Extreme Value Models

This document addresses continuous random variables that represent the maximum of several outcomes of another continuous random variable. For the sake of the following derivations, consider a continuous random variable Z. In any one realization the probability that the outcome is less than z is the CDF, $F_Z(z)$. Now consider a situation where there are n independent realizations of the random variable. In this situation it is often the maximum realization that is of interest, perhaps because it represents a loading value that may exceed some capacity. To address the probability that the maximum realization will not exceed a threshold, x, first define X as a new random variable, where

$$X = \max\left(Z_1, Z_2, \cdots, Z_n\right) \tag{1}$$

Under the condition that the random variables Z_i are independent and identically distributed the probability that the maximum X does not exceed the threshold x is

$$\begin{aligned} F_X(x) &= P\left(X \le x\right) \\ &= P\left(Z_1 \le x \cap Z_2 \le x \cap \cdots \cap Z_n \le x\right) \\ &= P\left(Z_1 \le x\right) \cdot P\left(Z_2 \le x\right) \cdots P\left(Z_n \le x\right) \\ &= F_Z(x)^n \end{aligned} \tag{2}$$

where it is emphasized that F_Z is the CDF for the original random variable Z. This is the CDF for the extreme value of Z in n "experiments." The corresponding PDF is obtained by differentiation:

$$\begin{aligned} f_X(x) &= \frac{d}{dx} F_X(x) \\ &= \frac{d}{dx} F_Z(x)^n \\ &= n \cdot F_Z(x)^{n-1} \cdot f_Z(x) \end{aligned} \tag{3}$$

where it is reiterated that it is the CDF and PDF of the original random variable Z that enter the expression.

When the number of experiments is large then the distribution for the extreme value, X, is mostly dependent on the tail behaviour of the underlying probability distribution for Z. It is rather insensitive to the overall behaviour of the actual underlying probability distribution. For these situations several asymptotic extreme-value distributions are developed. They cannot be synthesized into one distribution because the result is different for minimum and maximum values. Furthermore, the result is different for different types of tail-behaviour in the underlying probability distribution.

Common Distribution Types

Type I Distributions (Gumbel)

This distribution addresses the maximum value of many experiments. The "Type I" assumption is that the tail of the underlying distribution varies exponentially:

$$F_Z(z) = 1 - \exp(-h(z)) \tag{4}$$

The tails are unbounded. This type of tail is found in the normal, exponential, and gamma distributions. Application of this underlying tail distribution in extreme value theory yields the Type I Largest and Type I Smallest distributions, for the maximum and minimum of many realizations, respectively. The resulting distributions are named after Gumbel:

$$F(x) = \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) \tag{5}$$

Type II Distribution (Frechet)

Here the left tail of the underlying distribution is bounded at zero, while the upper tail varies according to

$$F_Z(z) = 1 - c_1 \cdot \left(\frac{1}{z}\right)^{c_2} \tag{6}$$

where c_1 and c_2 are constants. The derived Type II extreme value distribution is named after Frechet:

$$F(x) = \exp\left(-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}\right) \quad x>\mu \tag{7}$$

Type III Distribution (Weibull)

The fundamental assumption for Type III distributions is that the tail of the underlying random variable is bounded by a value z_0:

$$F_Z(z) = 1 - c_1 \cdot (z_0 - z)^{c_2} \tag{8}$$

This results in the extreme value distributions named after Weibull. A simple sign-change in Frechet's distribution yields the reversed Type III Weibull distribution:

$$F(x) = \exp\left(-\left(-\frac{\mu-x}{\sigma}\right)^{\alpha}\right) \quad x<\mu \tag{9}$$

The Weibull distribution is written:

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^{k}\right) \quad \text{for } x\geq 0 \tag{10}$$

$$f(x) = \frac{k}{\lambda} \cdot \left(\frac{x}{\lambda} \right)^{k-1} \cdot \exp\left(-\left(\frac{x}{\lambda} \right)^{k} \right) \qquad \text{for } x \geq 0 \tag{11}$$

where $k>0$ is the shape parameter and $\lambda>0$ is the scale parameter.

Generalized Extreme Value Distribution

$$F(x) = \exp\left(-\left(1 + \xi \cdot \left(\frac{x-\mu}{\sigma} \right) \right)^{-1/\xi} \right) \tag{12}$$

where μ is the location parameter, σ is the scale parameter, and ξ is the shape parameter.