varepsilon \geq 2\varepsilon^{-1}, \qquad \varphi_\varepsilon(n) = n^{\frac{\rho+1}{\rho+2}-\varepsilon}.$$

Note that proposed parameter estimation procedure gives estimator γ_n with convergence rate, with optimal (for $p=1$) convergence rate, see[6]. At the same time the sequences (6) in the definition of γ_n depend on the unknown model parameters. Then the adaptive estimation procedure should be constructed, e.g., on the presented scheme using some estimators of γ and ρ. The main aim is to get adaptive estimators with the optimal convergence rate.

Consider, for instance, the case of known ρ. Define the known deterministic sequence $(m_n)_{n\geq1}$, $m_n=n^\kappa$, $\kappa\in(0,1)$ and pilot estimator $\tilde\gamma_n=\tilde\gamma_n(\tilde X(m_n))$ of γ as follows

$$\tilde\gamma_n(\tilde X(m_n))$$

$$= \min\left\{\frac{\log(P_{m_n}(\tilde x_2(m_n))/P_{m_n}(\tilde x_1(m_n)))}{\log(\tilde x_1(m_n)/\tilde x_2(m_n))}\cdot\chi(P_{m_n}(\tilde x_1(m_n))\geq\log^{-1}m_n),\gamma_0\right\}, \tag{12}$$

where $\tilde X(n)=(\tilde x_1(n),\tilde x_2(n))$,

$$\tilde x_1(n) = \mathrm{e}\cdot\tilde x_2(n), \qquad \tilde x_2(n) = n^{\frac{p}{\gamma_0[2p(\rho+2)-1]}}, \quad \gamma_0=\beta_0^{-1}. \tag{13}$$

This estimator has the property

$$E(\tilde{\gamma}_n-\gamma)^{2p}\le C\cdot r_0^{-1}(n,p),\quad p\ge 1,$$

$$r_0(n,p)=n^{\frac{2p\gamma\kappa(\rho+1)}{\gamma_0[2p(\rho+1)+2p-1]}p},$$

which can be proved similar to (11) and is strongly consistent according to the Borel-Cantelli lemma with the following rate

$$n^{\nu}(\tilde{\gamma}_n-\gamma)\to 0\quad \text{a.s.} \tag{14}$$

for every

$$0<\nu<\frac{\gamma\kappa(\rho+1)}{\gamma_0(2\rho+3)}.$$

Indeed, for every $a>0$, ν defined above and p large enough

$$\sum_{n\ge 1}P(n^{\nu}(\tilde{\gamma}_n-\gamma)>a)\le a^{-2p}\sum_{n\ge 1}n^{2\nu p}E(\tilde{\gamma}_n-\gamma)^{2p}\le C\sum_{n\ge 1}\frac{n^{2\nu p}}{r_0(n,p)}<\infty.$$

Define the adaptive estimator of γ as follows

$$\hat{\gamma}_n=\frac{\log(\tilde{P}_n(\hat{x}_2(n))/\tilde{P}_n(\hat{x}_1(n)))}{\log(\hat{x}_1(n)/\hat{x}_2(n))}, \tag{15}$$

where $\tilde{P}_n$ is the empirical tail distribution function

$$\tilde{P}_n(x)=\frac{1}{n-m_n}\sum_{k=m_n+1}^{n}\chi(X_k\ge x)$$

and $\hat{X}(n)=(\hat{x}_1(n),\hat{x}_2(n))$,

$$\hat{x}_1(n)=\mathrm{e}\cdot\hat{x}_2(n),\qquad \hat{x}_2(n)=n^{\frac{p}{\tilde{\gamma}_n[2p(\rho+2)-1]}}. \tag{16}$$

The estimator $\hat{\gamma}_n$ has the property

$$E[(\hat{\gamma}_n-\gamma)^{2p}|\mathscr{F}_{m_n}]\le C\cdot\tilde{r}^{-1}(n,p),\quad p\ge 1,$$

where σ-algebra $\mathscr{F}_{m_n}=\sigma\{X_1,\dots,X_{m_n}\}$ and

$$\tilde{r}(n,p)=n^{\frac{2p\gamma(\rho+1)}{\tilde{\gamma}_n[2p(\rho+1)+2p-1]}p}.$$

Thus, using the Borel-Cantelli lemma and strong consistency of the pilot estimator $\tilde{\gamma}_n$ it is easy to prove the last property of Theorem 1 for the adaptive estimator $P_{\hat{\gamma}}$ in the Hall model.

To prove the strong consistency of $\varphi(n,p)(\hat{\gamma}_n-\gamma)^2$ and, as follows, Theorem 1, we establish first the convergence to zero of $\tilde{\varphi}(n,p)(\hat{\gamma}_n-\gamma)^2$, where $\tilde{\varphi}(n,p)=$

$(n^{-2}\tilde{r}(n,p))^{1/p}$:

$$\sum_{n\geq 1} P(\tilde{\varphi}(n,p)(\hat{\gamma}_n-\gamma)^2 > a) \leq a^{-2p} \sum_{n\geq 1} E\tilde{\varphi}^p(n,p)E[(\hat{\gamma}_n-\gamma)^{2p}|\mathscr{F}_{m_n}]$$

$$= a^{-2p} \sum_{n\geq 1} n^2 E\tilde{r}(n,p)E[(\hat{\gamma}_n-\gamma)^{2p}|\mathscr{F}_{m_n}] \leq C \sum_{n\geq 1} \frac{1}{n^2} < \infty.$$

Then as $n \to \infty$

$$\tilde{\varphi}(n,p)(\hat{\gamma}_n-\gamma)^2 \to 0 \quad \text{a.s.}$$

Using the property (14), we have

$$\log \tilde{r}(n,p) r^{-1}(n,p) \sim (\hat{\gamma}_n-\gamma)\log n \to 0 \quad \text{a.s.}$$

and, as follows, for the function $\varphi(n,p)$ defined after formula (11), as $n \to \infty$

$$\varphi(n,p)(\hat{\gamma}_n-\gamma)^2 \to 0 \quad \text{a.s.}$$

Thus Theorem 1 is proven with

$$p_\varepsilon > (\rho+1)\max\left[\frac{\gamma_0}{\kappa\gamma(\rho+1)}, \frac{2}{\rho+1-\varepsilon(\rho+2)}\right] \tag{17}$$

if in the Hall distribution estimator we use, according to the notation in Section 2 the adaptive parameter estimator $\Delta_n = \hat{\gamma}_n$ of $\Delta = \gamma$, defined in (15).

4 Simulation results

To establish the convergence of χ^2 divergence (1) one needs to check the condition (3), which is the key point to investigate the properties of estimators. In this section, to present some numerical results we define the quantity Λ_n as $L_{2p}-$norm of normalized deviation of estimator γ_n from the parameter γ

$$\Lambda_n = [r(n,p)E(\gamma_n-\gamma)^{2p}]^{\frac{1}{2p}} \tag{18}$$

in Hall's model [6]

$$F(x) = 1 - 2x^{-1} + x^{-2.5}, \quad x \geq 1, \;\; \rho = 0.5,$$

as a function of n. The values of Λ_n are given in Fig. 1,2 for different sample sizes n. Each coordinate is computed as an empirical average over 1000 Monte Carlo simulations of the experiment (for each value of n).

First the simulation was performed for the case when one can choose the sequences $x_1(n),\;\; x_2(n)$ according to (6) to get the estimators γ_n with the rate of convergence close to the optimal one, see [6]. The value of $p = p_\varepsilon$ was chosen as

$p_\varepsilon = [2/\varepsilon] + 1$. The results are presented in Fig.1 for $\varepsilon = 0.1$ and $\varepsilon = 0.05$. One can see that Λ_n remains bounded from above as n increases and therefore the condition (3) is fulfilled. Similar results were obtained for $p > p_\varepsilon$.

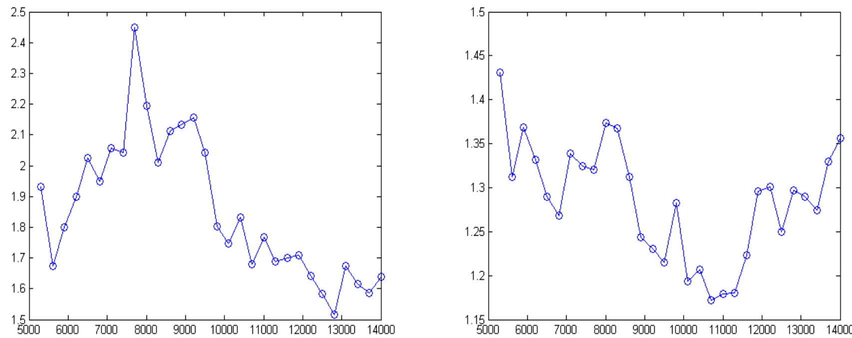

Fig. 1 Λ_n as a function of n in Hall's distribution with $\gamma = 1.0$. Left panel: $\varepsilon = 0.1$, Right panel: $\varepsilon = 0.05$

The results for adaptive estimator $\hat{\gamma}_n$ (15) with the sequences $\hat{x}_1(n)$, $\hat{x}_2(n)$ defined by (16) are given in Fig.2 for $\varepsilon = 0.1$ and $\varepsilon = 0.05$. The value of β_0 was equal 0.5, the sequences $\tilde{x}_1(n)$, $\tilde{x}_2(n)$ were defined by (13) with $m_n = n^\kappa$, $\kappa = 0.8$. The pilot estimator $\tilde{\gamma}_n$ was determined by (12), the power p was defined as the right hand side of inequality (17). The quantity Λ_n remains bounded from above as n increases as well.

Our numerical simulations in all cases give practical confirmation of the theoretical properties of the proposed estimators.

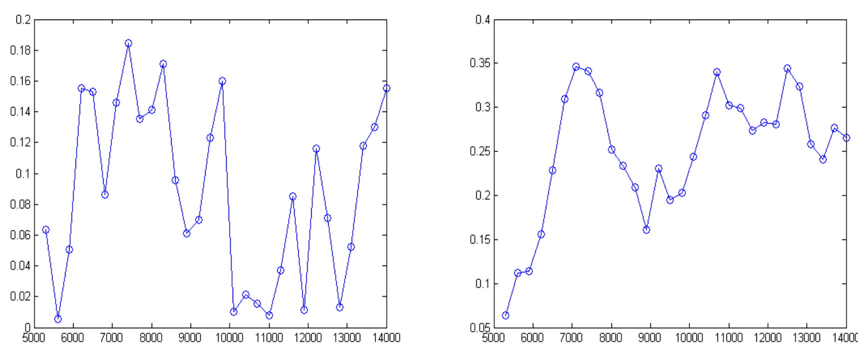

Fig. 2 Λ_n as a function of n in Hall's distribution with $\gamma = 1.0$. Left panel: $\varepsilon = 0.1$, Right panel: $\varepsilon = 0.05$

Acknowledgment

This study was supported by the NSF grant DMS 16-13026 and RNF Project No. 17-11-01049.

References

1. BEIRLANT, J., TEUGELS, J. L. and VYNCKIER, P. (1996). *Practical Analysis of Extreme Values*. Leuven Univ. Press.
2. BRITO, M., CAVALCANTE, L. and FREITAS, A.C.M. (2016). Bias-corrected geometric-type estimators of the tail index. *Journal of Physics A: Mathematical and theoretical*. **49** 1–38.
3. DREES, H. (1998). Optimal rates of convergence for estimates of the extreme value index. *The Annals of Statistics*. **26** 434–448.
4. EMBRECHTS, P., KLUPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
5. GOLDIE, C. M. and SMITH, R.L. (1987). Slow variation with remainder: theory and applications. *Quarterly Journal of Mathematics*. **38** 45–71.
6. GRAMA, I. and SPOKOINY, V. (2008). Statistics of extremes by oracle estimation. *Ann. Statist.* **36** 1619–1648.
7. HALL, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roj. Statist. Soc. Ser. B.* **44** 37-42. MR0655370
8. HALL, P. and WELSH, A. H. (1984). Best attainable rates of convergence for estimates of parameters of regular variation. *Ann. Statist.* **12** 1079-1084. MR0751294
9. HILL, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163–1174.
10. MARKOVICH, N. (2007). *Nonparametric Analysis of Univariate Heavy-Tailed Data Research and Practice.* John Wiley and Sons, Ltd.
11. NOVAK, S. Y. (1996) On the distribution of the ratio of sums of random variables. *Theory Probab. Appl.* **41** 479–503.
12. NOVAK, S. Y. (2000) On self-normalised sums. *Math. Methods Statist*. **9** 415-436.
13. NOVAK, S. Y. (2002) Inference of heavy tails from dependent data. *Siberian Advances in Mathematics*. **12** 73–96.
14. NOVAK, S. Y. (2011) EXTREME VALUE METHODS WITH APPLICATIONS TO FINANCE. London: Chapman and Hall/CRC Press. ISBN 9781439835746.
15. RAMLAU-HANSEN, H. (1988). A solvency study in non-life insurance. *Scand. Actuar. J.* **80** 3–34.
16. RESNICK, S.1997 SPECIAL INVITED PAPER. Heavy Tail. Modeling and teletraffic data. *The Annals of Statistics.* **25(5)** 1805-1869.
17. RESNICK, S.I. and STARICA, C. (1999) Smoothing the moment estimate of the extreme value parameter. *Extremes*. **1** 263–294.
18. SCHULTZE, J. and STEINEBACH, J. (1996) On least squares estimates of an exponential tail coefficient. *Stat. Decis.*. **14** 353–372.
19. VASILIEV, V. (2014) A truncated estimation method with guaranteed accuracy. *Annals of the Institute of Statistical Mathematics*. **66** 141–163.