Statistics of heteroscedastic extremes

Introduction

- Develop extreme value statistics to handle the case when observations are drawn from different distributions.
- It will turn out that extreme value statistics go through under mild variation of the underlying distributions and that we can quantify this variation which reflects the frequency of extreme events.

Model

- At time points $i = 1, \ldots, n$, we have independent observations $X_1^{(n)}, \ldots, X_n^{(n)}$ following various continuous distribution functions $F_{n,1}, \ldots, F_{n,n}$ that share a common right endpoint $x^* = \sup\{x : F_{n,i}(x) < 1\} \in (-\infty, \infty]$,
- there is a continuous distribution function F with the same right end point and a continuous positive function c defined on $[0, 1]$ such that

$$\lim_{x \to x^*} \frac{1 - F_{n,i}(x)}{1 - F(x)} = c\left(\frac{i}{n}\right),$$

uniformly for all $n \in \mathbb{N}$ and for all $1 \leq i \leq n$. We impose the condition

$$\int_0^1 c(s)ds = 1.$$

Introduction

- This not only makes the function c uniquely defined but also, similar to a density, c can now be interpreted as the frequency of extremes.
- We call this situation *heteroscedastic extremes* and we call c the *scedaias function*.
- For example, $c \equiv 1$ resembles the uniform or homogeneous density, $i.e.$ we have homoscedastic extremes.
- Note that the limit relation compares only the distribution tails and does not impose any assumption on the central parts of the distributions.

Introduction

In addition, we assume that $F \in D(G_\gamma)$. It then can be shown that

$$\lim_{t\to\infty} \frac{U_{n,i}(tx) - U_{n,i}(t)}{a(t)\{c(i/n)\}^\gamma} = \frac{x^\gamma - 1}{\gamma}. \qquad (1.4)$$

Hence $F_{n,i}$ belong to the domain of attraction of the same extreme value distribution. They have the same extreme value index γ but different scale function a.

Introduction

In this paper, we restrict on the heavy-tailed case, i.e. $\gamma > 0$. Then $x^* = \infty$ and the domain of attraction condition simpilies to

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^\gamma.$$

And the analogue of (1.4) is

$$\lim_{t \to \infty} \frac{U_{n,i}(tx)}{U(t)\{c(i/n)\}^\gamma} = x^\gamma.$$

Estimation

- We begin with estimating the integrated function c, which is defined by $C(s) = \int_0^s c(u)du$ for $s \in [0, 1]$.
- Intuitively, by focusing on the observations above a high threshold, the function C should be proportional to the number of exceedances of the threshold in the first $[ns]$ observations.
- This leads to the following estimator. Order the observations $X_1^{(n)}, \ldots, X_n^{(n)}$ as $X_{n,1} \leq \ldots, \leq X_{n,n}$. For a suitable intermediate sequence $k = k(n)$,

$$k \to \infty, k/n \to 0.$$

- We define the estimator

$$\hat{C}(s) := \frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} \geq X_{n,n-k}\}}.$$

Estimation

When the observations are all different, the estimator can be written in terms of the ranks

$$R_{n,i} = \sum_{j=1}^{n} 1_{\{X_i^{(n)} \geq X_j^{(n)}\}}, 1 \leq i \leq n.$$

as

$$\hat{C}(s) = (1/k) \sum_{i=1}^{[ns]} 1_{R_{n,i} > n-k}.$$

Next we define the Hill estimator as

$$\hat{\gamma}_{\mathrm{H}} := \frac{1}{k} \sum_{j=1}^{k} \log\left(X_{n,n-j+1}\right) - \log\left(X_{n,n-k}\right).$$

Conditions

- Second order condition. Suppose there is a function $A_1(t) \to 0$,

$$\sup_{n\in\mathbb{N}} \max_{1\leqslant i\leqslant n} \left| \frac{1-F_{n,i}(x)}{1-F(x)} - c\left(\frac{i}{n}\right) \right| = O\left[A_1\left\{\frac{1}{1-F(x)}\right\}\right].$$

- Second order condition, suppose there is a function A_2 and a $\rho < 0$ such that ,

$$\lim_{t\to\infty} \frac{U(tx)/U(t) - x^\gamma}{A_2(t)} = x^\gamma \frac{x^\rho - 1}{\rho},$$

- We require , as $n \to \infty$,

$$\sqrt{k}A_1(n/2k) \to 0, \quad \sqrt{k}A_2(n/k) \to 0.$$

- We further assume

$$\lim_{n\to\infty} \sqrt{k} \sup_{|u-v|\leqslant 1/n} |c(u) - c(v)| = 0.$$

Theorem 1

Under the above conditions and under a Skorokhod construction, we have that

$$\sup_{0 \leqslant s \leqslant 1} |\sqrt{k}\{\hat{C}(s) - C(s)\} - B\{C(s)\}| \to 0 \qquad a.s.$$

and

$$\sqrt{k}\,(\hat{\gamma}_{\mathrm{H}} - \gamma) \to \gamma N_0, \qquad a.s.$$

with B a standard Brownian bridge and N_0 standard normal random variable. In addition, B and N_0 are independent.

Kernel Estimation

Let G be a continuous, symmetric kernel function on $[-1, 1]$ such that $\int_{-1}^{1} G(s)ds = 1$; set $G(s) = 0$ for $|s| > 1$. Let $h := h_n$ be a bandwidth such that $h \to 0$ and $kh \to \infty$ as $n \to \infty$. Then the function c can be estimated non-parametrically by

$$\hat{c}(s) = \frac{1}{kh} \sum_{i=1}^{n} \mathbf{1}_{\left\{X_i^{(n)} > X_{n,n-k}\right\}} G\left(\frac{s - i/n}{h}\right).$$

Testing

$H_0 : c = c_0$ or

$H_0 : C = C_0$

We consider the KS test statistic

$$T_1 := \sup_{0 \leqslant s \leqslant 1} \left| \hat{C}(s) - C_0(s) \right|$$

and a Cramer-Von-Mises-type test statistic

$$T_2 := \int_0^1 \left\{ \hat{C}(s) - C_0(s) \right\}^2 \mathrm{d}C_0(s).$$

Corollary 1

Assume that the conditions of theorem 1 hold with $c = c_0$. Then, as $n \to \infty$,

$$\sqrt{k}\,T_1 \xrightarrow{\mathrm{d}} \sup_{0 \leqslant s \leqslant 1} |B(s)|$$

$$k T_2 \xrightarrow{\mathrm{d}} \int_0^1 B^2(s)\mathrm{d}s.$$

High Quantiles

High quantiles are the quantiles $U_{n,i}(1/p)$ with very small tail probability p. We have

$$p = 1 - F_{n,i}\left\{U_{n,i}\left(\frac{1}{p}\right)\right\} \approx c\left(\frac{i}{n}\right)\left[1 - F\left\{U_{n,i}\left(\frac{1}{p}\right)\right\}\right]$$

Hence, we obtain $U_{n,i}(1/p) \approx U(c(i/n)/p)$. Then

$$U_{n,i}\left(\frac{1}{p}\right) = X_{n,n-k}\left\{\frac{k\hat{c}(i/n)}{np}\right\}^{\hat{\gamma}_{\mathrm{H}}}.$$

The high quantile estimator can be extended to forecasting tail risks, i.e. we intend to estimate the high quantile of an unobserved random variable in the next period $X_{n+1}^{(n)}$.

High Quantiles

High quantile $U_{n,n+1}(1/p)$ of the unobserved $X_{n+1}^{(n)}$

$$\widehat{U_{n,n+1}}\left(\frac{1}{p}\right) = X_{n,n-k}\left\{\frac{k\hat{c}(1)}{np}\right\}^{\hat{\gamma}_{\mathrm{H}}}.$$

Since the estimator involves $\hat{c}$ at the boundary point 1, we use a boundary kernel as follows

$$\hat{c}(1) = \frac{1}{kh}\sum_{i=1}^{n} \mathbf{1}_{\left\{X_i^{(n)} > X_{n,n-k}\right\}} G_b\left(\frac{1-i/n}{h}\right),$$

with

$$G_b(x) = \frac{\int_0^1 u^2 G(u)\mathrm{d}u - x\int_0^1 uG(u)\mathrm{d}u}{\frac{1}{2}\int_0^1 u^2 G(u)\mathrm{d}u - \left\{\int_0^1 uG(u)\mathrm{d}u\right\}^2} G(x).$$

Testing

We test whether the extreme value index γ is constant over time.

Concretely, we write $\hat{\gamma}_{(s1,s2]}$ for the Hill estimator based on $X_{[ns_1]+1}, \ldots, X_{[ns_2]+1}$ for any $0 \leq s_1 < s_2 \leq 1$.

We would like to choose $k_{(s_1,s_2]} := k\{\hat{C}(s_2) - \hat{C}(s_1)\}$.

Theorem 3. Assume that the conditions of theorem 1 hold. Then, under a Skorokhod construction, we have that for any $\delta > 0$

$$\sup_{0 \leqslant s_1 < s_2 \leqslant 1, s_2 - s_1 \geqslant \delta} \left| \sqrt{k\left(\hat{\gamma}_{(s_1,s_2]} - \gamma\right)} - \gamma \frac{W\left\{C\left(s_2\right)\right\} - W\left\{C\left(s_1\right)\right\}}{C\left(s_2\right) - C\left(s_1\right)} \right| \to 0 \qquad a.$$

Testing

We can define the test statistic

$$T_3 := \sup_{0 \leqslant s_1 < s_2 \leqslant 1, \hat{C}(s_2) - \hat{C}(s_1) \geqslant \delta} \left| \frac{\hat{\gamma}_{(s_1, s_2]}}{\hat{\gamma}_{\mathrm{H}}} - 1 \right|,$$

or

$$T_4 := \frac{1}{m} \sum_{j=1}^{m} \left(\frac{\hat{\gamma}_{(l_{j-1}, l_j]}}{\hat{\gamma}_{\mathrm{H}}} - 1 \right)^2,$$

where $\hat{\gamma}_H = \hat{\gamma}_{(0,1]}$, $l_1, l_2, \ldots, l_{m-1}$ are cutoff values with $l_j := \sup\{s : \hat{C}(s) \leq j/m\}$; set $l_0 = 0, l_m = 1$.

Corollary 2

Assume that the conditions of theorem 1 hold. Then, we have that, as $n \to \infty$,

$$\sqrt{k}T_3 \xrightarrow{\mathrm{d}} \sup_{0\leqslant s_1 < s_2 \leqslant 1, s_2 - s_1 \geqslant \delta} \left| \frac{W(s_2) - W(s_1)}{s_2 - s_1} - W(1) \right|,$$

$$kT_4 \xrightarrow{\mathrm{d}} \chi^2_{m-1}.$$

Simulations

We consider four data-generating processes (DGPs) as follows.

- observations are IID and follow the standard Frechet distribution. $c \equiv 1$.
- observations are independent, $F_{n,i}(x) = \exp\{-(0.5 + i/n)x\}$. Hence $c(s) = 0.5 + s$.
- observations are independent, $F_{n,i}(x) = \exp\{-c(i/n)x\}$,with $c(s) = 2s + 0.5$ for $s \in [0, 0.5]$ and $c(s) = -2s + 2.5$ for $s \in (0.5, 1]$.
- observations are independent, $F^{(4)}_{n,i}(x) = \exp\{-c(i/n)/x\}$ with $c(s) = 0.8$ for $s \in [0, 0.4] \cup [0.6, 1]$ and $c(s) = 20s - 7.2$ for $s \in (0.4, 0.5]$ and $c(s) = -20s + 12.8$ for $s \in (0.5, 0.6)$.

For each DGP, we simulate 1000 samples of size $n = 5000$ and take $k = 400$.

Simulation

Table 1. Number of rejections out of 1000 simulated data sets

DGP	*Numbers of rejections for the following values of α and tests:*					
	$\alpha = 1\%$		$\alpha = 5\%$		$\alpha = 10\%$	
	T_1	T_2	T_1	T_2	T_1	T_2
1	8	12	44	47	95	98
2	990	998	998	999	1000	1000
3	455	570	838	921	941	987
4	663	521	930	903	979	978

Simulation

Table 2. Bias, variance and asymptotic variance for the forecasted high quantile for $p = 0.02$

DGP	*Bias*	*Variance*	*Asymptotic variance*
1	−0.028	0.137	0.128
2	−0.041	0.094	0.085
3	0.023	0.278	0.256
4	0.004	0.167	0.160

Application

Address the question 'Are financial crises nowadays more frequent than before?'

$H_0 : \gamma$ is constant over time

- Full sample $n = 6302$, from 1998 to 2012, $k = 160$,

  p is nearly zero.

- Subsample $n = 5043$, from 1988 to 2008, $k = 130$

  $p = 0.98$ for T_3 and $p = 0.76$ for T_4

  Then test whether c is constant over time: p is virtually 0.

Application

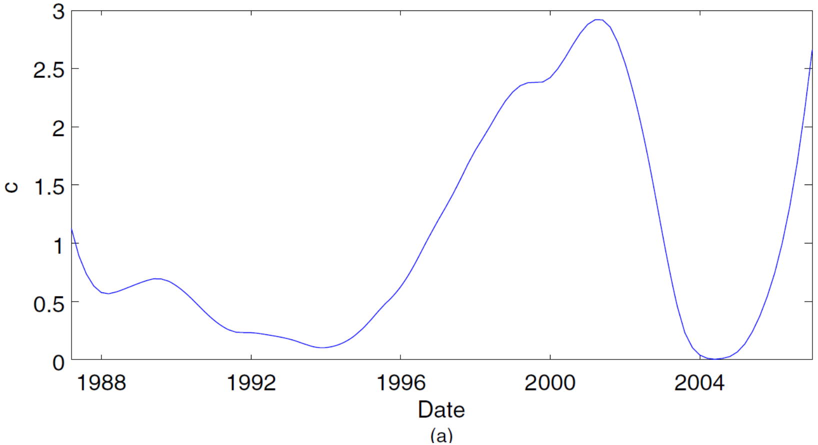

Figure 3: daily returns

Application

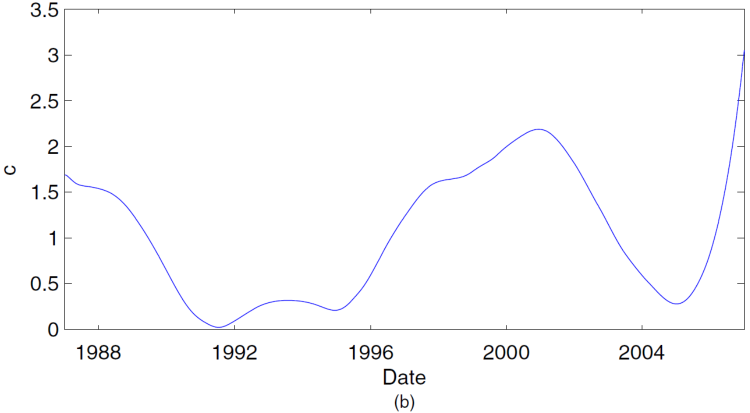

Figure 4: weekly returns