Extreme Value Theory and Estimation of Quantiles

Jussi Klemelä

October 10, 2014

1 Block Maxima Method

1.1 Limit Distributions of the Maximum

Let the real valued observations $X_1, \ldots, X_n$ be independent and identically distributed, and denote the maximum

$$M_n = \max\{X_1, \ldots, X_n\}.$$

We assume that there exists sequences (c_n) and (d_n) where $c_n > 0$ and $d_n \in \mathbf{R}$ so that

$$P\left(\frac{M_n - d_n}{c_n} \le x\right) \xrightarrow{d} H_\xi(x), \tag{1}$$

for all $x \in \mathbf{R}$, as $n \to \infty$, where H_ξ is a distribution function. The Fisher-Tippet-Gnedenko theorem states that if the convergence in (1) holds, then H_ξ can only be a Fréchet, Weibull, or Gumbel distribution. See Fisher and Tippett (1928), Gnedenko (1943), and Embrechts, Klüppelberg and Mikosch (1997, page 121). The Fréchet distribution functions are

$$\Phi_\alpha(x) = \begin{cases} 0, & x \le 0, \\ \exp\{-x^{-\alpha}\}, & x > 0, \end{cases} \tag{2}$$

where $\alpha > 0$. The Weibull distribution functions are

$$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^{\alpha}\}, & x \le 0, \\ 1, & x > 0, \end{cases}$$

where $\alpha > 0$. The Gumbel distribution function is

$$\Lambda_\alpha(x) = \exp\{-e^{-x}\}, \qquad x \in \mathbf{R}.$$

These distributions are called the extreme value distributions. Define

$$H_\xi = \begin{cases} \Phi_{1/\xi}, & \text{if } \xi > 0, \\ \Lambda, & \text{if } \xi = 0, \\ \Psi_{-1/\xi}, & \text{if } \xi > 0. \end{cases}$$

Then,

$$H_\xi(x) = \begin{cases} \exp\left\{-(1-\xi x)^{-1/\xi}\right\}, & \xi \neq 0, \\ \exp\left\{-e^{-x}\right\}, & \xi = 0, \end{cases}$$

where $1 + \xi x > 0$. This is called the Jenkinson-von Mises representation of the extreme value distributions or the generalized extreme value distribution; see Embrechts et al. (1997, page 152).

If the distribution which generated the observations $X_1, \ldots, X_n$ has polynomial tails, then (1) holds and the limit distribution of the maximum belongs to the Fréchet class. If the distribution which generated the observations is exponential, normal, or log-normal, then (1) holds and the limit distribution of the maximum is the Gumbel distribution. See Embrechts et al. (1997, pages 131, 135, 145).

1.2 An Expression for the Quantiles

When

$$P\left(\frac{M_n - d_n}{c_n} \leq x\right) \approx H_\xi(x),$$

as suggested by (1), then

$$P\left(M_n \leq x\right) \approx H_\xi\left(\frac{x - d_n}{c_n}\right). \tag{3}$$

Since $X_1, \ldots, X_n$ is an i.i.d. sample from the distribution of X,

$$\begin{aligned} P(M_n \leq x) &= P(X_1 \leq x, \ldots, X_n \leq x) \\ &= P(X_1 \leq x) \cdots P(X_n \leq x) \\ &= [P(X \leq x)]^n. \end{aligned} \tag{4}$$

Let $q_p = Q_p(X)$ be the quantile. Then,

$$P(X \leq q_p) = p, \tag{5}$$

when X has a continuous distribution. Thus, combining (4) and (5),

$$P(M_n \leq q_p) = [P(X \leq q_p)]^n = p^n. \tag{6}$$

Thus, combining (3) and (6),

$$H_\xi\left(\frac{q_p - d_n}{c_n}\right) \approx p^n$$

and we get

$$Q_p(X) \approx d_n + c_n H_\xi^{-1}(p^n). \tag{7}$$

1.3 Estimation of the Parameters

The expression (7) for a quantile contains unknown parameters ξ, d_n, and c_n, which we have to estimate. Let us denote

$$H_{\xi,\mu,\sigma}(x) = H_\xi\left(\frac{x-\mu}{\sigma}\right).$$

We consider the family of distributions $H_{\xi,\mu,\sigma}$, where ξ is the shape parameter, μ is the location parameter, and σ is the scale parameter. The parameters can be estimated using the block maxima method. Let $X_1, \ldots, X_T$ be i.i.d. observations. Since (3) holds, we could estimate the parameters if we would have several observations of the maxima. This can be achieved when we divide the observations into m blocks of size n, assuming for simplicity that $T = nm$. Denote by M_{ni}, $i = 1, \ldots, m$, the maximum of the ith block:

$$M_{ni} = \max\{X_{(i-1)n+1}, \ldots, X_{in}\}, \qquad i = 1, \ldots, m.$$

The maxima $M_{n1}, \ldots, M_{nm}$ are independent[1] and we define the likelihood function

$$L\left(\xi, \mu, \sigma; M_{n1}, \ldots, M_{nm}\right) = \prod_{i=1}^{m} H'_{\xi,\mu,\sigma}(M_{ni}),$$

where $H'_{\xi,\mu,\sigma}$ is the density function corresponding to the distribution function $H_{\xi,\mu,\sigma}$. We define the estimators $\hat{\xi}$, $\hat{\mu}$, and $\hat{\sigma}$ to be maximizers of the likelihood function. From (7) we get the estimator for a quantile

$$\hat{Q}_p(X) = \hat{\mu} + \hat{\sigma} H_{\hat{\xi}}^{-1}(p^n). \tag{8}$$

Note that in (8) the sample size is T and n is the block size.

For the calculation of the maximum likelihood estimator we need the density function. For example, the density of the Fréchet distribution in (2) is

$$\phi_\alpha(x) = \alpha x^{-(\alpha+1)} e^{-x^{-\alpha}},$$

where $x > 0$.

[1] Even when the original observations are not independent, the block maxima are approximately independent, for large block sizes.

2 Threshold Exceedances

2.1 Limit Distributions for Excess Distribution

Let $X \in \mathbf{R}$ be a random variable. We define the excess distribution with threshold u as the distribution with the distribution function

$$F_u(x) = P(X - u \le x \mid X > u) = \frac{F_X(x+u) - F_X(u)}{1 - F_X(u)}.$$

We can typically approximate the distribution function F_u with the distribution function of a generalized Pareto distribution. This follows from the Pickands-Balkema-de Haan theorem; see Embrechts et al. (1997, page 158). The distribution function of the generalized Pareto distribution is

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp\{-x/\beta\}, & \xi = 0, \end{cases}$$

where $\beta > 0$, $x \ge 0$, when $\xi \ge 0$, and $0 \le x \le -\beta/\xi$, when $\xi > 0$. Parameter ξ is a shape parameter and parameter β is a scale parameter.[2]

2.2 An Expression for the Quantiles

If the quantile satisfies $Q_p(X) \ge u$, then

$$Q_p(X) = u + F_u^{-1}\left(1 - \frac{1-p}{P(X > u)}\right). \tag{9}$$

Indeed, let $x \ge u$. Now

$$P(X > x \mid X > u) = \frac{P(X > x)}{P(X > u)}.$$

Thus,

$$\begin{aligned} P(X > x) &= P(X > u)\, P(X > x \mid X > u) \\ &= P(X > u)\, P(X - u > x - u \mid X > u) \\ &= P(X > u)\, [1 - F_u(x - u)]. \end{aligned}$$

[2]Note that sometimes the distribution function of the Pareto distribution is defined by $F(x) = 1 - (\kappa/(\kappa + x))^\alpha$ for $x \ge 0$, where $\alpha > 0$, $\kappa > 0$. Sometimes the distribution function of the Pareto distribution is defined by $F(x) = 1 - (c/x)^\alpha$ for $x \ge c$, where $\alpha > 0$, $c > 0$.

In particular,

$$P(X > Q_p(X)) = P(X > u)\left[1 - F_u(Q_p(X) - u)\right].$$

If the distribution of X is continuous, then

$$P(X > Q_p(X)) = 1 - p.$$

Combining the two previous equations gives the equation

$$1 - p = P(X > u)\left[1 - F_u(Q_p(X) - u)\right].$$

Solving this equation gives the result (9).

2.3 Estimation of the Parameters

Let $X_1, \ldots, X_n$ be an i.i.d. sample from the distribution of X and let

$$N_u = \#\{X_i > u\}.$$

Now we can estimate

$$P(X > u) = \frac{N_u}{n}.$$

Then we estimate the parameters ξ and β of the generalized Pareto distribution with the maximum likelihood method. Let us denote

$$\{Y_1, \ldots, Y_{N_u}\} = \{X_i : X_i > u\}.$$

Now $Y_1, \ldots, Y_{N_u}$ is a sample from the distribution F_u. Since $F_u \approx G_{\xi,\beta}$, we define the likelihood function

$$L(\xi, \beta; Y_1, \ldots, Y_{N_u}) = \prod_{i=1}^{N_u} g_{\beta,\xi}(Y_i),$$

where $g_{\beta,\xi} = G'_{\beta,\xi}$ is the density function of the generalized Pareto distribution. Define $\hat{\xi}$ and $\hat{\beta}$ as the maximizers of the likelihood function. We can choose parameter u as

$$u = \hat{F}^{-1}(p'),$$

where $p' < p$ and $\hat{F}$ is the empirical distribution function. This amounts to choosing u so that

$$\frac{N_u}{n} = 1 - p'.$$

Then the quantile estimator is

$$\hat{Q}_p(X) = u + F^{-1}_{\hat{\xi},\hat{\beta}}\left(1 - \frac{1-p}{N_u/n}\right).$$

References

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997), *Modelling Extremal Events*, Springer, Berlin.

Fisher, R. A. and Tippett, L. H. C. (1928), ‘Limiting forms of the frequency distribution of the largest or smallest member of a sample’, *Proc. Cambridge Philos. Soc.* **24**, 180–190.

Gnedenko, B. V. (1943), ‘Sur la distribution limité du terme d’une série aléatoire’, *Ann. Math.* **44**, 423–453.