Statistical Theory of Extremes
Homepage: http://www.gathacognition.com/book/gcb14
http://dx.doi.org/10.21523/gcb1

Part 2
Statistics for Univariate Extremes

Chapter 8

Analytic Statistical Choice for Univariate Extremes

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

Abstract

Method of statistical choice between Weibull, Gumbel and Fréchet distribution is described. Locally optimal approach and variational approach are discussed to illuminate the problem from two points of view. Analysis begins with samples of reduced values and extended for generalization. The variational approach is coincident with the locally optimal approach. However, there does not exist an optimal procedure for statistical choice based on Q_n.

Published Online

23 June 2017

Keywords

Weibull distribution;
Gumbel distribution;
Fréchet distribution;
Locally optimal approach;
Variational approach.

Editor(s)

J.C. Tiago de Oliveira

8.1 Introduction

As said before, the three limiting distribution of univariate maxima, under reduced form $\Psi_\alpha(z)$ (Weibull distribution), $\Lambda(z)$ (Gumbel distribution), and $\Phi_\alpha(z)$ (Fréchet distribution), can be imbedded in the general von Mises-Jenkinson formula

$$G(z|\theta) = \exp(-(1+\theta\, z)_+^{-1/\theta}), -\infty < \theta < +\infty.$$

It is immediate that for $\theta < 0$ we have $\Psi_{-1/\theta}(z) = G(-\frac{1+z}{\theta}|\theta)$, a Weibull distribution with $\alpha = -1/\theta$, and for $\theta > 0$ we have $\Phi_{1/\theta}(z) = G(\frac{z-1}{\theta}|\theta)$, a Fréchet distribution with $\alpha = 1/\theta$; evidently, for $\theta = 0$ we obtain a Gumbel $\Lambda(z)$ as $G(z|0^+) = G(z|0^-) = \Lambda(z)$.

Originally published in 'Statistical Analysis of Extremes', 1997, 2016
http://dx.doi.org/10.21523/gcb1.1713

We will describe first a method of statistical choice between the three models ($\theta < 0, \theta = 0$ or $\theta > 0$) which is asymptotically optimal and leads to a statistic that gives asymptotically a locally most powerful unbiased test of $\theta = 0$ vs. $\theta \neq 0$.

We will begin by giving a locally optimal approach, as in Tiago de Oliveira (1981) and (1982), enlarged to a more general set-up in Tiago de Oliveira (1983) and then a variational approach, the two ways being coincident in the statistical technique used, which illuminates the problem from two points of view; application to data comes from Tiago de Oliveira (1981) and Fransén and Tiago de Oliveira (1984).

We will begin with the analysis for samples of reduced values and then extend it to the general case.

For other details see Tiago de Oliveira (1981), (1982), (1984) and, (1986) and Fransén and Tiago de Oliveira (1984).

Afterwards, we will sketch other analyses of the problem by van Montfort (1973), van Montfort and Otten (1978), Galambos (1980), Gomes (1982) and Pickands (1975); the use of probability paper and of Q_n was detailed in the Chapter 4, “Quick Exploration of Extreme Data”, in this book.

8.2 The locally optimal approach

Consider a sample of (independent) reduced values $(z_1, \ldots, z_n)$ with distribution function $G(z|\theta)$, from which we want to decide for the Weibull model $(\theta < 0)$, the Gumbel model $(\theta = 0)$, or the Fréchet model $(\theta > 0)$.

Let $g(z|\theta) = G'(z|\theta)$ denote the probability density,

$$L_n(z|\theta) = L(z_1, \ldots, z_n|\theta) = \prod_1^n g(z_i|\theta)$$

the likelihood of the sample, and $v(z) = \frac{\partial \log g(z|\theta)}{\partial \theta}\Big|_{\theta=0} = \frac{z^2}{2} - z - \frac{z^2}{2}\, e^{-z}$.

The locally optimal test of Gumbel vs. Fréchet models $(\theta = 0 \text{ vs.} \theta > 0)$ is given by the acceptance region $V_n(z) = \sum_1^n v(z_i) < a_n'$, the locally optimal test of Gumbel vs. Weibull models $(\theta = 0 \text{ vs.} \theta < 0)$ is given by the acceptance region $V_n(z) = \sum_1^n v(z_i) > b_n'$, and the asymptotically unbiased locally optimal test of Gumbel vs. Fréchet or Weibull models $(\theta = 0 \text{ vs.} \theta \neq 0)$ is given by the acceptance region $b_n < V_n(z) < a_n$; see Tiago de Oliveira (1981) for the proof of the asymptotic unbiasedness using only $V_n(z)$, where it is proven that the part containing the second derivatives, coming from the Neyman-Pearson theorem, converges to zero: the two-sided test is thus an intersection of the two one-sided tests. Thus:

The procedure is to decide for the Weibull model if $V_n(z) < b_n$, *for the Gumbel model if* $b_n \leq V_n(z) \leq a_n$ *and for the Fréchet model if* $a_n < V_n(z)$ *and we should have* $Prob\,\{accept\ \theta = 0\,|\theta = 0\} \to 1 - \alpha$ *and*

$$\frac{d\,Prob\,\{accept\ \theta=0|\theta\}}{d\,\theta}/_{\theta=0} \to 0.$$

The first attempt — see Tiago de Oliveira (1981) and (1984) — was to seek quantities $b_n < a_n$ that asymptotically split, in a balanced way, the significance level α, i.e., such that

$$\text{Prob}\,\{V_n(z) \leq b_n\} \to \alpha/2,$$

$$\text{Prob}\,\{V_n(z) \geq a_n\} \to \alpha/2,$$

and $\quad \text{Prob}\,\{b_n < V_n(z) < a_n\} \to 1 - \alpha.$

As, for $\theta = 0, v(z)$, has mean value zero and variance $\sigma^2 = \Gamma^{(4)}(1)/4 + \Gamma^{(3)}(1) + \Gamma'(1) = 2.42361$, we know that $V_n(z)/\sqrt{n}\,\sigma$ is asymptotically standard normal.

Putting $c_n(\alpha) = \sqrt{n}\,\sigma\,\lambda\,(\alpha/2)$, where $\lambda(p)$ is the solution of the equation $N(x) = 1 - p, N(.)$ being the standard normal distribution function, we see immediately that $a_n \sim c_n(\alpha)$ and $b_n \sim -c_n(\alpha)$. The approximation $-b_n = a_n = c_n(\alpha)$ was used in the first papers on the subject and applications, such as those referred to before, and Fransén and Tiago de Oliveira (1984) was also written under this pattern. We will show that a considerable improvement can be made.

Let us denote by $\mu(\theta)$ and $\sigma^2(\theta)$ the mean value and the variance of $v(z)$ with respect to the distribution function $G(z|\theta)$; $\mu(0) = 0$ and $\sigma^2(0) = \sigma^2$. It can be shown that $\mu(\theta)$ exists only when $-1 < \theta < 1/2$ and the variance exists only when $-1 < \theta < 1/4$. As $\mu'(0) = \sigma^2 (> 0)$ we see that $\mu(\theta)$ increases in the neighbourhood of $\theta = 0$. We have $\mu(\theta) = 2.42361\,\theta + 4.15362\,\theta^2 + O(\theta^3)$ and $\sigma^2(\theta) = 2.42361 + 25.26470\,\theta + 275.03597\,\theta^2 + O(\theta^3)$.

For $-1 < \theta < 1/4$ *the Central Limit Theorem is valid and, thus,* $(V_n(z) - n\,\mu(\theta))/\sqrt{n}\,\sigma\,(\theta)$ *is asymptotically standard normal.*

Under this formulation, and denoting by $\widetilde{P}_n(\theta)$ the normal approximation based on the Central Limit Theorem, we obtain that the probabilities of correct decision are:

— *for the Weibull model:*

$$P_n(\theta) = Prob\,\{V_n(z) \leq -c_n(\alpha)|\theta\} =$$

$$Prob\{\frac{V_n(z) - n\,\mu\,(\theta)}{\sqrt{n}\;\sigma\,(\theta)} \le \frac{-c_n(\alpha) - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)}/\theta\}$$

If $-1 < \theta < 0$, *is approached asymptotically by*

$$\tilde{P}_n(\theta) = N(\frac{-c_n(\alpha) - n\,\mu(\theta)}{\sqrt{n}\,\sigma\,(\theta)})\,,$$

which, as $\mu(\theta) < 0$ *for small values of* $\theta < 0$, *leads to* $\tilde{P}_n(\theta) \to 1$ *;*

— *for the Gumbel model:*

$$P_n(0) = Prob\{-c_n(\alpha) < V_n(z) < c_n(\alpha)|\theta = 0\} =$$

$$= Prob\{\frac{c_n(\alpha)}{\sqrt{n}\,\sigma} < \frac{V_n(z)}{\sqrt{n}\,\sigma} < \frac{c_n(\alpha)}{\sqrt{n}\,\sigma}|\theta = 0\},$$

is approached asymptotically by

$$\tilde{P}_n(0) = N(\lambda(\alpha/2)) - N(-\lambda(\alpha/2)) = 1 - \alpha;$$

— *for the Fréchet model:*

$$P_n(\theta) = Prob\{V_n(z) \ge c_n(\alpha)|\theta\} =$$

$$Prob\{\frac{V_n(z)-n\,\mu(\theta)}{\sqrt{n}\,\sigma\,(\theta)} \ge \frac{c_n(\alpha)-n\,\mu(\theta)}{\sqrt{n}\,\sigma\,(\theta)}/\theta\}$$

if $0 < \theta < 1/4$, *is approached asymptotically by*

$$\tilde{P}_n(\theta) = 1 - N(\frac{c_n(\alpha) - n\,\mu(\theta)}{\sqrt{n}\,\sigma\,(\theta)}\theta),$$

which, as $\mu(\theta) > 0$ *for small values of* $\theta > 0$, *leads to* $\tilde{P}_n(\theta) \to 1$;

But there is a remark to be made. As

$$\mathrm{Prob}\{\mathrm{accept}\ \theta = 0|\theta\} = \mathrm{Prob}\{-\,c_n(\alpha) < V_n(z) < c_n(\alpha)|\theta\} \approx \tilde{P}_n(\theta)$$

$$= N(\frac{c_n(\alpha) - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)}) - N(\frac{-c_n(\alpha) - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)}),$$

we have

$$\tilde{P}_n(0) = N(\lambda(\alpha/2)) - N(-\lambda(\alpha/2)) = 1 - \alpha\,;$$

but $\frac{d\,\tilde{P}_n(\theta)}{d\,\theta}|_{\theta=0} = -2\,\lambda(\alpha/2)N'(\lambda(\alpha/2))\,\frac{\sigma'(0)}{\sigma(0)} < 0$

for the usual values of α and thus the symmetric test chosen is not asymptotically unbiased.

This statistical decision procedure is thus *consistent* in the same way as tests are said to be consistent, the Gumbel model ($\theta = 0$) having a central position here. Recall that the procedure is unbiased and locally optimal at $\theta = 0$.

It can be changed to be *strongly consistent*, i.e., to be such that the probability of correct decision converges to 1, *whatever* θ may be, for $-1 < \theta < 1/4$. But in this approach we do not have asymptotic unbiasedness. We will summarize it.

Consider then as sequence $\{d_n\}(d_n > 0)$ *and use the statistical decision procedure:*

decide for the Weibull model $(\theta < 0)$ *if* $V_n(z) \leq -\sqrt{n}\,\sigma\, d_n$,

decide for the Gumbel model $(\theta = 0)$ *if* $/\,V_n(z)/ < \sqrt{n}\ \sigma\, d_n$,

and *decide for the Fréchet model* $(\theta > 0)$ *if* $V_n(z) \geq \sqrt{n}\,\sigma\, d_n$.

A necessary and sufficient condition for strong consistency is that $d_n \to +\infty$ *but* $d_n/\sqrt{n} \to 0$.

Evidently the conditions $d_n \to \infty$ and $d_n/\sqrt{n}$ avoid the "over-rejection" of the Gumbel model of the previous procedure with $c_n(\alpha)$ as shown by the strong consistency.

Optimization of $\{d_n\}$, whatever it may mean, has not been solved.

A way to obtain $\{d_n\}$ is, after choosing a (moving) level of significance $2\alpha_n (0 < \alpha_n < 1/2, \alpha_n \to 0)$ to define d_n by

$$\text{Prob}\{|V_n(z)| < \sqrt{n}\,\sigma\, d_n | \theta = 0\} \approx N(d_n) - N(-d_n) = 1 - 2\alpha_n$$

or, asymptotically, as is known,

$$d_n = \sqrt{-2\log\alpha_n} - \frac{\log(-\log\alpha_n) + \log(4\pi)}{2\sqrt{-2\ \log\alpha_n}}\,.$$

As $\alpha_n \to 0$ we see that $d_n \to +\infty$; the condition $d_n/\sqrt{n} \to 0$ leads to $(\log\alpha_n)/n \to 0$. A sample solution is to take $\alpha_n = 1/n$, thus leading to a probability of incorrect rejection of $\theta = 0$ of order $2/n$; for details see Tiago de Oliveira (1981) and (1984).

But let us return to the initial question seeking asymptotic unbiasedness.

Let us study the decision rule: choose $b_n < a_n$ and

decide for the Weibull distribution if $V_n \leq b_n$,

decide for the Gumbel distribution if $b_n < V_n < a_n$, and decide for the Fréchet distribution if $a_n \leq V_n$.

We will seek the best (b_n, a_n) for unbiasedness and some useful asymptotic approximation, as done in the previous step.

We will study the rule centering it on the Gumbel distribution, which is the pivot of the system. We have

$$P_n(k) = \text{Prob}\{\text{accept Gumbel df}|\theta\} = \text{Prob}\{b_n < V_n < a_n|\theta\}$$

$$= \text{Prob}\{\frac{b_n - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)} < \frac{V_n - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)} < \frac{a_n - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)}\}$$

if $\mu(\theta)$ and $\sigma(\theta)$ exist, and by the Central Limit Theorem we have

$$P_n(\theta) \approx \widetilde{P}_n(\theta) = N(\frac{a_n - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)}) - N(\frac{b_n - n\,\mu(\theta)}{\sqrt{n}\,\sigma(\theta)}).$$

For simplicity of notation we will denote by $\mu_0 = 0$, $\mu'_0 = \sigma_0^2$, $\mu''_0, \ldots$ and $\sigma_0, \mu'_0, \mu''_0, \ldots$ the values of $\mu(0), \mu'(0), \mu''(0), \ldots$ and $\sigma(0), \sigma'(0), \sigma''(0), \ldots$.

Let us introduce the convenient notations $P = \sigma'_0/\sigma_0^2 = 3.34804$, $\overline{b}_n = b_n / \sqrt{n}\,\sigma_0$ and $\overline{a}_n = a_n / \sqrt{n}\,\sigma_0$.

By successive derivation we get

$$\widetilde{P}_n(0) = N(\overline{a}_n) - N(\overline{b}_n),$$

$$\widetilde{P}'_n(0) = -\sqrt{n}\,\sigma_0\{N'(\overline{a}_n)(1 + p\,\overline{a}_n/\sqrt{n}) - N'(\overline{b}_n)(1 + p\,\overline{b}_n/\sqrt{n})\},$$

and $$\widetilde{P}''_n(0) = n\,\sigma_0^2\,\{N''(\overline{a}_n)(1 + p\,\overline{a}_n/\sqrt{n})^2 - N''(\overline{b}_n)(1 + p\,\overline{b}_n/\sqrt{n})^2\}$$

$$-(\sqrt{n}/\sigma_0)\,\{N'(\overline{a}_n)\,(\mu''_0 + \sigma''_0\,p\,\overline{a}_n / \sqrt{n} - 2\,\sigma_0^3(1 + p\,\overline{a}_n / \sqrt{n}))$$

$$-\,N'(\overline{b}_n)(\mu''_0 + \sigma''_0\,p\,\overline{b}_n/\sqrt{n} - 2\,\sigma_0^3(1 + p\,\overline{b}_n/\sqrt{n}))\}$$

which will be needed to obtain the asymptotically optimal solution and a sequence of approximations.

For the limit solution, with significance level α, we have

$$\widetilde{P}_n(0) = N(\overline{a}_n) - N(\overline{b}_n) = 1 - \alpha\,,$$

$\widetilde{P}'_n(0) = -\sqrt{n}\,\sigma_0[N'(\overline{a}_n)(1 + p\,\overline{a}_n/\sqrt{n}) - N'(\overline{b}_n)(1 + p\,\overline{b}_n / \sqrt{n})] = 0$ which gives *the basic equations for the problem*

$$N(\overline{a}_n) - N(\overline{b}_n) = 1 - \alpha$$

$[N'(\bar{a}_n)(1 + p\,\bar{a}_n/\sqrt{n}) - N'(\bar{b}_n)(1 + p\,\bar{b}_n/\sqrt{n})] = 0.$

They must be solved numerically and do not lead to a symmetrical solution. We will get, first, a limiting solution and then a sequence of approximations that give a good simple solution.

Letting $n \to \infty$ in the basic equation with $\bar{b}_n \to \bar{b}$ and $\bar{a}_n \to \bar{a}$, we get

$N(\bar{a}) - N(\bar{b}) = 1 - \alpha$

$N'(\bar{a}) = N'(\bar{b})$

which lead to $\bar{b} = -\lambda$, $\bar{a} = \lambda$, with $\lambda = \lambda(\alpha/2)$ being defined by the equation $N(\lambda) = 1 - \alpha/2$.

A remark can now be made: as $\bar{a}_n \to \lambda$, $\bar{b}_n \to -\lambda$ we see that

$\widetilde{P}_n''(0)/n\,\sigma_0^2 \to 2\,N''(\lambda)$ or $\widetilde{P}_n''(0) \sim -2\,\sigma_o^2\,\lambda\,N'(\lambda) n < 0,$

showing that we have a maximum of $\widetilde{P}_n(\theta)$ at $\theta = 0$. But the decision procedure with the limiting solution $\bar{b} = -\lambda$, $\bar{a} = \lambda$, although verifying the condition $\widetilde{P}_n(0) = 1 - \alpha$, is slightly biased because $\widetilde{P}_n'(0) = -2\ \sigma_0\,p\,\lambda\,N'(\lambda) < 0$.

We will define approximations to $\bar{b}_n$ and $\bar{a}_n$ developing them in powers of $1/\sqrt{n}$ as

$\bar{b}_{nj} = -\lambda + \bar{b}_1'/n^{-1/2} + \bar{b}_2'/n + \cdots + \bar{b}_j'\,n^{-j/2}$

$\bar{a}_{nj} = \lambda + \bar{a}_1'/n^{-1/2} + \bar{a}_2'/n + \cdots + \bar{a}_j'\,n^{-j/2}$;

obviously we have

$\bar{b}_n = \bar{b}_{nj} + O(n^{-(j+1)/2}),$

$\bar{a}_n = \bar{a}_{nj} + O(n^{-(j+1)/2}),$

$\bar{b}_{n,j+1} = \bar{b}_{n,j} + \bar{b}_{j+1}'\,/n^{(j+1)/2},$

and $\bar{a}_{n,j+1} = \bar{a}_{n,j} + \bar{a}_{j+1}'/n^{(j+1)/2}.$

For convenience we will introduce the functions:

$$\bar{b}_j(\tau) = -\lambda + \sum_1^j \bar{b}_r'\,\tau^r$$

$$\overline{a}_j(\tau) = \lambda + \sum_1^j \overline{a}_r^{'}\,\tau^r;$$

we evidently have $\overline{b}_{nj} = \overline{b}_j(1\sqrt{n})$ and $\overline{a}_{nj} = \overline{a}_j(1\sqrt{n})$.

$\overline{b}_{nj}$ and $\overline{a}_{nj}$ will be the j -th approximations to $\overline{b}_n$ and $\overline{a}_n$; the 0 -th approximation is evidently $\overline{b} = -\lambda$,and $\overline{a} = \lambda$. We will denote by $\widetilde{P}_{nj}(0)$ and $\widetilde{P}_{nj}^{'}(0)$ the values of $\widetilde{P}_n(0)$ and $\widetilde{P}_n^{'}(0)$ when we use the j -th approximation (in the initial notation, when we take $\overline{b}_n = \overline{b}_{nj}$ and $\overline{a}_n = \overline{a}_{nj}$).

The deviation and the slope, at origin $(\theta = 0)$, of the decision procedure are

$\tilde{\Delta}_{nj} = \left|\widetilde{P}_{nj}(0) - (1-\alpha)\right|$ and $\widetilde{P}_{nj}^{'}(0)$.

As we have seen for $j = 0$ we have $\tilde{\Delta}_{n0} = 0$ and $\widetilde{P}_n^{'}(0) = -2\,\sigma_0 p\,\lambda\,N'(\lambda)$.

For each j, accepting the previous $\overline{b}_{ni}$ and $\overline{a}_{ni}$ $(i < j)$ we could develop in series of $1/\sqrt{n}$ the terms of the two basic equations, put equal to zero the first non-null term of the series in each equation, thus obtaining two linear equations in $\overline{b}_j^{'}$ and $\overline{a}_j^{'}$, and evaluate $\tilde{\Delta}_{nj}$ and $\widetilde{P}_{nj}(0)$ by the asymptotics of the next non-zero term. We will use a systematic method. Denote by

$$f_j(\tau) = N(\overline{a}_j(\tau)) - N(\overline{b}_j(\tau)) - (1-\alpha)$$

$$g_j(\tau) = N'(\overline{a}_j(\tau))(1 + p\,\tau\,\overline{a}_j(\tau)) - N'(\overline{b}_j(\tau))(1 + p\,\tau\,\overline{b}_j(\tau))\ ;$$

we have $\tilde{\Delta}_{nj} = |f_j(1/\sqrt{n})|$ and $P_{nj}^{'}(0) = -\sqrt{n}\,\sigma_0\,g_j(1/\sqrt{n})$.

For $j = 1$ we have

$$f_1(0) = 0, f_1^{'}(0) = (\overline{a}_1^{'} - \overline{b}_1^{'})N'(\lambda), \ldots$$

$$g_1(0) = 0, g_1^{'}(0) = (\overline{a}_1^{'} + \overline{b}_1^{'})N''(\lambda) + 2\,p\,\lambda\,N'(\lambda).$$

The equations defining $\overline{b}_1^{'}$ and $\overline{a}_1^{'}$ are then $f_1^{'}(0) = 0$ and $g_1^{'}(0) = 0$ and we get $\overline{a}_1^{'} = \overline{b}_1^{'} = p$ and so $\overline{b}_{n1} = -\lambda + p/\sqrt{n}$ and $\overline{a}_{n1} = \lambda + p/\sqrt{n}$. Then we get, using $\overline{a}_1^{'} = \overline{b}_1^{'} = p$,

$$f_1^{''}(0) = 2\,p^2\,N''(\lambda) = -2\,p^2\,N'(\lambda),\ \ g_1^{''}(0) = 0,$$

$$g_1^{'''}(0) = 4\,p^3(\lambda^3 - 2\,\lambda)\,N'(\lambda)\ \text{ and so}$$

$$\tilde{\Delta}_{n1} \sim \frac{|f_1''(0)|}{2}\left(\frac{1}{\sqrt{n}}\right)^2 = \lambda\, N'(\lambda)\, p^2/n$$

and $\tilde{P}'_{n1}(0) \sim -\sqrt{n}\ \sigma_0\ \frac{g_1'''(0)}{6}\left(\frac{1}{\sqrt{n}}\right)^3 = -2(\lambda^3 - 2\,\lambda)\, N'(\lambda)\, p^3\sigma_0/3n.$

The development of $[N(\overline{a}_{n1}) - N(\overline{b}_{n1}) - (1-\alpha)]$ up to terms of second order in it $1/\sqrt{n}$ shows that

$$\tilde{\Delta}_{n1} \le \frac{p^2}{n}\max_{\lambda}|N''(\lambda)| = \frac{p^2}{n}|N''(\pm 1)| = \frac{1}{\sqrt{2\,\pi\, e}}\cdot\frac{p^2}{n}$$

$= 0.2419707\ \frac{p^2}{n} = 2.7123396/n.$

Note that $\tilde{\Delta}_{n1}$ and $\tilde{P}_{n1}(0)$ are both $O(n^{-1})$. In the same way we can evaluate the next order approximations. For $j = 2$ we get through the analysis of $f_2(\tau)$ and $g_2(\tau)$, $\overline{b}_2' = -p^2\,\lambda/2$ and $\overline{a}_2' = p^2\,\lambda/2$ and then, as $f_2''(0) = 4\, p^2\lambda^2\, N'(\lambda)^2$ we get $\tilde{\Delta}_{n2} \sim 2\, p^2\lambda^2\, N'(\lambda)^2/n$ with a small reduction of the deviation measured by

$\tilde{\Delta}_{n2}/\tilde{\Delta}_{n1} \sim 2\,\lambda\, N'(\lambda) \le 2\,\lambda\, N'(\lambda)|_{\lambda=1} = .48394$

but with a deviation of the same order $O(n^{-1})$; and, as $g_2'''(0) = 2\, p^3(-\lambda^3 - 2\,\lambda^2 + \lambda + 4)\, N'(\lambda)$ and so

$$\tilde{P}''_{n2}(0) \sim -\sqrt{n}\,\sigma_0\ \frac{g'''(0)}{6}\left(\frac{1}{\sqrt{n}}\right)^3 = \sigma_0\frac{p^3\ (\lambda^3 + 2\,\lambda^2 - \lambda - 4)N'(\lambda)}{3n}$$

and

$$\tilde{P}''_{n2}(0)/\tilde{P}'_{n1}(0) \sim \frac{-\lambda^3 - 2\,\lambda^2 + \lambda + 4}{2(\lambda^3 - 2\,\lambda)}\ (\to -1/2 \text{ as } \lambda \to +\infty)$$

and so $\tilde{P}''_{n2}(0) = O(n^{-1})$ as $\tilde{P}'_{n1}(0)$; there is not a great improvement in the use of the second order approximation.

Consequently the procedure, asymptotically unbiased to order $O(n^{-1})$, is to take as bounds for $V_n(z)/n$ the values

$b_n = \sigma_0\,\overline{b}_n/\sqrt{n} = \sigma_0(-\lambda + p/\sqrt{n} - p^2\,\lambda/2n)/\sqrt{n}$

$a_n = \sigma_0\,\overline{a}_n/\sqrt{n} = \sigma_0(\lambda + p/\sqrt{n} + p^2\,\lambda/2n)/\sqrt{n}$

or to decide for the Gumbel distribution if

$$-\lambda + p/\sqrt{n} - p^2\,\lambda/2n \le \frac{V_n(z)}{\sqrt{n}\ \sigma_0} \le \lambda + p/\sqrt{n} + p^2\,\lambda/2n$$

and correspondingly for the Weibull or Fréchet distributions.

The acceptance interval for the Gumbel distribution (in terms of $V_n(z)/n$) thus increases from $a_0 - b_0 = 2\,\lambda$ to $a_2 - b_2 = 2\,\lambda\,(1 + \frac{p^2}{2n})$ from $O(1)$ to $O(n^{-1})$.

Let us now generalize this statistical choice procedure, when introducing location λ and dispersion $\delta(>0)$ parameters.

If the underlying distribution is the general Gumbel distribution $\Lambda((x-\lambda)/\delta)$, the maximum likelihood estimators $\hat{\lambda}$ and $\hat{\delta}$, for the sample $(x_1, \dots, x_n)$, are given by the equations

$$\hat{\delta} = \bar{x} - \frac{\sum_1^n x_i \exp(-x_i/\hat{\delta})}{\sum_1^n \exp(-x_i/\hat{\delta})}$$

and

$$\hat{\lambda} = -\hat{\delta}\log\Big(\frac{\sum_1^n \exp\big(-x_i/\hat{\delta}\big)}{n}\Big)\,.$$

Now we will use "estimated" reduced values $\hat{z}_i = (x_i - \hat{\lambda})/\hat{\delta}$ and compute the "estimated $V_n(z)$" as $V_n(\hat{z})$, given by

$$\hat{V}_n(x) = \sum_1^n v((x_i - \hat{\lambda})/\hat{\delta}) = \frac{1}{2\,\hat{\delta}^2}\Big\{\sum_1^n x_i^2 - 2\,\hat{\delta}\sum_1^n x_i - \frac{n\sum_1^n x_i^2 \exp\,(-x_i/\hat{\delta})}{\sum_1^n \exp(-x_i/\hat{\delta})}\Big\}$$

$$= \frac{1}{2}\Big\{\sum_1^n ((x_i - \bar{x})/\hat{\delta})^2 - n\,\frac{\sum_1^n((x_i-\bar{x})/\hat{\delta})^2\exp(-(x_i-\bar{x})/\hat{\delta})}{\sum_1^n \exp(-(x_i-\bar{x})/\hat{\delta})}\Big\}.$$

Either by the δ-method, Tiago de Oliveira, (1982), but with a lengthy algebra, or following Tiago de Oliveira (1981), it can be shown that, for the Gumbel model, $\hat{V}_n(x)/n$ is asymptotically normal with mean value zero and variance $\hat{\sigma}^2/n$ with $\hat{\sigma}^2 = 2.09797$; note that $\hat{\sigma}^2 < \sigma^2$, a reduction of 13.4%. More generally $\hat{V}_n(x)/n$ is asymptotically normal with mean value $\mu(\theta)$ and variance $\hat{\sigma}^2(\theta)/n$, if the underlying distribution is $G(z|\theta)$; $\hat{\mu}(0) = 0$. Also $\hat{\mu}'(0) = \hat{\sigma}^2(0) = \hat{\sigma}^2$ and also $\hat{\mu}(\theta)$ is increasing in the neighbourhood of $\theta = 0$. The expansions of $\hat{\mu}(\theta)$ and $\hat{\sigma}(\theta)$ are $\hat{\mu}(\theta) = 2.09797\,\theta + 4.58653\,\theta^2 + O(\theta^3)$ and $\hat{\sigma}^2(\theta) = 2.09797 + 19.91015\,\theta + 248.60225\,\theta^2 + O(\theta^3)$.

Thus we can apply the reasoning used previously with the reduced sample and formulate a consistent or a strongly consistent decision procedure.

The strongly consistent procedure is to choose $\{d_n\}$ *such that* $d_n \to \infty$ *and* $d_n/\sqrt{n} - 0$ *and decide for the Weibull model* $(\theta < 0)$ *if* $\hat{V}_n(x) \leq -\sqrt{n}\,\hat{\sigma}\,d_n$, *decide* *for the Gumbel model* $(\theta = 0)$ *if* $|\hat{V}_n(x)| < \sqrt{n}\,\hat{\sigma}\,d_n$, *and decide for the Fréchet model* $(\theta > 0)$ *if* $\hat{V}_n(x) \geq \sqrt{n}\,\hat{\sigma}\,d_n$.

Evidently for a (classical) consistent decision procedure we should use the asymptotically locally optimal approach given for the reduced case with the substitution of σ by $\hat{\sigma}$ and of p by $\hat{p} = \hat{\sigma}'/\hat{\sigma}_0^2 = 3.27601$.

The method of analysis described above can be related to the likelihood ratio test of $\theta = 0$ vs. $\theta \neq 0$.

For more details and other information see Tiago de Oliveira (1981), although connected with the initial symmetrical approach.

The example associated with Table 1 and Figure. 4.4 in Chapter 4 (with $\hat{\lambda} = 94.71$ and $\hat{\delta} = 12.49$) gives $\hat{V}_n/\sqrt{2.09797\,n} = -1.21$ leading, thus, to the acceptance of Gumbel model, as suggested by Figure. 4.4.

8.3 The variational approach

We begin by considering the reduced case. From a sample $(z_1, \ldots, z_n)$ we want to find a statistic $t_n(z) = t(z_1, \ldots, z_n)$ which approaches zero (0) in mean-square if $\theta = 0$ and deviates at the maximum positive speed if $\theta \neq 0$ in the neighbourhood of $\theta = 0$. The basic idea is to use $t_n(z)$ to decide for the Weibull model if $t_n(z) \leq B_n$, for the Gumbel model if $B_n < t_n(z) < A_n$ and for the Fréchet model if $A_n \leq t_n(z)$; naturally we can expect $B_n < 0 < A_n$. μ

Once more $L_n(z|\theta) = L_n(z_1, \ldots, z_n|\theta) = \prod_1^n g(z_i|\theta)$ denotes the likelihood of the sample; the mean value and variance of $t_n(z)$ for the value θ of the shape parameter are denoted by

$$v_n(\theta) = \int_{-\infty}^{+\infty} t_n(z)\, L_n(z|\theta)\, d\,z$$

and

$$\tau_n^2(\theta) = \int_{-\infty}^{+\infty} t_n^2(z)\, L_n(z|\theta) d\,z \; - v_n^2(\theta)$$

if they exist.

The conditions described above, with a scaling condition on the variance at $\theta = 0$, for convenience, are

$$v_n(0) = \int_{-\infty}^{+\infty} t_n(z)\, L_n(z|0)\, d\,z = 0$$

$\tau_n^2(0) = \int_{-\infty}^{+\infty} t_n^2(z)\, L_n(z|0)\, d\, z \; = C$ (a positive constant)

and we want

$$v_n'(0) = \frac{d\, v_n(\theta)}{d\, \theta}\,|_{\theta=0} = \int_{-\infty}^{+\infty} t_n(z) \sum_1^n v(z)\, L_n(z|0) d\, z$$

to be a positive maximum ; as before we have

$$v(z) = \frac{\partial \log g(z|\theta)}{\partial\, \theta}\,|_{\theta=0} = \frac{z^2}{2} - z - \frac{z^2}{2} e^{-z}.$$

The Lagrangean for the Calculus of Variations, with Lagrange multipliers A and B for the side conditions $v_n(0) = 0$ and $\tau_n^2(0) = C$, is

$\overline{L}(t_n) = \{t_n(z) + A\, t_n(z) + B\, t_n^2(z)\}\, L_n(z|0).$

The Euler-Lagrange equation for t_n is

$$\frac{\partial\, \overline{L}(t_n)}{\partial\, t_n} = 0$$

giving thus, as $L_n(z|0) > 0$ for $-\infty < z < +\infty$,

$\sum_1^n v(z_i) + A + 2\, B\, t_n(z) = 0.$

Multiplying by $L_n(z|0)$ and integrating the equation on z we get $A = 0$ and so $t_n(z)$ is proportional to $V_n(z) = \sum_1^n v(z_i)$.

Denoting, as before, by $\mu(\theta)$ and $\sigma^2(\theta)$ the mean value and the variance of $v(z)$ with respect to $g(z|\theta)$, we see that taking $t_n(z) = V_n(z)$ we have

$v_n(\theta) = n\, \mu(\theta)$ $\qquad$ $(v_n(0) = 0)$

and

$\tau_n^2(\theta) = n\, \sigma^2(\theta)$ $\qquad$ $(\tau_n^2(0) = n\, \sigma^2)$.

Then as $v_n'(0) = \tau_n^2(0) = n\, \sigma^2 > 0$ the variance increases with θ.

The variational approach is thus coincident with the locally optimal one and the analysis can proceed on the same lines as before.

The variational approach could be continued by seeking a statistic $t_n(x) = t(x_1, \ldots, x_n)$, independent of the location and dispersion parameters (i.e., $t(\alpha + \beta\, x_1, \ldots, \alpha + \beta\, x_n) = t(x_1, \ldots, x_n)$) with the property of quickest deviation from zero, as before, but the result does not seem workable, as happened before in the search of the best quasi-linear estimators and predictors in Chapters 5, 6 and 7.

8.4 Other analytic approaches

Let us describe, in a sketchy way, some approaches tried by Jenkinson (1955), van Montfort (1973), Pickands (1975), van Montfort and Otten (1978), Galambos (1980), Gomes (1982) and Tiago de Oliveira and Gomes (1984).

Jenkinson (1955) suggested the following decision statistic: for a sample of 2n, we split it in n pairs whose maxima are taken and use the ratio of the variances of the 2n observations and the (produced) sample of n maxima of pairs. This statistic is, clearly, independent of the location and dispersion parameters and converges to a function of the shape parameter. A modification of statistic was studied in Tiago de Oliveira and Gomes (1984).

van Monfort (1973) developed the following statistic W to test Gumbel vs. Fréchet distribution : define

$$W_n = -\frac{1}{2} \log \frac{1 + r_n}{1 - r_n}$$

where r_n is the correlation coefficient between $\Psi_n(i + 1/2) = -\log(-\log((i + 1/2)/n))$ and $L_i = (x'_{i+1} - x'_i)/(\eta_{i+1} - \eta_i)$ $(i = 1,2, \dots, n - 1)$ where the x'_i are the order statistics of a sample $(x_1, \dots, x_n)$ and η_i are the mean values of x'_i for the Gumbel distribution under reduced form ($\lambda = 0, \delta = 1$. Evidently r, and thus W, are location and dispersion free. Simulation has shown that W is approximately normal with the mean value and variance to be fitted to the simulated critical points. The rejection region is $W > a$.

In the statistical choice proposed by van Montfort and Otten (1978) there is defined the statistic

$$A_n = \sqrt{n} \sum_{2}^{n} \frac{(\Delta_i - \bar{\Delta})}{\sigma_\Delta} W_i$$

where

$$W_i = \frac{L_i}{L_2 + \dots + L_n}, L_i = (x'_i - x'_{i-1})/(\eta_i - \eta_{i-1}), x'_i,$$

and η_i with the same meaning as in van Montfort (1973),

$\Delta_i = -\gamma - (S^*_i - S^*_{i-1})/(S_i - S_{i-1})$ $(i = 2, \dots, n), \gamma = 0.57722$ (the Euler constant), $S_i = i\binom{n}{i} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \frac{\log(i+j)}{i+j}$, $S^*_i = i\binom{n}{i} \sum_{j=0}^{n-i} (-i)^j \binom{n-i}{j} \frac{\log^2(i+j)}{2(i+j)}$, $\bar{\Delta} = \frac{1}{n-1} \sum_2^n \Delta_i$ and $\sigma^2_\Delta = \frac{1}{n-1} \sum_2^n (\Delta_i - \bar{\Delta})^2$.

The rule is to decide for $\theta > 0$ (Fréchet) if A_n is large, for $\theta < 0$ (Weibull) if it is small and for $\theta = 0$ (Gumbel) if it is intermediate. To this statistic A_n, a normal distribution is fitted giving conservative values, as shown by simulation.

Pickands (1975) gives a procedure to estimate the upper tail of a distribution, with an approach close to the one of Galambos (1980) testing procedure, as follows: for large a he shows that the conditional distribution $\frac{1-F(x+a)}{1-F(a)}$, if F is continuous, is close to $1 - L(x/a\,|c) = \exp\{-\int_0^{x/a}((1+ct)_+)^{-1}\,dt\}$ if $F(.)$ is attracted to a limiting extreme value distribution, where $c < 0, c = 0, c > 0$ correspond to $\theta > 0$ (Fréchet), $\theta = 0$ (Gumbel) and $\theta < 0$ (Weibull) distributions. Then, letting $x_1'' \geq x_2'' \geq \cdots \geq x_k'' \geq \cdots \geq x_n''$ be the (descending) order statistics of a sample of n i.i.d. observations with distribution function $F(x)$, and taking the upper $4M$ observations $(x_1'' \geq x_2'' \geq \cdots x_{4M}'')$, we use

$$\hat{c} = \log((x_M'' - x_{2M}'')/(x_{2M}'' - x_{4M}''))/\log 2$$

and

$$\hat{a} = (x_{2M}'' - x_{4M}'')/\int_0^{\log 2} e^{\hat{c}\,u}\,d\,u$$

as estimators of c and a if $M \overset{P}{\to} \infty$ but $M/n \overset{P}{\to} 0$. A choice of M is to take $d_M = \min_{1\leq j\leq n/4} d_j$, with $d_j = \sup_{0\leq x<+\infty} \left|\hat{F}_j(x) - P_j(x|\hat{a},\hat{c})\right|$, $\hat{F}_j(x)$ being the usual empirical distribution function of $x_1'' - X_{4i}''(i = 1, 2, \ldots, 4j-1)$, and $P_j(x|\hat{a},\hat{c})$ is computed with $\hat{a}, \hat{c}$ corresponding to $M = j$. $1 - F(x)$ is then approximated by $1 - \overline{G}_M(x|\hat{a},\hat{c})$.

The distribution functions $P(.)$ are called *the generalized Pareto distributions*.

For more details connected to tail estimation see the Annex 4.

In Galambos (1980) another approach was presented, for the same type of decision for Gambel vs. Fréchet models. It can be shown that the maximum of an i.i.d. sequence of random variables $\{X_j\}$, unbounded to the right and with mean value, is attracted to $\Lambda(.)$ if

$$\text{Prob}\{X > t + \mu(t)x\}/\text{Prob}\{(X > t)\} \to e^{-x} \quad \text{as } t \to +\infty,$$

where, $\mu(t) = \int_t^{+\infty}(1 - F(y))\,dy/(1 - F(t))$ is the conditional mean of $X - t$ if $X > t$. Thus for large t, the $Y_j = (X_j - t)/\mu(t)$ are asymptotically standard exponential, and $\mu(t)$ can be estimated by $\mu^*(t)$, the average of the values $x_j - t > 0$. Then it is suggested to use for t the m-th maximum $(t = x_{n+1-m}' - x_m'')$ and to proceed to the common tests for exponentiality; in the case of rejection of $\Lambda(.)$ we will accept the Fréchet model.

The Gumbel statistic Q_n for estimation of θ for the Fréchet distribution has already been used for quick exploration of extremes (Chapter 4). Gomes (1982) studied, by simulation, the behaviour of the statistic $Q_n = \frac{x_n' - x_{[n/2]+1}'}{x_{[n/2]+1}' - x_1'}$ where x_n' and x_1' are obviously the maximum and the minimum of the sample and $x_{[n/2]+1}'$ the median of the sample. The asymptotic distribution of this statistic was obtained in Tiago de Oliveira and Gomes (1984) and has different behaviours according to the exact value of θ: a Gumbel distribution of minima if $\theta < 0$, a Gumbel distribution of maxima if $\theta = 0$ and Fréchet distribution of maxima if $\theta > 0$, as said in Chapter 4. But there does not exist an optimal procedure for statistical choice based on Q_n.

References

Fransén, A. and Tiago de Oliveira, J., 1984. Statistical choice of univariate extreme models, part II. in *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 373-394, D. Reidel, Dordrecht.

Galambos, J., 1980. A statistical test for extreme value distributions. *Coll. Math. Soc. Janos Bolyai*, 32, 221-230, Budapest.

Gomes, M. I., 1982. A note on statistical choice of extremal models. *Act. IX Jorn. Matem. Hispano-Lusas, Salamanca*, 2, 653-656.

Jenkinson, A. F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quat. J. Roy. Meteor. Soc.*, 81(348), 158-171.

Montfort, M. A. J. VAN. and Otten, A., 1978. On testing a shape parameter in the presence of a scale parameter. *Math. Operations Forsch. Statist.*, Ser. Statistics, 9, 91-104.

Montfort, M. A. J. VAN., 1973. An asymmetric test on the type of the distribution of extremes. *Medelelugen Landbbouwhoge-School Wageningen*, 73(18), 1-15.

Pickands, J. III., 1975. Statistical inference using extreme order statistics. *Ann. Statist.*, 3, 119-131.

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 Oliveira ed., D. Reidel, Dordrecht.

Tiago de Oliveira, J., 1981. Statistical choice of univariate extreme models. in *Statistical Distributions in Scientific work*, C. Taillie *et al*. eds., 6, 367-387, D. Reidel, Dordrecht.

Tiago de Oliveira, J., 1982. Efficient estimation for quantiles of Weibull distributions *Rev. Belge Statist. Rech. Oper.*, 22, 3-10.

Tiago de Oliveira, J., 1982b. The δ-method for obtention of asymptotic distributions; applications. *Publ. Inst. Statist. Univ. Paris*, 27, 49-70.

Tiago de Oliveira, J., 1983. Gumbel distribution. in *Encyclopedia of Statistical Sciences* III, N.L. Johnson and S. Katz eds., 552- 558, Wiley, New York.

Tiago de Oliveira, J., 1984. Initiation of statistical decision for Weibull distribution. *cuad Bioestad. Aplic. Informat.*, 2, 495-499.

Tiago de Oliveira, J., 1986. Statistical choice of univariate extreme models, (summary). *Mem. Acad. Ciências Lisboa, cl. Ciências*, 27, 107-110.

*****