C. R. Acad. Sci. Paris, Ser. I 334 (2002) 709–712

Statistique/*Statistics*
(Probabilités/*Probability Theory*)

Comparaison de deux indicateurs de la régularité en variation des queues de distributions

Rym Worms

Laboratoire Raphael Salem, Université de Rouen, UMR CNRS 6085, site Colbert, 76 821 Mont-saint-Aignant cedex, France

Reçu le 10 novembre 2001 ; accepté après révision le 18 février 2002

Note présentée par Paul Deheuvels.

Résumé Pour une loi de fonction de répartition (f.d.r.) F, nous comparons les hypothèses faites dans [4] pour étudier la vitesse de convergence vers 0, lorsque $u \to s_+(F)$, de $d(u) = \sup_{x \in [0, s_+(F)-u[} |\overline{F}_u(x) - \overline{G}_\gamma(\frac{x+u-\alpha(u)}{\sigma(u)})|$, où $\overline{F}_u$ est la fonction de survie de la loi des excès au-delà de u, $s_+(F) = \sup\{x, F(x) < 1\}$ est la borne supérieure du support de F et $\overline{G}_\gamma$ est la fonction de survie d'une loi de Pareto Généralisée, avec les hypothèses faites dans [2] pour étudier la vitesse de convergence vers 0, lorsque $n \to +\infty$, de $\hat{d}_n = \sup_{x \in \mathbb{R}} |F^n(x) - H_\gamma(\frac{x-\alpha_n}{\sigma_n})|$, où H_γ est la f.d.r. d'une loi des extrêmes. Dans chaque cas, avait été introduit un indicateur lié à des hypothèses de variation régulière. Nous caractérisons des situations où ces deux indicateurs coïncident et d'autres où ils diffèrent. ***Pour citer cet article : R. Worms, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 709–712.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Comparison between two indicators for the variation regularity of tails of distributions

Abstract We compare assumptions used in [4] in order to study the rate of convergence to 0, as $u \to s_+(F)$, of $d(u) = \sup_{x \in [0, s_+(F)-u[} |\overline{F}_u(x) - \overline{G}_\gamma(\frac{x+u-\alpha(u)}{\sigma(u)})|$, where $\overline{F}_u$ is the survival function of the excesses over u, $s_+(F) = \sup\{x, F(x) < 1\}$ is the upper end point of the distribution function (d.f.) F and $\overline{G}_\gamma$ is the survival function of the Generalized Pareto Distribution, with assumptions used in [2] in order to study the rate of convergence to 0, as $n \to +\infty$, of $\hat{d}_n = \sup_{x \in \mathbb{R}} |F^n(x) - H_\gamma(\frac{x-\alpha_n}{\sigma_n})|$, where H_γ is the d.f. of an extreme value distribution. In each case, an indicator linked to regular variation assumptions had been introduced. We characterize situations where these two indicators coincide, and others where they are different. ***To cite this article: R. Worms, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 709–712.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Abridged English version

Let F be a distribution function in the domain of attraction of an extreme value distribution H_γ. This means that there exist two normalizing sequences (α_n) and (σ_n) such that

$$\hat{d}_n = \sup_{x \in \mathbb{R}} \left| F^n(x) - H_\gamma\left(\frac{x - \alpha_n}{\sigma_n}\right) \right| \longrightarrow 0 \quad \text{as } n \to +\infty.$$

Adresse e-mail : rym.worms@univ-rouen.fr (R. Worms).

S1631-073X(02)02324-5/FLA

It is well known (*see* [3] or [1]) that this is equivalent to the existence of $\sigma(u) > 0$ such that

$$\sup_{x\in[0,s_+(F)-u[}\left|\overline{F}_u(x)-\overline{G}_\gamma\left(\frac{x}{\sigma(u)}\right)\right| \longrightarrow 0 \quad \text{as } u\to s_+(F),$$

where $\overline{F}_u$ is the survival function of the excesses over u, $s_+(F)=\sup\{x,\,F(x)<1\}$ is the upper end point of F and $\overline{G}_\gamma$ is the survival function of a Generalized Pareto distribution.

We have studied, in [4], the rate of convergence to 0, as $u\to s_+(F)$, of

$$d(u)=\sup_{x\in[0,s_+(F)-u[}\left|\overline{F}_u(x)-\overline{G}_\gamma\left(\frac{x+u-\alpha(u)}{\sigma(u)}\right)\right|,$$

for appropriate normalizing functions α and σ, under the following first and second order conditions: [1]

(H_1) F is twice differentiable and $\lim_{t\to+\infty}A(t)=0$,

(H_2) A is of constant sign near $+\infty$ and there exists $\rho\leqslant 0$ such that $|A|\in RV_\rho$,

where $A(t)=\frac{V''(\ln t)}{V'(\ln t)}-\gamma$ and $V(t)=\overline{F}^{-1}(\mathrm{e}^{-t})$.

In [2], De Haan and Resnick have made a similar study for $\hat{d}_n$, using assumptions:

($\widehat{\mathrm{H}}_1$) F is twice differentiable and $\lim_{t\to+\infty}\hat{A}(t)=0$,

($\widehat{\mathrm{H}}_2$) $\hat{A}$ is of constant sign at $+\infty$ and there exists $\hat{\rho}\leqslant 0$ such that $|A|\in RV_{\hat{\rho}}$,

where $\hat{A}(t)=\frac{\widehat{V}''(\ln t)}{\widehat{V}'(\ln t)}-\gamma$ and $\widehat{V}(t)=F^{-1}(\exp(-\mathrm{e}^{-t}))$.

These assumptions have the same form as ours, but function A is replaced by $\hat{A}$.

We obtain the two following results:

(i) $\lim_{t\to+\infty}\hat{A}(t)=0$ if and only if $\lim_{t\to+\infty}A(t)=0$ (i.e., (H_1) and ($\widehat{\mathrm{H}}_1$) are equivalent);

(ii) Let ρ_0 be equal to -1 if $\gamma\neq 1$ and -2 if $\gamma=1$; if ($\widehat{\mathrm{H}}_1$) and ($\widehat{\mathrm{H}}_2$) (resp. (H_1) and (H_2)) hold, and if $\hat{\rho}\neq\rho_0$ (resp. $\rho\neq\rho_0$) then (H_2) (resp. ($\widehat{\mathrm{H}}_2$)) holds with $\rho=\max(\hat{\rho},\rho_0)$ (resp. $\hat{\rho}=\max(\rho,\rho_0)$).

---

1. Introduction

Soit F la fonction de répartition d'une loi dans le domaine d'attraction de la loi des extrêmes H_γ, c'est-à-dire, qu'il existe deux suites normalisantes (α_n) et (σ_n) telles que

$$\hat{d}_n=\sup_{x\in\mathbb{R}}\left|F^n(x)-H_\gamma\left(\frac{x-\alpha_n}{\sigma_n}\right)\right| \longrightarrow 0 \quad \text{quand } n\to+\infty,$$

avec

$$H_\gamma(x)=\exp\bigl(-(1+\gamma x)^{-1/\gamma}\bigr)\ (\gamma\neq 0,\ 1+\gamma x>0) \quad \text{et} \quad H_0(x)=\exp\bigl(-\mathrm{e}^{-x}\bigr).$$

On sait (*voir* [3] ou [1]), que ceci équivaut à l'existence de $\sigma(u)>0$ tel que

$$\sup_{x\in[0,s_+(F)-u[}\left|\overline{F}_u(x)-\overline{G}_\gamma\left(\frac{x}{\sigma(u)}\right)\right| \longrightarrow 0 \quad \text{quand } u\to s_+(F),$$

où $\overline{F}_u$ est la fonction de survie de la loi des excès au-delà de u, $s_+(F)=\sup\{x,\,F(x)<1\}$ est la borne supérieure du support de F et $\overline{G}_\gamma$ est la fonction de survie d'une loi de Pareto Généralisée, avec

$$\overline{G}_\gamma(x)=(1+\gamma x)^{-1/\gamma}\ (\gamma\neq 0,\ 1+\gamma x>0) \quad \text{et} \quad \overline{G}_0(x)=\mathrm{e}^{-x}.$$

Nous avons étudié, dans [4], la vitesse de convergence vers 0, quand u tend vers $s_+(F)$, de

$$d(u)=\sup_{x\in[0,s_+(F)-u[}\left|\overline{F}_u(x)-\overline{G}_\gamma\left(\frac{x+u-\alpha(u)}{\sigma(u)}\right)\right|,$$

pour des fonctions normalisantes α et σ appropriées, sous les hypothèses de premier et second ordre : [2]

(H_1) F est deux fois dérivable et $\lim_{t\to+\infty}A(t)=0$,

(H_2) A est de signe constant à l'infini et il existe $\rho\leqslant 0$ tel que $|A|\in RV_\rho$,

où $A(t)=\frac{V''(\ln t)}{V'(\ln t)}-\gamma$ et $V(t)=\overline{F}^{-1}(\mathrm{e}^{-t})$.

Dans [2], De Haan et Resnick avaient fait une étude similaire pour $\hat{d}_n$, avec les hypothèses :
$(\widehat{\mathrm{H}}_1)$ F est deux fois dérivable et $\lim_{t\to+\infty} \hat{A}(t) = 0$,
$(\widehat{\mathrm{H}}_2)$ $\hat{A}$ est de signe constant à l'infini et il existe $\hat{\rho} \leqslant 0$ tel que $|\hat{A}| \in RV_{\hat{\rho}}$,
où $\hat{A}(t) = \frac{\widehat{V}''(\ln t)}{\widehat{V}'(\ln t)} - \gamma$ et $\widehat{V}(t) = F^{-1}(\exp(-\mathrm{e}^{-t}))$.

Leurs hypothèses s'écrivent ainsi de la même manière que les nôtres, mais la fonction A est remplacée par $\hat{A}$. Le but de cette Note est d'étudier le lien entre $(\mathrm{H}_1, \mathrm{H}_2)$ et $(\widehat{\mathrm{H}}_1, \widehat{\mathrm{H}}_2)$.

2. Lien entre ρ et $\hat{\rho}$

PROPOSITION 1. –
(i) $\lim_{t\to+\infty} \hat{A}(t) = 0$ *si et seulement si* $\lim_{t\to+\infty} A(t) = 0$ (*i.e.* (H_1) *et* $(\widehat{\mathrm{H}}_1)$ *sont équivalentes*).
(ii) *Soit* ρ_0 *égal à* -1 *si* $\gamma \neq 1$ *et à* -2 *si* $\gamma = 1$; *si* $(\widehat{\mathrm{H}}_1)$ *et* $(\widehat{\mathrm{H}}_2)$ (*resp.* (H_1) *et* (H_2)) *sont satisfaites, et si* $\hat{\rho} \neq \rho_0$ (*resp.* $\rho \neq \rho_0$) *alors* (H_2) (*resp.* $(\widehat{\mathrm{H}}_2)$) *est satisfaite avec* $\rho = \max(\hat{\rho}, \rho_0)$ (*resp.* $\hat{\rho} = \max(\rho, \rho_0)$).

Remarque 1. – Si $\hat{A} \in RV_{\rho_0}$, cette proposition ne permet pas d'en tirer une conclusion au sujet de A, et il en est de même en intervertissant les rôles de A et $\hat{A}$. Les situations qui sont susceptibles de se présenter si l'une des fonctions A et $\hat{A}$ est à variation régulière sont donc les suivantes
- $A \in RV_\rho$ et $\hat{A} \in RV_\rho$ avec $\rho_0 \leqslant \rho \leqslant 0$.
- $A \in RV_\rho$, avec $\rho < \rho_0$, et $\hat{A} \in RV_{\rho_0}$.
- $\hat{A} \in RV_{\hat{\rho}}$, avec $\hat{\rho} < \rho_0$, et $A \in RV_{\rho_0}$.
- $A \in RV_{\rho_0}$ et $\hat{A}$ n'est pas à variation régulière.
- $\hat{A} \in RV_{\rho_0}$ et A n'est pas à variation régulière.

Démonstration. – (i) Posons

$$C(x) = V(-\ln x) = F^{-1}(1-x) \quad \text{et} \quad \widehat{C}(z) = \widehat{V}(-\ln z) F^{-1}\big(\mathrm{e}^{-z}\big),$$
$$B(x) = A\left(\frac{1}{x}\right) \quad \text{et} \quad \widehat{B}(z) = \hat{A}\left(\frac{1}{z}\right),$$

où $x \in]0,1[$ et $z \in]1,+\infty[$. Alors $C(x) = \widehat{C}(-\ln(1-x))$ et on en déduit que

$$B(x) = -\gamma - \frac{1}{1-x} - (\gamma+1)\frac{x}{(1-x)\ln(1-x)} - \frac{x}{(1-x)\ln(1-x)}\widehat{B}\big(-\ln(1-x)\big), \tag{1}$$

$$\widehat{B}(z) = -\gamma - 1 + \frac{z}{\mathrm{e}^z - 1}\big(B\big(1-\mathrm{e}^{-z}\big) + \gamma + \mathrm{e}^z\big). \tag{2}$$

Il en résulte immédiatement que $\lim_{z\to 0} \widehat{B}(z) = 0$ équivaut à $\lim_{x\to 0} B(x) = 0$.

(ii) Supposons que $\hat{A} \in RV_{\hat{\rho}}$. Notons $\beta(x) = -\frac{x}{(1-x)\ln(1-x)}$; il vient $\beta(x) = 1 + \frac{x}{2} + \frac{5}{12}x^2 + \mathrm{o}(x^2)$ d'où, d'après (1),

$$\begin{aligned} B(x) &= -\gamma - (1-x)^{-1} + (\gamma+1)\beta(x) + \beta(x)\widehat{B}\big(-\ln(1-x)\big) \\ &= \frac{\gamma-1}{2}x + \frac{5\gamma-7}{12}x^2 + \mathrm{o}\big(x^2\big) + \widehat{B}\big(-\ln(1-x)\big)\big(1+\mathrm{o}(1)\big), \end{aligned}$$

autrement dit $A(t) = A_1(t) + A_2(t)$, où

$$A_1(t) = \frac{\gamma-1}{2}\frac{1}{t} + \frac{5\gamma-7}{12}\frac{1}{t^2} + \mathrm{o}\left(\frac{1}{t^2}\right), \qquad A_2(t) = \hat{A}\left(\left(-\ln\left(1-\frac{1}{t}\right)\right)^{-1}\right)\big(1+\mathrm{o}(1)\big).$$

On remarque que, au voisinage de $+\infty$, $(-\ln(1-\frac{1}{t}))^{-1} \sim t$ et donc $A_2 \in RV_{\hat{\rho}}$.

D'autre part, si $\gamma \neq 1$, $A_1(t) \in RV_{-1}$ et donc :
- si $\hat{\rho} > -1$, $A(t) \sim \hat{A}((-\ln(1-\frac{1}{t}))^{-1}) \in RV_{\hat{\rho}}$ (i.e. $B(x) \sim \widehat{B}(-\ln(1-x))$),
- si $\hat{\rho} < -1$, $A(t) \sim \frac{\gamma-1}{2}\frac{1}{t} \in RV_{-1}$ (i.e. $B(x) \sim \frac{\gamma-1}{2}x$),

– si $\hat{\rho} = -1$, on ne peut rien conclure sur A.

De même, si $\gamma = 1$, $A_1(t) \sim -\frac{1}{6}\frac{1}{t^2}$, d'où $A_1(t) \in RV_{-2}$ et :

– si $\rho > -2$, $A(t) \sim \hat{A}((-\ln(1-\frac{1}{t}))^{-1}) \in RV_\rho$ (i.e. $B(x) \sim \widehat{B}(-\ln(1-x))$),

– si $\rho < -2$, $A(t) \sim -\frac{1}{6}\frac{1}{t^2} \in RV_{-2}$ (i.e. $B(x) \sim -\frac{1}{6}x^2$),

– si $\rho = -2$, on ne peut rien conclure sur A.

Réciproquement, supposons que $A \in RV_\rho$ et notons $\hat{\beta}(z) = \frac{z}{e^z - 1}$; il vient $\hat{\beta}(z) = 1 - \frac{z}{2} + \frac{1}{12}z^2 + o(z^2)$, d'où

$$\begin{aligned}\widehat{B}(z) &= -\gamma - 1 + \hat{\beta}(z)\big(B\big(1-e^{-z}\big) + \gamma + e^z\big)\\ &= \frac{1-\gamma}{2}z + \frac{\gamma+1}{12}z^2 + o\big(z^2\big) + B\big(1-e^{-z}\big)\big(1+o(1)\big),\end{aligned}$$

autrement dit $\hat{A}(t) = \hat{A}_1(t) + \hat{A}_2(t)$, où

$$\hat{A}_1(t) = \frac{1-\gamma}{2}\frac{1}{t} + \frac{\gamma+1}{12}\frac{1}{t^2} + o\left(\frac{1}{t^2}\right), \qquad \hat{A}_2(t) = A\left(\left(1-\exp\left(\left(-\frac{1}{z}\right)\right)\right)^{-1}\right)\big(1+o(1)\big),$$

et la démonstration s'achève comme dans la partie directe.

3. Exemple

Soit F la fonction de répartition d'une loi dans le domaine d'attraction de H_γ, vérifiant les hypothèses (H_1) et (H_2) et F^* la fonction de répartition vérifiant, pour x au voisinage de $s_+(F) = s_+(F^*)$,

$$F^*(x) = \exp\big(-\overline{F}(x)\big). \tag{3}$$

D'une part, pour t assez grand, $\widehat{V}_{F^*}(t) = V_F(t)$. On en déduit que l'hypothèse $(\widehat{H}_1)$ associée à F^* est satisfaite et donc que F^* est également dans le domaine d'attraction de H_γ et $\hat{A}_{F^*}(t) = A_F(t)$. De plus, F^* vérifie l'hypothèse $(\widehat{H}_2)$ avec

$$\hat{\rho}_{F^*} = \rho_F.$$

D'autre part, pour t assez grand, $V_{F^*}(t) = \overline{F}^{-1}(-\ln(1-e^{-t})) = V_F(-\ln(-\ln(1-e^{-t})))$. On n'a donc pas entre F et F^* la « dualité » qu'aurait exprimée l'égalité $V_{F^*}(t) = \widehat{V}_F(t)$ et, si $\hat{\rho}_F \neq \rho_0$, F^* vérifie l'hypothèse (H_2) avec $\rho_{F^*} = \max(\rho_0, \hat{\rho}_F)$ (conséquence immédiate de $\hat{\rho}_{F^*} = \rho_F$ et de la Proposition 1).

Cas particulier des loi de Burr. – La fonction de répartition d'une loi de Burr s'écrit, pour tout $x > 0$,

$$F(x) = 1 - \big(1 + x^\tau\big)^{-\lambda},$$

où $\tau > 0$ et $\lambda > 0$. F est dans le domaine d'attraction de Fréchet avec $\gamma = 1/(\lambda\tau)$.

Si $\tau \neq 1$, F vérifie l'hypothèse (H_2) avec $\rho_F = -1/\lambda$ et l'hypothèse $(\widehat{H}_2)$ avec $\hat{\rho}_F = \max(\rho_0, -1/\lambda)$. Ainsi, $\hat{\rho}_F$ est borné alors que ρ_F ne l'est pas.

La loi de fonction de répartition F^* définie par (3) vérifie alors $(\widehat{H}_2)$ avec $\hat{\rho}_{F^*} = -1/\lambda$ et (H_2) avec $\rho_{F^*} = \max(\rho_0, -1/\lambda)$. Cette fois ρ_{F^*} est borné et non $\hat{\rho}_{F^*}$.

[1] $f \in RV_\alpha$ (i.e. regular variation of index α) means that, for all $x > 0$, $\lim_{t\to+\infty} f(tx)/f(t) = x^\alpha$.

[2] $f \in RV_\alpha$ (i.e., f de variation régulière d'indice α) veut dire que, pour tout $x > 0$, $\lim_{t\to+\infty} f(tx)/f(t) = x^\alpha$.

Références bibliographiques

[1] A. Balkema, L. de Haan, Residual life time at great age, Ann. Probab. 2 (1974) 792–801.

[2] L. de Haan, S.I. Resnick, Second-order regular variation and rates of convergence in extreme-value theory, Ann. Probab. 24 (1996) 97–124.

[3] J. Pickands III, Statistical inference using extreme order statistics, Ann. Statist. 3 (1975) 119–131.

[4] R. Worms, Vitesse de convergence de l'approximation de Pareto généralisée de la loi des excès, C. R. Acad. Sci. Paris, Série I 333 (2001) 65–70.