



# Statistical Theory of Extremes

Homepage: <http://www.gathacognition.com/book/gcb14>  
<http://dx.doi.org/10.21523/gcb14>



## Part 2

Statistics for Univariate Extremes

## Chapter 4

# A Quick Exploration of Extreme Data

**J. Tiago de Oliveira**

Academia das Ciências de Lisboa (Lisbon Academy of Sciences), Lisbon, Portugal.

### Abstract

The chapter is devoted to quick exploration techniques on how statistically decide for one of the three models, Fréchet, Gumbel or Weibull, assuming a distribution that is one of these three forms. First, it is considered a model from previous knowledge. Afterwards, the technique of the probability paper for Gumbel distribution is explored for quick statistical choice between Weibull, Gumbel and Fréchet distribution. A simple statistical choice between the extreme models is presented (trilemma). Moreover, short illustrating examples of quick exploration of data are considered.

### Published Online

23 June 2017

### Keywords

Extreme values trilemma;  
Probability paper;  
Statistical choice.

### Editor(s)

J.C. Tiago de Oliveira

## 4.1 Introduction

Statistical decision for univariate extremes is currently at a moderate stage of development. Some problems have been effectively solved, but not all of the essential ones. On the way, many problems to be solved will appear, constituting a happy hunting ground for researchers. Curiously, even, Bayesian methods have not been strongly developed for statistical extremes, except essentially for some Reliability questions.

The analysis that follows has the following steps: once it is *assumed* that the useful distribution is of one of the three forms — Fréchet, Gumbel or Weibull — we will discuss statistical decision using the *assumed* distribution.

Originally published in 'Statistical Analysis of Extremes', 1997, 2016

<http://dx.doi.org/10.21523/gcb1.1709>

© 2017 GATHA COGNITION® All rights reserved.

This chapter will be devoted to quick exploration techniques, or even as a rough guide on how to assume one of the three models. The last chapter will deal with optimal analytical techniques, comparing them, as far as possible, with the quick ones.

Note that as a matter of practical convenience we will deal with statistical decision for Gumbel and Fréchet distributions for maxima and for Weibull distribution for minima, as the latter is very important in Reliability.

The statistical techniques presented, in the all but the two last chapters, deal with complete samples and an assumed distribution; the penultimate chapter will deal with statistical choice of models and the last one with some cases of incomplete information.

Clearly the Gumbel distribution for maxima with  $\delta(> 0)$  known, the Fréchet distributions for maxima with  $\lambda$  and  $\alpha(> 0)$  known, and the Weibull distributions for minima with  $\lambda$  and  $\alpha(> 0)$  known are easily reduced to the exponential distribution (only with a scale parameter) by evident transformations, and the whole technique of the exponential distribution can be transformed to these special distributions of extremes.

## 4.2 Assumption of a model from previous knowledge

In fact, this section could also be called “assumption by tradition”. This essentially involves a variable mixture: a convenient usage of accumulated experimentation, of limiting results to be considered as being finitely capable of fitting data, as well as some intuition. The situation is completely analogous to the assumption of normal (multinormal) distribution, where the Central Limit Theorem plays an important role as does long-term accumulated experience with measurements (\*), to the assumption of binomial or Bernoulli distribution for use in repeatable trials with an approximately stable probability of the event under consideration, to the assumption of Poisson either as behaviour of small (random) numbers or as a limit of binomial distribution (\*\*), to the assumption of the exponential distribution in the (first)

---

(\*) Recall that Poincaré (a mathematician) said Lippman (a physicist) told him that physicists used normal distribution supposing it to be a theorem, while mathematicians studied it supposing it to be experimentally verified.

(\*\*) Recall the justification of the Poisson distribution used by “Student” for the counts of particles in a haemacytometers.

analysis of inter-arrival times (thus implying the Poisson distribution for the associated counting process), to the assumption of geometric, inverse binomial or Pascal distributions in stopping time experiments with the event probabilities approximately stable, etc... The situation here is very similar.

Let us suppose that we are dealing with maxima (or dually with minima) of i.i.d. samples, with the same size  $n$ , and underlying distribution  $F(\cdot)$ .

Suppose that  $F(\cdot) \in \mathcal{D}(\tilde{L}(\cdot))$ , that is  $F(\cdot)$  belongs to domain of attraction of  $\tilde{L}(\cdot)$  (Weibull, Gumbel or Fréchet distributions), i.e., there exist sequences  $\{\lambda_n, \delta_n; \delta_n > 0\}$  such that  $F^n(\lambda_n + \delta_n x) \rightarrow \tilde{L}(x)$ , the convergence being at every point of  $\mathbb{R}$  by the continuity of  $\tilde{L}(\cdot)$ , and the sequences not being unique, as said before.

Then as convergence is uniform, i.e.

$$\sup_x |F^n(\lambda_n + \delta_n x) - L(x)| = \sup_y |F^n(y) - \tilde{L}(\frac{y - \lambda_n}{\delta_n})| \rightarrow 0,$$

for large  $n$ , we can assume that there exists  $(\lambda, \delta; \delta > 0)$  such that  $F^n(y) \approx \tilde{L}(\frac{y - \lambda}{\delta})$ .

Also  $\tilde{L}(\cdot)$  satisfies the stability equation  $\tilde{L}^t(x) = \tilde{L}(\alpha_t + \beta_t x)$ ,  $\beta_t > 0$ .

The heuristic reasoning used by [Fisher and Tippett \(1927/28, p.180\)](#), shows *approximate* stability of maxima.

Thus, with the uniformity of convergence and the (approximate) stability of maxima, we can use the approximation  $F^n(y) \approx \tilde{L}((y - \lambda)/\delta)$ .

But what happens when  $F(\cdot)$  is one of those “awkward” distributions that are not attracted to one of the  $\tilde{L}$ ’s? This, in general, is a consequence of high tails (the right-tail for maxima), and the use of “slowing down” transformations as the  $\log(\cdot)$  or  $\log \log(\cdot)$  can be useful; recall the example given in Part-1.

Note that the approximation to  $F^n(y)$  depends on the size of the sample ( $n$ ); accumulated experience and simulation can give some idea when the fit is good or even reasonable.

It should be noted that in the heuristics given before we assumed i.i.d. samples of the same size. If the samples are i.i.d. but not of the same size, a not uncommon situation, a correction can be made and the data fit analysed.

*Essentially if two samples are of sizes  $n$  and  $n'$  we have under general conditions:*

$\lambda_{n'} = \lambda_n + \delta_n \log \frac{n'}{n}, \delta_{n'} = \delta_n$       *for the Gumbel distribution,*

$\lambda_{n'} = \lambda_n = 0, \delta_{n'} = \delta_n (n'/n)^{1/\alpha}$       *for the Fréchet distribution and*

$\lambda_{n'} = \lambda_n = \bar{w}, \delta_{n'} = \delta_n (n/n')^{1/\alpha}$       *for the Weibull distribution*

*for maxima; see [Tiago de Oliveira \(1977\)](#) and [Ramachandran \(1974\)](#). If we have more than two samples we will shift their parameters to those of a sample of average size.*

If the i. or i. d. on both conditions fail we have seen (Part-1) that usually the asymptotic results are still valid and the same approximating technique can be used; but, in general, the sample sizes should be larger chiefly when correlation is positive.

Some rules of thumb, integrating accumulated experience, can be suggested.

**Table 4.1**

For maxima:	
Floods	Gumbel and Fréchet distribution; see <a href="#">Gumbel (1954)</a> , (1958) and <a href="#">Thom (1960)</a>
Extreme winds	Gumbel and Fréchet distribution; see <a href="#">Thom (1960)</a>
Largest waves	Fréchet distribution; see <a href="#">Thom (1971)</a>
Longest lives	Gumbel distribution; see <a href="#">Gumbel (1958)</a> .
For minima:	
Droughts	Weibull distribution; See <a href="#">Gumbel (1958)</a> and (1962)
Breaking strength of materials	Weibull distribution; See <a href="#">Freudhental and Gumbel (1953)</a> and <a href="#">Weibull (1939)</a> .

As a consequence, and with these guidelines coming from accumulated experience, we can give the first steps in the approximation of  $F^n(y)$  by  $\tilde{L}((y - \lambda)/\delta)$  and of  $1 - (1 - F(y))^n$  by  $1 - \tilde{L}(-(y - \lambda)/\delta)$  where  $\tilde{L}(\cdot)$  can be  $\Psi_\alpha(\cdot)$ ,  $\Lambda(\cdot)$  or  $\Phi_\alpha(\cdot)$ .

But, also, other quick and more experimental techniques are available, as follows.

### 4.3 The probability paper for Gumbel distribution

Consider that it is assumed that the i.i.d. sample has the underlying Gumbel distribution

$$\Lambda(x|\lambda, \delta) = \exp(-e^{-(x-\lambda)/\delta})$$

which can be written by taking logarithms twice:

$$z = \frac{x - \lambda}{\delta} = -\log(-\log \Lambda)$$

or

$$z = \lambda + \delta[-\log(-\log \Lambda)] ;$$

$z$  is the reduced variate and  $x$  the observed one.

Recall that the return period or return time for the level  $x$  is

$$T(x|\lambda, \delta) = (1 - \Lambda((x - \lambda)/\delta))^{-1} .$$

For the probability paper, the iterated logarithms can be built on one of the scales of the paper so that we need only to plot  $(x, \Lambda)$ . We do not have only scales for  $x$  and  $\Lambda$  but also for  $T$  as been in the usual probability paper that follows.

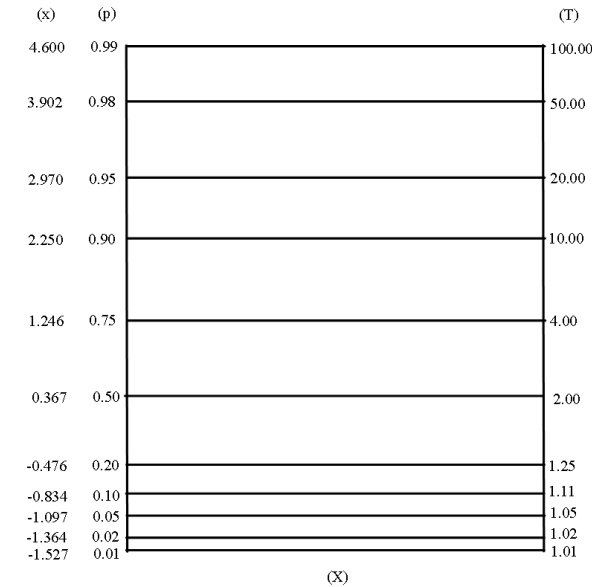


Figure 4.1. Gumbel maxima probability paper

The next figure illustrates the reverse probability paper where the scales of  $(x, \Lambda)$  are exchanged, as engineers often do.

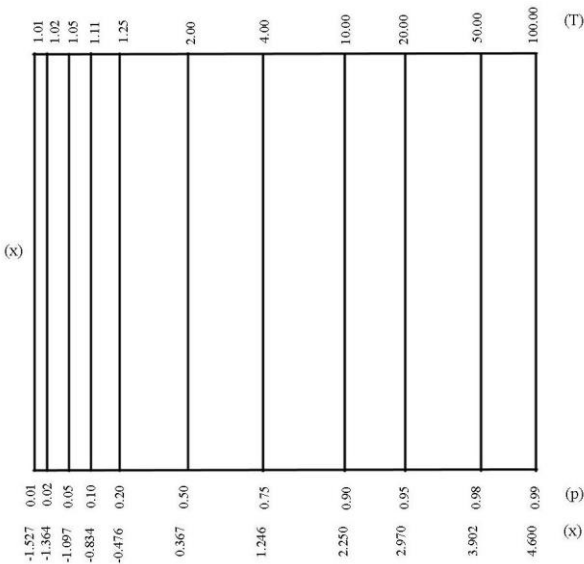


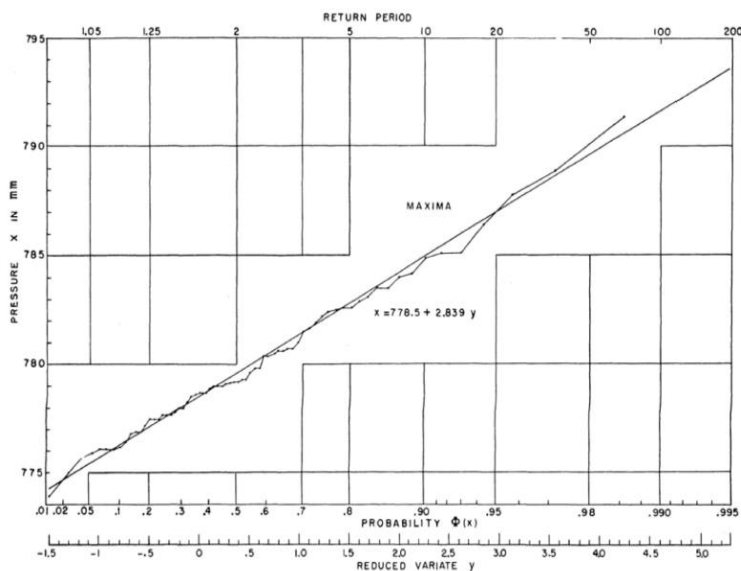
Figure 4.2. Gumbel maxima probability paper (reverse)

So far we have ignored the problem of determining an estimate of  $\Lambda((x'_i - \lambda)/\delta)$  to plot. One could use,  $i/n$  as plotting position to go with  $x'_i$ , the  $i^{\text{th}}$  order statistic in a sample of  $n$ , which is the usual plotting position for the sample distribution function, used for example in the Kolmogorov-Smirnov bounds. This is equivalent to estimating  $\Lambda((x'_i - \lambda)/\delta)$  as the proportion of  $x'_j$ 's less than or equal to  $x'_i$ . Since 1 (and also 0) do not appear on the  $\Lambda$ -scale, we cannot plot  $\Lambda((x'_n - \lambda)/\delta)$ . If we estimate  $\Lambda((x'_i - \lambda)/\delta)$  by  $(i - 1)/n$ , the proportion of  $x'_j$ 's less than  $x'_i$ , we cannot plot  $\Lambda((x'_1 - \lambda)/\delta)$ . As a compromise it was suggested to “split the difference” and use  $(i - 1/2)/n$  as the estimate of  $\Lambda((x'_i - \lambda)/\delta)$ . Another procedure is to use the mean value of  $\Lambda((x'_i - \lambda)/\delta)$  as a plotting position, which leads to the value  $i/(n + 1)$ . Blom (1958) devoted almost an entire book to the problem of obtaining “optimum” plotting positions, based on the idea of Chernoff and Lieberman (1954) that the plotting position should depend on the quantity to be estimated. For example Blom (1958, p.145) shows that for the *normal distribution*, plotting position  $(i - 3/8)/(n + 1/4) = (i - .375)/(n + .25)$  leads to a practically unbiased estimate of  $\sigma$  with a mean square error about the same as that of the unbiased best linear estimate, while the plotting position  $(i - .5)/n$  leads to a biased estimate of  $\sigma$  with nearly minimum mean square error.

Note that some symmetry of plotting positions must be verified: if  $p_{i,n}$  is the plotting position for the  $x'_i$  then we should have  $p_{i,n} + p_{n+1-i,n} = 1$  which does not happen with the plotting positions  $i/n$  and  $(i - 1)/n$  but happens with the two other ones. Another perspective, to be used further in the next section, leads to the plotting position  $p_i = (i - 0.3)/(n + 0.4)$ , also symmetrical; for details see Tiago de Oliveira (1983). Taking  $p_{i,n} = (i + A)/(n + B)$  the symmetry condition leads to  $B = 2A + 1$  and  $p_{i,n} = (i + A)/(n + 1 + 2A)$  and as  $0 < p_{i,n} < 1$  for each  $i$  and  $n$  we must have  $-1 < A$ ; for  $A = -1/2, A = 0, A = 3/8$  and  $A = -.3$  we obtain the given plotting positions. We have  $\max_i \left| \frac{i+A}{n+1+2A} - \frac{i+A'}{x+1+2A'} \right| \leq \frac{|A-A'|}{n}$ . Note that for very large  $A$  all the plotting positions are close to  $1/2$  which means that it is acceptable to use  $|A| < 1$  to have a reasonable spread of the plotting positions in  $]0,1[$ ; in that case the upper bound  $|A - A'|/n$  of the maximum

error is bounded above by  $2/n$  and is practically irrelevant for sample sizes  $n \geq 50$ , and all  $p_{i,n}$  are practically equivalent (\*).

In the graph that follows we will use the classical Gumbel plotting position  $i/(n+1)$ . Figure 4.3 that follows was taken from [Gumbel and Lieblein \(1954\)](#) and illustrates the procedure for estimating  $\lambda^* = 778.5$  and  $\delta^* = 2.839$  for maximum atmospheric pressure (in mm) at Bergen, Norway for the period between 1857 and 1926. If the fitted straight line is extrapolated beyond the fitted values, one would “predict”, for example, that a pressure of 793 mm corresponds to a probability of 0.994. That is, pressures of this magnitude have less than 1 chance in 100 of being exceeded in any particular year.



**Figure 4.3. Annual maxima of atmospheric pressures, Bergen, Norway, 1857-1926 (reverse probability paper)**

---

(\*) Also if we choose  $A$  such that for  $F(x'_n) = p_{n,n}, p_{n,n}^n \rightarrow 1/2$  (i.e.,  $x'_n$  is approximately the median of  $F^n(x)$ ) we get  $A = \log 2 - 1$  and  $p_{i,n} = (i - .307)/(n + .386)$ ; we get the same result for minima with  $1 - (1 - p_{i,n})^n \rightarrow 1/2$ .



#### 4.4 The use of Gumbel probability paper for quick statistical choice between Weibull, Gumbel and Fréchet distributions for maxima

Here we will describe a graphical technique taken from [Tiago de Oliveira \(1983\)](#). As was shown in the previous section,  $Z$  being a reduced random variable, the probability paper for Gumbel distribution was made as follows: from  $p = \Lambda(z)$ , which can be written as  $\Lambda^{-1}(p) = z$ , if we note by  $y = \Lambda^{-1}(p)$  and graduate the  $y$ -axis in the functional scale  $p$ , the relation  $p = \Lambda(z)$  can be written as the first diagonal  $y = z$ . In addition, if the true distribution is  $G(z|\theta) = \exp\{-(1 + \theta z)_+^{-1/\theta}\}$ , the von Mises-Jenkinson formula, the relation  $p = \Lambda(y) = G(z|\theta)$  may be written as  $y = \Lambda^{-1}(G(z|\theta))$  and the relation between  $(z, y)$  (or  $(z, p)$ ) is the curve in the plane  $y = \log(1 + \theta z)/\theta$ , for  $1 + \theta z > 0$ .

Let us now return to the plotting positions. If  $p_{i,n}$  is the plotting position for the order statistic  $z'_i (= z'_{i,n}$  with greater rigour) the distance between the plotted point  $(z'_i, p'_{i,n})$  and the first diagonal is  $|\Lambda^{-1}(p_{i,n}) - z'_i|$  horizontally or vertically or  $|\Lambda^{-1}(p_{i,n}) - z'_i|/\sqrt{2}$  orthogonally. The distance  $|\Lambda^{-1}(p_{i,n}) - z'_i|$  is minimized in mean if  $\Lambda^{-1}(p_{i,n})$  is the median of  $z'_i$  (with Gumbel distribution) or  $p_{i,n}$  is the median of Beta  $(i, n - 1 - i)$  distribution; a good approximation is to take the one given before,  $p_{i,n} = \frac{i-0.3}{n+0.4}$ ; note that we are not using the least squares approach but we are minimizing  $\max_i |\Lambda^{-1}(p_{i,n}) - z'_i|$ .

Having chosen one of these plotting positions, let us consider the general case, i.e. observations with unknown location and dispersion parameters  $\lambda$  and  $\delta$ ,  $z_i = (x_i - \lambda)/\delta$  being the reduces values. Then if  $\chi_\theta(p)$  is the  $p$ -quantile of  $G(z|\theta)$ , i.e.  $G(\chi_\theta(p)|\theta) = p$ , the  $p$ -quantile of  $G((x - \lambda)/\delta|\theta)$  is, evidently,  $\lambda + \delta\chi_\theta(p)$ . Consider then the random variable

$$v = \frac{x - (\lambda + \delta\chi_\theta(r))}{(\lambda + \delta\chi_\theta(s)) - (\lambda + \delta\chi_\theta(r))} = \frac{z - \chi_\theta(r)}{\chi_\theta(s) - \chi_\theta(r)}$$

with  $0 < r < s < 1$ ;  $v$  is obviously independent of the location and dispersion parameters.

Clearly we are going to plot *estimated*  $\hat{v}_i'$  where the  $r$  and  $s$  quantiles are estimated by the sample quantiles by  $x'_{[nr]+1}$  and  $x'_{[ns]+1}$ .

The relation  $p = \Lambda(y) = G(z|\theta)$  can then be written as  $y = \Lambda^{-1}(G(\chi_\theta(r) + (\chi_\theta(s) - \chi_\theta(r))v|\theta) = y_\theta(v)$ , for  $r$  and  $s$  fixed. Taking  $v = 0$  and  $v = 1$  we see that all curves  $y_\theta(v)$  pass through the points  $(0, \chi_0(r))$  and  $(1, \chi_0(s))$ , or  $(0, r)$  and  $(0, s)$  if we use the functional scale for  $y$ . Thus between  $v = 0$  and  $v = 1$  the curves  $y_\theta(v)$  will with difficulty be separated from the first diagonal, chiefly taking into account that the plotted points, even using real and not estimated reduced observations, do not fall exactly on the first diagonal for Gumbel distribution. In summary, the points between  $v = 0$  and  $v = 1$  are *lost* for separation of models; it is, thus, natural to use about  $1/3$  of the observations for the zone  $v \in [0, 1]$  and  $2/3$  outside; a smaller percentage in this zone would introduce instability in the implicit estimation of  $\lambda$  and  $\delta$  by the quantiles, as the denominator would be small.

With this rule of thumb, we have taken  $\chi_0(r) = 0$  and,  $\chi_0(s) = 1$ , or  $r = \Lambda^{-1}(0) = e^{-1} = 0.3678794$  and  $s = \Lambda^{-1}(1) = \exp(-e^{-1}) = 0.6922006$  (note that, as said before,  $r \approx 1/3$  and  $s \approx 2/3$ ); practically we can take  $r = 0.368$  or even  $r = 0.37$  and  $s = 0.692$  or even  $s = 0.69$ .

Let us now compute the curves  $y_\theta(v)$ . As  $\chi_\theta(p) = \frac{(-\log p)^{-\theta}-1}{\theta}$  and so  $\chi_\theta(e^{-1}) = 0$  and  $\chi_\theta(\exp(-e^{-1})) = (e^\theta - 1)/\theta$  we get

$$y_\theta(v) = \frac{\log(1 + (e^\theta - 1)v)}{\theta}$$

which is evidently defined when  $1 + (e^\theta - 1)v > 0$ ; note that  $y_\theta(v)$  is convex for  $\theta < 0$  (Weibull model) and concave for  $\theta > 0$  (Fréchet model),  $y_0(v)$  (Gumbel model) being a straight line.

Thus in a Gumbel probability paper we plot a bundle of curves  $(v, y_\theta(v))$  for different values of  $\theta$  and the quick statistical choice is made according to the location of the plotted points  $(\hat{v}_i', p_i)$  close to the diagonal (Gumbel distribution) or in the Fréchet or Weibull zones as shown in the Figure 4.4. Estimation of the parameters is to be made by the analytical methods of the next three chapters.

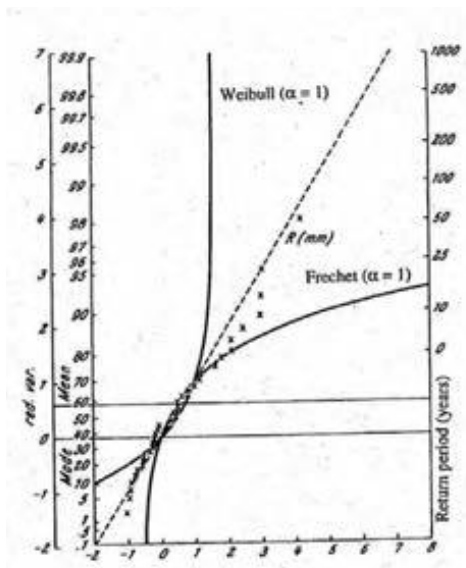
As an example consider the Table 1 of maximum wind velocities in km/h between 1941-1970 in Lisbon.

**Table 4.2.**

Year	Wind Velocities (km/h)	Year	Wind Velocities (km/h)
1941	129.0	1956	108.0
1942	117.0	1957	102.0
1943	100.0	1958	102.0
1944	100.0	1959	112.0
1945	132.0	1960	107.0
1946	94.0	1961	86.0
1947	108.0	1962	91.0
1948	113.0	1963	96.0
1949	96.0	1964	89.0
1950	113.0	1965	90.0
1951	96.0	1966	89.0
1952	72.0	1967	89.0
1953	98.0	1968	84.0
1954	85.0	1969	107.0
1955	124.0	1970	111.0

The values of the maximum likelihood estimators are  $\hat{\lambda} = 94.71$  and  $\hat{\delta} = 12.49$ ; the  $1/e$ -quantile and  $\exp(-1/e)$ -quantile needed for computation of the  $v_i'$  are  $Q(.367) = 96.0$  and  $Q(.692) = 108.0$ ; we took, as said before  $p_i = i/(n + 1)$ .

As regards  $m^{\text{th}}$  extreme, the simplest way is to try the statistical choice for maxima (or minima) because then the limiting distribution is well known and was given in Part-1.



**Figure 4.4.** Plotting of  $(v_i, p_i)$  for the maximum wind velocities in km/h in Lisbon, 1941-1970; Fréchet and Weibull zones shown.

#### 4.5 Simple statistical choice between the extreme models

We will use a simple statistic proposed in [Gumbel \(1965\)](#) for the study of the Fréchet distribution. In [Tiago de Oliveira and Gomes \(1984\)](#) we developed the asymptotic theory of this statistic for all distributions. Let us denote for a sample  $(x_1, \dots, x_n)$ , from the distribution  $G(z|\theta)$  by  $x'_1 = \min\{x_i\}$ ,  $x'_{[n/2]+1} = \text{med}\{x_i\}$  and  $x'_n = \max\{x_i\}$  the maximum, the medium and the minimum of the sample. We consider the statistic

$$Q_n = Q_{n,\theta} = \frac{x'_n - x'_{[n/2]+1}}{x'_{[n/2]+1} - x'_1} = \frac{\max\{x_i\} - \text{med}\{x_i\}}{\text{med}\{x_i\} - \min\{x_i\}}$$

which is, evidently, location and dispersion parameter-free.

We have shown in [Tiago de Oliveira and Gomes \(1984\)](#) that if  $\theta = 0$ ,  $G(z|0^+) = G(z|0^-) = \Lambda(z)$ , there exist  $\beta_{n,0}$  and  $\alpha_{n,0} > 0$  such that  $\frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} \xrightarrow{w} \Lambda(z)$ ; a choice is  $\beta_{n,0} = \frac{\log n + \log \log 2}{\log \log n - \log \log 2}$ ,  $\alpha_{n,0} = 1/\log \log n$ ; if  $\theta > 0$ ,  $G(z|\theta) = \Phi_{1/\theta}(\theta z + 1)$ , there exist  $\beta_{n,\theta}$  and  $\alpha_{n,\theta} > 0$  such that

$\frac{Q_n - \beta_{n,\theta}}{\alpha_{n,\theta}} \xrightarrow{w} G(z|\theta) = \Phi_{1/\theta}(\theta z + 1);$  a choice is

$\beta_{n,\theta} = \frac{n^\theta - (\log 2)^{-\theta}}{(\log 2)^{-\theta} - (\log n)^{-\theta}}; \alpha_{n,\theta} = \theta(n \log 2)^\theta$  or more simply  $\beta'_{n,\theta} = (n \log 2)^\theta, \alpha'_{n,\theta} = \theta \beta'_{n,\theta};$  if  $\theta < 0$ ,  $G(z|\theta) = \Psi_{-1/\theta}(-\theta z - 1)$ , there exist  $\beta_{n,\theta}$  and  $\alpha_{n,\theta} > 0$  such that  $\frac{Q_n - \beta_{n,\theta}}{\alpha_{n,\theta}} \xrightarrow{w} 1 - \Lambda(-z);$  a choice is  $\beta_{n,\theta} = \frac{n^\theta - (\log 2)^{-\theta}}{(\log 2)^{-\theta} - (\log n)^{-\theta}}; \alpha_{n,\theta} = -\theta(\log 2)^{-\theta}(\log n)^{\theta-1}$  or even more simply for  $\theta < 1, \beta'_{n,\theta} = (\frac{\log n}{\log 2})^\theta$  and similarly  $\alpha'_{n,\theta} = \alpha_{n,\theta} = -\frac{\theta}{\log n} \beta'_{n,\theta}.$

Two intuitive conclusions may be drawn from these results:

1. There is a “discontinuity” between  $\theta \geq 0$  and  $\theta < 0$ ; for  $\theta \geq 0$  the influent term is the maximum but for  $\theta < 0$  the influent term is the minimum; this explains, to a certain extent, why we cannot have a simple common expression for  $\beta_{n,\theta}$  and  $\alpha_{n,\theta}$ .
2. As for  $\theta \geq 0, \alpha_{n,\theta} \rightarrow \infty$  and for  $\theta < 0, \beta_{n,\theta} \rightarrow \infty$  it seems natural to decide for the Fréchet distribution ( $\theta > 0$ ) if  $Q_n$  is large and for the Weibull distribution ( $\theta < 0$ ) if  $Q_n$  is small; then the decision rule will be:

*Choose  $0 < b < a < +\infty$  and decide for the Gumbel distribution when  $b \leq \frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} \leq a$ , for the Fréchet distribution when  $\frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} > a$ , and for the Weibull distribution when  $\frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} < b$ , where  $\beta_{n,\theta} = \frac{\log n + \log \log 2}{\log \log n - \log \log 2}$  and  $\alpha_{n,\theta} = (\log \log n)^{-1}.$*

Let us compute, for  $b < a$ , the probabilities of correct decision; using the Khintchine convergence of types theorem:

$$1) \theta = 0: \quad \text{Prob}\{b \leq \frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} < a\} \rightarrow \Lambda(a) - \Lambda(b);$$

$$2) \theta > 0: \quad \text{Prob}\{\frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} > a\} \rightarrow 1 - G(-1/\theta|\theta) = 1;$$

$$3) \theta < 0: \quad \text{Prob}\{\frac{Q_n - \beta_{n,0}}{\alpha_{n,0}} < b\} \rightarrow 1 - G(-1/\theta|\theta).$$

The technique is thus consistent in the usual sense, whatever  $0 < b < a < +\infty$  may be; in fact if the decision technique is such that for  $\theta = 0$  the limit probability of correct decision is  $\Lambda(a) - \Lambda(b)$ , but if  $\theta \neq 0$  the limit probability of correct decision is 1. We could improve the situation by putting the significance level  $\alpha_n = 1 - (\Lambda(a_n) - \Lambda(b_n)) \rightarrow 0$  with convenient restrictions in  $b_n$  and  $a_n$ . But this does not seem to be needed because it is well known that for large  $\alpha$  (i.e, for small  $\theta$  in von Mises-Jenkinson form) both the Fréchet and the Weibull distributions are very close to the Gumbel distribution. Following on from the previous remark, it seems natural to solve only the statistical trilemma ( $\theta < 0, \theta = 0$  or  $\theta > 0$ ) and not a statistical dilemma or one-sided test of hypothesis of  $\theta = 0$  vs.  $\theta > 0$  or  $\theta = 0$  vs.  $\theta < 0$ . The latter would correspond to the partial acceptance of  $\theta \geq 0$  or  $\theta \leq 0$  which lead to a one-sided test as a consequence of assuming stronger previous knowledge or a stronger plausibility. Thus we will not deal with one-sided tests.

It remains now to obtain the one equation for  $(b, a)$  and the best seems, from what has been said, that the length of the Gumbel decision interval  $a - b$  should be the smallest possible with the condition  $\Lambda(a) - \Lambda(b) = 1 - \alpha$ . The use of Lagrange multipliers gives the equations

$$\Lambda(a) - \Lambda(b) = 1 - \alpha$$

$$\Lambda'(a) = \Lambda'(b).$$

The values corresponding to the usual significance levels are:

**Table 4.3**

$\alpha$	$b$	$a$
.050	- 1.561334	3.161461
.025	- 1.719620.	3.841321
.010	- 1.893530	4.740459
.001	- 2.222951	7.010001

Examples of quick exploration of data, as well as the analysis of univariate models and statistical choice, can be found in the “Some Case Studies” chapter.

## References

- Blom, G., 1958. *Statistical Estimation and Transformed Bet-Variables*, Wiley, New York.
- Chernoff, H., and Lieberman, G. J., 1954. Use of normal probability paper. *J. Amer. Statist. Soc.*, 49, 778-785.
- Fisher, R. A., and Tippett, L. H. C., 1928. Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Camb. Phil. Soc.*, 24, 180-190.
- Freudenthal, A. M., and Gumbel, E. J., 1953. Statistical interpretation of fatigue tests. *Proc. Roy. Soc. London*, 216, 309-332.
- Gumbel, E. J., 1954. Statistical Theory of Extreme Values and Some Practical Applications. *Nat. Bur. Standards, Appl. Math. Series*, 33, Washington.
- Gumbel, E. J., and Lieblein, J., 1954. Some Applications of Extreme-Value Methods. *Amer. Statist.*, 8(5), 14-17.
- Gumbel, E. J., 1955. The calculated risk in flood control. *Appl. Sc. Research (Holland)*, 5, A, 273-280.
- Gumbel, E. J., 1958. *Statistics of Extremes*, Columbia University Press, New York.
- Gumbel, E. J., 1962. Statistical theory of extreme values (main results). in *Contributions to Order Statistics*, Ahmed E. Sarhan and Bernard G. Greenberg eds., 56-93, Wiley, New York.
- Gumbel, E. J., 1965. A quick estimation of the parameters in Fréchet distribution. *Rev. Int. Statist. Inst.*, 33, 349-363.
- Ramachandran, G., 1974. Extreme value theory and large fire losses. *The Astin Bull.*, 7(3), 293-310.
- Thom, H. C. S., 1960. Distribution of extreme winds in the United States. *J. Structural Div., Proc. ASCE*, 86.
- Thom, H. C. S., 1971. Asymptotic extreme-value distributions of wave heights in the open ocean. *J. Marine Res.*, 29, 19-27.
- Tiago de Oliveira, J., 1977. Asymptotic distributions of univariate and bivariate m-th extremes. in *Recent Developments in Statistics*, J. R. Barra et al. eds., 613-619, North-Holland, Amsterdam.
- Tiago de Oliveira, J., 1983. Gumbel distribution. in *Encyclopedia of Statistical Sciences III*, N. L. Johnson and S. Kotz eds., 552- 558, Wiley, New York.
- Tiago de Oliveira, J., and Gomes, M. I., 1984. Two test statistics for choice of univariate extremes models. in *Statistical Extremes and Applications*, J. Tiago de ed., D. Reidel, Dordrecht.
- Weibull, W., 1939. A statistical theory of the strength of materials. *Proc. Roy. Swed. Inst. Eng. Res.*, 151, Stockholm.

\*\*\*\*\*