Biometrika (2018), **xx**, x, pp. 1–4 *Printed in Great Britain*

Supplementary Material for Discrete Extremes

The following auxiliary lemma is elementary (as the sum can be sandwiched between two integrals).

LEMMA 1. If $\xi > 0$, then

$$u^{1/\xi}H_{1+1/\xi,u} \to \xi_{1}$$

5

 \Box

as $u \to \infty$, where $H_{s,q} = \sum_{i=0}^{\infty} (q+i)^{-s}$ is the Hurwitz-Zeta function.

Proof of Proposition 1. Suppose first that $\xi > 0$. Then

$$\frac{p_{\text{D-GPD}}(k;\sigma,\xi)}{f_{\text{GPD}}(k;\sigma,\xi)} = \frac{\left(1+\xi\frac{k}{\sigma}\right)^{-1/\xi} - \left(1+\xi\frac{k+1}{\sigma}\right)^{-1/\xi}}{\frac{1}{\sigma}\left(1+\xi\frac{k}{\sigma}\right)^{-1/\xi-1}} \\ = \left\{1 - \left(1+\frac{\xi}{\sigma+\xi k}\right)^{-1/\xi}\right\}(\sigma+\xi k) \to 1$$

uniformly in k = 0, 1, 2, ... as $\sigma \to \infty$. Furthermore,

$$\sup_{k=0,1,2,\dots} \frac{f_{\text{GPD}}(k;\sigma,\xi)}{p_{\text{GZD}}(k;\sigma,\xi)} = \sigma^{-1} \sum_{i=0}^{\infty} (1+\xi i/\sigma)^{-1/\xi-1} \to 1,$$

as $\sigma \to \infty$ by Lemma 1. In the case $\xi = 0$,

$$p_{\text{D-GPD}}(k;\sigma,0)/f_{\text{GPD}}(k;\sigma,0) = p_{\text{GZD}}(k;\sigma,0)/f_{\text{GPD}}(k;\sigma,0) = \sigma(1-e^{-1/\sigma}) \to 1.$$

Proof of Proposition 2. By assumption, there exists a random variable $Y \in \text{MDA}_{\xi}$ for $\xi \ge 0$ and a positive function $(\tilde{a}_u, u > 0)$ such that $X = \lfloor Y \rfloor$ in distribution and the sequence of functions pr $\{\tilde{a}_u^{-1}(Y - u) \ge x \mid Y \ge u , x \ge 0$, converges uniformly, as $u \to \infty$, to the function $\bar{F}_{\text{GPD}}(x; \sigma, \xi), x \ge 0$, for some $\sigma > 0$ and $\xi \ge 0$. For a positive integer u we let $a_u = \tilde{a}_u \sigma$. Then

$$\begin{split} \sup_{k=0,1,2,\dots} |\operatorname{pr}(X = u + k \mid X \ge u) - p_{\text{D-GPD}}(k; a_u, \xi)| \\ = \sup_{k=0,1,2,\dots} \operatorname{pr}\left\{\tilde{a}_u^{-1}(Y - u) \ge \tilde{a}_u^{-1}k \mid Y \ge u - \operatorname{pr}\left\{\tilde{a}_u^{-1}(Y - u) \ge \tilde{a}_u^{-1}(k + 1) \mid Y \ge u - \bar{F}_{\text{GPD}}(k; a_u, \xi) + \bar{F}_{\text{GPD}}(k + 1; a_u, \xi) \right. \\ & \left. \le 2 \sup_{x \ge 0} \operatorname{pr}\left\{\tilde{a}_u^{-1}(Y - u) \ge x \mid Y \ge u - \bar{F}_{\text{GPD}}(x; \sigma, \xi) \to 0 \right. \end{split}$$

as $u \to \infty$ over the integers.

The proof of Theorem 1 relies on properties of regularly varying functions. Recall that a positive and measurable function f on $[1, \infty)$ is regularly varying if there exists $\alpha \in \mathbb{R}$ such that

$$\lim_{u \to \infty} \frac{f(ux)}{f(u)} \to x^{\alpha}, \quad x \ge 1,$$

c 2016 Biometrika Trust

A. S. HITZ, R. DAVIS AND G. SAMORODNITSKY

and we write $f \in RV_{\alpha}$ (see e.g. Bingham et al. (1989)). If $f \in RV_{-\alpha}$ for $\alpha \ge 0$, then

$$\lim_{u \to \infty} \sup_{x \in [1,b]} \frac{f(ux)}{f(u)} - x^{-\alpha} \to 0,$$
(S.1)

for $b = \infty$ if $\alpha > 0$, and for any $b < \infty$ if $\alpha = 0$. If $f \in RV_{-\alpha}$ for $\alpha > 0$, then by Potter's bounds (see e.g. Resnick (1987)) for any $\epsilon > 0$ there is $u_{\epsilon} \in (0, \infty)$ such that

$$e^{-\epsilon}x^{-\alpha-\epsilon} \le \frac{f(ux)}{f(u)} \le e^{\epsilon}x^{-\alpha+\epsilon}, \quad x \ge 1,$$
 (S.2)

for $u \ge u_{\epsilon}$. We say that X is regularly varying if $\overline{F}_X \in \mathrm{RV}_{-\alpha}$ for some $\alpha > 0$, a necessary and sufficient condition for $X \in \mathrm{MDA}_{1/\alpha}$.

Proof of Theorem 1. We start by proving the first part of the theorem. By assumption, there exists a survival function \overline{F} such that $\overline{F}(k) = c p_X(k)$ for c > 0, k large enough and $\overline{F} \in \mathrm{RV}_{-1/\xi-1}$. The last condition is equivalent to $\overline{F}(\lfloor \cdot \rfloor) \in \mathrm{RV}_{-1/\xi-1}$ (Shimura, 2012). It follows from results on integrals of monotone regularly functions in Bingham et al. (1989) that $\overline{F}_X \in \mathrm{RV}_{-1/\xi}$, and thus $X \in \mathrm{MDA}_{\xi}$. We now show that (8) holds. Thanks to Proposition 1, it suffices to provide a proof for $q = p_{\text{GZD}}$. We have

$$\frac{\operatorname{pr}(X = k_u + u \mid X \ge u)}{p_{\operatorname{GZD}}(k_u; \xi u, \xi)} = \frac{\bar{F}(u + k_u)/\bar{F}(u)}{(1 + k_u/u)^{-1/\xi - 1}} \frac{\sum_{i=0}^{\infty} (1 + i/u)^{-1/\xi - 1}}{\sum_{i=0}^{\infty} \bar{F}(u + i)/\bar{F}(u)}$$

First, by the uniform convergence (S.1) and the fact that k_u grows at most linearly fast, we conclude that

$$\frac{\bar{F}(u+k_u)/\bar{F}(u)}{(1+k_u/u)^{-1/\xi-1}} \to 1,$$

as $u \to \infty$ over the integers. Second, Lemma 1 yields

$$u^{-1} \sum_{i=0}^{\infty} (1+i/u)^{-1/\xi-1} \to \xi.$$

Third, it follows from (S.2) that for $\epsilon \in (0, 1/\xi)$, there exists $u_{\epsilon} > 0$ such that for $u \ge u_{\epsilon}$,

$$u^{-1}\sum_{i=0}^{\infty}\bar{F}(u+i)/\bar{F}(u) \le u^{-1}e^{\epsilon}\sum_{i=0}^{\infty}\left(1+\frac{i}{u}\right)^{-1-1/\xi+\epsilon} \to \frac{\xi e^{\epsilon}}{1-\xi\epsilon},$$

using Lemma 1 once again. A similar lower bound can be found in the same manner. Now letting $\epsilon \to 0$, this completes the proof of (8).

Let us now prove the second part of the theorem. For large integers u,

$$pr(X = k + u \mid X \ge u) = \frac{\bar{F}(k + u)/\bar{F}(u)}{\sum_{i=0}^{\infty} \bar{F}(i + u)/\bar{F}(u)}$$

We have for every i = 0, 1, 2, ...,

$$\bar{F}(i+u)/\bar{F}(u) = \frac{c(i+u)}{c(u)} \exp\left\{-\int_0^i 1/a(u+y)dy\right\} \to e^{-i/\sigma}$$

as $u \to \infty$. Since $\sigma > 0$, the dominated convergence theorem gives us

$$\sum_{i=0}^{\infty} \bar{F}(i+u)/\bar{F}(u) \to \sum_{i=0}^{\infty} e^{-i/\sigma} = 1/(1-e^{-1/\sigma}),$$

20



Number of extreme tornadoes per outbreak

Fig. 1: Frequency plot of the number of extreme tornadoes per outbreak for the 435 outbreaks with 12 or more extreme tornadoes in the United States between 1965 and 2015.

showing (9). Finally, it follows from

$$p_X(n) = c(n) \exp\left\{-\int_0^n \frac{1}{a(y)} dy\right\}$$

for all n and $a(y) \to \sigma \in (0, \infty)$ that

$$\lim_{n \to \infty} \frac{p_X(n)}{\operatorname{pr}(X \ge n)} = 1 - e^{-1/\sigma} \in (0, \infty)$$

which immediately implies that $X \in D-MDA_0$ as well.

Example 2. The probability mass function p_X of a Poisson distribution with rate $\lambda > 0$ coincides on k = 0, 1, 2, ... with the function

$$g(x) = \frac{\lambda^x e^{-\lambda}}{\Gamma(x+1)}$$

a continuous function on \mathbb{R}_+ satisfying $\lim_{x\to\infty} g(x) = 0$. Moreover,

$$\frac{d}{dx}\log g(x) = -\psi_0(x+1) + \log \lambda,$$

where ψ_0 is the polygamma function of order 0. Since $\psi_0(x) \to \infty$ as $x \to \infty$, we see that g'(x) < 0 for x sufficiently large. Therefore, $\bar{F}_Y(x) = g(x)/g(d)$ is a survival function on $[d, \infty)$ for some $d \ge 0$. Furthermore,

$$\frac{d}{dx}\left(-\frac{1}{g'(x)}\right) = -\frac{\psi_1(x+1)}{\{\psi_0(x+1) - \log \lambda\}^2},$$

where $\psi_1 = \psi'_0$ is is the polygamma function of order 1. Since $\psi_1(x) \to 0$ as $x \to \infty$, we conclude that F_Y satisfies the von Mises condition, with the auxiliary function $a(x) = \{\psi_0(x + 1) - \log \lambda\}^{-1} \to 0$ as $x \to \infty$. Therefore, the Poisson probability mass function is in D-MDA₀.

Similarly, the probability mass function of the negative binomial distribution with probability of success $p \in (0, 1)$ and number of successes r > 0 is also in D-MDA₀ because it coincides on $\{0, 1, 2, ...\}$ with the function

$$g(x) = \frac{p^r}{\Gamma(r)} \frac{\Gamma(x+r)}{\Gamma(x+1)} (1-p)^x$$

a continuous function on \mathbb{R}_+ . It is simple to check that $\lim_{x\to\infty} g(x) = 0$, and g'(x) < 0 for x large enough, so that $\bar{F}_Y(x) = g(x)/g(d)$ is a survival function on $[d,\infty)$ for some $d \ge 0$. Furthermore, $g(x) \sim cx^{r-1}(1-p)^x$ for large x, where c is a positive constant. Therefore, \bar{F}_Y is of the form (7) with the auxiliary function

$$a(x) = \frac{1}{-\log(1-p) - (r-1)/x}, \ x \text{ large,}$$

and so it converges to $-1/\log(1-p)$ as $x \to \infty$.

REFERENCES

BINGHAM, N. H., GOLDIE, C. M. & TEUGLES, J. L. (1989). *Regular Variation*. Cambridge University Press.
⁴⁰ RESNICK, S. I. (1987). *Extreme values, regular variation, and point processes*. Springer, New York.
SHIMURA, T. (2012). Discretization of distributions in the maximum domain of attraction. *Extremes* 15, 299–317.