Notes on the Application of Extreme Value Theory

16 September 2002

This note describes how to apply certain results from extreme value statistical theory (EVT) to problems arising at CFSI. The basic idea is the following: If one is concerned with the tail of a random variable's distribution, then more robust predictions can be obtained by fitting an appropriate tail distribution function to just the tail of the data available. By more robust we mean that predictions are unaffected by mis-specifying the true distribution from which the observations are drawn.

1. Statistical theory

The result we use is a theorem of Pickands (1975) and Balkema and de Haan (1974) that under moderately general conditions the distribution of the *excess* $y \equiv X - u$ of a random variable X over a high threshold u, conditional on the threshold being exceeded, converges to the generalized Pareto distribution (GPD).[1] The GPD has two parameters, k and σ. Its cumulative distribution function is

$$G(y;\sigma,k) = \begin{cases} 1-(1-ky/\sigma)^{1/k} & k \neq 0 \\ 1-e^{-y/\sigma} & k = 0 \end{cases} \tag{1}$$

For $k > 0$, the above applies for $0 \leq y < \sigma/k$, with $G = 1$ for $y \geq \sigma/k$ and $G = 0$ for $y < 0$ (i.e., the range of y is bounded, as when sampling from a uniform distribution). The case $k = 0$ will occurs if sampling is from normal, lognormal and several other common distributions. The case $k < 0$ occurs if sampling from fatter tailed distributions.[2] Differentiating G with respect to y gives the probability density function[3]

$$g(y;\sigma,k) = \frac{1}{\sigma}(1-ky/\sigma)^{-1+1/k} \tag{2}$$

in the relevant domain. This forms the basis for estimation of k, σ in the next section.

A result useful in applications is that the *mean excess* of X over any specified higher threshold $L > u$ is[4]

$$\mathbf{E}[\, X - L \mid X > L \,] = \frac{\sigma - kL}{1+k} \tag{3}$$

[1] This convergence takes place as the threshold u increases toward the right endpoint of the support of X (which may be infinite).

[2] There does not seem to be uniform conventions in how the parameters of the GPD are expressed. I have here followed Neftci (2000). Embrechts et al (1997) use β in place of σ and ξ in place of $-k$.

[3] The pdf is for $y \geq 0$ when $k \leq 0$, zero otherwise; and for $0 \leq y < \sigma/k$ when $k > 0$, zero otherwise.

[4] See Embrechts et al (1997, p.165) Theorem 3.4.13(e).

2. Estimation of tail distribution parameters

Suppose we have a large set of N independent observations of X. Choose a tail estimation threshold u large enough so that the observations exceeding u are credibly in the upper tail, but not so large that the number of observations is insufficient to confidently identify k, σ. Let n be the number of observations exceeding u. There is a tradeoff here between bias, from not being sufficiently far in the tail for the asymptotic distribution result to hold, and variance, from having too few observations to fit k, σ. A good starting point would be to choose u such that n/N lay between 1% and 5%. A careful illustration of how one might analyze the data is provided by McNeil (1996).

Let X_i, $i = 1 \ldots n$ denote the observations above the u selected, and $y_i = X_i - u$ the corresponding exceedances. Using the iid assumption, the log of the likelihood function for this data is the sum of the log density functions from (2):[5]

$$\mathcal{L} = -n \ln \sigma - (1 - \frac{1}{k}) \sum_{i=1}^{n} \ln(1 - k y_i / \sigma) \tag{4}$$

Maximum likelihood estimates are the values of k, σ maximizing this expression.[6] First order conditions for a maximum of $\mathcal{L}$ with respect to k, σ are

$$\mathcal{L}_\sigma = 0 = -n/\sigma - (1 - \frac{1}{k}) \sum_i \frac{k y_i / \sigma^2}{1 - k y_i / \sigma} \tag{5}$$

$$\mathcal{L}_k = 0 = -\frac{1}{k^2} \sum_i \ln(1 - k y_i / \sigma) + (1 - \frac{1}{k}) \sum_i \frac{y_i / \sigma}{1 - k y_i / \sigma} \tag{6}$$

Define the new variable $c \equiv k/\sigma$. Substitute into the first order conditions. The pair of equations can be partially solved to

$$k = -\frac{1}{n} \sum_i \ln(1 - c y_i) \tag{7}$$

$$0 = n + \left(1 + \frac{n}{\sum_i \ln(1 - c y_i)}\right) \sum_i \frac{c y_i}{1 - c y_i} \tag{8}$$

Note that the last equation contains only the single unknown c. Although not analytically solvable, it is readily numerically solvable for $\hat{c}$ (e.g., by Newton's method or the secant method). Then $\hat{k}$ can be computed from (7), and $\hat{\sigma} = \hat{k}/\hat{c}$.

3. Mean option payoffs

Suppose our object of interest is the expected value of a call option-like payout, with X as underlying and strike price $L > u$ in the upper tail of the distribution of possibilities. I.e.,

[5] $\mathcal{L}$ is $-\infty$ for $\max y_i \geq \sigma/k$ when $k > 0$.

[6] Embrechts et al (1997 p.356) state this is valid only for $k < 1/2$.

we wish an estimate of

$$V \equiv \mathbf{E}[\ \max\{\ 0\ ,\ X - L\ \}\ ] \tag{9}$$

for a single X draw. Proceed as follows. From our total data set and the iid assumption, an unbiased estimate of the probability that X exceeds u is given by

$$\mathbf{p}(X > u)\ =\ \frac{n}{N} \tag{10}$$

Next, the probability that X exceeds L *conditional* on it exceeding u is one minus the estimated GPD cumulative distribution function $G(L - u, \hat{\sigma}, \hat{k})$.

$$\mathbf{p}(X > L \,|\, X > u)\ =\ (1 - k(L - u)/\sigma)^{1/k} \tag{11}$$

Finally, the expected value of $X - L$ conditional on X exceeding L is given by equation (3). The expectation V is the product of these three terms. Hence

$$\hat{V}\ =\ \frac{n}{N}\,\frac{\sigma + ku}{1 - k}\,(1 - k(L - u)/\sigma)^{1/k} \tag{12}$$

Note that the probability that the option makes a non-zero payoff at all is the product of the first two expressions (10) and (11) alone.

References and Bibliography

Davison, A.C., and R. L. Smith, "Models for Exceedences of High Thresholds," *Journal of the Royal Statistical Society*, 1990.

Embrechts, Paul, Claudia Kluppelberg, and Thomas Mikosh, *Modelling Extremal Events*, Berlin: Springer-Verlag, 1997.

Kotz, Samuel, and Saralees Nadarajah, *Extreme Value Distributions*, London: Imperial College Press, 2000.

McNeil, Alexander J., "Estimating the Tails of Loss Severity Distributions using Extreme Value Theory," mimeo, Department of Mathematics, ETH Zentrum (Zurich), December 1996.

Neftci, Salih N., "Value at Risk Calculations, Extreme Events, and Tail Estimation," *Journal of Derivatives*, 7 (Spring 2000), 23-38.

Pickands, James, "Statistical Inference Using Extreme Order Statistics," *Annals of Statistics*, 3 (January 1975), 119-131.

Smith, R. L., "Estimating Tails of Probability Distributions," *Annals of Statistics*, 1987.