ORF 505: Statistical Analysis of Financial Data

Tuesday Precept 2: Generalized Pareto Distributions (GPDs)

Sep 17, 2024

Lecturer: Qishuo Yin *Scribe: Qishuo Yin*

1 Generalized Pareto Distributions (GPDs)

The classical Pareto distribution is a distribution on the positive axis $[0, \infty)$ with density function:

$$f_\alpha(x) = \begin{cases} (1+\frac{x}{\alpha})^{-(1+\alpha)} = \frac{1}{(1+\frac{x}{\alpha})^{1+\alpha}} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

with a positive real number $\alpha > 0$.

In finance applications, we write $\xi = \frac{1}{\alpha}$ as the shape parameter of the distribution to get the *Ordinary Pareto Distribution*:

$$f_\xi(x) = \begin{cases} (1+\xi x)^{-(1+1/\xi)} = \frac{1}{(1+\xi x)^{1+1/\xi}} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then the cdf of the distribution will be written as:

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(x')dx' = \begin{cases} 1-(1+\xi x)^{-1/\xi} = 1 - \frac{1}{(1+\xi x)^{1/\xi}} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

If we change the location and the scale of the distribution, let's say we write $x' \leftarrow \frac{1}{\lambda}(x-m)$ with location parameter $m \in \mathbb{R}$ and scale parameter $\lambda > 0$, the density function $f_{m,\lambda,\xi}$ can be written as:

$$f_{m,\lambda,\xi}(x) = \begin{cases} \frac{1}{\lambda}\left(1+\frac{\xi}{\lambda}(x-m)\right)^{-(1+1/\xi)} & \text{if } x \geq m, \\ 0 & \text{otherwise.} \end{cases}$$

And the cdf of the distribution will be written as:

$$F_{m,\lambda,\xi}(x) = \begin{cases} 1-(1+\xi\frac{x-m}{\lambda}))^{-1/\xi} = 1 - \frac{1}{(1+\xi\frac{x-m}{\lambda})^{1/\xi}} & \text{if } x > m, \\ 0 & \text{otherwise.} \end{cases}$$

If you are looking for more general information about the Generalized Pareto Distributions (GPDs), please refer to section 2.1.2 of [1].

2 Problem 2.8 from Textbook

Problem 2.8. The goal of this problem is to highlight some of the properties of the estimates obtained with the command *fit.gpd* when fitting a GPD to a data sample $x_1, \cdots, x_n$. We assume that the distribution of the data has two tails (one extending to $-\infty$ and the other one to $+\infty$), and we are interested in understanding the effect of the choice of the thresholds lower and upper.

For your convenience, I remind you that a distribution with an upper tail is a GPD if its density $f(x)$ is well approximated in the tail by a function of the form:

$$f_{\lambda_+,m_+,\xi_+}(x) = \frac{1}{\lambda_+\xi_+}\left(1 + \frac{(x-m)}{\xi_+}\right)^{-(1+1/\xi_+)}$$

at least when $x > m_+$ for some large enough threshold m_+, where λ_+ is interpreted as a scale parameter, and where $\xi_+ > 0$ is called the shape parameter governing the size of the upper tail. If the distribution has a lower tail, one requires a similar behavior for $x < m_-$ for possibly different parameters m_-, λ_-, and ξ_-.

For the purpose of the problem, we assume that the true density of the sample $x_1, \cdots, x_n$ is given in figure 1. It is exactly equal to the function $f_{\xi_+,m_+,\lambda_+}(x)$ for $x > 2$ with $m_+ = 2$ and some value $\xi_+ > 0$ (to be estimated), and equal to the corresponding function $f_{\xi_-,m_-,\lambda_-}(x)$ for $x < 2$ with $m = 2$ and some $\lambda > 0$ (to be estimated as well).

1. What should you expect from the estimate $\widehat{\xi}_+$ given by the function *fit.gpd* if you use a threshold *upper*

1.1. Exactly equal to 2.
1.2. Greater than 5.
1.3. Between 0 and 1.

2. What should you expect from the estimate $\widehat{\xi}_-$ given by the function *fit.gpd* if you use a threshold *lower*

2.1. exactly equal to -2.
2.2. smaller than -8.
2.3. between 0 and -1.

and in each case, say how the estimate $\hat{\pi}_{0.01}$ of the one percentile compares to the true value $\pi_{0.01}$.

[Solution]

As hinted, ξ controls the rate of decay of the GPD distribution. For bigger (smaller) values of ξ the density decays slower (faster) and thus the tails become heavier (thinner).

1.1 $\hat{\xi}_+ \approx \xi_+$ since the actual and the chosen threshold coincide.

1.2 This case is the most difficult to predict. There is only one thing one can be sure of: the lack of points in the estimation, especially if the threshold is too large (i.e much larger than 5). So we should expect the estimator to be quite imprecise, and unstable with respect to any change. In some cases we will have $\hat{\xi}_+ < \xi_+$ but it could also be the case that $\hat{\xi}_+ > \xi_+$.

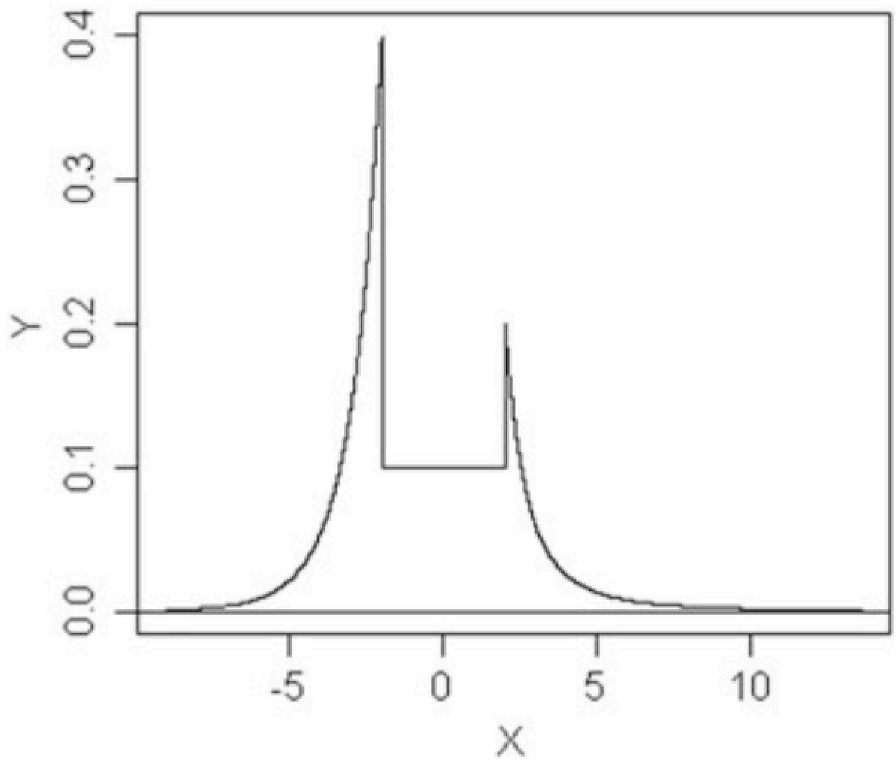

Figure 1: Density of the GPD from which the sample $x_1, \cdots, x_n$ is generated

1.3 $\hat{\xi}_+ < \xi_+$,. The peak in the density suggests that there should be a higher concentration of points around 2 in any sample from this distribution. So including these points in the tail will force the estimated density to decay toward 0 from a higher level, and it is reasonable to expect that the chosen value for the estimate of ξ_+ will be lower than what it should be (so $1/\xi_+$ will be larger and give a faster decay).

2. The conclusions for the estimate ξ_- are completely symmetric, so we do not repeat them. Concerning the estimate of the $1-$ percentile $\pi_{0.01}$ we have:

2.1 $\widehat{\pi}_{0.01} \approx \pi_{0.01}$ since the estimate should be giving the right value of the decay of the density.

2.2 $\widehat{\pi}_{0.01} < \pi_{0.01}$ if the left tail is thinner than the estimated one, and $\widehat{\pi}_{0.01} > \pi_{0.01}$, since the left tail is actually heavier than the estimated one.

2.3 $\widehat{\pi}_{0.01} > \pi_{0.01}$, since the left tail is actually heavier than the estimated one.

References

[1] R. Carmona. *Statistical analysis of financial data in R*, volume 2. Springer, 2014.