Financial Risk
4-th quarter 2022/23
Lecture 5: Extreme value statistics wrap up

The big recession 2009

"As an alternative to the traditional 30-year mortgage, we also offer an interest-only mortgage, balloon mortgage, reverse mortgage, upside down mortgage, inside out mortgage, loop-de-loop mortgage, and the spinning double axel mortgage with a triple lutz."

Gudrun January 2005
326 MEuro loss
72 % due to forest losses
4 times larger than second largest

Read

S. Coles: An Introduction To Statistical Modelling of Extremes, pages 30-34, 45-53, 74-91, 92-104

S. Lauridsen: Estimation of Value At Risk By Extreme Value Methods

H. Rootzen, N. Tajvidi: Extreme value statistics and wind storm losses, A case study

Course lecture slides

Use

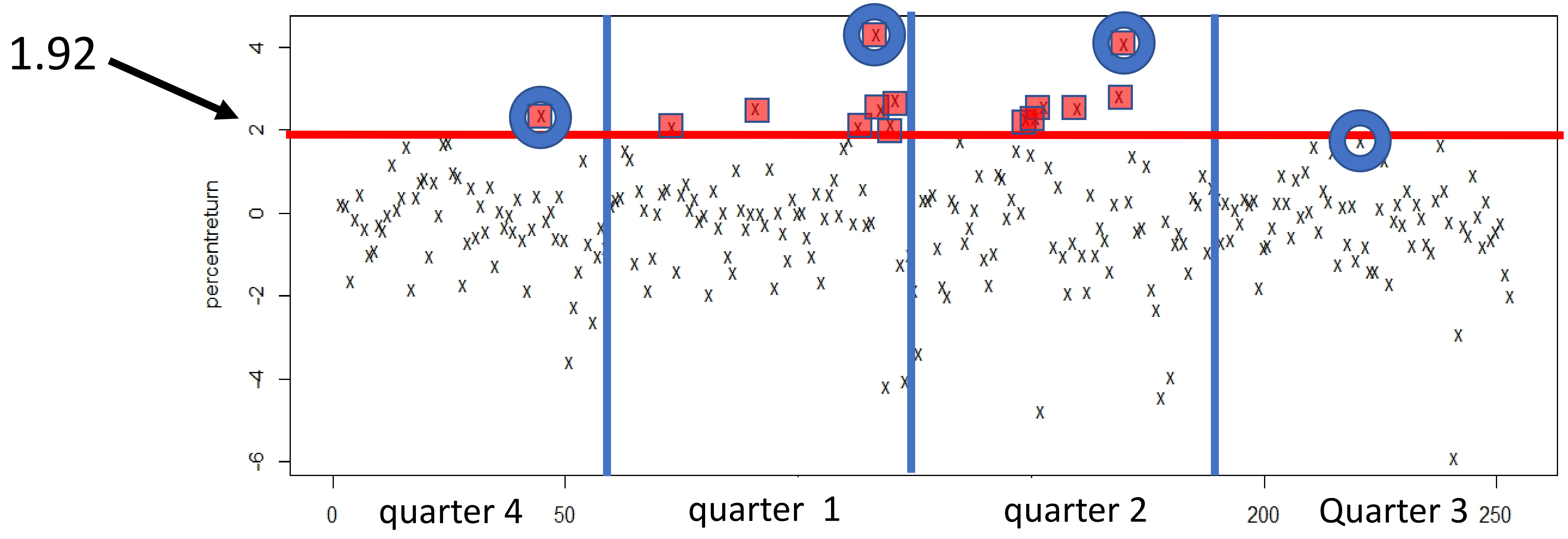
E. Gilleland: “Computing Software” as an aid in choosing what software to use

Links are available under “Anslag” in Canvas: remember that you need to hook up to Chalmers/GU with vpn when you want to follow the links

Extreme value statistics (EVS) is the branch of statistics developed to handle extreme risks

The philosophy of EVS is simple: extreme events, perhaps extreme water levels or extreme financial losses, are often quite different from ordinary everyday behavior, and ordinary behavior then has little to say about extremes, so that only other extreme events give useful information about future extreme events.

Apple losses (= - $100 \times \frac{\text{price tomorrow} - \text{price today}}{\text{price today}}$) one year back



Maximum quarterly loss   excess of the level $u = 1.92$

How large is the risk of a big quarterly loss? BM

How large is the risk of a big loss tomorrow? PoT

The block maxima method (Coles p. 45-53)

Obtain observations $x_1, \dots, x_n$ of block maxima (e.g. weekly or yearly maxima)

- Assume observations are i.i.d and have a GEV distribution
- Use $x, \dots, x_n$ to estimate the GEV parameters
- Use the fitted GEV to compute estimates and confidence intervals for, e.g., quantiles of yearly maximum distribution (= VaR) or of Expected Shortfall

The asymptotic motivation: What does

$$P\left(\tfrac{M_n-a_n}{b_n}\leq x\right)\to G(x),\ \ \text{as}\ \ n\to\infty\ \ \text{for all}\ \ x,$$

mean in practice? That $P\left(\frac{M_n-a_n}{b_n}\leq x\right)\approx G(x)$, for large n, or,

with $y=b_nx+a_n$ and $G(x)=\exp\{-(1+\gamma\frac{x-\mu'}{\sigma'}))^{-1/\gamma}\}$, that

$$\begin{aligned} P\big(M_n\leq y\big) &\approx G\Big(\frac{y-a_n}{b_n}\Big)=\exp\{-(1+\gamma\frac{y-(a_n+b_n\mu')}{b_n\sigma'})^{-1/\gamma}\} \\ &= \exp\{-(1+\gamma\frac{y-\mu}{\sigma})^{-1/\gamma}\},\ \text{for}\ \mu=a_n+b_n\mu',\ \sigma=b_n\sigma. \end{aligned}$$

Since all the parameters are unknown anyway, we are left with the problem of estimating μ,σ,γ from data, *i.e.* to use the Block Maxima method.

The Peaks over thresholds (PoT) method (Coles p. 74-91)

Choose a (high) threshold u and obtain observations of the N excesses of the threshold and of the times $t_1, \dots, t_N$ of exceedance

- Assume excesses are i.i.d and GP distributed and that $t_1, \dots, t_N$ are occurrence times of an independent Poisson process, so that N has a Poisson distribution
- Use the observed excesses to estimate the GP parameters and N to estimate the mean of the Poisson distribution
- Estimate tail $\bar{F}(x) = 1 - F(x) = \bar{F}(u)\bar{F}_u(x-u)$, where $\bar{F}_u(x-u)$ = the conditional distribution function of threshold excesses, by

$$\hat{\bar{F}}(x) = \frac{N}{n}\,\hat{\bar{F}}_u(x-u)$$

Why not just estimate $\bar{F}(x)$ by $N(x)/n$? Because if x is a very high level then $N(x)$ is small or zero, and then this estimator is useless -- and it is such very large x-es we are interested in.

Assume we have computed estimators $\hat{\sigma}, \widehat{\gamma}$ of the parameters in the GP distribution $\bar{F}_u(x) = H(x)$. Then

$$\hat{\bar{F}}(x) = \frac{N}{n} \hat{\bar{H}}(x-u) = \frac{N}{n}\left(1 + \frac{\hat{\gamma}}{\hat{\sigma}}(x-u)\right)_+^{-1/\widehat{\gamma}}$$

Solving $1 - \hat{\bar{F}}(x_p) = p$ for x_p we get an estimator of the p-th quantile of X:

$$x_p = u + \frac{\hat{\sigma}}{\hat{\gamma}}\left(\left(\frac{n}{N}(1-p)\right)^{-\widehat{\gamma}} - 1\right), \qquad \text{for } p > 1 - \frac{N}{n}$$

the distribution of Block Maxima can be obtained from a PoT model

- suppose that values larger than u occur as a Poisson process with intensity λ and that this process is independent of the sizes of the excesses
- suppose excesses are i.i.d. and have GP distribution $H(x) = 1 - \left(1 + \frac{\gamma}{\sigma}(x-u)\right)^{-1/\hat{\gamma}}$
- Let $M_T =$ the largest of the observations in the time interval $[0, T]$.

Then,

$$P(M_T \leq x) = \exp\left\{-\left(1 + \gamma\frac{x - u - \left((\lambda T)^{\gamma} - 1\right)\sigma/\gamma)}{\sigma(\lambda T)^{\gamma}}\right)^{-\frac{1}{\gamma}}\right\}, \qquad \text{for } x > 0$$

Maximum likelihood inference (Coles p. 30-34)

- parameters are estimated by (numerically) maximizing the likelihood function
- functions of parameters are estimated by plugging estimated parameters into the function
- simplest confidence intervals for parameters obtained from the inverse of the observed information matrix
- confidence intervals for functions of parameters, e.g. Var or ES, obtained from the delta method
- profile likelihood confidence intervals often more accurate
- likelihood ratio test to check if simpler models are OK, e.g. if one can use the same value of γ for two series of returns

Dependence: extremes come in small clusters (Coles p. 92-104)

- extremal index $\theta = 1/\text{"mean cluster length“}$
- typically $P(M_n \leq x) = F(x)^{\theta n}$ where M_n is the maximum of n variables and F(x) is the distribution function of one variable
- typically clusters approximately i.i.d. but with dependence within clusters
- typically tail of cluster maxima approximately same as tail of F(x) (strange but true)
- typically the GEV distributions the only possible limit distributions
- block maxima method often works just as for i.i.d observations

stationarity??

The PoT method for stationary time series

1. *Decluster:* identify approximately i.i.d clusters of large values by

a) Block method: divide observations up into blocks of a fixed length *r*, all values in a block which exceed the level *u* is a cluster

b) Blocks-runs method: the first cluster starts at first exceedance of *u* and contains all excesses of *u* within a fixed length *r* thereafter. The second cluster starts at the next exceedance of *u* and contains all excesses of *u* within *r* thereafter, and so on. . .

c) Runs method: the first cluster starts with the first exceedance of *u* and stops as soon as there is a value below *u,* the second cluster starts with the next exceedance of *u,* and so on …

2. $\hat{\theta} = \dfrac{\text{no. of clusters}}{\text{no. of exceedances}}$ estimate of the *extremal index*

3. ***PoT:*** Use standard i.i.d. PoT model, but with excesses replaced by cluster maxima, and excedance times replaced by the times when cluster maxima occur. (A bit of a miracle this works. Proof not given here.)

4. Use $P\big(M_n \leq x\big) \approx F(x)^{\theta n}$ to switch between block maxima and PoT or

5. Use formula for i.i.d. variables with excesses replaced by excesses by cluster maxima and the number of excesses replaced by the number of clusters

The i.i.d. formula: Suppose excesses are GP distributed and occur as a Poisson process which is independent of the sizes of excesses. Let M_T be the maximum in the interval $[0, T]$ and $x > 0$. Then

$$
\begin{aligned}
P\big(M_T \le u + x\big) &= \sum_{k=0}^{\infty} P\big(M_T \le u + x,\ \text{there are k exceedances in } [0,\ \mathrm{T}]\big) \\
&= \sum_{k=0}^{\infty} H(x)^k \frac{(\lambda T)^k}{k!} \exp\{-\lambda T\} \\
&= \sum_{k=0}^{\infty} (1 - (1 + \frac{\gamma}{\sigma} x)_+^{-1/\gamma})^k \frac{(\lambda T)^k}{k!} \exp\{-\lambda T\} \\
&= \exp\{(1 - (1 + \frac{\gamma}{\sigma} x)_+^{-1/\gamma})\lambda T\} \exp\{-\lambda T\} \\
&= \exp\{-(1 + \frac{\gamma}{\sigma} x)_+^{-1/\gamma} \lambda T\} \\
&= \exp\{-(1 + \gamma \frac{x - ((\lambda T)^\gamma - 1)\sigma/\gamma}{\sigma(\lambda T)^\gamma})_+^{-1/\gamma}\}
\end{aligned}
$$

Remember

If one does not understand the real-world situation well enough, the best quantitative tools will not help. Taleb's Turkey example: