EXTREMES, RISK and RESAMPLING TECHNIQUES

V.E.X.T.R.A.

CEAUL and DEIO Tomar, 2003 FCT / POCTI / FEDER

Extremes, Risk And Resampling Techniques

Extremes, Risk and Resampling Techniques[1]

Extended Abstracts

[1] Edition partially supported by FCT / POCTI / FEDER.

ISBN: 972-8628-91-9
Depósito legal: 203583/03
Number of copies: 60
November 2003

Organizing and Scientific Committee

- M. Ivette Gomes — *Lisbon University, Portugal*
- Haan, Laurens de — *Erasmus University, Rotterdam, the Netherlands*
- Dinis Pestana — *Lisbon University, Portugal*
- Luisa Canto e Castro — *Lisbon University, Portugal*
- M. Isabel Fraga Alves — *Lisbon University, Portugal*

Sponsors

- *Fundação para a Ciência e a Tecnologia* (FCT)
- *Centro de Estatística e Aplicações* (CEAUL)
- *Departamento de Estatística e Investigação Operacional* (DEIO)
- *Faculdade de Ciências de Lisboa* (FCUL)

Contents

Anderson, Clive — Continuous time extremes from discrete time observations 1

Brito, Margarida — Tail bootstrap and the estimation of tail parameters 3

Caeiro, Frederico (jointly with M. Ivette Gomes) — Bias reduction for light tail models — the Generalizaed Jackknife methodology 7

Dias, Sandra (jointly with Luisa Canto e Castro) — Ratios of differences of order statistics in max-semistable models . . 13

Egídio Reis, Alfredo — The compound binomial model revisited 17

Falk, Michael (jointly with René Michel) — Testing for tail independence in multivariate extreme value models 19

Figueiredo, Fernanda (jointly with M. Ivette Gomes) — Scale estimators in Statistical Quality Control 25

Gomes, M. Ivette (jointly with Frederico Caeiro, Fernanda Figueiredo, Laurens de Haan and Dinis Pestana) — Linear combinations and bias reduction in the estimation of heavy tails 31

Guillou, Armelle (jointly with Jan Beirlant and Emmanuel Delafosse) — Estimation of the extreme value index and high quantiles under random censoring 37

Haan, Laurens de (jointly with Teresa Themido Pereira) — Spatial extremes: the stationary case 39

Hall, Andreia (jointly with Manuel G. Scotto and Helena Ferreira) — Extremes of generalized periodic sub-sampled moving average sequences 41

Martins, Ana Paula (jointly with Helena Ferreira) — Extremes of periodic moving averages of random variables with regularly varying tail probabilities 43
Neves, Cláudia (jointly with Jan Picek and Isabel Fraga Alves) — Revisiting the trilemma problem for max-domains of attraction 45
Neves, Manuela (jointly with M. João Martins and M. Ivette Gomes) — Kernel estimators of the tail index 51
Pestana, Dinis (jointly with Fátima Brilhante, José Rocha, Sandra Mendonça e Rita Vasconcelos) — Summing up or Randomization — how far can we go? 55
Prata Gomes, Dora (jointly with Manuela Neves) — Block bootstrap methods in extremal index estimation 63
Ramos, Alexandra (jointly with Anthony Ledford) — Joint tail modelling and properties 69
Rootzén, Holger (jointly with Anne-Laure Fougères, Sture Holm and John Nolan) — Extreme value models for pitting corrosion 75
Segers, Johan — Clusters of extremes of stationary sequences in general state space 77
Stărică, Cătălin — Is GARCH as good a model as the Nobel prize accolades would imply? 79
Temido, Graça (jointly with Andreia Hall) — On the maximum term of moving average and max-autoregressive models with margins in Anderson's class 81
Teugels, Jef (jointly with Hansjoerg Albrecher) — Light and heavy tails in Ruin Theory 87
Valente de Freitas, Adelaide — Using the auxiliary function ϕ... 89
Velosa, Sílvio (jointly with Dinis Pestana) — Stopped sums with generalized Panjer subordinators 91
List of Participants 97

CONTINUOUS TIME EXTREMES FROM DISCRETE TIME OBSERVATIONS

Clive Anderson
Dept. of Probability and Statistics
University of Sheffield
United Kingdom

In many applications of statistical extreme value theory the primary object of interest is the maximum of some continuous-time variable over a specified period. It often is true, however, that measurements of the variable can be made only at a discrete set of times. For example, in the estimation of the maximum wave height expected to occur over a given number of years at a specified point in the ocean, data are likely to be available at best only from observations made at intervals of several hours. The true continuous-time maximum over any period will generally be larger than the largest discrete-time observation seen during the same period, the difference depending on the smoothness of the variable being observed and the frequency of discrete-time observation. For estimation of extremes of the continuous-time process based on discrete-time data we therefore need to make an adjustment to allow for the effect of discrete sampling. The paper discusses how the size of this adjustment might be estimated if some broad characteristics of the continuous time variable can be assumed.

Key words and phrases: Continuous-time extremes, discrete-time extremes, adjustment.

TAIL BOOTSTRAP AND THE ESTIMATION OF TAIL PARAMETERS

Margarida Brito
Faculdade de Ciências & Centro de Matemática
Universidade do Porto
Portugal

Key words and phrases: Adjustment coefficient, Edgeworth expansions, regular variation, tail bootstrap, tail indices, universal asymptotic normality.

Let X_1, X_2,..., be independent non negative random variables with common distribution function (d.f.) F satisfying

$$1 - F(x) = x^{-1/\alpha} L(x), \quad \text{for } x > 0,$$

where L is a slowly varying function at infinity. The problem of estimating the exponent α or other related tail parameters, from a finite sample $X_1, X_2, \ldots, X_n$, has received much attention and several estimators have been proposed (see e.g. Csörgő and Viharos (1998)). One of the most commonly used estimators is the Hill estimator (1975), given as follows. Let $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ denote the order statistics (o.s.) of the sample and assume that (k_n) is a sequence of positive integers such that $1 \leq k_n < n$, $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. The Hill estimator, say H_n, is defined by

$$H_n = k_n^{-1} \sum_{i=1}^{k_n} \log X_{n-i+1,n} - \log X_{n-k_n,n}\,.$$

In the last years, there has been considerable interest in the application of resampling procedures to the problem of estimating tail parameters (see e.g. Gomes (1999)). We shall consider here a particular bootstrap procedure, called tail bootstrap, introduced by Bacro and Brito (1998). Since it is assumed that only the upper tail behaviour of F is specified, it seems clear that estimators of α should use the upper order statistics of the sample. In the same way, it seems quite natural that resampling could be performed only over these statistics. The basic idea of the tail bootstrap is, rather than bootstrapping the original sample X_1, X_2,..., X_n, to resample the sample W_1, W_2, ..., W_{ℓ_n} where

$$W_i = \log X_{n-\ell_n+i,n} - \log X_{n-\ell_n,n}, \;\; 1 \leq i \leq \ell_n,$$

and (ℓ_n) is a sequence of positive integers.

With an appropriate choice of the sequence ℓ_n, several known estimators can be easily written in terms of these random variables. In particular, for $\ell_n = k_n$, the Hill estimator is the associated sample mean, that is, $H_n = (1/k_n)\sum_{i=1}^{k_n} W_i$. Denoting by $(W_1^*, \ldots, W_{\ell_n}^*)$ a sample drawn with replacement from $W_1, \ldots, W_{\ell_n}$, the corresponding tail bootstrap version is then defined by $H_n^* = (1/k_n)\sum_{i=1}^{k_n} W_i^*$.

Moreover, under usual regularity conditions on the slowly varying function L, the distribution of the variables W_1, W_2, ..., W_{ℓ_n} can be related to the one of an i.i.d. sample of ℓ_n standard exponential variables. The properties of the tail bootstrap procedure are then governed by those of an usual bootstrap of an i.i.d. standard exponential sample.

The above estimation problem is equivalent to the estimation of an exponential tail coefficient. Setting $Z_i = \log X_i$, with X_i, $i = 1, 2, \ldots$ as above, then

$$1 - G(z) = P(Z_1 > z) = r(z)e^{-Rz}, \quad z > 0,$$

where $r(z) = L(e^z)$ is regularly varying at infinity and $R = 1/\alpha$. This equivalent formulation is more convenient for some applications. Following Csörgő and Steinebach (1991), we have in mind an important application in risk theory, the estimation of the adjustment coefficient.

In this context, several estimators have been proposed. We shall give particular attention to a geometric type estimator (Brito and Moreira (2001)). In order to simplify the notation, we consider here the estimator for $1/R$, defined by

$$\widehat{R}_n = \frac{1}{\sqrt{i_n(k)}} \left(\sum_{i=1}^{k_n} Z_{n-i+1,n}^2 - \frac{1}{k_n} \left(\sum_{i=1}^{k_n} Z_{n-i+1,n} \right)^2 \right)^{1/2},$$

where $i_n(k) = \sum_{i=1}^{k_n} \log^2(n/i) - \frac{1}{k_n} \left(\sum_{i=1}^{k_n} \log(n/i) \right)^2$.

This estimator is directly related to the least squares estimators introduced by Schultze and Steinebach (1996) and its weak limiting behaviour has been investigated in Brito and Moreira Freitas (2003). In particular, it is shown there, that $\widehat{R}_n$ has in common with Schultze and Steinebach estimators the specific property, of being universally asymptotically normal over the whole family, in the usual sense, that is, with deterministic centering sequences converging to $1/R$. The norming sequence is also the ideal factor $(k_n)^{1/2}$.

We turn now to the tail bootstrap for approximating the distribution of the normalised $\widehat{R}_n$. Remarking that

$$\widehat{R}_n = \frac{1}{\sqrt{i_n(k)}} \left(\sum_{i=1}^{k_n} W_i^2 - \frac{1}{k_n} \left(\sum_{i=1}^{k_n} W_i \right)^2 \right)^{1/2},$$

the tail bootstrap version of $\widehat{R}_n$ is then defined by

$$\widehat{R}_n^* = \frac{1}{\sqrt{i_n(k)}} \left(\sum_{i=1}^{k_n} W_{i,k_n}^{*2} - \frac{1}{k_n} \left(\sum_{i=1}^{k_n} W_{i,k_n}^* \right)^2 \right)^{1/2}.$$

We shall discuss here the asymptotic validity of the tail bootstrap in this case. Since the main term of $\widehat{R}_n^2$ is the variance of the W_i-sample, it is more convenient to consider the root

$$RR_n^* = \frac{\sqrt{k_n}}{\widehat{R}_n^2} \left(\widehat{R}_n^{*2} - \widehat{R}_n^2 \right).$$

Under the usual assumptions and, conditionally on the sample tail, RR_n^* converges weakly in probability to the same limit as the original root $RR_n = R^2(k_n)^{1/2} \left(\widehat{R}_n^2 - (1/R)^2 \right)$, that is, to the normal distribution.

We will make use of the representation

$$W_i = E_i + II_n, \;\; 1 \leq i \leq k_n,$$

where $(E_i)_{1\leq i\leq k_n}$ is distributed as the vector of the o.s. of an exponential i.i.d. k_n-sample and II_n represents an error term. This representation used in conjunction with the bootstrap CLT, was the fundamental tool for establishing the asymptotic validity of the tail bootstrap for H_n (cf. Bacro and Brito (1998)). Here, we will use Edgeworth expansions, in order to investigate the accuracy of the approximation. We also give the corresponding expansions for the Hill estimator H_n and H_n^*. In this way, we can explore the theoretical properties of the tail bootstrap procedure.

References

[1] Bacro, J. N. and Brito, M. (1998). A tail bootstrap procedure for estimating the tail Pareto-index. *J. Statist. Plan. Inf.* 71, 245-260.

[2] Brito, M. and Moreira, A.C. (2001). Estimação do Coeficiente de Cauda Exponencial. In: *"Um Olhar sobre a Estatística"* (P. Oliveira and E. Athayde, eds), Ed. SPE, 223-237.

[3] Brito, M. and Moreira Freitas (2003), Limiting behaviour of a geometric-type estimator for tail indices. *Insurance: Mathematics and Economics* 33, 211-226.

[4] Csörgő, M. and Steinebach, J. (1991). On the estimation of the adjustment coefficient in risk theory via intermediate order statistics. *Insurance: Mathematics and Economics* 10, 37-50.

[5] Csörgő, S. and Viharos, L. (1998). Estimating the tail index. In: *Asymptotic Methods in Probability and Statistics*, (B. Szyszkowicz, ed), North Holland, Amesterdam, 833-881.

[6] Gomes, M. I. (1999). The Jackknife and the bootstrap methodologies in the estimation of parameters of rare events. *Revista de Estatstica 1. Quadrimestre*, 7-23.

[7] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* 3, 1663-1674.

[8] Schultze, J. and Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient. *Statistics & Decisions* 14, 353-372.

BIAS REDUCTION FOR LIGHT TAIL MODELS — THE GENERALIZED JACKKNIFE METHODOLOGY[2]

Frederico Caeiro
Departamento de Matemática
Universidade Nova de Lisboa
Portugal

M. Ivette Gomes
CEAUL e DEIO (FCUL)
Universidade de Lisboa
Portugal

Abstract: In this paper we propose a new estimator of a light tail index $\gamma < 0$. Under a second order condition on the tail $1 - F$, this new class is "asymptotically unbiased". Using simulation techniques for the Extreme Value (EV) model, we validate the asymptotic results for small finite samples.

Key words and phrases: Statistical Theory of Extremes, The Generalized Jackknife Methodology.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be independent random variables (r.v.'s) with common distribution function (d.f.) F in the domain of attraction of an EV d.f. with shape parameter $\gamma < 0$: $G_\gamma(x) = \exp[-(1+\gamma x)^{-\frac{1}{\gamma}}]$, $1+\gamma x > 0$. The previous condition, denoted $F \in D(G_\gamma)$, is equivalent to:

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \quad \forall\, x > 0,$$

[2]Research partially supported by FCT/POCTI/FEDER.

where $a(t) > 0$ and $U(t) := F^{\leftarrow}(1-1/t),\ t \geq 1$.

Some of the classical estimators, like Moment's estimator or Pickand's estimator, have high variance for small values of k and high bias for large values of k. This leads to a very sharp Mean Square Error (MSE) pattern that makes difficult the choice of the best k. Some research has been done in the reduction of bias, but essentially for heavy tail models. Such a reduction allows us to use alternative estimators, with stable sample paths and "bath-tube" MSE patterns, where the choice of k is "almost" irrelevant. This is the main objective of this work, but now for light tail models.

The class of semi-parametric estimators of $\gamma_- = \min(0, \gamma)$, herewith introduced, is dependent on a tuning parameter α, and given by:

$$\gamma_{n,\alpha}^{M^-}(k) := 1 - \frac{\alpha}{\alpha+1}\left(1 - \frac{M_n^{(\alpha)}(k) M_n^{(1)}(k)}{M_n^{(\alpha+1)}(k)}\right)^{-1}, \quad \alpha \geq 1, \tag{1}$$

where $M_n^{(\alpha)}(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\ln\left(X_{n-i+1:n}/X_{n-k:n}\right)\right)^{\alpha}, \quad \alpha \geq 0$.

Remark 1. *The moment estimator* $\gamma_n^M(k)$ *is given by* $\gamma_n^M(k) = \gamma_n^H(k) + \gamma_{n,1}^{M^-}(k)$*, where* $\gamma_n^H(k)$ *is the classical Hill estimator.*

2. Asymptotic properties

We need a second order condition to get the estimators' asymptotic distribution: there is a function $A(\cdot) \longrightarrow 0$ such that, for every $x > 0$ and $\rho \leq 0$:

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^{\gamma}-1}{\gamma}}{A(t)} = H(x,\gamma,\rho) = \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^{\gamma}-1}{\gamma}\right), \tag{2}$$

where ρ is the second order parameter.

Remark 2. *Note that if* $\rho < 0$, $a(t) \sim a_1 t^{\gamma}$*, and* $A(t) \sim \rho(\gamma+\rho)a_2 t^{\rho}/a_1$*, (2) is equivalent to:* $U(t) = a_0 + a_1\frac{x^{\gamma}-1}{\gamma} + a_2 t^{\gamma+\rho} + o(t^{\gamma+\rho})$ *as* $t \longrightarrow \infty$*,* $a_1 > 0,\ a_2 \neq 0$.

Proposition 1. *Under the validity of the second order condition* (2) *and for $k = k(n)$ intermediate:*

$$\frac{M_n^{(\alpha)}(k)}{(a_0(t))^\alpha} \stackrel{d}{=} \mu(\alpha,\gamma_-)+\frac{\sigma(\alpha,\gamma_-)}{\sqrt{k}} Z_n^{(\alpha)}(k)+b(\alpha,\gamma_-,\rho^*)A^*(n/k)\left(1+o_p(1)\right),$$

where $Z_n^{(\alpha)}(k) = \sqrt{k}\frac{\frac{1}{k}\sum_{i=1}^{k}\left[\left(\frac{Y_i^{\gamma_-}-1}{\gamma_-}\right)^\alpha - \mu(\alpha,\gamma_-)\right]}{\sigma(\alpha,\gamma_-)}$ *is an asymptotically standard normal r.v., Y_i are Pareto r.v.'s, with d.f. $F_Y(x) = 1 - 1/x$, $x \geq 1$ and:*

$$\mu(\alpha,\gamma_-) = \mathbb{E}\left[\left(\frac{Y_i^{\gamma_-}-1}{\gamma_-}\right)^\alpha\right] = \begin{cases} \frac{\Gamma(\alpha+1)\Gamma(-1/\gamma)}{(-\gamma)^{\alpha+1}\Gamma(\alpha+1-1/\gamma)}, & \gamma < 0 \\ \Gamma(\alpha+1), & \gamma \geq 0 \end{cases}$$

$$\sigma^2(\alpha,\gamma_-) = V\left[\left(\frac{Y_i^{\gamma_-}-1}{\gamma_-}\right)^\alpha\right] = \mu(2\alpha,\gamma_-) - \mu^2(\alpha,\gamma_-)$$

$$b(\alpha,\gamma_-,\rho^*) = \mathbb{E}\left[\alpha\left(\frac{Y_i^{\gamma_-}-1}{\gamma_-}\right)^{\alpha-1} H(Y_i,\gamma_-,\rho^*)\right]$$

$$\rho^* = \begin{cases} -\gamma, & 0<\gamma<-\rho \wedge \lim_{t\to\infty} \frac{U(t)-a(t)}{\gamma} \neq 0 \\ \gamma, & \rho < \gamma \leq 0 \\ \rho, & (0<\gamma<-\rho \wedge \lim_{t\to\infty} \frac{U(t)-a(t)}{\gamma} = 0) \;\vee\; \gamma<\rho \;\vee\; \gamma \geq -\rho \end{cases}$$

Proposition 2. *Under the conditions and notations of the previous proposition:*

$$\gamma_{n,\alpha}^{M^-}(k) \stackrel{d}{=} \gamma_- + \frac{(1-\gamma_-)[(\alpha+1)\gamma_- - 1]}{\alpha} \times \left[\frac{N_n^{(\alpha)}(k)}{\sqrt{k}} + B(\alpha,\gamma_-,\rho^*)A^*(n/k)\left(1+o_p(1)\right)+o_p\left(\frac{1}{\sqrt{k}}\right)\right]$$

where:

$$N_n^{(\alpha)}(k) = \frac{\sigma(\alpha,\gamma_-)}{\mu(\alpha,\gamma_-)} Z_n^{(\alpha)}(k) + \frac{\sigma(1,\gamma_-)}{\mu(1,\gamma_-)} Z_n^{(1)}(k) - \frac{\sigma(\alpha+1,\gamma_-)}{\mu(\alpha+1,\gamma_-)} Z_n^{(\alpha+1)}(k),$$

$$B(\alpha,\gamma_-,\rho^*) = \frac{b(\alpha,\gamma_-,\rho^*)}{\mu(\alpha,\gamma_-)} + \frac{b(1,\gamma_-,\rho^*)}{\mu(1,\gamma_-)} - \frac{b(\alpha+1,\gamma_-,\rho^*)}{\mu(\alpha+1,\gamma_-)}.$$

Corollary 1. *Under the conditions and notations of the previous proposition and assuming that* $\sqrt{k}\ A^*(n/k) \longrightarrow \lambda$,

$$\sqrt{k}\left(\gamma_{n,\alpha}^{M^-}(k) - \gamma_-\right) \xrightarrow{d} N\left(c(\gamma_-)B(\alpha,\gamma_-,\rho^*)\lambda, (c(\gamma_-)\sigma_N)^2\right).$$

where $V[N_n^{(\alpha)}(k)] = \sigma_N^2 \qquad c(\gamma_-) = \frac{(1-\gamma_-)[(\alpha+1)\gamma_- - 1]}{\alpha}$.

Since $\lambda \neq 0$ and $\nexists\ \alpha_0 : c(\gamma_-)B(\alpha_0,\gamma_-,\rho^*) = 0$, $\gamma_{n,\alpha}^{M^-}(k)$ has always a non null asymptotic bias. To achieve our objective (bias reduction), we shall use next the generalized jackknife methodology.

3. The Generalized Jackknife estimator

Asymptotic properties

Let $\widehat{\theta}_1$ and $\widehat{\theta}_2$ be two unbiased estimators of θ such that for $i = 1,2$, $\mathbb{E}(\widehat{\theta}_i) = \theta + b(\theta)\ d_i(n)$. Then, the Generalized Jackknife estimator associated to $(\widehat{\theta}_1,\widehat{\theta}_2)$ is:

$$\widehat{\theta}^G = \frac{\widehat{\theta}_1 - q\widehat{\theta}_2}{1-q}, \qquad \text{where} \quad q = \frac{\text{Bias}[\widehat{\theta}_1]}{\text{Bias}[\widehat{\theta}_2]} = \frac{d_1(n)}{d_2(n)}.$$

Let us make $\widehat{\theta}_1 = \gamma_{n,2}^{M^-}(k)$ and $\widehat{\theta}_2 = \gamma_{n,1}^{M^-}(k)$ ($\gamma_{n,i}^{M^-}(k)$, both in the class $\gamma_{n,\alpha}^{M^-}(k)$). We now have the following Generalized Jackknife estimator (with null asymptotic bias):

$$\gamma_n^{G(q)}(k) \equiv \gamma_n^G(k) = \frac{\gamma_{n,2}^{M^-}(k) - q\gamma_{n,1}^{M^-}(k)}{1-q},$$

$$q = \frac{(1-3\gamma_-)(1-\rho^*/2-2\gamma_-)}{(1-\rho^*-3\gamma_-)(1-2\gamma_-)},$$

with $0.5 \le q \le 1$ if $\rho^* < 0$.

Note that q needs to be properly estimated (to get a null asymptotic bias). Since $q = f(\gamma,\rho*)$ and we do not have, for light tail models, any good estimator for ρ^*, the value q can result from a heuristic estimation (using the stability of the sample path). This has also been done with success for other estimators of a heavy tail index.

Finite sample behavior — a simulation study for the EV model

The results were based on a multi-sample simulation of size 10×1000 for the EV model,

$$\begin{aligned} U(t) &= \frac{(-\ln(1-1/t))^{-\gamma}-1}{\gamma} \\ &= \begin{cases} \frac{t^{\gamma}-1}{\gamma} - \frac{1}{2}t^{\gamma-1} + o(t^{\gamma-1}),\ t \longrightarrow \infty\ (\gamma \neq 1) \\ -\frac{1}{2} + (t-1) - \frac{t^{-1}}{12} + o(t^{-1}),\ t \longrightarrow \infty\ (\gamma = 1) \end{cases} \end{aligned}$$

We shall present in Figure 1 and Table 1 the results for $\gamma = -0.5$ and $\rho^* = -0.5$.

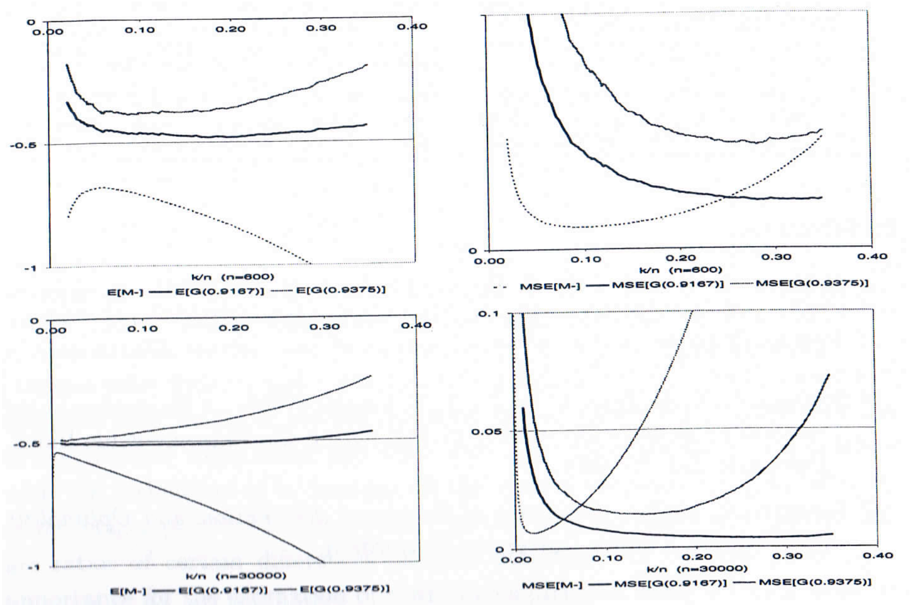

Figure1: Mean values (left), MSE (right) of $\gamma_n^{G(q)}(k)$ and $\gamma_{n,1}^{M^-}(k)$ from the model EV with $\gamma = -0.5$.

Table 1: Simulated optimal sample fraction, mean value and MSE for the *EV* model.

$\gamma = -0.5$

	150	600	2000	3000	6000	15000	30000	60000
	$\gamma_{n,1}^{M^-}(k)$							
k_0/n	0.1767	0.0987	0.0571	0.0518	0.0382	0.0243	0.0198	0.0139
$\mathbb{E}[\hat{\gamma}(k_0)]$	**-0.8993**	**-0.7067**	**-0.6253**	**-0.6146**	**-0.5904**	-0.5660	-0.5566	-0.5446
$MSE[\hat{\gamma}(k_0)]$	**0.3261**	**0.0914**	**0.0364**	**0.0278**	**0.0177**	0.0099	0.0066	0.0043
	$\gamma_n^{G(0.9167)}$							
k_0/n	0.3047	0.2923	0.2945	0.3077	0.2897	0.2849	0.2739	0.2656
$\mathbb{E}[\hat{\gamma}(k_0)]$	-0.3580	-0.4645	-0.4821	-0.4789	-0.4887	**-0.4915**	**-0.4949**	**-0.4973**
$MSE[\hat{\gamma}(k_0)]$	0.9630	0.1920	0.0550	0.0372	0.0184	**0.0073**	**0.0037**	**0.0019**
	$\gamma_n^{G(0.9375)}$							
k_0/n	0.2947	0.2787	0.2312	0.2180	0.1906	0.1639	0.1401	0.1113
$\mathbb{E}[\hat{\gamma}(k_0)]$	-0.1237	-0.2957	-0.3708	-0.3884	-0.4114	-0.4302	-0.4433	-0.4584
$MSE[\hat{\gamma}(k_0)]$	2.0133	0.4205	0.1331	0.0917	0.0498	0.0233	0.0136	0.0079

References

[1] Dekkers, A. L. M., J. H. J. Einmahl e L. de Haan (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statit.* **17**, 1833-1855.

[2] Draisma, G. L. de Haan, L. Peng e T. T. Pereira (1999). A Bootstrap-based Method to Achieve Optimality in Estimating the Extreme-value Index. *Extremes.* **2:4**, 367-404.

[3] Ferreira, A. (2002). *Statistics of Extremes: Estimation and Optimality.* Ph.D. thesis, Tilburg University, Rotterdam.

RATIOS OF DIFFERENCES OF ORDER STATISTICS IN MAX-SEMISTABLE MODELS[3]

Sandra Dias
Departamento de Matemática
Universidade de Trás-os-Montes e Alto Douro
Portugal

Luisa Canto e Castro
CEAUL e DEIO (FCUL)
Universidade de Lisboa
Portugal

Abstract: According with the results in Canto e Castro et al. (2000) the max-semistable models can be characterized by a parameter $r(\geq 1)$, by the extreme value index γ and a real function y defined in $[0, 1]$. The parameter r identifies each family of models and must be estimated before the estimation of the extreme value index and the function y. However there is a problem with the estimation of r, because all the results are valid for r, or any one of its integer powers (Temido, 2000). In this talk sequences of statistics that are ratios of certain differences of order statistics will be considered, their importance for the estimation of r will be emphasized along with the analysis of the asymptotic behaviour of their trajectories.

Key words and phrases: max-stable laws, convergent estimators.

[3]Research partilly supported by FCT / POCTI / FEDER.

1. Introduction

Let $\{k_n\}$ be a nondecreasing positive integer sequence, satisfying

$$\lim_{n\to+\infty} \frac{k_{n+1}}{k_n} = r, \tag{1}$$

with $r \geq 1$.

Assume there exist sequences of norming constants $\{a_n\}$, $a_n > 0$, and $\{b_n\}$ such that, for all $x \in C(G)$,

$$\lim_{n\to+\infty} F^{k_n}(a_n x + b_n) = G(x), \tag{2}$$

where G is a nondegenerate distribution function. Here $C(G)$ denotes the set of continuity points of the distribution function G.

Grinevich (1993) [3] proved that the limit distribution functions G in (2) verify

$$G(x) = G^r(ax + b), \tag{3}$$

$a > 0$, $b \in \mathbb{R}$ e $r > 1$, leading to the max-semistable models.

The generalized standard form (analogous to the Von Mises-Jenkinson generalised form) is

$$G_{\nu,\gamma}(x) = exp\left\{-(1+\gamma x)^{-\frac{1}{\gamma}}\nu(log(1+\gamma x))\right\},$$

$x \in \mathbb{R}$, $1 + \gamma x > 0$ and $\gamma \neq 0$.

When $\gamma \to 0$ we obtain

$$G_{\nu,0}(x) = exp\{-e^{-x}\nu(x)\},$$

where ν is a positive, bounded and periodic function. The parameters γ and the period p of the function ν are related with the parameters a, b and r in (3) in the following way:

$$\gamma = \frac{log\ a}{log\ r}, \quad p = |log\ a|, \gamma \neq 0 \ \text{ and } \ p = b, \gamma = 0.$$

In Canto e Castro *et al* (2000) [1] the following condition, equivalent to (2), is presented,

$$\lim_{n\to+\infty} \frac{V(x + log\ k_n) - b_n}{a_n} = \Psi(x),\ \forall x \in C(G), \tag{4}$$

with $V(x) := (-log(-log\ F))^{\leftarrow}(x)$ e $\Psi(x) := (-log(-log\ G))^{\leftarrow}(x)$, where $f^{\leftarrow}$ represents the generalized inverse function of f.

Using (4) and supposing, without loss of generality,

$$\begin{cases} G(0) = e^{-1} \\ G(1) = exp(-r^{-1}) \\ G \text{ is continuous at } x = 0 \end{cases},$$

these authors present the following representation of these functions

$$\Psi(nlog\ r + x) = s_n + a^n w(x), \forall\ x \in [0, log\ r],\ n \in \mathbb{Z}$$

where $w : [0, log\ r] \to [0, 1]$ is non-decreasing, left-continuous and continuous at $x = 0$, and $s_n = \frac{a^n-1}{a-1}$ if $a \neq 1$ and $a > 0$ or $s_n = n$ if $a = 1$, for any integer n.

It is easy to see in this representation that if the condition holds for one value of r then it also holds for any of its integer powers.

Using the previous representation the authors proved that the following conditions are necessary and sufficient to (2)

$$\lim_{n\to+\infty} \frac{V(log\ k_{n+1}) - V(log\ k_n)}{V(log\ k_n) - V(log\ ; k_{n-1})} = a \tag{5}$$

and

$$\lim_{n\to+\infty} \frac{V(log\ k_n + x) - V(log\ k_n)}{V(log\ k_{n+1}) - V(log\ k_n)} = w(x),\ x \in [0, log\ r]. \tag{6}$$

2. Ratios of differences of order statistics in max-semistable models

In analogy to what has been done in the max-stable case, we are interested in the asymptotic behavior of the following sequence of statistics

$$Z_s(m_n) = \frac{X_{(\frac{m_n}{s})} - X_{(m_n)}}{X_{(m_n)} - X_{(m_n s)}}, \tag{7}$$

where $X_{(m_n)} := X_{N-m_n+1:N}$ are order statistics of a sample of size N from any random variable X, with $\lim\limits_{n\to+\infty} m_n = +\infty$ and $\lim\limits_{n\to+\infty} \frac{m_n}{N} = 0$.

In this work we have proved the next theorem that in some sense shows that the trajectory of the statistics only stabilizes if $s = r^c$, $c \in \mathbb{N}$

Theorem 1. *Assume that* $(X_1, X_2, ..., X_N)$ *is a sample of independent identically distributed random variables with continuous distribution function* F*, where* F *verifies (2) with* $\{k_n\}$ *verifying (1),* $r > 1$*. Let* $n = n(N) = max\{p \in \mathbb{N} : r^p \leq N\}$*,* $h_n = \frac{N}{r^{n-t_n}} \left(r^{t_{n+1}-t_n-1} - 1\right) - 1$*,* $\{t_n\}$ *a sequence of positive integers verifying*

$$\lim_{n\to+\infty} t_n = \lim_{n\to+\infty} n - t_n = +\infty, \ \lim_{n\to+\infty} t_{n+1} - t_n = 0$$

and

$$m_n = \frac{N}{r^{n-t_n}}.$$

Consider the sequence of statistics in (7)*. Then,*

i) $Z_s(m_n)$ *converges* $\forall\, s > 1$.

ii) If $s = r^c, c \in \mathbb{N}$ *then*

$$\lim_{n\to+\infty} Z_s(m_n + l_n) \overset{P}{=} a^c,$$

for all sequences $\{l_n\}$ *with elements in* $\{0, 1, 2, ..., h_n\}$*.*

iii) If $s \neq r^c, c \in \mathbb{N}$ *then there exist sequences* $l_{1,n}, l_{2,n}$ *with elements in* $\{0, 1, 2, ..., h_n\}$*,* $l_{1,n} \neq l_{2,n}$ *such that*

$$\lim_{n\to+\infty} Z_s(m_n + l_{1,n}) \overset{P}{\neq} \lim_{n\to+\infty} Z_s(m_n + l_{2,n}).$$

References

[1] Canto e Castro, L., de Haan, L., and Temido, M. G. (2000). Rarely observed sample maxima. *Theory of Probab. Appl.*, **45**, 779-782.

[2] Grinevich, I.V. (1992). Max-semistable laws corresponding to linear and power normalizations. *Theory of Probab. Appl.*, **37**, 720-721.

[3] Grinevich, I.V. (1993). Domains of attraction of the max-semistable laws under linear and power normalizations. *Theory of Probab. Appl.*, **38**, 640-650.

[4] Temido, M.G. (2000) *Classes de Leis Limite em Teoria de Valores Extremos — Estabilidade e Semiestabilidade*. PhD Thesis, Coimbra University.

THE COMPOUND BINOMIAL MODEL REVISITED[4]

Alfredo Egídio Reis
CEMAPRE and ISEG
Universidade Técnica de Lisboa
Portugal

For Actuarial Science purposes, or Risk Theory in particular, we consider a discrete time risk model where the aggregate claim process is compound binomial. In each period there is a claim with probability p, or no claim with probability $1 - p$ and that there is an independence in claim occurrence in different time periods. The sequence of individual claims is a sequence of i.i.d. random variables, distributed on the positive integers, and independent of the claim number process. We follow the model formulation presented by Gerber (1988) and Dickson (1994).

Starting from a non-negative integer initial surplus, there is a positive probability that the risk process is ruined, i.e., it drops to negative values. If ruin occurs, it happens at the instant of a claim, then we can address the ruin probability problem, either finite or infinite time, by considering the *number of claims necessary to get ruined.* We then consider the calculation of the distribution of the number of claims up to ruin, if it occurs. Besides, since the process once ruined will recover to positive levels some time in the future with probability one, we also consider the

[4]Support from *Fundação para a Ciência e a Tecnologia* — FCT/POCTI is gratefully aknowledged.

distribution of the number of claims ocurring during the recovery time period.

An interesting result is achieved concerning the particular case when the initial surplus is zero, which is the fact that the two discrete random variables above have the same distribution and that the distribution belongs to the *Lagrangian-type* family, and a closed form for the distribution is found.

From that, it is possible to find a recursion that allows the computation of the distribution of the number of claims up to ruin, considering any positive integer initial surplus. For these cases, we also find a formula for the distribution of the number of claim during a recovery time period.

From the results above, we are also able to find expressions for a key point in ruin theory, which is the computation of ruin probabilities, either infinite or finite time.

Besides, with this model we will be able to compute approximations for the related quantities in the classical compound Poisson risk model.

References

[1] Dickson, D. (1994). Some comments on the compound binomial model. *ASTIN Bulletin* **24**, 33–45.

[2] Gerber, E. (1988). Mathematical fun with the compound binomial process. *ASTIN Bulletin* **18**, 161–168.

TESTING FOR TAIL INDEPENDENCE IN MULTIVARIATE EXTREME VALUE MODELS

Michael Falk
Inst. of Applied Math. & Statistics
University of Würzburg
Germany

René Michel
Inst. of Applied Math. & Statistics
University of Würzburg
Germany

Abstract: Let (X, Y) be a random vector which follows in its upper tail a bivariate extreme value distribution with reverse exponential margins. We show that the conditional distribution function (df) of $X + Y$, given that $X + Y > c$, converges to the df $F(t) = t^2$, $t \in [0, 1]$, as $c \uparrow 0$ if and only if X and Y are tail independent. If they are not tail independent, the limit df is $F(t) = t$, $t \in [0, 1]$.

This result is utilized to test for tail independence of X and Y via various tests, including the one suggested by the Neyman–Pearson lemma and goodness–of–fit tests such as a suitable version of Fisher's κ–test, the Kolmogorov–Smirnov test and the chi–square goodness–of–fit test, applied to a sample of independent copies of (X, Y). Simulations show that the Neyman–Pearson test performs best if the threshold c is close to 0, whereas otherwise the Kolmogorov–Smirnov test performs best. Concerning the type II error rate, Fisher's κ fails.

Key words and phrases: Bivariate extremes, Pickands dependence function, tail independence, tail dependence parameter, Neyman–Pearson lemma, Kolmogorov–Smirnov test, Fisher's κ, chi–square goodness–of–fit test.

1. Introduction

Let (X, Y) be a random vector with values in $(-\infty, 0]^2$, whose distribution function (df) $H(x, y)$ coincides for $x, y \le 0$ close to 0 with a max–stable or extreme value df (EV) G with reverse exponential margins, i.e., $G(x, 0) = G(0, x) = \exp(x)$, $x \le 0$, and $G^n(x/n, y/n) = G(x, y)$, $x, y \le 0$, $n \in \mathbb{N}$.

It is well–known that G can be represented as

$$G(x, y) = \exp\left((x + y)D\left(\frac{x}{x + y}\right)\right), \qquad x, y \le 0,$$

where $D : [0, 1] \to [1/2, 1]$ is the *Pickands dependence function* ([11], [4], Theorem 5.4.5, [13], Proposition 5.11). We refer to Section 2 of [3] for elementary derivations of basic properties of D.

Let now $(X_1, Y_1), \dots, (X_n, Y_n)$ be independent copies of (X, Y). If diagnostic checks of $(X_1, Y_1), \dots, (X_n, Y_n)$ suggest X, Y to be independent in their upper tail, modelling with dependencies leads to the over estimation of probabilities of extreme joint events. Some inference problems caused by model mis–specification are, for example, exploited in [2]. Testing for tail independence is, therefore, mandatory in a data analysis of extreme values.

In Section 2 we will establish the fact that the conditional distribution of $X + Y$, given $X + Y > c$, has limiting df $F(t) = t^2$, $t \in [0, 1]$, as $c \uparrow 0$ if and only if $D(z) = 1$, $z \in [0, 1]$, i.e., if and only if X and Y are tail independent. If D is not the constant function 1, then the limiting df is that of the uniform distribution on $[0, 1]$: $F(t) = t$, $t \in [0, 1]$.

This result will be utilized to define tests for tail independence of X and Y which are suggested by the Neyman–Pearson lemma as well as via goodness–of–fit tests that are based on Fisher's κ, on the Kolmogorov–Smirnov test as well as on the chi–square goodness–of–fit test, applied to the exceedances $X_i + Y_i > c$ among the sample $(X_1, Y_1), \dots, (X_n, Y_n)$. Numerous simulations which we did indicate that the Neyman–Pearson test has the smallest type II error rate, closely followed by the Kolmogorov–Smirnov test and the chi–square test, whereas Fisher's κ almost fails. The Neyman–Pearson test does not, however, control the type I error rate if the threshold c is too far away from 0. The other three tests control the type I error rate for any c. Note that a multivariate EV with arbitrary onedimensional margins is transformed to an EV G with reverse exponential margins by simple corresponding transformations of its margins.

Starting with the work by Geffroy ([5], [6]) and Sibuya ([14]), X and Y are said to be *tail independent* or *asymptotically independent* if the *tail dependence parameter*

$$\chi := \lim_{c \uparrow 0} P(Y > c | X > c)$$

equals 0. Note that $\chi = 2(1 - D(1/2))$ and, thus, the convexity of $D(z)$ implies that $\chi = 0$ is equivalent to the condition $D(z) = 1$, $z \in [0, 1]$.

The recent attention given to statistical properties of asymptotically independent distributions is largely a result of a series of articles by Ledford and Tawn ([8], [9], [10]); [1] gives an elementary synthesis of the theory. For a directory of coefficients of tail dependence such as χ we refer to [7].

2. The Bivariate Case

We assume in the following that the random vector (rv) (X, Y) has a df $H(x, y)$, which coincides for x, y close to 0 with the EV $G(x, y) = \exp((x + y)D(x/(x + y))$, where D is an arbitrary Pickands dependence function.

Lemma 1. *We have for $c < 0$ close to 0*

$$P(X + Y \leq c) = \exp(c) - c \int_0^1 \exp(cD(z)) \big(D(z) + D'(z)(1 - z)\big)\, dz.$$

Lemma 2. *We have uniformly for $t \in [0, 1]$ as $c \uparrow 0$*

$$P(X + Y > ct | X + Y > c) = \begin{cases} t^2(1 + O(c)), & \text{if } D(z) = 1,\ z \in [0, 1], \\ t(1 + O(c)) & \text{elsewhere.} \end{cases}$$

If X and Y are tail independent, then $(X + Y)/c$, conditional on $X + Y > c$, has by Lemma 1 and 2 for c close to 0 the df

$$\begin{aligned} F_c(t) := P(X + Y > tc | X + Y > c) &= \frac{1 - (1 - tc)\exp(tc)}{1 - (1 - c)\exp(c)} \\ &= t^2(1 + O(c)), \end{aligned} \tag{1}$$

$0 \le t \le 1$. Otherwise, the conditional df converges to the uniform df on [0,1].

Suppose now that we have n independent copies $(X_1, Y_1), \dots, (X_n, Y_n)$ of (X, Y). Fix $c < 0$ and consider only those observations $X_i + Y_i$ among the sample with $X_i + Y_i > c$. Denote these by $C_1, C_2, \dots, C_{\tau(n)}$ in the order of their outcome. Then C_i/c, $i = 1, 2, \dots$ are iid with common df F_c, if c is large enough, and they are independent of $\tau(n)$, which is binomial $B(n, q)$–distributed with $q = 1 - (1 - c)\exp(c)$. This is a consequence of Theorem 1.4.1 in [3].

The first test which we will consider is suggested by the Neyman–Pearson lemma. We have to decide, roughly, whether the df of $V_i := C_i/c$, $i = 1, 2, \dots$ is either equal to the null hypothesis $F_{(0)}(t) = t^2$ or the alternative $F_{(1)}(t) = t$, $0 \le t \le 1$. Assuming that these approximations of the df of $V_i := C_i/c$ were exact and that $\tau(n) = m > 0$, the optimal test for testing $F_{(0)}$ against $F_{(1)}$ is then based on the loglikelihood ratio

$$T(V_1, \dots, V_m) := \log\left(\prod_{i=1}^{m} \frac{1}{2V_i}\right) = -\sum_{i=1}^{m} \log(V_i) - m\log(2),$$

and $F_{(0)}$ is rejected if $T(V_1, \dots, V_m)$ gets too large. Note that $-2\log(V_i)$ has df $1 - \exp(-x)$, $x \ge 0$, under $F_{(0)}$, and hence, $2(T(V_1, \dots, V_m) + m\log(2))$ has df $1 - \exp(-x)\sum_{0 \le j \le m-1} x^j/j!$, $x \ge 0$, under $F_{(0)}$.

The p–value of the optimal test derived from the Neyman–Pearson lemma is, therefore,

$$\begin{aligned} p_{NP} &= \exp\left(2\sum_{i=1}^{m} \log(V_i)\right) \sum_{0 \le j \le m-1} \frac{\left(-2\sum_{i=1}^{m} \log(V_i)\right)^j}{j!} \\ &\approx \Phi\left(\frac{2\sum_{i=1}^{m} \log(V_i) + m}{m^{1/2}}\right) \end{aligned}$$

if m is large by the central limit theorem, where Φ denotes the df of the standard normal df.

Next we consider goodness–of–fit tests based on C_i/c for testing for tail independence of X and Y. Conditional on $\tau(n) = m > 0$, the rvs

$$U_i := F_c(C_i/c) = \frac{1 - (1 - C_i)\exp(C_i)}{1 - (1 - c)\exp(c)}, \qquad 1 \le i \le m,$$

are by equation (1) independent and uniformly on $(0, 1)$ distributed, if X and Y are tail independent and c is close to 0. We test this hypothesis by

Fisher's κ–statistic, conditional on $\tau(n) = m$, the Kolmogorov–Smirnov test and the chi–square goodness–of–fit test. Based on numerous simulations, it turns out that Neyman–Pearson test for independence has the smallest type II error rate, followed by the Kolmogorov–Smirnov test. The test for independence of X and Y based on Fisher's κ, however, fails.

References

[1] Coles, S., Heffernan, J. and Tawn, J.A. (1999). Dependence measures for extreme value analysis. *Extremes* **2**, 339-365.

[2] Dupuis, D.J. and Tawn, J.A. (2001). Effects of mis–specification in bivariate extreme value problems. *Extremes* **4**, 315-330.

[3] Falk, M. and Reiss, R.–D. (2003). Efficient estimation of the canonical dependence function. *Extremes* **6**, 61-82.

[4] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics.*, 2nd edition, Krieger, Malabar.

[5] Geffroy, J. (1958). Contribution à la théorie des valeurs extrêmes.*Publl. l'Inst. Statist. l'Univ. Paris* **7**, 37-121.

[6] Geffroy, J. (1959). Contribution à la théorie des valeurs extrêmes, II.*Publl. l'Inst. Statist. l'Univ. Paris* **8**, 3-65.

[7] Heffernan, J.E. (2000). A directory of coefficients of tail dependence. *Extremes* **3**, 279-290.

[8] Ledford, A. and Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169-187.

[9] Ledford, A. and Tawn, J.A. (1997). Modelling dependence within joint tail regions. *J. R. Stat. Soc.* **B 59**, 475-499.

[10] Ledford, A. and Tawn, J.A. (1998). Concomitant tail behaviour for extremes. *Adv. Appl. Probab.* **30**, 197-215.

[11] Pickands, J. (1981). Multivariate extreme value distributions. *Proc. 43th Session ISI* (Buenos Aires) 859-878.

[12] Reiss, R.–D. (1993). *A Course on Point Processes.* Springer Series in Statistics, Springer, New York.

[13] Resnick S.I. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer Series in Applied Probability, Springer, New York.

[14] Sibuya, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.* **11**, 195-210.

SCALE ESTIMATORS IN STATISTICAL QUALITY CONTROL[5]

Fernanda Figueiredo
CEAUL e Faculdade de Economia
Universidade do Porto
Portugal

M. Ivette Gomes
CEAUL e DEIO (FCUL)
Universidade de Lisboa
Portugal

Abstract: In many practical situations we assume normal data due to the advantages of the use of this assumption. However, even if in potential normal situations, there is some possibility of having disturbances in the data, and the classic procedures of inference may be inappropriate, as well as the traditional control charts. In this paper we propose some efficient and robust estimators for the standard deviation, and we analyze the performance of some control charts to monitor the standard deviation based on these robust statistics.

Key words and phrases: Bootstrap Sample, Nonparametric Inference, Robust Estimators, Statistical Process Control (SPC).

1. Introduction

The classic procedures of inference based on the assumption of normal data may be inappropriate for non-normal distributions, or even when we have some disturbances in the data, such as contamination or outliers. The same problem arises when we are implementing a control chart to monitor an industrial process: the traditional control charts for normal

[5]Research partially supported by FCT/POCTI/FEDER.

data may exhibit a bad performance specially for skewed and/or heavy-tailed distributions. It is thus important to find efficient and robust estimators of the process parameters. In this paper we focus on the estimation and monitoring of σ, the population standard deviation. The problem of estimating the population mean μ has already been addressed in Figueiredo and Gomes (2003a).

Although there are several robust location and scale estimators, including the M and R estimators (see, for instance, Lax (1985) and Tatum(1997)), we propose other estimators for the standard deviation that can be considered robust alternatives to the usual ones for contaminated distributions or in the presence of outliers. These estimators are the total range, TR, and two modified versions of the sample standard deviation, S^* and S^{**}. In Section 2, we define the proposed estimators for the standard deviation, and in Section 3 we describe the study we have done to evaluate their efficiency and robustness; we also analyze the robustness of some control charts to monitor the standard deviation. Finally, in Section 4 we present some conclusions.

2. Definition of some scale estimators

Let $(X_1, X_2, ..., X_n)$ be a random sample of size n from a distribution F and let us denote $X_{i:n}$, $1 \leq i \leq n$, the random sample of the associated ascending order statistics ($o.s.$). The bootstrap sample, $(X_1^*, X_2^*, ..., X_n^*)$, is obtained by randomly sampling n times, with replacement, from the observed sample $(x_1, x_2, ..., x_n)$. These variables X_i^* are independent and identically distributed ($i.i.d.$) replicates from a random variable X^*, with distribution function ($d.f.$) equal to the empirical $d.f.$ of our observed sample.

The *total range* is the statistic

$$TR = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} \left(X_{j:n} - X_{i:n}\right) := \sum_{i=1}^{n} b_i X_{i:n},$$
$$\beta_{ij} = P(BR = x_{j:n} - x_{i:n} | BR \neq 0),\ 1 \leq i < j \leq n,$$

where BR denotes the range of the bootstrap sample, defined by

$$BR = X_{n:n}^* - X_{1:n}^*.$$

The modified versions of the sample standard deviation that we have here considered are the statistics

$$S^* = \sqrt{\sum_{i=1}^{n} a_i \left(X_{i:n} - \overline{X}\right)^2} \qquad and \qquad S^{**} = \sqrt{\sum_{i=1}^{n} a_i \left(X_{i:n} - TMd\right)^2},$$

where TMd denotes the *total median* of the bootstrap sample, defined by

$$TMd = \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{ij} \frac{X_{i:n}+X_{j:n}}{2} =: \sum_{i=1}^{n} a_i X_{i:n},$$
$$\alpha_{ij} = P(BMd = \tfrac{x_{i:n}+x_{j:n}}{2}),\ 1 \le i \le j \le n.$$

Here BMd denotes the median of the bootstrap sample.

The probabilities β_{ij} can be obtained in terms of multinomial distributions, and are given by

$$\beta_{ij} = \begin{cases} \sum_{k=1}^{n-1} \sum_{r=1}^{n-k} \frac{n!}{k!r!(n-k-r)!} \left(\frac{1}{j-i-1}\right)^{k+r} \left(\frac{j-i-1}{n}\right)^n, \ 1 \le i < j \le n \\ \left(\frac{1}{n}\right)^n, \ \ 1 \le i = j \le n \end{cases}.$$

The coefficients b_i are related with the probabilities β_{ij} through the expression

$$b_i = \sum_{j=1}^{i-1} \beta_{ji} - \sum_{j=i+1}^{n} \beta_{ij},\ 1 \le i \le n,$$

with $\sum_{j=k}^{r} \beta_{ij} = 0$, for $k > r$. These coefficients b_i as well as the coefficients a_i are "*distribution-free*", i.e., they are independent of the underlying model F, they depend only on the sample size n, and they verify the conditions

$$\begin{aligned} a_i &= a_{n-i+1},\ 1 \le i \le n, \quad 0 < a_1 \le a_2 \le ... \le a_{[n/2]}, \quad \sum_{i=1}^{n} a_i = 1, \\ b_i &= -b_{n-i+1},\ 1 \le i \le n, \quad b_1 \le b_2 \le ... \le b_{[n/2]} < 0, \quad \sum_{i=1}^{[n/2]} b_i = -1. \end{aligned}$$

Details about the probabilities α_{ij} and β_{ij}, and about the coefficients a_i and b_i can be found in Figueiredo (2002).

3. Efficiency and robustness of the proposed estimators. Quasi-robust control charts.

To compare the overall performance of the proposed estimators for the population standard deviation a large scale simulation study of size 25×2500 has been undertaken. For comparison, we have also considered the *median absolute deviation from the median* defined by

$$MAD = Md|X_i - Md|, \ 1 \le i \le n,$$

where Md denotes the usual sample median, as well as the sample standard deviation and the sample range estimators, defined by

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \qquad and \qquad R = X_{n:n} - X_{1:n}.$$

To describe the data we have considered a global set of symmetric and asymmetric distributions (F) related to the normal, with different levels of skew (γ_F),

$$\gamma_F = \frac{\mu_3}{\mu_2^{3/2}}$$

and with different tail-weight (τ_F),

$$\tau_F = \frac{1}{2}\left(\frac{F^{-1}(0.99)-F^{-1}(0.5)}{F^{-1}(0.75)-F^{-1}(0.5)} + \frac{F^{-1}(0.5)-F^{-1}(0.01)}{F^{-1}(0.5)-F^{-1}(0.25)}\right)\left(\frac{\Phi^{-1}(0.99)-\Phi^{-1}(0.5)}{\Phi^{-1}(0.75)-\Phi^{-1}(0.5)}\right)^{-1},$$

where μ_r denotes the r-th central moment of F, and F^{-1} and Φ^{-1} denote the inverse functions of F and of the standard normal distribution function Φ, respectively. For the considered distributions we have got values γ_F between 0 and 33.5, and values τ_F between 1 and 3.5, and the set includes the standard normal, the lognormal, the logistic, the chi-square, the Student-t, the gamma, the Weibull and some contaminated normal distributions.

To estimate the standard deviation we have standardized the above mentioned statistics, through the divison by a scale constant, so that the statistics have a unit mean value whenever the underlying model is normal. We have compared their efficiency, and the most efficient estimator among the ones considered is the one with the smallest value of $Var(\log W)$. To select a robust estimator in a global set of possible estimators we have applied the $MaxMin$ criterion. Then, we have implemented upper control charts to monitor the standard deviation of the process, and we have evaluated the alarm rates of these charts for rational subgroups of size $n = 5, 10, 15$ and 20. Details can be found in Figueiredo (2002) and Figueiredo and Gomes (2003b).

4. Conclusions

The main conclusions of this study can be summarized as follows. The total range TR is an efficient estimator of the standard deviation of an asymmetric distribution for sample sizes smaller than 15, and turns out to be the most robust estimator (among the ones considered) for sample sizes smaller than 25. For symmetric contaminated distributions S^* and S^{**} are the most efficient and robust estimators for small-to-moderate sample sizes (up to 17), being the TR estimator preferable only for sample sizes $n \leq 5$. The sample standard deviation is more efficient than the others only for symmetric distributions very close to the normal, such as, the logistic and the Student-t distribution, although it is not at all robust. Finally, the median absolute deviation from the median is the most efficient and robust estimator among the ones considered only for large sample sizes.

For the detection of an increase in the standard deviation of the process we may say that the MAD-chart is quite "robust", in the sense of resistant to changes in the model, not a long way from the S^{**}-chart when the rational subgroup size increases. The TR-chart does not exhibit a "robust" behaviour to departures from the normal model. Anyway, all these charts are preferable to the classical S and R charts in terms of robustness, and they also present a reasonable performance.

References

[1] Figueiredo, F. (2002). *Controlo Estatístico da Qualidade e Métodos Robustos. Ph. D. Thesis.* Faculdade de Ciências de Lisboa.

[2] Figueiredo, F. and Gomes, M. I. (2003a). *The total median in Statistical Quality Control.* Notas e Comunicações CEAUL 6/2003. Submitted.

[3] Figueiredo, F. and Gomes, M. I. (2003b). *Box-Cox transformations versus data modelling — robust charts in Statistical Process Control.* Notas e Comunicações CEAUL 13/2003. Submitted.

[4] Lax, D. A. (1985). Robust Estimators of Scale: Finite-sample Performance in Long-tailed Symmetric Distributions. *JASA-Theory and Methods,* vol.80, n.391, pp. 736-741.

[5] Tatum, L. G. (1997). Robust Estimation of the Process Standard Deviation for Control Charts. *Technometrics*, vol.39, n.2, pp.127-141).

LINEAR COMBINATIONS AND BIAS REDUCTION IN THE ESTIMATION OF HEAVY TAILS[6]

M. Ivette Gomes
C.E.A.U.L. e D.E.I.O., Universidade de Lisboa, Portugal

Abstract: We are here essentially interested in the estimation of a positive tail index γ through "asymptotically best linear unbiased" combinations of Hill's estimators and of scaled log-spacings. We shall additionally consider a weighted log-excesses' estimator of γ. The main objective of all these estimators is the accomodation of the dominant component of asymptotic bias. Apart from the derivation of the asymptotic distributional properties of the estimators, we shall study their distributional properties for finite samples, though Monte Carlo techniques.

Key words and phrases: Statistics of Extremes; Heavy tails; Best Linear Unbiased (BLUE) estimation; Log-excesses, Scaled log-spacings.

1. Introduction and preliminaries

F is a heavy tail model whenever the *tail function*, $\overline{F} = 1 - F$, is a regularly varying function with a negative index of regular variation $\alpha = -1/\gamma$, where γ is the *tail index*. Equivalently, if we consider $U(t) = F^{\leftarrow}(1 - 1/t)$, $t \geq 1$, $F^{\leftarrow}$ the generalized inverse function of F,

[6]Joint work with Frederico Caeiro, Fernanda Figueiredo, Laurens de Haan and Dinis Pestana, partially supported by FCT/POCTI/FEDER.

and with the usual notation RV_α for the class of regularly varying functions with index of regularly variation α,

$$F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma) \quad \text{iff} \quad 1-F \in RV_{-1/\gamma} \quad \text{iff} \quad U \in RV_\gamma. \tag{1}$$

The second order parameter ρ (≤ 0) is the non-positive value which appears in the limiting relation

$$\lim_{t\to\infty} \left\{\ln U(tx) - \ln U(t) - \gamma \ln x\right\}/A(t) = \left(x^\rho - 1\right)/\rho, \tag{2}$$

which we assume to hold for every $x > 0$, and where $|A| \in RV_\rho$, $\rho \leq 0$.

In the general theory of Statistics, whenever we ask the question whether the combination of information may improve the performance of an estimator, we are led to think on *Best Linear Unbiased Estimators* ($BLUE$), i.e., on unbiased linear combinations of an adequate set of statistics, with minimum variance among the class of such linear combinations. The basic theorem underlying this theory is due to Aitken (1935): *If* $\mathbf{X}$ *is a vector of observations with mean values* $\mathbb{E}\mathbf{X} = \mathbf{A}\theta$ *depending linearly on the unknown vector of parameters* θ*, with a known coefficient matrix* $\mathbf{A}$*, and with a covariance matrix* $\delta^2\Sigma$*, known up to a scale factor* δ^2*, the least-squares estimator of* θ *is the vector* θ^* *which minimizes the quadratic form* $(\mathbf{X} - \mathbf{A}\ \theta)'\,\mathbf{\Sigma^{-1}}\,(\mathbf{X} - \mathbf{A}\ \theta)$*. Such a vector is thus the vector of solutions of the "normal equations",* $\mathbf{A'\Sigma^{-1}A}\ \theta^* = \mathbf{A'\Sigma^{-1}X}$*. This solution is explicitly given by* $\theta^* = \left(\mathbf{A'\Sigma^{-1}A}\right)^{-1}\mathbf{A'\Sigma^{-1}X}$*, and its variance matrix is* $\mathbf{Var}(\theta^*) = \delta^{\mathbf{2}}\left(\mathbf{A'\Sigma^{-1}A}\right)^{\mathbf{-1}}$.

If we think on "asymptotically best linear unbiased" combinations of Hill's estimators (Hill, 1975), given by

$$H(i) := \tfrac{1}{i}\sum_{j=1}^{i}\left\{\ln X_{n-j+1:n} - \ln X_{n-i:n}\right\},\ 1 \leq i \leq k,$$

where $X_{i:n}$ denotes, as usual, the i-th ascending order statistic associated to the random sample X_i, $1 \leq i \leq n$, we shall come to the estimator

$$BL_{H}^{(\widehat{\rho})}(k) := \sum_{i=1}^{k} a_i^H(\widehat{\rho})\ H(i), \tag{3}$$

where $\widehat{\rho}$ is any adequate estimator of the second order parameter ρ in (2), and for $1 \le i \le k-1$,

$$\begin{aligned} a_i^H(\rho) &= \frac{k^{2\rho}}{k(P_{kk}-k)} \left(-i(i-1)S_{i-1} + 2i^2 S_i - i(i+1)S_{i+1}\right), \\ a_k^H(\rho) &= \frac{k^{2\rho}}{k(P_{kk}-k)} \left(-k(k-1)S_{k-1} + k^2 S_k\right), \end{aligned}$$

with

$$S_i = \sum_{j=1}^{k-1} j \left(j^{-\rho} - i^{-\rho}\right)\left(2\, j^{1-\rho} - (j-1)^{1-\rho} - (j+1)^{1-\rho}\right) \\ + k\left(k^{-\rho} - i^{-\rho}\right)\left(k^{1-\rho} - (k-1)^{1-\rho}\right),\ 1 \le i \le k.$$

Since these formulas are a bit heavy, computationally, we have decided to consider the scaled log-spacings $U_i := i\ \{\ln X_{n-i+1:n} - \ln X_{n-i:n}\}$, $1 \le i \le k$, as the basic statistics. We then come to the weights

$$a_i^U(\rho) = \frac{\sum_{j=1}^k (j/k)^{-2\rho} - (i/k)^{-\rho}\ \sum_{j=1}^k (j/k)^{-\rho}}{k\ \sum_{j=1}^k (j/k)^{-2\rho} - \left(\sum_{j=1}^k (j/k)^{-\rho}\right)^2},\ 1 \le i \le k,$$

and to the linear estimator

$$BL_U^{(\widehat{\rho})}(k) := \sum_{i=1}^k a_i^U(\widehat{\rho})\ U_i. \tag{4}$$

The same type of derivation based on the log-excesses

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n},\ 1 \le i \le k, \tag{5}$$

was also considered in Gomes and Figueiredo (2003) and led us to terribly high computer time consuming estimators.

We are here going to consider also an alternative way to accommodate bias in the log-excesses, a by-product for heavy tails, from a possible accomodation of bias in the excesses, for a general tail index γ, developed together with Laurens de Haan. Such an accomodation is similar to the one used to accomodate bias in the scaled log-spacings. Under the first order framework in (1), and denoting Y and E standard Pareto and exponential random variables (rv's), respectively,

$$V_{ik} \stackrel{d}{=} \ln U(Y_{n-i+1:n}) - \ln U(Y_{n-k:n}) \;\sim\; \gamma \;\ln Y_{k-i+1:k} \stackrel{d}{=} \gamma \; E_{k-i+1:k},$$

i.e., the V_{ik}'s, $1 \le i \le k$, are, approximately, the k o.s. from an exponential random sample with mean value γ. Under the second order framework in (2), we may say that for intermediate k, i.e., if $k = k_n \to \infty$, $k = o(n)$ as $n \to \infty$, and for $1 \le i \le k$,

$$V_{ik} - \gamma \; e^{\frac{A(n/k)}{\gamma} \frac{Y_{k-i+1:k}^{\rho}-1}{\rho \ln Y_{k-i+1:k}}} \; E_{k-i+1:k} = o_p \left(V_{ik} - \gamma \; E_{k-i+1:k}\right),$$

and since

$$\frac{Y_{k-i+1:k}^{\rho} - 1}{\rho \ln Y_{k-i+1:k}} \sim -\frac{(i/k)^{-\rho} - 1}{\rho \ln(i/k)} =: \psi_i \equiv \psi_{ik}(\rho) \; [\psi_k \equiv 1],$$

with ψ_i a limited function, we expect to get a less biased estimator if we assume that the random log-excess V_{ik} comes from an exponential model with mean value not equal to γ, as it is done to give rise to the Hill estimator, but dependent on i (and k), more specifically given by

$$\gamma_i = \gamma \; e^{A^*(n/k)\psi_i} = \gamma \; e^{\beta \; \left(\frac{n}{k}\right)^{\rho} \psi_i}, \quad A^*(t) = \beta \; t^{\rho} = A(t)/\gamma, \; 1 \le i \le k,$$

assuming thus that we are in Hall's class of models (Hall and Welsh, 1985), with $A(t) = D \; t^{\rho}$, $\rho < 0$, $\beta = D/\gamma$.

We shall thus consider the following weighted combination of the log-excesses,

$$WLE(k) \equiv WLE_{\widehat{\beta},\widehat{\rho}}(k) := \frac{1}{k} \sum_{i=1}^{k} e^{-\widehat{\beta} \; (n/k)^{\widehat{\rho}} \; \widehat{\psi}_i} \; V_{ik}, \tag{6}$$

where $\widehat{\beta}$ and $\widehat{\rho}$ are adequate estimators of the second order parameters β and ρ, respectively, to be estimated externally. Such a decision is related to the discussion in Gomes and Martins (2002b) on the advantages of an external estimation of the second order parameters versus an internal estimation at the same level k, as done in Beirlant et al. (1999) and Feuerverger and Hall (1999). In a joint work with F. Caeiro and D. Pestana we are now also dealing with the direct removal of bias to Hill's estimateor, in the same spirit as before.

2. Motivation for the WLE estimator

Let us assume everything is known, apart from γ. Then,

Theorem 1. *Under the second order framework in (2), and for intermediate levels k, we get,*

$$WLE_{\beta,\rho}(k) \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}} N_k + o_p(A(n/k)), \tag{7}$$

with N_k an asymptotically standard normal rv's. Consequently $\sqrt{k}\,(WH(k) - \gamma)$ is asymptotically normal with null mean value not only when $\sqrt{k}\ A(n/k) \longrightarrow 0$, but also when $\sqrt{k}\ A(n/k) \longrightarrow \lambda \neq 0$, finite, as $n \to \infty$. More than that: the asymptotic variance of this r.v. is equal to the asymptotic variance of the Hill estimator.

Proof: The rv $WLE_{\beta,\rho}(k)$ may be written as

$$\frac{\gamma}{k} \sum_{i=1}^{k} E_{k-i+1:k} \left(1 + \frac{A(n/k)}{\gamma}\ \frac{Y_{k-i+1:k}^{\rho} - 1}{\rho\, E_{k-i+1:k}} (1 + o_p(1))\right) \\ \times \left(1 - \frac{A(n/k)}{\gamma} \psi_i (1 + o(1))\right).$$

Consequently, with $R_k := \frac{1}{k}\sum_{i=1}^{k} \frac{Y_{k-i+1:k}^{\rho} - 1}{\rho} - \frac{1}{k}\sum_{i=1}^{k} \psi_i\ E_{k-i+1:k} =: Q_k^{(1)} - P_k^{(1)}$ we have $WH(k) = \frac{\gamma}{k}\sum_{i=1}^{k} E_i + A(n/k)\ R_k\ (1 + o_p(1))$. The weak law of large numbers enables to say that both $Q_k^{(1)}$ and $P_k^{(1)}$ converge in probability towards their mean values, and we have $E[Q_k^{(1)}] = E[(Y^{\rho} - 1)/\rho] = 1/(1 - \rho) = E[P_k^{(1)}]$. Consequently R_k converges in probability towards 0, as $k \to \infty$, and, with $N_k = \sqrt{k}\left(\sum\limits_{i=1}^{k} E_i/k - 1\right)$, (7) follows, i.e., the usual dominant component of bias, which is for the classical estimators of the tail index of the order of $A(n/k)$ is now of smaller order. □

We may also state the following:

Corollary 1. *Under the conditions of Theorem 1 the same distributional representation (7) holds true if we consider the tail index estimator $WH_{\widehat{\beta},\widehat{\rho}}(k)$ for any estimators $\widehat{\beta}$ and $\widehat{\rho}$ such that both $\widehat{\beta} - \beta$ and $\widehat{\rho} - \rho$ are $o_p(1)$, for the k-values on which we base the estimation of γ.*

As an illustration we present Figure 1. Notice particularly the pattern of mean square error of the new estimator, which is below the one of the Hill estimator for every k.

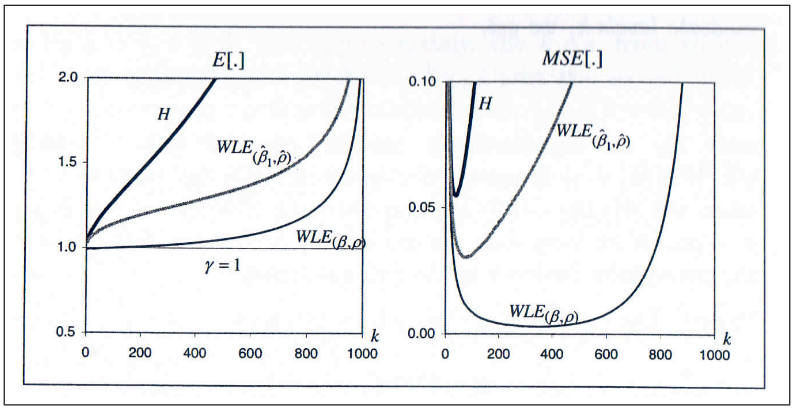

Figure 1: Mean values and mean Square Errors of WLE-rv and estimator for a sample size $n = 1000$, from a Burr parent with $\rho = -0.5(\gamma = 1)$.

References

[1] Aitken, A. C. (1935). On least squares and linear combinations of observations. *Proc. Roy. Soc. Edin.* **55**, 42-48.

[2] Beirlant, J., Dierckx, G., Goegebeur, Y. and G. Matthys (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177-200.

[3] Feuerverger, A. and P. Hall (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.

[4] Gomes, M.I. and Martins, M.J. (2002b). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5**:1, 5-31.

[5] Hall, P. and A.H. Welsh (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13**, 331-341.

[6] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.

ESTIMATION OF THE EXTREME VALUE INDEX AND HIGH QUANTILES UNDER RANDOM CENSORING

Jan Beirlant
Department of Mathematics
Katholieke Universiteit Leuven
Belgium

Emmanuel Delafosse
Department of Mathematics
Katholieke Universiteit Leuven
Belgium

Armelle Guillou
Laboratoire de Statistique Théorique et Appliquée
Université Pierre et Marie Curie
France

In this paper we consider the estimation problem of the extreme value index and extreme quantiles in the presence of censoring. Taking into account the fact that our main motivation is application in insurance, we focus on Fréchet and Gumbel domains of attraction. In the case of no-censoring, the most famous estimator of the Pareto index is the classical Hill estimator (1975). Some adaptations of this estimator in the case of censoring are proposed and used to build extreme quantile estimators. A theoretical study of the asymptotic properties of such estimators is started. The finite sample behaviour is illustrated in a small simulation study and also in a practical insurance example.

References

[1] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.

SPATIAL EXTREMES: THE STATIONARY CASE[7]

Laurens de Haan
Econometric Institute
Universiteit Rotterdam
The Netherlands

Teresa Themido Pereira
CEAUL e DEIO (FCUL)
Universidade de Lisboa
Portugal

The aim of the paper is to provide models for spatial extremes in the case of stationarity. The spatial dependence at extreme levels of a stationary process is modelled using the theory of max-stable processes of de Haan and Pickands (1986). We propose three models. These models depend on just one parameter that measures the strength of tail dependence as a function of the distance between locations, although in one case a model depending on three parameters is considered. We propose an estimator for this parameter, prove consistency under domain of attraction conditions and asymptotic normality under appropriate further conditions.

References

[1] Haan, L. de and Pickands, J. (1986). Stationary min-stable stochastic processes. *Probab. Th. Rel. Fields.* **72**, 477-492.

[7]Research partially supported by FCT/POCTI/FEDER.

EXTREMES OF GENERALIZED PERIODIC SUB-SAMPLED MOVING AVERAGE SEQUENCES[8]

Andreia Hall
Departamento de Matemática
Universidade de Aveiro
Portugal

Let $\{X_k\}$ be a stationary moving average sequence of the form $X_k = \sum_{j=-\infty}^{\infty} \beta_j * Z_{k-j}$ where $\{Z_k\}$ is and iid sequence of random variables with regularly varying tails, and the operator $*$ denotes multiplication if Z_k is continuous and binomial thinning if Z_k is a non-negative integer-valued. Let $g(k) : \mathbb{N}_0 \to \mathbb{N}_0$ be a strictly increasing sequence with a periodic pattern of the form $g(k+I) = g(k) + M$ for some fixed integers I and M verifying $1 \leq I \leq M$. Define $Y_k = X_{g(k)}$ as the generalised periodic sub-sampled sequence. In this work we look at the extremal properties of $\{Y_k\}$. In particular, we investigate the limiting distribution of the sample maxima and the corresponding extremal index. Motivation comes from the comparison of schemes for monitoring a variety of medical, finance, environmental, and social science data sets.

[8]Joint work with Manuel G. Scotto and Helena Ferreira, partially supported by FCT/POCTI/FEDER.

EXTREMES OF PERIODIC MOVING AVERAGES OF RANDOM VARIABLES WITH REGULARLY VARYING TAIL PROBABILITIES[9]

Ana Paula Martins
Departamento de Matemática
Universidade da Beira Interior
Portugal

Helena Ferreira
Departamento de Matemática
Universidade da Beira Interior
Portugal

Let $\mathbf{Z} = \{Z_n\}_{n\geq 1}$ be a T-periodic sequence of independent real-valued variables, such that $\overline{F}_i(x) = P\left(|Z_i| > x\right),\ i = 1, 2, ..., T$, are regularly varying with exponent $-\alpha$, i.e.,

$$\overline{F}_i(x) = x^{-\alpha}\ L_i(x), \quad x > 0, \quad i = 1, 2, ..., T,$$

for some $\alpha > 0$ and $L_i : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ slowly varying functions, satisfying the tail balance conditions

$$\lim_{x\to\infty} \frac{P(Z_i > x)}{\overline{F}_i(x)} = p_i, \quad \lim_{x\to\infty} \frac{P(Z_i < -x)}{\overline{F}_i(x)} = q_i, \quad i = 1, 2, ..., T,$$

for some p_i and $q_i \in [0, 1]$ such that $p_i + q_i = 1,\ i = 1, 2, ..., T$, and tail equivalence in the following way: for every $i, j = 1, 2, ..., T$,

$$\lim_{x\to\infty} \frac{P(Z_i > x)}{P(Z_j > x)} = \gamma_{i,j}^{(+)} > 0,\ \lim_{x\to\infty} \frac{P(Z_i < -x)}{P(Z_j < -x)} = \gamma_{i,j}^{(-)} > 0.$$

We study the extreme value behaviour of the T-periodic moving average $\mathbf{X} = \{X_n\}_{n\geq 1}$ of the form

$$X_n = \sum_{j=-\infty}^{\infty} c_j\ Z_{n-j}, \quad n \geq 1,$$

[9]Research partially supported by FCT/POCTI/FEDER.

where $\mathbf{c} = \{c_j\}_{j\in\mathbb{Z}}$ is a sequence of constants such that $\sum_{j=-\infty}^{\infty} |c_j|^{\delta} < \infty$, for some $\delta < \min\{\alpha, 1\}$.

We define a family of local mixing conditions $D_T^{(k)}(\mathbf{u}),\ k \geq 1$, that enable the computation of the extremal index of periodic sequences from the joint distributions of k consecutive variables. By applying results, under local and global mixing conditions, to the $(2m-1)$-dependent periodic sequence $X_n^{(m)} = \sum_{j=-m}^{m-1} c_j Z_{n-j},\ n \geq 1$, we compute the extremal index of the periodic moving average $\mathbf{X}$.

This presentation generalizes the theory for extremes of stationary moving averages with regularly varying tail probabilities.

References

[1] Alpuim, M. T. (1988). *Contribuies Teoria de Extremos em Sucesses Dependentes.* Ph. D. Thesis, DEIOC, University of Lisbon.

[2] Chernick, M. R., Hsing, T. and McCormick, W. (1991). Calculating the extremal index for a class of stationary sequences. *Adv. Appl. Probab.*, 23, 845-850.

[3] Cline, D. (1983). Infinite series of random variables with regularly varying tails. *Tech. Report, Univ. British Columbia*, 83-24.

[4] Davis, R. and Resnick, S. I. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.*, 13, 179-195.

[5] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). Modelling Extremal Events. Springer-Verlag.

[6] Ferreira, H. (1994). Multivariate extreme values in T-periodic random sequences under mild oscilation restrictions. Stochastic Process. Appl., 14, 111-125.

[7] Ferreira, H. and Martins, A. P. (2003). The extremal index of sub-sampled periodic sequences with strong local dependence. To appear in Revstat.

[8] Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). *Extremes and related properties of random sequences and series.* New York: Springer Verlag.

[9] Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* New York: Springer-Verlag.

REVISITING THE TRILEMMA PROBLEM FOR MAX-DOMAINS OF ATTRACTION[10]

Cláudia Neves
Departamento de Matemática
Universidade de Aveiro
Portugal

The order statistics (o.s.) $X_{1:n} \leq ... \leq X_{n:n}$ are the legacy of independent random variables $X_1, ..., X_n$ with common distribution function (d.f.) F, after arranging these in nondecreasing order. Due to their nature, semi-parametric models are never specified in detail by hand. Instead, the only assumption made is that, for some index $\gamma \in \mathbb{R}$, F is in the domain of attraction of an extreme value distribution G_γ (notation: $F \in \mathcal{D}(\mathcal{G}_\gamma)$), i.e.,

$$\begin{aligned} \exists_{b_n \in \mathbb{R}}^{a_n > 0} \quad : \quad & \lim_{n \to \infty} F^n(a_n x + b_n) = G_\gamma(x) \\ := \quad & \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}), & 1 + \gamma x > 0 \quad \text{if} \quad \gamma \neq 0 \\ \exp(-e^{-x}), & x \in \mathbb{R} \quad \text{if} \quad \gamma = 0 \end{cases}, \end{aligned}$$

for all x. Hence, G_γ denotes the universally known Generalized Extreme Value (GEV(γ)) distribution in the Von Mises parametrization. Gnedenko (1943) established that the class $\{G_\gamma\}_{\gamma \in \mathbb{R}}$ represents, in an unified version, all possible non-degenerate weak limits of the maximum $X_{n:n}$, up to location/scale parameters. For $\gamma < 0$, $\gamma = 0$ and $\gamma > 0$, GEV(γ) d.f. reduces to Weibull, Gumbel and Fréchet distributions, respectively.

[10] Joint work with Jan Picek and Isabel Fraga Alves, partially supported by FCT / POCTI / FEDER.

The following necessary and sufficient condition for $F \in \mathcal{D}(G_\gamma)$ was established by de Haan (1984) (*first order extended regular variation property*):

$$\lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = D_\gamma := \int_1^x y^{\gamma-1}dy = \begin{cases} \frac{x^\gamma-1}{\gamma}, & \gamma \neq 0 \\ \log x, & \gamma = 0 \end{cases}, \qquad (1)$$

for every $x > 0$ and some positive mensurable function a, with $U(t) := \big(1/(1-F)\big)^{\leftarrow}(t)$ standing for the tail quantile function (q.f.) pertaining to F. The arrow indicates the generalized inverse function, meaning that $U(t) := \inf\{x : (1/(1-F))(x) \geq t\}$.

Observe that the limit function $(x^\gamma - 1)/\gamma$ is the tail q.f. of the Generalized Pareto (GP) distribution

$$\begin{aligned} W_\gamma(x) &:= 1 + \log G_\gamma(x) \\ &= 1 - (1+\gamma x)^{-\frac{1}{\gamma}} \quad \text{for} \begin{cases} x \geq 0 & \text{if} \quad \gamma \geq 0 \\ 0 \leq x \leq -\frac{1}{\gamma} & \text{if} \quad \gamma < 0 \end{cases}. \end{aligned}$$

This fact reflects its exceptional role in Extreme Value Theory (cf. Pickands(1975), Balkema and de Haan (1974)) and appels to the appropriateness of classifying the tails of all possible distributions in $\mathcal{D}(G_\gamma)$ into three classes, discriminated by the tail index sign.

For positive γ, the power-law behavior in the tail of the underlying distribution F has important implications since it may suggest, for instance, the presence of infinite moments. Because the first order condition (1) can be reformulated as $\lim_{t\to\infty} U(tx)/U(t) = x^\gamma$, for all $x > 0$, that is, U is γ- regular varying at infinity (notation: $U \in RV_\gamma$), Karamata's Theorem for integration of regularly varying functions yields that $E(X_1^+)^p$ is infinite for $p > 1/\gamma$, where $X_1^+ = \max(0, X_1)$. This class of heavy tailed distributions includes Pareto, Cauchy and Student's t distributions.

A relevant fact is that all d.f.'s belonging to $\mathcal{D}(G_\gamma)$ with $\gamma < 0$ — Weibull domain of attraction — have finite endpoint. Such domain of attraction encloses Uniform and Beta distributions.

The intermediate case $\gamma = 0$ is of particular interest in many applied sciences where extremes are relevant, not only because of greater simplicity of inference within the Gumbel domain G_0 but also for the great variety of distributions possessing an exponential tail whether having infinite endpoint or not. Distributions ranging from moderately heavy tailed (such as Lognormal) to light tailed (such as Normal or Gamma) can be found in Gumbel class.

Taking all into consideration, it has become clear the advantage of looking for the most propitious type of tail when fitting empirical distributions at high quantiles. Effectively, separating statistical inference procedures according the most suitable domain of attraction for the underlying d.f. F has become an usual practice.

A test for Gumbel domain against Fréchet or Weibull max-domain has received the general designation of statistical choice of extreme domains of attraction (see e.g. Castillo et al. (1989), Hasofer and Wang (1992), Fraga Alves and Gomes (1996) and Marohn (1998)). Among these, Hasofer and Wang's may be pointed out as one of the most commonly used testing procedure, in particular, Reiss and Thomas (2001, p.154) have incorporated it in "XTREMES" software. This test is based on a *location/scale* invariant statistic, function of the excesses over a random threshold $X_{n-k:n}$. The asymptotic statements of the referred authors settle on a fixed k whereas n goes to infinity, bearing on results due to Weissman(1978). Nevertheless, in the last part of the paper there is an attempt to extend the range of the test by allowing k to increase with the sample size n.

Pursuing the same objective, Segers and Teugels (2001) have recently suggested a large sample test for the Gumbel domain hypothesis. After deriving the asymptotic distribution of Galton's ratio, provided condition (1), the authors appealed for Rao's (see Serfling (1980)) test statistic for simple null hypothesis in order to establish a decision rule. In the process, they were confronted with the need of blocking the original sample into k subsamples, each of size $n_i, i = 1, \cdots, k$ also under pledge of largeness.

The present work deals with the two-sided problem of testing Gumbel domain against Fréchet or Weibull domains, i.e., $F \in \mathcal{D}(G_0)$ versus $F \in \mathcal{D}(G_\gamma)_{\gamma \neq 0}$. Considering k upper order statistics in a way that these might present a satisfactory picture of the tail of F, we introduce a new test statistic which is simply the ratio between the maximum and the mean of the excesses above a random threshold $X_{n-k:n}$:

$$T_{k,n} := \frac{X_{n:n} - X_{n-k:n}}{\frac{1}{k}\sum_{i=1}^{k}(X_{n-i+1:n} - X_{n-k:n})}. \tag{2}$$

where $k = k_n$ is an intermediate sequence of positive integers, meaning that $k_n \to \infty$ and $k_n/n \to 0$ as the sample size n tends to infinity.

The exact distribution $T_n(k)$ does not depend on location or scale parameters whereas its discriminant asymptotic behavior towards heavy or light tailed distributions proves to be basically governed by the sample maximum.

The speed of convergence of the partial maxima to its limit distribution is controlled by a *second order extended regular variation property* (cf. de Haan and Stadtmüller (1996)):

$$\begin{aligned}\lim_{t\to\infty}\frac{\frac{U(tx)-U(t)}{a(t)}-\frac{x^\gamma-1}{\gamma}}{A(t)} &= H_{\gamma,\rho}(x) := \int_1^x y^{\gamma-1}\int_1^y u^{\rho-1}du dy\\ &= \frac{1}{\rho}\left\{\frac{x^{\gamma+\rho}-1}{\gamma+\rho}-\frac{x^\gamma-1}{\gamma}\right\} \qquad (3)\end{aligned}$$

for all $x>0$ with ρ the non-positive second order parameter, $a>0$ and A a suitable positive or negative function. As before, for $\gamma=0$ and/or $\rho=0$, function $H_{\gamma,\rho}$ is understood in a limiting sense as $\log x$. The function $|A|$, which is regularly varying of order ρ and tends to zero, represents the speed of convergence. Condition (3) will be needed to obtain the asymptotic distribution of $T_{k,n}$ (after suitable normalization to avoid degeneracy), if the sampled distribution belongs to Gumbel domain of attraction.

Theorem 1. *Suppose $F\in\mathcal{D}(G_0)$ and that (3) holds for $\gamma=0$. Let $k=k_n$ be an intermediate sequence of integers such that $A(\frac{n}{k})\log^2 k\to 0$, as $n\to\infty$,*

$$T^*_{k,n} := T_{k,n}-\log k \xrightarrow{d} G, \quad G\sim \textit{Gumbel}.$$

The consistency of the test relies on the following theorem:

Theorem 2. *Suppose $F\in\mathcal{D}(\mathcal{G}_\gamma)$, meaning that condition (1) holds for some $\gamma\in\mathbb{R}$. Let $k=k_n$ be an intermediate sequence of integers such that $k_n\to\infty$ and $k_n/n\to 0$ as $n\to\infty$. Then, as $n\to\infty$,*

(i) if $\gamma<0$, $T^*_{k,n}\xrightarrow{P}-\infty$;

(ii) if $\gamma>0$, $T^*_{k,n}\xrightarrow{P}+\infty$.

Thus, one-sided testing problems $F \in \mathcal{D}(G_0)$ versus $F \in \mathcal{D}(\mathcal{G}_\gamma)_{\gamma<\prime}$ (or $F \in \mathcal{D}(\mathcal{G}_\gamma)_{\gamma>\prime}$) may also be treated by our results.

The normalized version $T^*_{k,n}$ is eventually our test statistic. According to the asymptotic results stated in Theorems 1 and 2, the null hypothesis $F \in \mathcal{D}(G_0)$ is to be rejected, against the bilateral alternative $F \in \mathcal{D}(\mathcal{G}_\gamma)_{\gamma\neq\prime}$, at an asymptotic significance level $\alpha \in (0,1)$ if $T^*_{k,n} < g_{\alpha/2}$ or $T^*_{k,n} > g_{1-\alpha/2}$, where g_ε denotes the ε-quantile of the Gumbel distribution, i.e., $g_\varepsilon = -\log(-\log\varepsilon)$. On the other hand, we reject the null hypothesis in favor of either unilateral alternatives $H_1^{'} : F \in \mathcal{D}(\mathcal{G}_\gamma)_{\gamma<\prime}$ or $H_1^{''} : F \in \mathcal{D}(\mathcal{G}_\gamma)_{\gamma>\prime}$ on either side $T^*_{k,n} < g_\alpha$ or $T^*_{k,n} > g_{1-\alpha}$, respectively.

On the light of a simulation study, Type I error probability and Power of the test have been evaluated. In addition, comparisons with the test introduced by Hasofer and Wang and the recently proposed test by Segers and Teugels have been carried out. As an example, Figure 1 displays the simulated power functions of T^*_k, Hasofer and Wang (W_k) and Segers and Teugels (S_m) tests, for a number $k = 150$ of top observations from a sample of size $n = 1000$, taken from GP(γ) model.

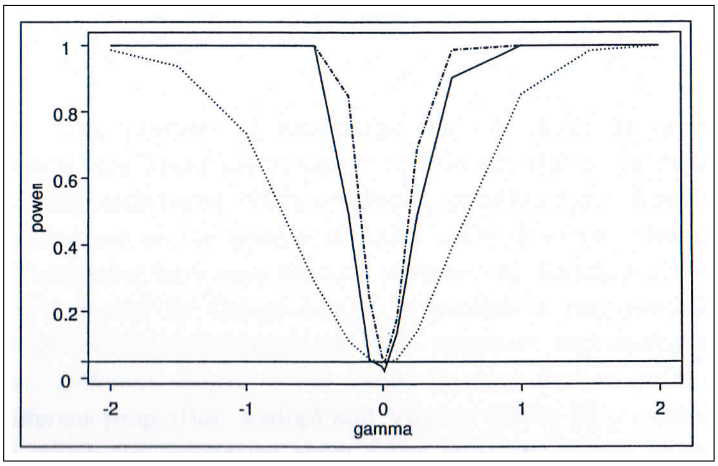

Figure 1: Power of T^*_k (solid), W_k (dashed) and S_{50} (dotted) tests as a function of γ, with $k = 150$ for GP(γ) samples of size $n = 1000$.

References

[1] Neves, C., Picek, J., Fraga Alves, M.I. (2003). A test for Gumbel domain of attraction — ratio of maximum to the mean of excesses. *Notas e Comunicações CEAUL, Faculty of Sciences, University of Lisbon*, **14**.

KERNEL ESTIMATORS OF THE TAIL INDEX[11]

M. João Martins
Departamento de Matemática, ISA
Universidade Técnica de Lisboa
Portugal

M. Manuela Neves
Departamento de Matemática, ISA
Universidade Técnica de Lisboa
Portugal

M. Ivette Gomes
CEAUL e DEIO (FCUL)
Universidade de Lisboa
Portugal

Abstract: The problem of estimating the tail index in heavy tailed-distributions is very important in many applications. One of the most popular semi-parametric estimators, Hill's estimator, presents some drawbacks such as high dependence on the number of upper order statistics. Generalizations of the Hill estimator have been recently proposed by several authors. Kernel estimators, proposed by Csörgó *et.al.* [1] provide a very general class of tail index estimators that contains Hill's estimator and averages of Hill's estimators. Different choices of the kernel function lead to estimators that present different properties. Csörgó and Viharos (1998) [2] proposed a kernel family for which the estimators show a low volatility or/and an increase in efficiency. However these estimators cannot be compared directly to Hill's estimator because at the level k, they use more than k order statistics. We herewith present alternative estimators, derived for different choices of kernel functions, that can be compared to Hill's estimator. Their properties

[11] Research partially supported by FCT/POCTI/FEDER.

are studied and comparisons are performed.

Key words and phrases: Kernel estimators, simulation, tail index, bias, volatility, mean squared error.

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be positive, independent and identically distributed (i.i.d.) random variables (r.v.'s), with common distribution function (d.f.) F. We assume that F has a heavy upper tail, i.e., $1-F$ is regularly varying at infinity of order $-1/\gamma$, that means

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma} \qquad x>0 \quad \text{and} \quad \gamma>0. \tag{1}$$

The tail index γ, directly related to the tail weight of the model F, is a central parameter in *Statistical Extremes.* Kernel estimators, proposed by Csörgó *et.al.* [1], are derived from a class of kernel functions. Let $\mathcal{K}$ be a class of non-negative, non-increasing, left-continuous functions $K(.)$ defined on $(0,\infty)$ such that

- $\int_0^\infty K(t)dt = 1, \qquad \int_0^\infty t^{-1/2}K(t)dt < \infty$ and
- there exists $\Lambda_K \in (0,\infty)$ for which $K(t) = 0$ if $t > \Lambda_K$.

Csörgó *et.al.* [1] considered the kernel estimator defined as

$$\hat{\gamma}_n^{(K)}(k) := \frac{\sum_{j=1}^n \frac{j}{k} K\left(\frac{j}{k}\right)[\ln X_{j:n} - \ln X_{j+1:n}]}{\frac{1}{k}\sum_{j=1}^n K\left(\frac{j}{k}\right)} \tag{2}$$

where $X_{1:n} \geq X_{2:n} \geq \cdots \geq X_{n:n}$ and $k \equiv k_n$ is an intermediate sequence, i.e.,

$$k_n \to \infty \quad \text{and} \quad k_n/n \to 0 \quad \text{as} \quad n \to \infty \tag{3}$$

It is easy to see that Hill's estimator is obtained for the choice $K(.) = K^{(H)}(.) := I_{(0,1]}(.)$ The estimator (2) can be written as

$$\hat{\gamma}_n^{(K)}(k) := \frac{\sum_{j=1}^{n} K\left(\frac{j}{k}\right) D_j}{\sum_{j=1}^{n} K\left(\frac{j}{k}\right)} \quad \text{with} \quad D_j = j[\ln X_{j:n} - \ln X_{j+1:n}], \tag{4}$$

i.e., the estimate $\hat{\gamma}_n^{(K)}(k)$ is the weighted average of spacings D_j.

Csörgó *et.al.* [1] showed that (2) converges in probability to γ whenever (1) and (3) hold and $K \in \mathcal{K}$.

For the special case when F belongs to the Hall's class, i.e, F may be written as $1 - F(x) = Cx^{-1/\gamma}\left(1 + Dx^{\rho/\gamma} + o\left(x^{\rho/\gamma}\right)\right),\ x \to \infty,\ C > 0,\ D \neq 0,\ \gamma > 0,\ \rho < 0$, Csörgó *et.al.* [1] obtained the "optimal" kernel function for which $MSE_\infty[\hat{\gamma}_n^{(K)}(k_0)]$ is minimized at the optimal level, k_0:

$$K_\rho^{(1)}(t) = \begin{cases} \left(\frac{\rho-1}{\rho}\right) \Lambda_{K_\rho^{(1)}}^{\rho-1} \left[\Lambda_{K_\rho^{(1)}}^{-\rho} - t^{-\rho}\right] & \text{if } \ 0 < t \leqslant \Lambda_{K_\rho^{(1)}} \\ 0 & \text{if } \ t > \Lambda_{K_\rho^{(1)}} = \frac{2-2\rho}{1-2\rho} > 1 \end{cases}. \tag{5}$$

Csörgó and Viharos (1998) [2] presented a kernel family parametrized in $\xi<0$:

$$K_\xi^{(2)}(t) = \begin{cases} \left(\frac{\xi-1}{\xi}\right) \Lambda_{K_\xi^{(2)}}^{\xi-1} \left[\Lambda_{K_\xi^{(2)}}^{-\xi} - t^{-\xi}\right] & \text{if } \ 0 < t \leqslant \Lambda_{K_\xi^{(2)}} \\ 0 & \text{if } \ t > \Lambda_{K_\xi^{(2)}} = \frac{2-2\xi}{1-2\xi} > 1 \end{cases}. \tag{6}$$

In this case $\lim_{\xi \to -\infty} K_\xi^{(2)}(t) = K^{(H)}(t) \quad \forall t \neq 1$.

For this family $Bias_\infty\left[\hat{\gamma}_n^{(2)}(k)\right]$ and $MSE_\infty\left[\hat{\gamma}_n^{(2)}(k_0^{(2)})\right]$ are minimized with $\xi = \rho$.

However, if one intends to perform comparisons to Hill's estimator, such is not possible directly because, for each k, $\gamma_n^{(2)}(k)$ needs more than k order statistics.

2. Alternative Kernel Estimators

Let us consider two alternative kernel functions. The first one is

$$K_\xi^{(3)}(t) = \begin{cases} \left(\frac{\xi-1}{\xi}\right)\left(1-t^{-\xi}\right) & \text{if } \ 0 < t \leqslant 1 \\ 0 & \text{if } \ t > 1 \end{cases} \qquad \Lambda_{K_\xi^{(3)}} = 1. \quad (7)$$

It can be seen that $\hat{\gamma}_n^{(3)}$ and $\hat{\gamma}_n^{(2)}$ are asymptotically equivalent at "optimal" levels, however $\hat{\gamma}_n^{(3)}(k)$ only uses k order statistics whatever ξ is, and has a higher, although slight, asymptotic efficiency relatively to Hill's estimator.

Another kernel function is

$$K_\xi^{(4)}(t) = \begin{cases} (1-\xi)\,(2-\xi)\,t^{-\xi}\,(1-t) & \text{if } \ 0 < t \leqslant 1 \\ 0 & \text{if } \ t > 1 \end{cases} \qquad \Lambda_{K_\xi^{(4)}} = 1. \quad (8)$$

The estimator $\hat{\gamma}_n^{(4)}$ that uses function (8) presents a minimum $MSE_\infty\left[\gamma_n^{(4)}(k_0^{(4)})\right]$ for $\xi = 0$, so the corresponding kernel function is

$$K_0^{(4)}(t) = \begin{cases} 2(1-t) & \text{if } \ 0 < t \leqslant 1 \\ 0 & \text{if } \ t > 1 \end{cases} .$$

The estimator $\hat{\gamma}_n^{(4)}\left(k_0^{(4)}\right)$ has an asymptotic relative efficiency better than Hill's estimator but slightly worse than that of $\hat{\gamma}_n^{(1)}\left(k_\rho^{(1)}\right)$. However, it does not depend on ρ.

References

[1] Csörgó, S., Deheuvels, P. and Mason, D.M. (1985). Kernel Estimates of the tail index of a distribution *Ann. Stat.*, **13**, 1050-1077.

[2] Csörgó, S. and Viharos L. (1998). Estimating the tail Index *In:* "Asymptotic Methods in Probability and Statistics (Szyszkowicz,ed.), North-Holland, Amsterdam.

[3] Hill, B.M. (1975). A simple approach to inference about the tail of a distribution. *Ann. Stat.*, **3**, 1163-1174.

SUMMING UP
OR
RANDOMIZATION – HOW FAR TO GO?[12]

Ftima Brilhante and **José Rocha**
Departamento de Matemática, Universidade doa Açores
CEAUL, Universidade de Lisboa, Portugal

Sandra Mendonça and **Rita Vasconcelos**
Departameto de Matemática, Universidade da Madeira
CEAUL, Universidade de Lisboa, Portugal

Dinis Pestana
CEAUL e DEIO (FCUL), Universidade de Lisboa
Portugal

Key words and phrases: Heteroscedasticity, studentized statistics, analysis of scale, integral transforms, randomization, $\mathcal{N}-$infinite divisibility, stopped sums, discrete models.

1. Heterocedasticity in gaussian samples

Comparison of mean effects in heteroscedastic gaussian populations can be solved using Smith–Welch–Satterthwaite's estimate fo the (fractional) number of degrees of freedom in an approximate $t-$test. Scheffé (1943), in the early stages of the development of computational statistics, by

[12]Research partially supported by FCT/POCTI/FEDER.

randomly selecting elements in the bigger sample to match to the elements of the smaller sample, devised a clever test to use a simple paired t–test.

The problem lost importance for a while, but meta-analysis has shown that Fisher's apraisal on the irrelevance of this question was wrong — *"Statisticsl methodology for meta-analysis has been moving away from approaches focused only on fitting a common estimate toward approaches that include estimating the extent and sources of heterogeneity among studies."*, a quote from the introduction to the field by Stangl and Berry (2000). Velosa (2001) tried to use more efficiently the information available, by matching each element in the *bigger* sample to an element obtained by simulating elements using the smaller sample. Results compare favourably to what Scheeffé's test achieves, but they are not as good as foreseen.

2. Studentization and scale analysis in non-gaussian populations

When comparing means using paired samples, $D = X - Y$ is symmetric. Expressing the vector $\left(\overline{X}_{n+1},\, nS^2_{n+1},\, X_{n+1}\right)$ in terms of the vector $\left(\overline{X}_n,\, (n-1)S^2_n,\, \overline{X}_n - X_{n+1}\right)$ leads to

$$f_{(\overline{X}_{n+1},\, nS^2_{n+1})}(w,s) = \sqrt{\tfrac{n+1}{n}\, s} \times$$

$$\times \int_{-1}^{+1} f_{(\overline{X}_n,\, (n-1)S^2_n)}\left(w + \sqrt{\frac{s}{n(n+1)}}\, v,\, s\left(1-v^2\right)\right) f_X\left(w - \sqrt{\frac{ns}{n+1}}\, v\right) \mathrm{d}v \tag{9}$$

with initial condition $f_2(w,s) = \sqrt{\frac{2}{s}}\, f\big(w + \sqrt{\frac{s}{2}}\big)\, f\big(w - \sqrt{\frac{s}{2}}\big),\ w \in \mathbb{R},\ s > 0$. Observing that this can be rewritten as

$$\begin{aligned} f_{(\overline{X}_{n+1},\, nS^2_{n+1})}(w,s) &= 2^{n-2}\sqrt{n-1}\left[\prod_{i=3}^{n}\left(1-\xi_i^2\right)^{\frac{i-4}{2}}\right] s^{\frac{n+1-3}{2}} \times \\ &\times \int_0^{+\infty}\left(1-v^2\right)^{\frac{n-3}{2}} \prod_{i=1}^{n+1} f_X\big(w + \alpha_{i,n+1}(v)\sqrt{s}\big)\, \mathrm{d}v, \end{aligned}$$

where $\alpha_{i,n+1} = \alpha_{i,n+1}(v) = \frac{v}{\sqrt{n(n+1)}} + \alpha_{in}\sqrt{1-v_n^2},\ i = 1,\ldots,n$ and

$\alpha_{n+1,n+1} = \alpha_{n+1,n+1}(v) = -\sqrt{\frac{n}{n+1}}\ v$, Rocha (2001) obtained the approximation

$$\begin{aligned} f_{(\overline{X}_{n+1},\, nS^2_{n+1})}(w,s) &= 2^{n+1-2}\sqrt{n+1}\left[\prod_{i=3}^{n+1}\left(1-\xi_i^2\right)^{\frac{i-4}{2}}\right] s^{\frac{n+1-3}{2}} \times \\ &\times \prod_{i=1}^{n+1} f_X(w+\alpha_{i,n+1}(\xi_{n+1})\sqrt{s}) \end{aligned}$$

with $\alpha_{i,n+1} = \alpha_{i,n+1}(\xi_{n+1}) = \frac{\xi_{n+1}}{\sqrt{n(n+1)}} + \alpha_{in}\sqrt{1-\xi^2_{n+1}}$, $i = 1,\ldots,n$; $\alpha_{n+1,n+1} = \alpha_{n+1,n+1}(\xi_{n+1}) = -\sqrt{\frac{n}{n+1}}\ \xi_{n+1}$, and $\sum_{i=1}^{n+1}\alpha_{i,n+1}(\xi_{n+1}) = 0$, $\sum_{i=1}^{n+1}\alpha^2_{i,n+1}(\xi_{n+1}) = 1$. From this

$$f_{T_{(n-1)}}(t) = \frac{2^{n-1}}{\sqrt{n-1}}\left[\prod_{i=3}^{n}\left(1-\xi_i^2\right)^{\frac{i-4}{2}}\right]\int_0^{+\infty} u^{n-1}\prod_{i=1}^{n} f\left(\left(\frac{t}{\sqrt{n(n-1)}}+\alpha_{in}\right)u\right)\mathrm{d}u \tag{10}$$

which is exact when the parent distribution is gaussian.

The non-symmetric case is much more difficult; however, for more general studentized statistics, Brilhante, Rocha and Pestana obtained exact results using inverse integral transforms, for gamma and beta distributions. In the case of beta populations, the inversion formula has been obtained with fractional derivatives, which are quite simple whenever it is possible to express the function in terms of geeneralized hypergeometric functions.

Pestana and Rocha developed, for $X_k \frown Exponential(\lambda_k, \delta_k)$, an analysis of scale similar to the analysis of variance for gaussian populations, using the independence between spacings; curiously, under the null hypothesis os homegeneity, the test statistic has an exact F distribution. Brilhante, analyzing the structure dependence of spacings, worked out results for Laplace and for generalized Pareto distributions.

3. Nonclassical limit distributions

The study of weak limit distributions for sums and for extreme order statistics has striking similarities; Zolotarev (1957) and his students developed a similar theory of products of independent random variables, and Bingham (1971) the abstract study of domains of attraction for generalized convolution algebras. Rnyi's (1956) work on elementary rarefaction has been fully developed by Kovalenko (1965) and by Gnedenko (1970), characterizing the Laplace transforms of stable geometric stopped sums of positive summands, and Kozubosvki (1994) characterized the class of stable of geometric stoppd sums. In this class, the Laplace law shares some of the most important structural properties of the gaussian law in the classical summation scheme, and this led to the investigation of general $\mathcal{N}-$gaussian laws, and of $\mathcal{N}-$ infinite divisibility. Other generalizations deal with relaxation of the hypothesis on tailweight, inependence and/or distributional homogeneity. Stable laws have also an interesting relation to self-reciprocal characteristi functions (Teugels, 1971).

Weak limit theorems in all these cases deal with convergence of classes, in the strict sense of Khinchine's types; following Logan, Mallows, Rice and Shepp, Mendona (2001) investigated limit theorems for self-normalized sequences. Mendona and Malva are also involved in the study of large deviations and tools to improve approximations, namely using saddle point methods. Work on the use of pseudo-mements for centering and scaling is also under development.

Infinitely divisible laws are limit laws of Poisson stopped sums, though they are more often described as imits of asimptotically null triangular arrays. As we said before, Kozubosvki (1994) arrived at definitive results on stoped geometric sums. Pestana and Velosa (2002, 2003), investigated the more general setting of Panjer stopped sums, with Panjer subordinator $N_{\alpha,\beta}$ with probability mass function satisfying $\frac{p_{\alpha,\beta}(n+1)}{p_{\alpha,\beta}(n)} = \alpha + \frac{\beta}{n+1}$, and extended Panjer subordinators with probability mass function satisfying, for $\alpha,\ \beta \in \mathbb{R},\ n = 0, 1, \ldots,$

$$\frac{p_{\alpha,\beta,\gamma}(n+1)}{p_{\alpha,\beta,\gamma}(n)} = \alpha + \beta\,\mathbb{E}(U_0^n) = \alpha + \beta\,\frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)}\,, \qquad (11)$$

where $U_\gamma \frown Uniform(\gamma, 1)$. Panjer classes have an important role in risk theory and finance cf. Klugman, Panjer and Willmot (1998) and Rlski, Schmidli, Schmidt and Teugels (1999).

4. Discrete models

Marques, Pestana and Velosa (2002) used Panjer classes and their extensions by Sundt and Jewell (1981) and by Willmot (1987) to elaborate a coordinated presentation of discrete models, and they have shown that more sophisticated models provide a much better fit to biological phenomena such as adultery in birds than the traditional hypergeometric, binomial and Poisson models. Truncation of logarithmic random variables, for instance, leads to substantial broadening of the parameter space, and is an interesting way of arriving at Zipf–Mandelbrot's models. The biological interpretation is striking, in terms of equilibrium between two conflicting interests, genetic diversity on one side, the need for a cooperating companion to raise the nestlings on the other. Zipf–Mandelbrot's hyperbolic laws are widely used for self-oganizing structures (Mandelbrot, 1997; Adamic, 2001). An interesting alternative is to investigate the possibility of organizing patterns with scale, and the discrete lognormal law provides a promising model.

Discrete probability mass functions of the form $\left\{p_k = \frac{1}{C(\theta)} \frac{1}{k^{1+\theta(k)}}\right\}_{k\geq 1}$ provide an even wider choice of models, among which Zipf–Mandelbrot's, corresponding to $\theta(k) = \rho$, is just the easier to deal with. The general case, on the other hand, may provide appropriate models where location may influence scale effects.

A particular choice $\theta(k) = \ln(\sqrt{k})$ leads to $p_k \propto \frac{1}{k^{1+\ln\sqrt{k}}}$, $k = 1, 2, \ldots$, i.e. to the *discrete lognormal* random variable

$$X_{0,1} = \begin{cases} k & k = 1, 2, \ldots \\ p_k = \frac{1}{C(0,1)\,k} \exp\left(-\frac{(\ln k)^2}{2}\right) \end{cases}$$

where $C(0,1) = \sum_{k=1}^{\infty} \frac{1}{k} \exp\left(-\frac{(\ln k)^2}{2}\right)$ is the norming constant.

More generally, the discrete lognormal law with parameters μ and σ

$$X_{\mu,\sigma} = \begin{cases} k & k = 1, 2, \ldots \\ p_k = \frac{1}{C(\mu,\sigma)\,k} \exp\left[-\frac{1}{2}\left(\frac{\ln k - \mu}{\sigma}\right)^2\right] \end{cases}$$

with $C(\mu, \sigma)$ the appropriate norming constant, seems a likely model for discrete skew populational data. The parameters μ and σ may be estimated by numerical maximization of the likelihood function.

Further observe that when $\mu \to -\infty$, the discrete lognormal law may be considered a Zipf-Mandelbrot law with shape parameter $\rho = \frac{\mu}{\sigma^2}$. In fact, if $\ln k \ll |\mu|$,

$$p_k = \frac{1}{C(\mu,\sigma)\,k}\ \exp\left[-\frac{1}{2}\left(\frac{\ln k-\mu}{\sigma}\right)^2\right]$$

$$\propto \frac{1}{k}\ \exp\left[-\frac{\ln k\,(\ln k-2\mu)}{2\,\sigma^2}\right] \approx \frac{1}{C^*(\mu,\sigma)}\ \frac{1}{k^{1-\frac{\mu}{\sigma^2}}}\,.$$

Fit to our biological data has been excellent. It provided as well an opportunity to revisit general power divergence statistics (Cressie and Read, 1984; Pestana and Vasconcelos, 1999), $T_\lambda = \frac{2}{\lambda(\lambda+1)} \sum_{k=1}^{N} O_k \left[\left(\frac{O_k}{e_k}\right)^\lambda - 1\right]$ and to investigate the influence of λ in messy situations.

References

[1] Adamic, L. A. (2001). Zipf, Power-Laws and Pareto — a ranking tutorial. http://ginger.hpl.hp.com/shl/ papers/ranking/

[2] Bingham (1971). Factorization theorems and domains of attraction for gneralized convolution algebras, *Proc. London Mth. Soc* **23**, 16–30.

[3] Brilhante, M. F. (1999). A resistant test statistic for testing exponentiality versus generalized Pareto. In *Extreme Values and Additive Laws*, 11–14, Lisboa.

[4] Brilhante, M. F. (2001). Inference on the location parameter of exponential populations — externally studentized statistics. *Rev. Estat.* **11**, 65-66.

[5] Brilhante, M. F., Pestana, D., Rocha, J., e Velosa, S. (2001). *Inferncia Estatstica Sobre Localizao e Escala*, Sociedade Portuguesa de Estatstica, Ponta Delgada.

[6] Cressie, N., and Read, T. R. C. (1984). Multinomial Goodness-of-fit tests, *J. Roy. Statist. Soc.* **B 46**, 440–464.

[7] Gnedenko, B. V. (1970). Limit theorems for sums of a random number of positive independent random variables, *Proc. 6th Berkeley Symp. Math. Statist. Probab.*, vol. 2, 537–549, California University Press, Berkeley.

[8] Kovalenko, I. N. (1965). On a class of limit distributions for rarefied flows of homogeneous events. In: *Selected Translations Mathematical Statististics and Probability* **9**, Providence, Rhode Island, 1971, pp. 75– 81.

[9] Logan, B. F., Mallows, C.L., Rice, S.O. and Shep, L. A. (1973). Limit distributions of self-normalized sums. *Ann. Probab.* **1**, 788–809.

[10] Kozubowski, T. J. (1994). Representation and properties of geometric stable laws, *Approximation, Probability, and Related Fields*, Plenum, New York. 321–337.

[11] Mandelbrot, B. (1997). *Fractals and Scaling in Finance*, Springer, New York.

[12] Marques, T. A., and Pestana, D. D. (2002). *Extra-Pair Fertilizations and the Distribution of Extra-Pair Young Within and Among Broods: A Case Study on Count Data Models in Biometry*, CEAUL, Lisboa.

[13] Mendonça, S. (2001). *Weak Convergence of Functions of Sums and Maxima of Sequences of Random Variables*, Ph. D. Thesis, Un. Madeira.

[14] Mendonça, S. (2001a). On sums and extremes of random variables. *Revista de estatística*, 2001:2, II, 209–212.

[15] Panjer, H. H. (1981). Recursive evaluation of a family of compound distributions, *ASTIN Bulletin* **12**, 22–26.

[16] Pestana, D., Brilhante, F., and Rocha, J. (1999). The analysis of variance revisited. In *Extreme Values and Additive Laws*, 73–77, Lisboa.

[17] Pestana, D. D., Sequeira, F. and Velosa, S. F. (2001). Parseval's relation and self-reciprocal characteristic functions, *Rev. Estat./Statist. Rev. — 23^{rd} European Meeting of Statisticians, Contributed Papers* II, 315–316.

[18] Pestana, D. D., and Vasconcelos R. (1999). Power divergence statistics, *Extreme Values and Additive Laws*, CEAUL, Lisboa, 78–81.

[19] Pestana, D. D., and Velosa, S. F. (2003). $\mathcal{N}$-infinitely divisible classes, In: *Proceedings Xth SPE Annual Congress*, pp. 511–522.

[20] Pestana, D. D., and Velosa, S. F. (2003a). Random extensions of Katz-Panjer families of discrete distributions, *Seventh International Congress on Insurance: Mathematics and Economics*, Lyon.

[21] Rényi, A., 1956. A characterization of the Poisson process. In: *Selected Papers of Alfred Rényi,* **1***, 1948–1956*, Akadémiai Kiadó, Budapest pp. 622–629.

[22] Rocha, J. (2001b). Inference on location parameters — internally studentized statistics. *Revista de Estatística* , 355-356.

[23] Rólski, T., Schmidli, H., Schmidt, V., and Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*, Wiley, New York.

[24] Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* **2**, 110-114.

[25] Scheffé, H. (1943). On solutions of the Behrens-Fisher problem based on the $t-$distribution. *Ann. Math. Stat.* **14**, 35-44.

[26] Smith, H. (1936). The problem of comparing the results of two experiments with unequal means. *J. Council Sci. Industr. Res.* **9**, 211-212.

[27] Stangl, D. K., and Berry, D. A. (2000). Meta-analysis: past and present challenges, in Stangl and Berry, eds. (2000). *Meta-Analysis in Medicine and Health Policy*, M. Dekker, New York, 1–28.

[28] Sundt, B. and Jewell, W. S. (1981). Further results on recursive evaluation of compound distrbutions, *ASTIN Bulletin* **12**, 27–39.

[29] Teugels, J. (1971). Probability density functions which are their own characteristic functions, *Bull. Soc. Math. Belg.* **23**, 236–262.

[30] Velosa, S. (2001). *O Problema de Behrens-Fisher*, Dissertação de Mestrado, DEIO, FCUL, Lisboa.

[31] Velosa, S. (2001a). Modificação do teste de Scheffé para comparação de valores médios de populações gaussianas com variâncias diferentes. *IX Congresso Anual da SPE*, Ponta Delgada.

[32] Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika* **29**, 350-361.

[33] Welch, B. L. (1943). The generalization of "Student's" problem when several different population variances are involved. *Biometrika* **34**, 28-35.

[34] Willmot, G. E. (1987). Sundt and Jewell's familly of discrete distributions, *ASTIN Bulletin* **18**, 17–29.

[35] Zolotarev, V. M. (1957). Melin-Stieljes transforms in Probability Theory, *Theor. Probab. Appl.* **2**, 433-460.

BLOCK BOOTSTRAP METHODS IN EXTREMAL INDEX ESTIMATION[13]

Dora Prata Gomes
Departamento de Matemática
Universidade Nova de Lisboa
Portugal

Manuela Neves
Departamento de Matemática, ISA
Universidade Técnica de Lisboa
Portugal

Abstract: The bootstrap is a computer-intensive method that provides answers to a large class of statistical inference problems without strong assumptions on the underlying random process generating the data. The methods that are available for implementing the bootstrap and the accuracy of bootstrap estimates depend on whether the data are a sample random from a distribution or dependent. Here, the application of the bootstrap to dependent data is considered. Block bootstrap methods, proposed for implementing the bootstrap in that situation are reviewed. However the performance of block bootstrap methods depends on the block size. For the method to be practical, a good choice of block size is necessary. We present a data-based method for choosing the optimal block length following the work of Hall, Horowitz, and Jing (1996) [6]. One drawback of the method is that it depends on the block length b and on the length s_n of subsamples which has to be chosen by the user. In extreme value theory, when we are in a situation of dependence, the extremal index is a parameter of great importance. The estimators considered in the literature, despite of having good asymptotic properties, present strong bias and dependence on the high threshold u, for finite samples. Block bootstrap methods will be considered for extremal index estimators.

[13]Research partially supported by FCT/POCTI/FEDER.

Key words and phrases: Bootstrap, Block Bootstrap, Extremal Index, Simulation.

1. Introduction

Bootstrap methodology was first introduced by Efron(1979) [2] in the context of "i.i.d." data. If the data are a random sample, the bootstrap can be implemented by sampling the data randomly with replacement or by sampling a parametric model of the data distribution. Shing(1981) [15] showed the inadequacy of the method under dependence. In fact the situation is more complicated when the data are dependent because bootstrap sampling must be carried out in a way that suitably captures the dependence structure of the data generation process. A general approach is to apply resampling to the original data sequence by considering blocks of data and placing them end to end rather than single data points as in the "i.i.d." setup. The motivation is that within each block the dependence structure of the underlying model is preserved and if the block size is allowed to tend to infinity with the sample size, asymptotically correct inference can ensue.

Here, we illustrate these approach to the extremal index estimation.

2. Extremal Index

Let $\{X_n\}_{n\geq 1}$ be a stationary sequence of random variables that has a marginal distribution function F and define $M_n = \max(X_i : i = 1, \cdots, n)$. Then if $\{a_n > 0\}$ and $\{b_n\}$ are sequences of constantes such that as $n \to \infty$,

$$P\left\{(M_n - b_n)/a_n \leq x\right\} \to G(x),$$

where G is a non-degenerate distribution function and $D(u_n)$ condition, Leadbetter *et.al.*(1983) [9] holds with $u_n = a_n x + b_n$ for every real x, G is a member of the generalized extreme value family of distributions. Then the maxima of stationary series follow the same distributional limit laws as those of independent series. However, the parameters of the limit distribution are affected by the dependence in the series.

Theorem 1. *Let* $X_1, X_2, \cdots$ *be a stationary process and* $X_1^*, X_2^*, \cdots$ *be a sequence of independent variables with the same marginal distribution. Define* $M_n^* = \max(X_1^*, \cdots, X_n^*)$*. Under suitable regularity conditions,*

$$P\left\{\left(M_n^* - b_n\right)/a_n \le x\right\} \to G_1\left(x\right),$$

as $n \to \infty$ *for normalizing sequences* $\{a_n > 0\}$ *and* $\{b_n\}$*, where* G_1 *is a non-degenerate distribution function, if and only if*

$$P\left\{\left(M_n - b_n\right)/a_n \le x\right\} \to G_2\left(x\right),$$

where

$$G_2(x) = G_1^{\theta}(x)$$

for a constant θ *such that* $0 < \theta \le 1$*.*

The quantity θ is termed the **extremal index**. For a more precise definition see Leadbetter *et.al.*(1983) [9]. The extremal index, θ, is the key parameter when extending discussions of the limiting behavior of the extreme values from "i.i.d." sequences to stationary sequences. One of the most popular estimators of the extremal index is the *up-crossing estimator*, Gomes(1990, 92, 94) [3], [4], [5] and Nandagopalan (1990) [11], defined, for u_n a high threshold, by

$$\widehat{\theta}_n^{CA} := \frac{\sum_{i=1}^{n-1} I(X_i \le u_n < X_{i+1})}{\sum_{i=1}^{n} I(X_i > u_n)}.$$

3. Block Bootstrap Methods

A pioneering work to extend the bootstrap method to the dependent case was provided by Carlstein(1986) [1], who used a blocks scheme to approximate the variance of a general statistic. His idea was to divide the original sequence in (nonoverlapping) blocks of size b, recompute the statistic of interest on these blocks, and use the sample variance of the block statistics, after some suitable normalization. Later Künsch(1989) [7] and Liu and Singh(1992) [10] independently introduced the "moving blocks" bootstrap which, besides variation estimation, can also be used to estimate the sampling distribution of a statistic so that confidence intervals or regions for unknown parameters can be constructed. The key difference between this method and Efron's "i.i.d." bootstrap is that the blocks of size $b < n$ observations, resampled with replacement from the data, are concatenated to form such a pseudo-sequence rather than single data points. In contrast to Carlstein(1986) [1] approach, the "moving blocks" bootstrap uses overlapping blocks. Politis and Romano(1992, 1994) [13] and [14] proposed variants of the moving blocks bootstrap that possess interesting features such as random block sizes and the circular bootstrap, that require periodic extension of the original data sequence. More details on these methods can be found in the Lahiri(2003) [8].

However the performance of block bootstrap methods depends on the block size. Hall, Horowitz and Jing (1995) [6] address the problem of optimal block choice when the block bootstrap is used in a variety of different contexts. They identify three different settings of practical importance: estimation of bias or variance, estimation of an one-sided distribution function and estimation of a two-sided distribution function. Here we considered their suggestion to use empirical methods to choose the block size for a subseries of the original data set of length $s_n < n$ and then estimating the optimal block size by $\widehat{b}_{opt,n} = \widehat{b}_{opt,s_n}(n/s_n)^{1/3}$.

The estimators $\widehat{\theta}_n^{CA}$ present strong dependence on the high threshold. The choice of the optimal level u, or $k_0 : \ u = X_{k_0:n},$, is carried out following an heuristic method, see Oliveira(2002) [12], after estimating the optimal block size. The idea is the following: use the knowledge of the optimal block length, $\widehat{b}_{opt,n}$, select a auxiliary threshold k_{aux}, we compute $\widehat{\theta}(k_{aux})$; divide the sample into blocks of size $\widehat{b}_{opt,n}$; resample B times from the sample and for $1 \leq k \leq n-1$, compute $\widehat{\theta}^*(k)$. Select the threshold which minimize mean squared error of $\widehat{\theta}(k_{aux})$ over a suitable set of thresholds and replace the auxiliary threshold by this threshold for the second iteration. Repeat this process until convergence. We will present some simulation results.

There is still a lot of work that needs to be done to carefully apply the bootstrap in other estimators of the extremal index, *blocks estimator* and the *runs estimator*, which depends also on the clusters identification parameters, r_n(the size of the block) and ℓ(number of observations below u_n that are separating two distinct clusters), respectively.

References

[1] Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Ann. of Stat.* 14, 1171-1179.

[2] Efron, B. (1979). Bootstrap methods: another look at the Jackknife. *Ann. Stat.* 7, 1-26.

[3] Gomes, M.I. (1990). Statistical inference in an extremal markovian model.*COMPSTAT*, 257-262.

[4] Gomes, M.I. (1992). Modelos extremais em esquemas de dependência. *I Congresso Ibero-Americano de Esdadistica e Investigacion Operativa.* 209-220.

[5] Gomes, M.I. (1994). Metodologias Jackknife e Bootstrap em Estatística de Extremos *Actas do II Congresso Anual da Sociedade Portuguesa de Estatística.* 31-46.

[6] Hall, P., Horowitz, J. L. and Jing, B.-Y. (1996). On blocking rules for the bootstrap with dependent data. *Biometrika* 50, 561-574.

[7] K*ü*nsch, H. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Math.*, 17, 1217-1241.

[8] Lahiri, S.N. (2003). *Resampling Methods for Dependent data.* Springer Verlag, Inc.

[9] Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). *Extremes and related properties of random sequences and series.* Springer Verlag. New York.

[10] Liu, R. Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence, in R. Lepage and L. Billard, eds, *Exploring the Limits of Bootstrap*, Wiley, New York, pp. 225-248.

[11] Nandagopalan, S. (1990). *Multivariate Extremes and Estimation of the Extremal Index.* Ph.D. Thesis. Techn. Report 315, Center for Stochastic Processes, Univ. North-Caroline.

[12] Oliveira, O. A. (2002). *In Extremis.* Ph.D. Thesis DEIO. Faculdade de Ciências da Universidade de Lisboa.

[13] Politis, D. N. and Romano, J. P. (1992). A circular block resampling procedure for stationary data, in R. Lepage and L. Billard, eds,

Exploring the Limits of Bootstrap, Wiley, New York, pp. 263-270.

[14] Politis, D. N. and Romano, J. P. (1994). The Sattionary Bootstrap, *J. of Amer. Stat. Assoc.*, 89, 1303-1313.

[15] Singh, K. (1981). On the asymptotic accuracy of the Efron's bootstrap. *Ann. of Stat.*, 9, 345-362.

JOINT TAIL MODELLING AND PROPERTIES[14]

Alexandra Ramos
Faculdade de Economia do Porto
Universidade do Porto
Portugal

Anthony Ledford
Research Group
Man Investment Products Ltd, London
United Kingdom

Abstract: A fundamental issue in applied multivariate extreme values (MEV) analysis is modelling dependence within joint tail regions. The primary focus of this work was to develop a pseudo-polar framework for modelling extremal dependence that extended the existing classical results for multivariate extremes to encompass asymptotically independent tails. A constructional procedure for obtaining parametric asymptotically independent joint tail models was then developed and properties of these models were studied. In contrast to the classical MEV approach, which concentrates on the distribution of the normalised componentwise maxima, our framework is based on modelling joint tails and focuses directly on the tail structure of the joint survivor function. For simplicity, we concentrate here on the bivariate case.

The practical application of such models is shown to give good results, even in the presence of negative dependence between the marginal extremes. Inference under our models is then examined and tests of extremal asymptotic independence and asymmetry are derived which are useful within model selection. Finally, methods for simulating from two of our bivariate parametric models are provided.

Key words and phrases: Asymptotic independence, Joint tail dependence models, Likelihood ratio tests, Simulation methods.

1. Modelling dependence within bivariate joint tails

Both standard parametric and non-parametric estimation techniques present some problems in the particular case where data are asymptotically independent, i.e. when their distribution is in the domain of attraction of independence. However, asymptotic independence is an important case in practice both for applications and for theoretical development, see e.g. de Haan and Ronde (1998) [1], and arises for most classical families of distribution, as listed in Capéraà *et al.* (2000) [1].

Consider a bivariate random variable (X, Y) with unit Fréchet marginal distributions that satisfies

$$\Pr(X > x, Y > y) = \overline{F}_{XY}(x, y) = \frac{\mathcal{L}(x, y)}{(xy)^{1/(2\eta)}} \tag{1}$$

where $\mathcal{L}$ is a bivariate slowly varying (BSV) function and $\eta \in (0, 1]$ is the coefficient of tail dependence (see Ledford and Tawn (1997) [3]).

Let g denote the limit function of $\mathcal{L}$, so that for all $(x, y) \in \mathbb{R}^2_+$ and $c > 0$

$$g(x, y) = \lim_{r\to\infty} \frac{\mathcal{L}(rx, ry)}{\mathcal{L}(r, r)} \quad \text{and} \quad g(cx, cy) = g(x, y). \tag{2}$$

Using this result, the limit function g can be written as $g(x, y) = g_* \{x/(x + y)\} = g_*(w)$ for $w = x/(x + y) \in (0, 1)$. This g measures the asymptotic ray dependence of the BSV function $\mathcal{L}$, and $\mathcal{L}$ can be defined as *asymptotically ray dependent* if $g_*(w)$ varies with w and as *asymptotically ray independent* if $g_*(w)$ is constant over different rays.

Let u denote a high threshold. We consider the behaviour of the bivariate conditional random variable (S, T) defined by $\lim_{u\to\infty} (X/u, Y/u) \mid (X > u, Y > u)$, in the sense that, for all $(s, t) \in [1, \infty) \times [1, \infty)$,

$$\begin{aligned} \Pr(S > s, T > t) &= \lim_{u\to\infty} \frac{\Pr(X > su, Y > tu)}{\Pr(X > u, Y > u)} \\ &= \lim_{u\to\infty} \frac{\mathcal{L}(us, ut)}{\mathcal{L}(u, u)\,(st)^{1/(2\eta)}} = \frac{g(s, t)}{(st)^{1/(2\eta)}}. \end{aligned} \tag{3}$$

Transforming to the pseudo-radial and angular coordinates defined by $R = S + T$ and $W = S/R$, the density of (R, W) can be shown to satisfy

$$f(r,w) = r^{-(1+1/\eta)} h_\eta(w), \tag{4}$$

where the function h_η is a non-negative measure density[15] on $[0,1]$ determined by g_* and η. Now, reconstructing $\overline{F}_{ST}$ from this density and letting $r^* = \max\{s/w, t/(1-w)\}$, we have

$$\begin{aligned}
\Pr(S>s, T>t) &= \int_{w=0}^{1} \int_{r^*}^{\infty} r^{-(1+1/\eta)} h_\eta(w)\, \mathrm{d}r\ \mathrm{d}w \\
&= \eta \int_0^{\frac{s}{s+t}} \left(\frac{w}{s}\right)^{1/\eta} h_\eta(w)\, \mathrm{d}w && (5) \\
&\quad + \eta \int_{\frac{s}{s+t}}^{1} \left(\frac{1-w}{t}\right)^{1/\eta} h_\eta(w)\, \mathrm{d}w, && (6)
\end{aligned}$$

for $(s,t) \in [1,\infty) \times [1,\infty)$. Providing h_η is known, this representation can be used to obtain parametric models for the joint survivor function of (S,T) and hence for g.

Writing $s = t = t_0$ in equation (6) we obtain the normalisation condition

$$\eta^{-1} = \int_0^{1/2} w^{1/\eta} h_\eta(w)\, \mathrm{d}w + \int_{1/2}^{1} (1-w)^{1/\eta} h_\eta(w)\, \mathrm{d}w. \tag{7}$$

The joint tail model for the original bivariate variable (X,Y) is then given by

$$\overline{F}_{XY}(x,y) = \lambda \overline{F}_{ST}(x/u, y/u) \tag{8}$$

for $x > u$ and $y > u$, where $\lambda = \Pr(X > u, Y > u)$.

2. Examples of parametric models

Example A: A joint tail model based on a modification of the logistic dependence structure. This parametric model is a particular case, when $\varrho = 0$, of that in Example B.

[15]Function h_η is such that $h_\eta(w) = \frac{\mathrm{d}H_\eta(w)}{\mathrm{d}w}$ if the measure H_η is differentiable; it has atomic masses otherwise.

Example B: A joint tail model based on a modification of the asymmetric logistic dependence structure.

Elementary integration shows that the measure density defined by

$$h_\eta(w) = \frac{\eta-\alpha}{\alpha\eta^2\varrho N_\varrho}\left\{w^{-1/\alpha}+\left(\frac{1-w}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta-2}\left(w\,\frac{1-w}{\varrho}\right)^{-(1+1/\alpha)},$$

for $w \in (0,1)$ and where $\eta \in (0,1]$ is the coefficient of tail dependence and $\alpha, \varrho > 0$ are dependence parameters, satisfies the normalisation condition (7). So, using equation (6), we have

$$\overline{F}_{ST}(s,t) = N_\varrho^{-1}\left[s^{-1/\eta}+\left(\frac{t}{\varrho}\right)^{-1/\eta}-\left\{s^{-1/\alpha}+\left(\frac{t}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right], \quad (9)$$

for $(s,t) \in [1,\infty)\times[1,\infty)$. The associated limit function g is given by

$$\begin{aligned} g(s,t) \equiv g_*(w) = {} & \frac{\{w(1-w)\}^{1/(2\eta)}}{N_\varrho} \\ & \times\left[w^{-1/\eta}+\left(\frac{1-w}{\varrho}\right)^{-1/\eta}-\left\{w^{-1/\alpha}+\left(\frac{1-w}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right] \end{aligned}$$

and, from equation (8), the model for the original bivariate (X,Y) variable is given by

$$\overline{F}_{XY}(x,y) = \frac{\lambda u^{1/\eta}}{N_\varrho}\left[x^{-1/\eta}+\left(\frac{y}{\varrho}\right)^{-1/\eta}-\left\{x^{-1/\alpha}+\left(\frac{y}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right]. \quad (10)$$

3. Properties

It is easy to check, for both parametric models defined above, that $g_*(w) \equiv 1$ for all $w \in (0,1)$ both when $\alpha = 2\eta$, and when $\alpha \to \infty$, independently of the value of the parameter ϱ. It is also apparent that g_* is concave ray dependent when $\alpha < 2\eta$ (which corresponds to $\mathcal{L}$ having a limit function which exhibits *concave ray dependence*) and is convex ray dependent when $\alpha > 2\eta$.

Using these parametric models, marginal properties of $\overline{F}_{ST}$, defined in equation (3), can be examined. Also, we verify that the joint tail models obtained for (X, Y) using the logistic and asymmetric logistic BEV dependence structure are identical respectively to those obtained from Examples A and B by setting $\eta = 1$.

A natural framework for testing asymptotic dependence ($\eta = 1$) against asymptotic independence is provided by our parametric models. Let $L_n(\hat{\alpha}, \hat{\eta})$ be the maximum of the likelihood obtained for our logistic model, taken over the dependence parameters $\alpha > 0$ and $\eta \in (0, 1]$ and write $L_n(\tilde{\alpha}, 1)$ for the corresponding maximised likelihood under the constraint $\eta = 1$. Then, under the condition $\eta = 1$, we have

$$2 \log \{L_n(\hat{\alpha}, \hat{\eta})/L_n(\tilde{\alpha}, 1)\} \xrightarrow{w} Z^2 \qquad \text{as } n \to \infty \tag{11}$$

where the non-negative random variable Z has law

$$\Pr(Z \leq z) = h^*(z)\Phi(z) \tag{12}$$

for $h^*(\cdot)$ the Heaviside step function and $\Phi(\cdot)$ the standard normal distribution function.

Exploiting our parametric models, likelihood ratio tests for 'near' extremal independence ($\eta = 1/2$), ray independence ($\alpha = 2\eta$) and asymmetry ($\varrho = 1$) can also be developed. Such tests provide an aid to model selection, as they inform whether the use of an asymptotically independent joint tail model or an asymmetric joint tail model is merited.

Finally, in this talk, methods for simulating points from the symmetric and asymmetric bivariate logistic distribution function F_{ST}, defined above, are given. The approach considered here uses transformations to derive random variables with a joint distribution from which simulation is straightforward. Methods for simulating from these distributions are vital for undertaking Monte Carlo integration when calculating expectations with respect to the underlying model, and also are useful for testing modelling and estimation approaches on simulated data.

References

[1] Capéraà, P., Fougères, A. -L., and C. Genest (2000). Bivariate distributions with given extreme value attractor. *J. Mult. Anal*, **72**, 30-49.

[2] de Haan, L. and de Ronde, J. (1998). Sea and wind: multivariate extremes at work. *Extremes*, **1**, 7-45.

[3] Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *J. R. Statist. Soc.*, B, **59**, 475-499.

EXTREME VALUE MODELS FOR PITTING CORROSION[16]

Holger Rootzén
Chalmers University of Technology
Gothenburg University
Sweden

An important way to reduce emmissions by cars is to decrease vehicle mass since each kg means about 0.12 g CO2/km of emmission. However, localized, or "pitting", corrosion can limit the usefulness of aluminum, magnesium and other new lightweight materials. Interest focusses on extreme pit depths, since these may cause fatigue problems in loaded areas, leakages in housings and perforation of car bodies. All of these can lead to warranty claims. The car industry resorts to extensive experimentation to study such problems.

This talk presents a careful selection and development of methods for analysis of designed experiments with extreme value distributed responses. The methods are aimed at maximal pit depth data. In contrast to standard ANOVA they are aimed at tail behaviour, do not assume data are heteroscedatic, and give easy prediction of reliability of assemblies. Main steps are to check experimental conditions, to analyze corrosion behavior for individual treatments, and to compare results from different treatments. We further introduce new mixture models for analysis of extreme value distributed data from blocks with varying quality

[16] Joint work with Anne-Laure Fougéres, Sture Holm and John Nolan.

of corrosion treatment. The models are based on new conjugate distributions for the Gumbel and extreme value distribution. The methods are applied to laboratory and field corrosion tests performed by Volvo Car Company, and were developed for this purpose. However we believe that they are widely useful also outside of the field of corrosion testing.

CLUSTERS OF EXTREMES OF STATIONARY SEQUENCES IN GENERAL STATE SPACE

Joahan Segers
Tilburg University
The Netherlands

Consider a triangular array of random variables in an arbitrary state space. Assume that every row of the array is a stationary random vector. Also, consider a certain subset of the state space, to be thought of as a failure set, which is hit with positive but small probability. A general limit theory is developed for the distribution of a row vector conditionally on the event that at least one variable in the row hits the failure set.

Applications include the limit distribution of clusters of extremes of a multivariate stationary time series of which all finite-dimensional distributions are in the domain of attraction of a multivariate extreme-value distribution. Equivalently, the time series must be such that its conditional distribution given the event that it starts at an extreme level converges to a limit.

IS GARCH AS GOOD A MODEL AS THE NOBEL PRIZE ACCOLADES WOULD IMPLY?

Cătălin Stărică
Chalmers University of Technology
Gothenburg University
Sweden

This paper investigates the relevance of the stationary, conditional, parametric model paradigm embodied by Garch(1,1) process to describing and forecasting the dynamics of returns of the Standard & Poors 500 (S&P 500) stock market index.

A detailed analysis of the series of S&P returns featured in the illustration of the use of the Garch(1,1) model in estimating and forecasting volatility given in Section 3.2 of the Advanced Information note on the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel reveals that the Garch(1,1) model severely over-estimated the unconditional variance of returns during the period under study. For example, the annualized implied Garch(1,1) unconditional sd of the sample is 35% while the sample sd estimate is a mere 19%. Over-estimation of the unconditional variance leads to poor volatility forecasts during the period under discussion with the MSE of the Garch(1,1) 1-year ahead volatility more than 4 times bigger than the MSE of a forecast based on historic volatility.

We test and reject the hypothesis that a Garch(1,1) process is the true data generating process of the longer sample of returns on the S&P

500 stock market index between March 4, 1957 and October 9, 2003. We investigate then the alternative use of the Garch(1,1) process as a local, stationary approximation of the data and find that the Garch(1,1) model fails during significantly long periods to provide a good local description to the time series of returns on the S&P 500 and Dow Jones Industrial Average indexes.

Since the estimated coefficients of the Garch model change significantly through time, it is not clear how the Garch(1,1) model can be used for volatility forecasting over longer horizons. A comparison between the Garch(1,1) volatility forecasts and a simple approach based on historical volatility questions the relevance of the Garch(1,1) dynamics for longer horizon volatility forecasting for both the S&P 500 and dow Jones Industrial Average indexes.

ON THE MAXIMUM TERM OF MOVING AVERAGE AND MAX-AUTOREGRESSIVE MODELS WITH MARGINS IN ANDERSON's CLASS[17]

M. Graça Temido
Departamento de Matemática
Universidade de Coimbra
Portugal

Andreia Hall
Departamento de Matemática
Universidade de Aveiro
Portugal

Let $\{X_n\}$ be an iid sequence with marginal distribution F. We shall be concerned with the limiting distribution of the maximum term, after appropriate normalization. First recall that if

$$\lim_{x \to w_F^-} F(x) = 1, \tag{1}$$

where $w_F = \sup\{x : F(x) < 1\}$, a necessary and sufficient condition for the existence of a real sequence $\{u_n\}$ such that $\lim\limits_{n \to +\infty} n(1 - F(u_n)) = \tau$, $\tau > 0$, is given by

$$\lim_{x \to w_F^-} \frac{1 - F(x)}{1 - F(x^-)} = 1. \tag{2}$$

Conditions (1) and (2), which hold trivially for any continuous distribution in a left neighborhood of w_F, exclude the Binomial, Negative

[17] Research partially supported by FCT/POCTI/FEDER.

Binomial and other discrete distributions. While in the Binomial case we don't even have (1), in the Negative Binomial case (2) is replaced by

$$\lim_{n \to +\infty} \frac{1 - F(n-1)}{1 - F(n)} = r, \tag{3}$$

with r in $]1, +\infty[$.

Although (2) holds trivially for all continuous distribution in a left neighborhood of w_F, there are many such situations where it is not possible to find a linear normalization for which there is a limiting non-degenerate distribution for the maximum term. As an example take the multi-modal distribution $F(x) = 1 - \exp(-x - senx)$, for $x \geq 0$.

In order to overcome the limited applicability of the Leadbetter's Extremal Types Theorem to discrete data, Anderson (1970) found an almost stable limiting distribution for the maximum term of discrete sequences verifying (3). He proved that (3) is a necessary and sufficient condition for the existence of normalizing constants $\{b_n\}$ such that $\lim_{n\to\infty}\{F^n(x+b_n)\}$ is bounded between $exp(-r^{-x})$ and $exp(-r^{-(x-1)})$ for any real x.

As an attempt to cover dependent discrete data, Hall (1998) introduces several discrete moving average models obtained by replacing multiplication with random thinning in well known continuous moving average models. For some of these models she proves that whenever the marginal distribution verifies (3), then the linearized maximum term M_n, behaves as

$$\begin{cases} \limsup\limits_{n \to +\infty} P(M_n \leq x + b_n) \leq \exp(-\theta e^{-\alpha x}) \\ \\ \liminf\limits_{n \to +\infty} P(M_n \leq x + b_n) \geq \exp(-\theta e^{-\alpha(x-1)}), \end{cases} \tag{4}$$

with $\theta = 1$ or $0 \leq \theta \leq 1$ whenever $D'(x + b_n)$ or $D^{(2)}(x + b_n)$ from Chernick *et al.* (1991) holds.

Looking back at the stated problems and examples we can say that the class of max-stable distributions seems to be rather restrictive for many applications in statistics. Therefore, looking for possible non-degenerate limiting distributions of either discrete distributions or multi-modal continuous distributions seems to be an interesting issue. The class of max-semistable (MSS) distributions proposed by Pancheva (1992) and Grinevich (1993) provide an answer for these questions.

This class arouse in the context of the limiting distribution of k_n iid random variables, where $\{k_n\}$ is a nondecreasing integer-valued sequence satisfying

$$\lim_{n\to+\infty} \frac{k_{n+1}}{k_n} = r, \quad r \in [1, +\infty[. \tag{5}$$

Following Grinevich (1992) we will say that a distribution function (d.f.) G on $\mathbb{R}$ is MSS if there are reals $r > 1$, $\gamma > 0$ and β such that

$$G(x) = G^r(x/\gamma + \beta),\ x \in \mathbb{R}, \tag{6}$$

or equivalently, if there is a sequence of iid random variables with d.f. F and two real sequences $\{a_n > 0\}$ and $\{b_n\}$ for which

$$\lim_{n\to+\infty} F^{k_n}(x/a_n + b_n) = G(x), \tag{7}$$

for each continuity point of G, with $\{k_n\}$ satisfying (5).

Analytically, the MSS laws can be written as

$$\begin{aligned} G_1(x) &= \exp\{-(t-x)^{\alpha}\nu(\ln(t-x))\} \quad if\ x \in]-\infty, t[, \\ G_2(x) &= \exp\{-(x-t)^{-\alpha}\nu(\ln(x-t))\} \quad if\ x \in]t, +\infty[, \\ G_2(x) &= \exp\{e^{-\alpha x}\nu(x)\} \quad if\ x \in \mathbb{R}, \end{aligned}$$

where ν is periodic positive and bounded, and α is related with r and the period of ν. The general definition of MSS laws is given in Grinevich (1993). If limit (7) holds we shall say that F belongs to the domain of attraction of G.

Temido (2000) proved that for any discrete distribution defined over the set $\{x_n\}$, a necessary and sufficient condition for the existence of a real sequence $\{u_n\}$ such that $F^{k_n}(u_n)$ has a non-degenerate limit is given by

$$\lim_{n\to+\infty} \frac{1 - F(x_{n-1})}{1 - F(x_n)} = r \tag{8}$$

where r and $\{k_n\}$ are related by (5).

This results allow us to prove that any distribution function in Anderson's class, that verifies (3), can only belong to the domain of attraction of a MSS distribution. Furthermore, the limit lies in the discrete Gumbel family. As an example, the Negative Binomial and von Misès distribution given by $F(x) = 1 - exp(-x - senx)$ for $x \geq 0$, belong to the domain of attraction of the MSS distributions $G(x) = exp(-p^{[x]})$ and $G(x) = exp(-exp(-x - senx))$, with real x, respectively.

In Temido and Canto e Castro (2000), all stationary sequences, $\{X_n\}$, satisfying a new dependence restriction, $D_{k_n}(u_n)$, which generalizes Leadbetter's condition $D(u_n)$, are considered. They prove that when $\{X_n\}$ satisfies $D_{k_n}(u_n)$ the limiting behaviour of the maximum M_{k_n} can be inferred from the limiting behaviour of the corresponding maximum of the associated independent sequence using an extension of the well known extremal index. Thus, the corresponding classical result and Leadbetter's Extremal Types Theorem are generalized.

Moreover, in Temido (2000) a suitable adaptations of the local dependence conditions of Chernich and of Leadbetter, denoted by $D_{k_n}^{(m)}(u_n)$ and by $D'_{k_n}(u_n)$, respectively, are introduced.

In this work we consider a particular moving average model and the well known INAR(1) model (see Hall (1998)), both with marginal distribution NB(m,p). We prove that these models verify $D_{k_n}(u_n)$ and $D'_{k_n}(u_n)$ and consequently there exist both a sequence $\{k_n\}$ as above and a real sequence $\{b_n\}$ such that

$$\lim_{n\to+\infty} P(M_{k_n} \leq x + b_n) = exp(-\beta r^{-[x]}) \tag{9}$$

with $\beta > 0$ and $x \in \mathbb{R}$. Similar results are presented considering the max-autoregressive model of second order firstly studied in Hall (1998), which, in this new context, verifies the $D_{k_n}^{(3)}(u_n)$ condition and has extremal index $\theta \neq 1$.

References

[1] Anderson, C.W. (1970). Extreme value theory for a class of discrete distribution with applications to some stochastic processes. *J. Appl. Prob.*, Vol.7, p. 99-113.

[2] Chernick, M.R., Hsing,T. and McCormick, W.P. (1991), Calculating the extremal index for a class of stationary sequences, *Adv. Appl. Prob.*, Vol. 23, 835-850.

[3] Grinevich, I.V. (1993). Domains of attraction of the max-semistable laws under linear and power normalizations. *Theory Probab. Appl.*, Vol. 38, p. 640-650.

[4] Hall, A. (1998). *Extremos de sucesses de contagem - Do outro lado do espelho.* Ph.D. Thesis. University of Lisbon.

[5] Temido, M. G. (2000). *Classes de leis limite em teoria de valores extremos-estabilidade e semiestabilidade.* Ph.D. Thesis. University of Coimbra.

[6] Temido, M.G. and Canto e Castro, L., Max-semistable laws in extremes of stationary random sequences. To appear in *Probability Theory and its Applications.*

LIGHT AND HEAVY TAILS IN RUIN THEORY[18]

Jef Teugels
Katholieke Universiteit Leuven
Belgium

Following a random walk approach, we provide a link between ruin problems in finite time and the behavior of the maximum of a random walk.

For the case of exponentially bounded claims we obtain sharp asymptotic approximations for the ruin in finite time. Thanks to the flexibility of the random walk setting, we even allow copula generated dependencies between claim times and claim sizes.

For the heavy tailed case the picture still looks incomplete. Nevertheless, partial results due to Baltrunas and Klueppelberg support the hopes that also in this case, sharp asymptotic results can be obtained.

[18] Joint work with Hansjoerg Albrecher.

USING THE AUXILIARY FUNCTION ϕ...[19]

Adelaide Valente de Freitas
Departmento de Matemática
Universidade de Aveiro
Portugal

Based on the auxiliary function $\phi(x) = 1/(-ln - ln\ F)'(x)$, where F is a distribution function, we propose a new estimator of the tail index of the generalised extreme value distribution. Simulation results are presented concerning bias and variability.

Again based on the auxiliary function ϕ and the convergence of the maxima of one particular triangular array, we discuss a characterization of the domains of geometric parcial attraction of max-semistable laws.

[19]Research partially supported by FCT/POCTI/FEDER.

STOPPED SUMS WITH GENERALIZED PANJER SUBORDINATORS[20]

Dinis Pestana
CEAUL e DEIO (FCUL)
Universidade de Lisboa
Portugal

Sílvio Velosa
Departamento de Matemática
Universidade da Madeira e CEAUL
Portugal

Key words and phrases: Poisson stopped sums (compound Poisson), geometric stopped sums (compound geometric), Panjer's algorithm.

Let $N^{(k)}_{\alpha,\,\beta}$ denote a counting random variable with probability mass function $\left\{p^{(k)}_n\right\}_{n\in\mathbb{N}} = \left\{p^{(k)}_{\alpha,\,\beta}(n)\right\}_{n\in\mathbb{N}}$ such that

$$p^{(k)}_{\alpha,\,\beta}(n+1) = \left(\alpha + \frac{\beta}{n+1}\right) p^{(k)}_{\alpha,\,\beta}(n)\,, \qquad \alpha,\,\beta \in \mathbb{R},\ n = k, k+1, \ldots \tag{1}$$

The class $\Pi^{(k)}_{\alpha,\,\beta}$ of such random variables has recently been described by Hess, Lewald and Schmidt (2002) in terms of what they call *basic claim number distributions*. Aside from the left k-truncated binomial, Poisson, and negative binomial distributions, the other basic claim number distributions are $\left\{p^{(k)}_{\alpha,\,\beta}(n) = \frac{1}{C(k,\theta)}\,\frac{\theta^n}{\binom{n}{k}}\right\}_{n=k}^{\infty}$, where $C(k,\,\theta) = \sum\limits_{j=k}^{\infty} \theta^j/\binom{j}{k}$, for $k = 1, 2, \ldots$ and $\theta \in (0,\,1)$ — in that case we say that $N \frown \mathit{ExtendedLogarithmic}(k,\,\theta)$

[20]Research partially supported by FCT/POCTI/FEDER.

—, and $\left\{p^{(k)}_{\alpha,\beta}(n)=\frac{1}{C(k,\beta,\theta)}\binom{\beta+n-1}{\beta-1}\theta^n\right\}_{n=k}^{\infty}$, where $C(k,\beta,\theta) = (1-\theta)^{-\beta} - \sum_{j=0}^{k-1}\binom{\beta+n-1}{\beta-1}\ \theta^j$ — and in this case we say that $N \frown ExtendedNegativeBinomial(k,\beta,\theta)$, $k=1,2,\ldots$, $\beta\in(-k,-k+1)$ and $\theta\in(0,1]$. The general description of $\Pi^{(k)}_{\alpha,\beta}$ can be done in term of $\{0,1,\ldots,k-1\}$-modified basic count random variables, or in terms of generalized *hurdle processes.*

This generalizes Sundt and Jewell (1981) investigation of class $\Pi^{(1)}_{\alpha,\beta}$, completed by Willmot (1987). The class $\Pi^{(1)}_{\alpha,\beta}$ is, in turn, a nontrivial extension of Panjer's (1981) class $\Pi^{(0)}_{\alpha,\beta}$ — which we shall denote simply $\Pi_{\alpha,\beta}$, and whose members will be denoted $N_{\alpha,\beta}$ — that started important theoretical and practical developments in Risk Theory.

Panjer (1981) fully described the nondegenerate members $N_{\alpha,\beta}$ of $\Pi_{\alpha,\beta}$: they are $N_{0,\beta} \frown Poisson(\beta)$, $\beta>0$, $N_{\alpha,\beta} \frown Binomial\left(-1-\frac{\beta}{\alpha},\frac{\alpha}{\alpha-1}\right)$, in case $\alpha<0$ and $-\frac{\beta}{\alpha}\in\mathbb{N}^+$, andv $N_{\alpha,\beta} \frown NegativeBinomial\left(\frac{\alpha+\beta}{\alpha},1-\alpha\right)$ when $\alpha\in(0,1)$ and $\alpha+\beta>0$. See also Klugman, Panjer and Willmot (1998) and Rlski, Schmidli, Schmidt and Teugels (1999).

The dispersion index $\frac{\mathrm{var}(N_{\alpha,\beta})}{\mathbb{E}(N_{\alpha,\beta})}=\frac{1}{1-\alpha}$ is less than 1 (underdispersion) for the binomial and greater than 1 (overdispersion) for the negative binomial. $N_{0,\beta}\frown Poisson(\beta)$ is a yardstick, with dispersion index 1. These random variables play an important role as subordinators in randomly stopped sums $S_{N_{\alpha,\beta}}=\sum_{k=1}^{N_{\alpha,\beta}}Y_k$ (with $\mathbb{P}\left[S_{N_{\alpha,\beta}}=0\mid N_{\alpha,\beta}=0\right]=1$, and therefore $\mathbb{P}\left[S_{N_{\alpha,\beta}}=0\right]=\mathbb{P}\left[N_{\alpha,\beta}=0\right]=p_{\alpha,\beta}(0)$ whenever $\mathbb{P}\left[Y_k=0\right]=0$), where the Y_k are i.i.d. random variables independent of the subordinator $N_{\alpha,\beta}$. Compound or generalized random variables (other names traditionally given to $S_{N_{\alpha,\beta}}$, cf. the discussion on terminology in Johnson, Kotz and Kemp, 1992) are at the core of branching processes and many other subjects where the issue is to obtain the distribution of randomly stopped sums, namely in the study of aggregate claims in the risk process. Panjer (1981) devised an algorithm for iterative evaluation of the density of stopped sums with subordinator $N_{\alpha,\beta}$, generalized by Sundt and Jewell (1981); roughly, the iterative expression satisfied by the probability mass function of the subordinator implies a similar structure for the density of the randomly stopped sum, cf. Rlski *et al.* (1999, pp. 118–120), and therefore its density can be evaluated iteratively.

Other extensions of Panjer's class, using more elaborate recursive expressions $p_{n+1} = g\left(p_n, p_{n-1}, \ldots, p_{n-r}\right)$, have been considered by Sundt (1992). In here, starting from the observation that Panjer's recursive expression can be rewritten

$$\frac{p_{\alpha,\beta}(n+1)}{p_{\alpha,\beta}(n)} = \alpha + \beta\,\mathbb{E}(U_0^n) = \alpha + \beta\,\frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_1^n)}, \qquad \alpha,\,\beta \in \mathbb{R},\ n = 0, 1, \ldots, \tag{2}$$

where $U_\gamma \frown Uniform(\gamma, 1)$, we shall investigate the classes $\Pi_{\alpha,\beta,\gamma}$ of counting random variables $N_{\alpha,\beta,\gamma}$ whose probabity mass functions $\{p_n\}_{n\in\mathbb{N}} = \left\{p_{\alpha,\beta,\gamma}(n)\right\}_{n\in\mathbb{N}}$ satisfy the recursive expression

$$\frac{p_{\alpha,\beta,\gamma}(n+1)}{p_{\alpha,\beta,\gamma}(n)} = \alpha + \beta\,\frac{\mathbb{E}\left(U_0^n\right)}{\mathbb{E}\left(U_\gamma^n\right)}, \qquad \alpha,\,\beta \in \mathbb{R},\ \gamma \in (-1, 1]\,,\ n = 0, 1, \ldots, \tag{3}$$

The issues we dwelve in are:

1. Panjer's algorithm, discrete infinite divisibility and discrete geo-infinite divisibility — Panjer's expression is used to obtain representation theorems for the probability generating function of Poisson stoped sums and of geometric stopped sums.

2. For which triplets (α, β, γ) is $\left\{p_{\alpha,\beta,\gamma}(n)\right\}_{n\in\mathbb{N}}$ a proper probability mass function? — from the formal expression of $\Pi_{\alpha,\beta,\gamma}$ probability generating functions

$$\mathcal{G}_{\alpha,\beta,\gamma}(s) = \prod_{k=0}^{\infty} \frac{1-\alpha\gamma^{k+1}s}{1-\alpha\gamma^{k+1}}\;\frac{1-[\alpha+\beta(1-\gamma)]\,\gamma^k}{1-[\alpha+\beta(1-\gamma)]\,\gamma^k s}\,. \tag{4}$$

   we establish in which cases the corresponding infinite convolutions make sensee, and interpret the results in terms of randomly adding 1 to each summand of infinite summation of independent geometric random variables.

3. How far can we go in the generalization of Panjer's iterative algorithm for the computation of the probability mass function of discrete stopped sums with $N_{\alpha,\beta,\gamma}$-subordinator? — We establish that the probability mass function of a discrete stopped sum with subordinator $N_{0,\beta,\gamma}$ can be iteratively computed using an extension of Panjer's algorithm:

$$\frac{1-\gamma^{n+1}}{1-\gamma}\,\tilde{p}_{0,\beta,\gamma}\,(n+1)=\sum_{k=0}^{n}\tilde{p}_{0,\beta,\gamma}\,(k)\;r_{0,\beta,\gamma}\,(n-k)\,,\quad r_{0,\beta,\gamma}\,(n)>0, \tag{5}$$

where $\left\{r_{0,\beta,\gamma}\,(n)\right\}_{n\in\mathbb{N}}$ has generating function

$$\mathcal{H}^{\tilde{\mathcal{G}}}_{0,\beta,\gamma}=\frac{1-\frac{\tilde{\mathcal{G}}_{0,\beta,\gamma}(\gamma s)}{\tilde{\mathcal{G}}_{0,\beta,\gamma}(s)}}{s\;(1-\gamma)}\,, \tag{6}$$

where $\tilde{\mathcal{G}}_{0,\beta,\gamma}(s)=\sum\limits_{k=0}^{\infty}\tilde{p}_{0,\beta,\gamma}s^n$, for $s\le 1$, is the probability generating function of $\left\{\tilde{p}_{0,\beta,\gamma}\,(n)\right\}_{n\in\mathbb{N}}$. Representation theorems of probability generating functions of these classes are also given.

4. Investigation of the special case $\alpha=0$, both when $\gamma\in[0,\,1]$ and when $\gamma\in(-1,\,0)$, and discussion of sensible extension for $\gamma=-1$, in view of theoretical studies on $\mathcal{N}-$infinite divisibility.

Other extensions using moments of other beta a inverse beta random variables as multipliers are under investigation, but so far we have no neat results worth mentioning.

References

[1] Hess, K. Th., Lewald, A., and Schmidt, K. D. (2002). An extenion of Panjer's recursion, *ASTIN Bulletin* **32**, 283–297.

[2] Johnson, N. L., Kotz, S., and Kemp, A. W. (1992). *Univariate Discrete Distributions*, Wiley, New York.

[3] Klugman, S. A., Panjer, H. H., and Willmot, G. E. (1998). *Loss Models: From Data to Decisions*, Wiley, New York.

[4] Panjer, H. H. (1981). Recursive evaluation of a family of compound distributions, *ASTIN Bulletin* **12**, 22–26.

[5] Rlski, T., Schmidli, H., Schmidt, V., and Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*, Wiley, New

York.

[6] Sundt, B. (1992). On some extensions of Panjer's class of counting distributions, *ASTIN Bulletin* **22**, 61–80.

[7] Sundt, B., and Jewell, W. S. (1981). Further results on recursive evaluation of compound distributions, *ASTIN Bulletin* **12**, 27–39.

[8] Willmot, G. E. (1987). Sundt and Jewell's familly of discrete distributions, *ASTIN Bulletin* **18**, 17–29.

List of Participants

Anderson, Clive	C.W.Anderson@sheffield.ac.uk	University of Sheffield	UNITED KINGDOM
Brilhante, Fátima	fbrilhante@notes.uac.pt	Universidade dos Açores	PORTUGAL
Brito, Margarida	mabrito@fc.up.pt	Universidade do Porto	PORTUGAL
Caeiro, Frederico	fac@fct.unl.pt	Universidade Nova de Lisboa	PORTUGAL
Canto e Castro, Luisa	luisacc@fc.ul.pt	Universidade de Lisboa	PORTUGAL
Colaço, Manuel	manuel.colaco@estig.ipbeja.pt	Instituto Politécnico de Beja	PORTUGAL
Dias, Sandra	sdias@utad.pt	Universidade de Trás-os-Montes e Alto Douro	PORTUGAL
Egídio dos Reis, Alfredo	alfredo@iseg.utl.pt	Universidade Técnica de Lisboa	PORTUGAL
Falk, Michael	falk@mathematik.uni-wuerzburg.de	Wurzburg University	GERMANY
Ferreira, Ana	anafh@isa.utl.pt	Universidade Técnica de Lisboa	PORTUGAL
Ferreira, Helena	ferreira@ubista.ubi.pt	Universidade da Beira Interior	PORTUGAL
Ferreira, Marta	msferreira@math.uminho.pt	Universidade do Minho	PORTUGAL
Figueiredo, Fernanda	otilia@fep.up.pt	Universidade do Porto	PORTUGAL
Fraga Alves, M. Isabel	isabel.alves@fc.ul.pt	Universidade de Lisboa	PORTUGAL
Gomes, M. Ivette	ivette.gomes@fc.ul.pt	Universidade de Lisboa	PORTUGAL
Guillou, Armelle	guillou@ccr.jussieu.fr	Université Pierre et Marie Curie	FRANCE
Haan, Laurens de	Ldehaan@few.eur.nl	Erasmus Universiteit Rotterdam	THE NETHERLANDS
Hall, Andreia	andreia@mat.ua.pt	Universidade de Aveiro	PORTUGAL
Henriques Rodrigues, Lígia	lhenriques@ipt.pt	Instituto Politécnico de Tomar	PORTUGAL
Malva, Madalena	madalena@mat.estv.ipv.pt	Escola Superior de Tecnologia de Viseu	PORTUGAL
Martins, Ana Paula	amartins@noe.ubi.pt	Universidade da Beira Interior	PORTUGAL
Martins, M. João	mjmartins@isa.utl.pt	Universidade Tcnica de Lisboa	PORTUGAL

List of Participants (continued)

Mendonça, Sandra	sandra@dragoeiro.uma.pt	Universidade da Madeira	PORTUGAL
Moreira de Freitas, Ana Cristina	amoreira@fep.up.pt	Universidade do Porto	PORTUGAL
Neves, Cláudia	claudia@mat.ua.pt	Universidade de Aveiro	PORTUGAL
Neves, Manuela	Manela@isa.utl.pt	Universidade Técnica de Lisboa	PORTUGAL
Oliveira Martins, João Paulo	jotapmartins@portugalmail.pt		PORTUGAL
Pereira, Luísa	Lpereira@noe.ubi.pt	Universidade da Beira Interior	PORTUGAL
Pestana, Dinis	dinis.pestana@fc.ul.pt	Universidade de Lisboa	PORTUGAL
Prata Gomes, Dora	dsrp@mail.fct.unl.pt	Universidade Nova de Lisboa	PORTUGAL
Ramos, Alexandra	aramos@fep.up.pt	Universidade do Porto	PORTUGAL
Rootzén, Holger	rootzen@math.chalmers.se	Chalmers University Gottenburg	SWEDEN
Santos, Rui	rsantos@estg.ipleiria.pt	Instituto Politécnico de Leiria	PORTUGAL
Segers, Johan	johan.segers@wis.kuleuven.ac.be	Tilburg Universiteit	THE NETHERLANDS
Sequeira, Fernando	fjsequeira@fc.ul.pt	Universidade de Lisboa	PORTUGAL
Souto Miranda, Cristina	cristina.miranda@isca.ua.pt	Universidade de Aveiro	PORTUGAL
Starica, Catalin	starica@math.chalmers.se	Chalmers University Gottenburg	SWEDEN
Temido, M. Graça	mgtm@mat.uc.pt	Universidade de Coimbra	PORTUGAL
Teugels, Jef	jef.teugels@wis.kuleuven.ac.be	Katolieke Universiteit Leuven	BELGIUM
Turkman, K. Feridun	kamil.turkman@fc.ul.pt	Universidade de Lisboa	PORTUGAL
Valente de Freitas, Adelaide	adelaide@mat.ua.pt	Universidade de Aveiro	PORTUGAL
Velosa, S'ilvio	sfilipe@uma.pt	Universidade da Madeira	PORTUGAL
Viseu, Clara	claraviseu@hotmail.com	ISCAC, Coimbra	PORTUGAL