HW Problems for Assignment 4 - Part I
Due 6:30 PM Wednesday, November 11, 2020

1. (10 Points, 5 points each) GEV for Discrete Distributions. Here, we will see that the limiting statement $\lim_{n\to\infty} F(c_n x + d_n)^n = H(x)$ for some non-trivial cdf H is a bit more restrictive than it might seem.

(a) Let X be a discrete random variable taking values in the finite set $\{1, 2, ..., K\}$ for some fixed integer K. X has pmf

$$p(k) = \mathbb{P}[X = k] > 0, k = 1, ..., K$$

and hence cdf

$$F(x) = \sum_{k=1}^{\lfloor x \rfloor} p(k), \qquad \lfloor x \rfloor = \max\left\{ j \text{ an integer } \;\middle|\; j \le x \right\}.$$

Show that there are NO $c_n > 0, d_n$ such that $F(c_n x + d_n)^n \to H(x)$ for a non-degenerate cdf H. Here, by "non-degenerate" we mean that H does NOT take the form $H(x) = 1_{x \ge x_0}$ for some $x_0 \in \mathbb{R}$, as this cdf corresponds the non-random constant x_0.

(b) Now, assume X has pmf

$$p(k) = \int_k^{k+1} \frac{1}{y^2} dy = \frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}, \qquad k = 1, 2, \ldots.$$

Show that there is $c_n > 0, d_n$ such that $F(c_n x + d_n)^n \to H(x)$ for a GEV H. Identify the parameters ξ, μ, σ for the GEV distribution.

(c) **Extra Credit: 10 points** Assume $X \sim \text{Geom}(p)$ for $p < 1$ so that

$$p(k) = (1-p)^{k-1} p, \qquad k = 1, 2, \ldots.$$

Are there $c_n > 0, d_n$ and non-degenerate H such that $F(c_n x + d_n)^n \to H(x)$?

2. (10 Points) GP Distributions for $U(0,1)$ random variables. Let $X \sim U(0,1)$. Find a function $\beta(u)$ so that

$$\lim_{u \uparrow 1} \sup_{0 \le x \le 1-u} \left| F_u(x) - G_{-1,\beta(u)}(x) \right| = 0.$$

3. (30 Points) Could we have predicted the 1987 crash using extreme value theory? The data file

"SP500_Log_Returns.csv"

contains daily log return data for the S%P 500 index from 6/10/1960 until 10/16/1987. Specifically, column 1 stores the date in numeric format (oldest to newest) and column 2 the log return.

As you may recall, 10/16/1987 was the day before the famous crash in the markets. In this exercise we will see if our extreme value theory-based risk measures could have predicted the crash.

Consider a hypothetical portfolio of $\$1,000,000$ in the S%P 500 index. The losses are estimated via the linearized returns so that $L' = -VX$ where V is the portfolio value, and X is the log return. For each of the methods described below, determine the VaR_α, as a function of α, for the portfolio losses.

(1) Using the empirical distribution for the log returns.

(2) Assuming that as of $t = 10/16/1987$ the log returns $X_{t+\Delta}$ over the next business day are normally distributed with mean $\mu_{t+\Delta}$ and variance $\sigma^2_{t+\Delta}$. To estimate $\mu_{t+\Delta}, \sigma^2_{t+\Delta}$ use an EWMA procedure with parameter $\lambda = \theta = .97$ and a 500 day initialization time (as we have done before).

(3) Assuming a GEV distribution for the maximum and the block maximum method. For the N days of data ($N = 6875$), break the data into blocks of $n = 125$ days for a total of $m = 55$ blocks. Estimate the resultant GEV parameters: in matlab this is done using the 'gevfit' command.

(4) Assume a GP excess loss distribution. Here, take $u = \mathrm{VaR}_{.95}$ as estimated in (1) above. To estimate the GP parameters, you will have to select only those losses above the u threshold, and compute the loss minus u for these selected days. With this data you can then fit the GP distribution. For example, in matlab this is the 'gpfit' command. With the fitted parameters, estimate the VaR_α in terms of α and $F(u)$, which in this case is .95 by construction.

For each of the four methods, produce a plot of $\alpha \mapsto \mathrm{VaR}_\alpha$ for α between .99 and .9999 in increments of .000099 (100 values). For $\alpha = .9999$ what are the four VaR_α values? Which is highest?

The actual log return over Monday October 19th, 1987 was $X = -0.099452258$, so that our linearized losses would have been $L' = \$99,452$. Could any of the above methods predicted such a loss via the VaR_α? If so, for what α?