HW Problems for Assignment 4 - Part 2
Due 6:30 PM Tuesday, November 9, 2021

1. (10 Points) GEV Distributions for Finite Discrete Distributions. 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

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

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 $H(x) = 1_{x \geq x_0}$ for some $x_0 \in \mathbb{R}$, as this cdf corresponds the non-random constant x_0. Thus, the limiting statement $\lim_{n\to\infty} F(c_n x + d_n)^n = H(x)$ for some non-trivial cdf H does not apply to discrete random variables taking a finite number of values.

2. (20 Points, 10 points each part) GP Distributions for $U(0,1)$ and $N(0,1)$ random variables.

(a) Let $X \sim U(0,1)$. Find a function $\beta(u)$ so that

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

(b) Let $X \sim N(0,1)$. Show that for $\beta(u) = 1/u$ and $\delta > 0$ that

$$\lim_{u \uparrow \infty} \sup_{\delta \leq x < \infty} \left| F_u(x) - G_{0,\beta(u)}(x) \right| = 0.$$

In fact, the above also holds taking the maximum over $[0, \infty)$, rather than $[\delta, \infty)$, but this is very technical to show.

Hint: To answer (b) above you may use without proof that for $x > 0$:

$$\frac{x}{1+x^2} e^{-x^2/2} \leq \int_x^\infty e^{-y^2/2} dy \leq \frac{1}{x} e^{-x^2/2}.$$

3. (20 Points). How bad can things get? The file

"SP500ClosingPrices.xlsx"

contains daily closing prices for the S%P 500 index from 11/1/2011 until 11/1/2021. Column 1 stores the date and column 2 the log closing price. Data is sorted oldest to newest.

In this exercise we will what our extreme value theory-based risk measures have to say about the potential for a crash.

Consider a hypothetical portfolio where each day we start with $\$1,000,000$

in the S%P 500 index. The daily losses are estimated using linearized returns so that $L^{lin} = -VX$ where V is the portfolio value, and X is the daily 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 100 day initialization time.

(3) Assuming a GEV distribution for the maximum and the block maximum method. For the $N + 1 = 2556$ days of data (which generate $N = 2555$ log returns), break the data into blocks of size (i) $n = 73$ days for a total of $m = 35$ blocks and (ii) $n = 35$ days for a total of $m = 73$ blocks. Estimate the resultant GEV parameters in each case: 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 method, 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 theVaR_α values? Which is highest? How do the two EVT estimates compare (is there a meaningful difference)?