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 7

Statistical Decision for Weibull Distribution (For Minima)

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

Abstract

The Weibull distribution for minima is explained. Weibull random variable (for minima) has a finite location parameter. The Weibull distribution deals with modeling of failure and helps to determine 'time-to-failure'. Characteristics are presented from definition of distribution function. Statistical decision for the 3-parameters case, estimation, testing and point prediction are discussed. Prediction difficulties are depending on the interplay of the location and shape parameter of the distribution.

Published Online

23 June 2017

Keywords

Weibull distribution;
Minima;
Point prediction;
Time-to-failure.

Editor(s)

J.C. Tiago de Oliveira

7.1 Introduction

The Weibull distribution for minima is one of the asymptotic distributions for minima when the initial distribution has a left-end point (recall that the other possible asymptotic distribution is the Gumbel distribution for minima). Thus any Weibull random variable (for minima) has a finite location parameter λ (a *threshold)* and it takes the form

$$W_\alpha(x|\lambda,\delta) = W_\alpha((x-\lambda)/\delta) = 1 - \exp\{-(\frac{x-\lambda}{\delta})^\alpha\} \quad \text{for} \quad x \geq \lambda$$
$$W_\alpha(x|\lambda,\delta) = W_\alpha((x-\lambda)/\delta) = 0 \quad \text{for} \quad x \leq \lambda.$$

Originally published in 'Statistical Analysis of Extremes', 1997, 2016
http://dx.doi.org/10.21523/gcb1.1712

For $\alpha = 1$, we have the exponential distribution and for $\alpha = 2$ we have the Rayleigh distribution important in delay-times and Oceanography, respectively. Evidently, $W_\alpha(z) = 1 - \Psi_\alpha(-z)$.

It is immediate that $Y = ((X - \lambda)/\delta)^\alpha$ has the standard exponential distribution. Also if $\lambda = \lambda_0$ is known (it can be taken as zero, for convenience, by subtraction) $Y = -\log(X - \lambda_0)$ has a Gumbel distribution for maxima with the location and dispersion parameters $-\log\delta$ and $1/\alpha$. Thus the theory for Gumbel distribution can be applied and the maximum likelihood estimators are (if $\lambda_o = 0$)

$$\frac{n}{\hat{\alpha}} - \sum_1^n \log x_i - n\,\frac{\sum_1^n x_i^{\hat{\alpha}} \cdot \log x_i}{\sum_1^n x_i^{\hat{\alpha}}} = 0$$

$$\hat{\delta} = \frac{1}{n}\sum_1^n x_i^{\hat{\alpha}}.$$

The solution of the first equation needs an initial seed which can be obtained either using the fact that $\{-\log x_i\}$ has a general Gumbel distribution where $1/\alpha$ is the dispersion parameter among other ways.

The random pair $(\hat{\delta}, \hat{\alpha})$ *has a binormal asymptotic behaviour with mean values* (δ, α) , *variances* $\left(1 + \frac{6}{\pi^2}(1-\gamma)^2\right)\delta^2/n\alpha^2 and\ 6\,\alpha^2/n\,\pi^2$, *covariance* $\frac{6(1-\gamma)\delta}{n\,\pi^2}$, *and correlation coefficient* $\rho = (1 + \frac{\pi^2}{6(1-\gamma)^2})^{-1/2} =$ .31307.

The confidence intervals; tests, estimators of quantiles, predictors, etc., can be dealt with by the usual techniques, as done before in previous chapters.

Weibull (1939) gave an empirical derivation of this distribution in an analysis of dynamic breaking strengths by requiring only that is should meet certain practical criteria. That paper and other related papers by Weibull dealing with the modelling of failure (or survival) data seem to have found a large audience chiefly among those concerned with reliability analysis. The use of the Weibull distribution as a model for "time-to-failure" has thus been widespread. It arises naturally also in the analysis of droughts, in Oceanography, in Survival Analysis; see, for instance, Gumbel (1954).

The Weibull distribution for maxima has been used to model maximum temperatures, maximum wind speeds and maximum earthquake magnitudes; see Jenkinson (1955) and Yegulalp and Kuo (1974).

Other methods of estimation, as the ones given for the Gumbel distribution, can be used by the reduction to $y_i = -\log x_i$.

Recall that if X has a Fréchet distribution with dispersion and shape parameters δ and α, respectively, then $1/X$ has a two-parameter Weibull distribution with shape parameter α and dispersion parameter δ^{-1}. Thus, when data have not been censored, estimation procedures that have been derived for the two-parameter Fréchet distribution can be used to estimate parameters of the Weibull distribution by simply taking reciprocals of the observed values or directly when appropriate.

Methods applicable for Weibull censoring on the right will be appropriate for censoring Fréchet data sets on the left, i.e., the smallest data values, etc.

For these censorings, and for obtaining best invariant estimators of $-\log\delta$ and $1/\alpha$ among those which are linear in the Gumbel order statistics (logarithms of the Weibull order statistics in reverse order), weights have been calculated and tabulated by Laue (1974) for $n = 1(1)25,\ r = 2(1)n$. For zero censoring they are equal to the weights calculated and tabulated by Mann (1967) for best linearly invariant estimation from the smallest $r \leq n$ Gumbel order statistics.

From best linearly invariant estimators of $-\log\delta$ and α^{-1} one can obtain estimators of the location and dispersion parameters and best linearly estimators of any quantile of the distribution. Invariant estimators and maximum likelihood equations yield estimators of α^{-1} and $-\log\delta$ that tend to be very nearly equal in value, and the constants in the variance-covariance matrix of the best linearly unbiased estimators can be used effectively to remove the bias from maximum likelihood estimators if approximate unbiased estimators are desired. All these types of estimators are asymptotically unbiased, asymptotically efficient and asymptotically normal.

7.2 Characteristics of the Weibull distribution

From the definition of the distribution function

$W_\alpha(x|\lambda,\delta) = 1 - \Psi_\alpha(-x|\lambda,\delta)$

we can get, when they exist, the usual coefficients. We will obtain them for the reduced random variable $Z = (X-\lambda)/\delta$, the transfer of the general case being immediate.

The computation of β_1 and β_2, also depending on α, is left as an exercise.

The graphs of probability densities $W'_\alpha(z) = \alpha\, z^{\alpha-1} \exp(-z^\alpha)(z \geq 0)$ for different values of $\alpha > 0$ are given in Figure 7.1.

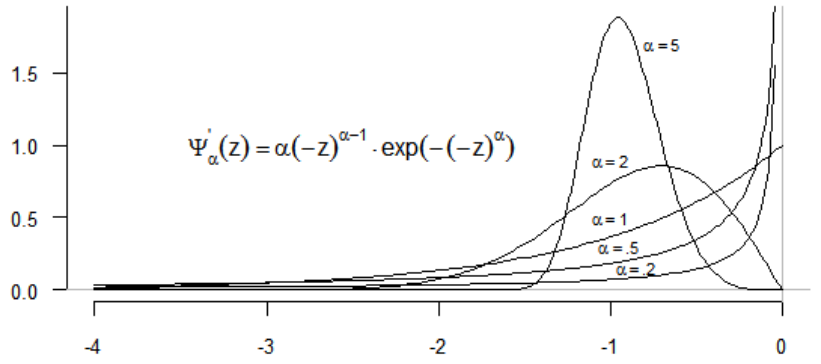

Figure 7.1 Reduced Weibull densities (for minima).

Evidently as $W_\alpha(z) = 1 - \Psi_\alpha(-z)$ the density $W'_\alpha(z)$ is the mirror image of $\Psi'_\alpha(z)$.

We have

$\mu'_r = M(Z^r) = \Gamma(1 + r/\alpha)$

and so the mean value is $\mu = \Gamma(1 + r/\alpha)$, the variance $\sigma^2 = \Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)$, the corresponding values for X being obvious.

Also we have the mode $\mu' = 0$ if $\alpha \leq 1$, $\mu' = (1 - 1/\alpha)^{1/\alpha}$ if $\alpha \geq 1$, the median $\tilde{\mu} = (\log 2)^{1/\alpha}$ and the p -quantil $\chi_p = (-\log(1-p))^{1/\alpha}$ with $\chi_{1-1/c} = 1$.

Note that $\alpha(1-Z)$, with Z having the distribution function $W_\alpha(z)$, has mean value $\alpha(1-\Gamma(1+1/\alpha)) = \gamma - (\frac{\pi^2}{6}+\gamma^2)/2\alpha + O(\alpha^{-2})$ and variance $\alpha^2(\Gamma(1+2/\alpha)) - \Gamma^2(1+1/\alpha) = \frac{\pi^2}{6} + O(\alpha^{-1})$ which suggests that $\alpha(1-Z)$ converges to a Gumbel distribution of maxima when $\alpha \to \infty$ and so, for large α, we can use it as an approximation. The exact proof is left as an exercise.

The moment generating function of Z does not have an explicit expression but that of $\log Z$ (which is a Gumbel random variable of minima) is

$$M\left(e^{\theta \log Z}\right) = M\left(Z^{\theta}\right) = \Gamma(1+\theta/\alpha).$$

The computation of β_1 and β_2, also depending on α, is left as an exercise.

7.3 Estimation and testing

Some difficulties can appear for the Weibull distribution depending on the interplay of the location (λ) and shape parameter (α).

If $\lambda = \lambda_0$ is known, as said before, there is no problem with its reduction to the Gumbel distribution for maxima.

Let us consider first the situation when α is known and (λ, δ) unknown — a real location-dispersion case. It is very easy to verify that if $\alpha < 1$, $\hat{\lambda} = x'_1 = \min(x_1, \dots, x_n)$ and *any positive statistic* $\delta^*(x_1, \dots, x_n) > 0$ are maximum likelihood estimators of (λ, δ)!

For $\alpha = 1$ (exponential distribution) we have $\hat{\lambda} = x'_1 = \min(x_1, \dots, x_n)$ and $\hat{\delta} = \bar{x} - \hat{\lambda}$, but owing to the influence of the minimum $\hat{\lambda}$ the asymptotic behaviour of $(\hat{\lambda}, \hat{\delta})$ is not binormal, because $\hat{\lambda}$ also has the same exponential distribution with location parameter λ and dispersion parameter δ/n. Note, also, that $P(\hat{\lambda} > \lambda) = 1$.

It can remarked that the minimum-variance unbiased estimators of λ and δ are

$$\lambda^* = (n\,x'_1 - \bar{x})/(n-1) = \hat{\lambda} - \hat{\delta}/(n-1)$$

and

$\delta^* = n(\bar{x} - x'_1)/(n-1) = n\,\hat{\delta}/(n-1);$

See Epstein and Sobel (1954) for details.

For $1 \le \alpha \le 2$ the maximum likelihood estimators exist but the Cramér-Rao bound, using the second derivatives does not. Other bounds, valid throughout the range of α, were found in Tiago de Oliveira and Littauer (1976). Are smaller than the Cramér-Rao bound in the range $\alpha > 2$, as shown in the paper by Tiago de Oliveira (1968).

In prediction results we will develop these bounds.

For $1 \le \alpha \le 2$ the maximum likelihood estimator $\hat{\lambda}$ is not asymptotically normal and whether or not it is efficient is an open question; see Woodroofe (1974). For known $\alpha < 2$, Tiago de Oliveira (1983) has shown that $\delta^{*\alpha} = \sum_{i=1}^{n}(x_i - x'_1)^{\alpha/n}$ has all the asymptotic properties of the maximum likelihood estimator of α. If α is known but λ and δ are unknown, the tables of Harter and Dubey (1967) can be used to test hypotheses concerning the mean value and variance.

The case where $\alpha > 2$ is simple and regular; see Cramér (1946). The maximum likelihood estimators of (λ, δ) for $\alpha = \alpha_0 (> 2)$ known, are given by the equations

$$\frac{\sum_1^n (x_i - \hat{\lambda})^{\alpha_0 - 1}}{\sum_1^n (x_i - \hat{\lambda})^{\alpha_0}} = \frac{\alpha_0 - 1}{n\,\alpha_0} \sum_1^n \frac{1}{x_i - \hat{\lambda}}$$

$$\hat{\delta} = \left(\frac{1}{n}\sum_1^n (x_i - \hat{\lambda})\right)^{1/\alpha_0}.$$

The random pair $(\hat{\lambda}, \hat{\delta})$ is asymptotically normal with mean values (λ, δ), variances $\alpha_0^2\,\delta^2/n\Delta, (\alpha_0 - 1)^2\,\Gamma(1 - 2/\alpha_0)\delta^2/n\Delta$, covariance $-\alpha_0(\alpha_0 - 1)\,\Gamma(1 - 1/\alpha_0)\delta^2/n\Delta$, and correlation coefficient $\rho = -\Gamma(1 - 1/\alpha_0)/\sqrt{\Gamma(1 - 2/\alpha_0)}$, where Δ denotes

$\alpha_0^2\,(\alpha_0 - 1)^2[\Gamma(1 - 2/\alpha_0) - \Gamma^2(1 - 1/\alpha_0)].$

From these results the usual tests and confidence intervals can be made. The estimator $\hat{\lambda} + \chi_p\hat{\delta}$ of the p-quantile $\lambda + \chi_p\,\delta$ is asymptotically normal with mean value $\lambda + \chi_p\,\delta$ and variance

$$((\alpha_0-(\alpha_0-1)\,\Gamma(1-1/\alpha_0)\chi_p)^2+(\alpha_0-1)^2[\Gamma(1-2/\alpha_0)-\Gamma^2(1-1/\alpha_0)]\,\chi_p)\delta^2/n\Delta$$

$$=\left(\frac{(\alpha_0-(\alpha_0-1)\,\Gamma(1-1/\alpha_0)\chi_p)^2}{\Delta}+\frac{\chi_p^2}{\alpha_0^2}\right)\delta^2/n$$

$$=\left(\frac{(1-\Gamma(2-1/\alpha_0)\chi_p)^2}{(\alpha_0-1)^2(\Gamma(1-2/\alpha_0)-\Gamma^2(1-1/\alpha_0))}+\frac{\chi_p^2}{\alpha_0^2}\right)\delta^2/n.$$

For prediction, the mean value of the minimum of the next m observations is $\lambda+\Gamma(1+1/\alpha_0)/m^{1/\alpha_0}\,\delta$. As before we use as predictor $\hat{\lambda}+\Gamma(1+1/\alpha_0)/m^{1/\alpha_0}\,\hat{\delta}$ and the mean square error of the predictor is, as we can take $\chi_p=\Gamma(1+1/\alpha_0)$,

$$(\Gamma(1+2/\alpha_0)-\Gamma^2\big(1+1(\alpha_0)\big)\,\delta^2+$$

$$\left(\frac{1-\Gamma(2-1/\alpha_0)\,\Gamma\big(1+1(\alpha_0)\big)^2}{(\alpha_0-1)^2\,(\Gamma(1-2/\alpha_0)-\Gamma^2(1-1/\alpha_0))}+\frac{\Gamma^2(1+1/\alpha_0)}{\alpha_0^2}\right)\delta^2/n.$$

Let us now consider the 3-parameter Weibull distribution. The state of affairs in this case is very awkward, chiefly because we have to estimate α with a global behaviour changing according to its real value; see Dubey (1967) and (1967') and Mann (1968) to get the first feelings of the difficulties of the problem, chiefly connected with the non-existence of maximum likelihood estimators for all values of α. Cohen and Whitten (1982) advise that one should attempt to obtain maximum likelihood estimators for the three-parameter Weibull distribution unless there is reason to expect $\alpha>2.2$.

With all these restrictions, let us write the maximum likelihood equations. From the likelihood of the sample $(x_1,\dots,x_n)$

$$L(\lambda,\delta,\alpha|x_i)=0 \qquad \text{if} \qquad \min(x_1,\dots,x_n)\le\lambda$$

$$L(\lambda,\delta,\alpha|x_i)=\frac{\alpha^n}{\delta^{\alpha n}}\prod_1^n(x_i-\lambda)^{\alpha-1}\exp\{-\sum_1^n(\tfrac{x_i-\lambda}{\delta})^{\alpha}\}\ \text{ if}\ \min(x_1,\dots,x_n)>\lambda$$

we get

$$\hat{\delta}\,\hat{\alpha}=\frac{1}{n}\sum_1^n(x_i-\hat{\lambda})^{\hat{\alpha}}$$

$$\frac{\hat{\alpha}-1}{\hat{\alpha}}\sum_{1}^{n}\frac{1}{x_i-\hat{\lambda}}=\frac{n\sum_1^n\left(x_i-\hat{\lambda}\right)^{\hat{\alpha}-1}}{\sum_1^n(x_i-\hat{\lambda})^{\hat{\alpha}}}$$

$$\frac{1}{\hat{\alpha}}+\frac{1}{n}\sum_{1}^{n}\log\left(x_i-\hat{\lambda}\right)=\sum_{1}^{n}(x_i-\hat{\lambda})^{\hat{\alpha}}\log(x_i-\hat{\lambda})/\sum_{1}^{n}(x_i-\hat{\lambda})^{\hat{\alpha}}.$$

For this system we have to solve for $(\hat{\lambda},\hat{\alpha})$, as a first step, the last two equations and then compute $\hat{\delta}$. If $\alpha>2$ the system is asymptotically trinormal and the previous techniques applied to the Fréchet distribution can be used. Those equations can be compared and are similar, to those for the Fréchet distribution of maxima, as could be expected. Also it can be shown that:

The random triple $(\hat{\lambda},\hat{\delta},\hat{\alpha})$ is asymptotically trinormal with mean values $(\hat{\lambda},\hat{\delta},\hat{\alpha})$ if $\alpha>2$ and variance-covariance matrix analogous to the one for Fréchet distribution.

The real important situations are considered below, particularly in the subsection "The maximum likelihood estimator when $\alpha_0>2$".

The usual techniques of confidence interval estimation and hypothesis testing can be applied as before. Prediction will be discussed in the next section.

Lemon (1975) modified the likelihood equations so that one need iteratively solve only two equations for estimates of λ and α, which together then specify an estimate of δ. Cohen and Whitten (1982) suggest other modifications to the maximum likelihood methods which involve simplification of one of the equations.

For estimation of the three Weibull parameters, the method of moments seems to work well, in general, as estimators can always be obtained. The tables of Dubey (1967) applying to the two-parameter distribution and the Monte Carlo results of Cohen and Whitten (1982) applying to the three-parameter distribution, indicate that the moment estimators are very efficient for $\alpha=4.7$ and become progressively less efficient as α converges towards 2 or towards $+\infty$. Cohen and Whitten (1982) provide a table which facilitates estimation of α from the distribution skewness coefficient β_1. These authors also give several modifications to the moment equations which simplify estimation of all three parameters.

Another iterative procedure for estimating λ is one by Mann and Fertig (1973). If α is not too large, this method also provides a lower confidence bound for λ. Also we can use the proposed statistic to test $\lambda = 0$ vs. $\lambda > 0$, which is very important in fatigue failure research. The basic statistic for this test is a ratio of selected sums of differences of successive ordered logarithms of sample observations, each difference being divided by its expected value. To obtain a point estimate of λ (which is median unbiased) or a lower confidence bound, the initial guess of λ must be subtracted from each observation and iterated until the modified statistic is numerically equal to a percentile of the statistics distribution. The method has two advantages: the test statistic, under $\alpha = 0$, has a Beta distribution and it is appropriate for use with censored data. The later advantage can also be claimed for a similar statistic by Weissman (1978) and for certain linear estimators of $-\log \delta$ and α^{-1} that can be applied to the ordered logarithms of the observed data after an estimate of λ is subtracted from each of the original values.

Wyckoff, Bain and Engelhardt (1980) suggest the use of a simplified linear estimator of α^{-1}, after having substituted λ by a simple estimator, as an approximation of the best linear unbiased estimator. This seems to yield appropriate and quite efficient estimates for α^{-1}, and ultimately for $\log \delta$, if $\alpha \leq 3.0$ when there is no censoring (or if α is somewhat less than 3.0 when there is censoring). For larger values of α, estimates of α^{-1} obtained by this method have a substantial negative bias unless sample size is extremely large.

Besides the linear estimators of $\log \delta$ and α^{-1}, there are other simple estimators of the Weibull parameters that are appropriate for use under certain conditions. For $\alpha \leq 1$, a special case suggested by Zanakis (1979) of the estimator of λ previously proposed by Dubey (1966), namely, $\lambda' = (x_1' \; x_n' - x_2'^2)/(x_1' + x_n' - 2\,x_2')$, has relatively small bias and mean-square error. For $\alpha > 1$, a statistic investigated by Kappenam (1981) has smaller bias and mean-square error than the last one for estimating λ, but is also considerably more difficult to compute, involving a weighted average of all the order statistics.

Zanakis and Mann (1982) compared two simple estimators of α based on a few order statistics:

$$\alpha_1^* = \log\,[(x_k' - x_j')/(x_j' - x_i')]/3.643$$

and $\alpha_2^* = \log[(x_m' - \lambda')/(x_h' - \lambda')]/2.989$

with $1 \le i < j < k \le n$, $1 \le h < m \le n$ defined appropriately and λ' given above.

They found both estimators to be good for small values of α, but having more negative bias than estimators based on the procedures prescribed by Wyckoff, Bain and Engelhardt (1980). The simple estimators are useful for first guesses in maximum-likelihood estimation and in other situations in which quick results are necessary.

7.4 Point prediction

As before, the mean-square point predictor of the minimum of the next m observations is evidently the estimator of the mean value of $W = \min(Y_1, \dots, Y_n)$. The distribution of W is, as seen before, $W_\alpha(\frac{x-\lambda}{\delta\, m^{-1/\alpha}})$ if the distributions of Y_j, as well as those of x_j, are $W_\alpha((x-\lambda)/\delta)$. The mean value of W is

$$M(W) = \lambda + m^{-1/\alpha}\,\Gamma(1 + 1/\alpha)\,\delta,$$

and thus a point predictor of W can be obtained substituting $(\lambda, \delta, \alpha)$ by their estimators. But difficulties exist when λ is not known and α may be ≤ 2. If $\alpha = \alpha_0$ is known and $\alpha_0 > 2$ (practically speaking $\alpha_0 > 2.1$), there are no difficulties with the maximum likelihood estimation; likewise if $(\lambda, \delta, \alpha)$ are all unknown but it is known that $\alpha > 2$. The general theory used before for the Fréchet distribution can also be applied here.

If $\lambda = \lambda_0$ is known then $-\log(x - \lambda_0)$ has a Gumbel distribution with location and dispersion parameters $-\log\delta$ and α^{-1} and estimation can be dealt with by the methods of Chapter 5.

The theory of best quasi-linear prediction when $\alpha = \alpha_0$ leads to an unmanageable formula for the predictor, following Tiago de Oliveira (1966) as we will recall briefly.

Later we will make considerable use of Tiago de Oliveira and Littauer (1976).

The likelihood of the observed sample is

$$\bar{L}_{\alpha_0}(\lambda, \delta | x_i) = 0 \qquad \text{if} \quad \min(x_i) > \lambda$$

$$\bar{L}_{\alpha_0}(\lambda,\delta|x_i) = \prod_1^n \{\frac{1}{\delta} W'_{\alpha_0}(\frac{x_i-\lambda}{\delta})\} \quad \text{if} \quad \min(x_i) > \lambda$$

and the likelihood of the sample and the future minimum W from m observations is

$$L_{\alpha_0}(\lambda,\delta|x_i,w) = \bar{L}_{\alpha_0}(\lambda,\delta|x_i,\dots,x_n)\frac{1}{\delta m^{-1/\alpha_0}} W'_{\alpha_0}(\frac{w-\lambda}{\delta m^{-1/\alpha_0}})$$

where $W_\alpha(z)$ was defined above.

According to the paper previously cited, the mean-square predictor has the expression

$$p_m(x_i) = \frac{\int_0^{+\infty} d\xi\, \xi^{n-1} \int_{-\infty}^{+\infty} d\eta\, \bar{L}(\eta+\xi x_1,\dots,\eta+\xi x_n)[\mu_m(\eta+\xi x_1,\dots,\eta+\xi x_n)-\eta]}{\int_0^{+\infty} d\xi\, \xi^n \int_{-\infty}^{+\infty} d\eta\, \bar{L}(\eta+\xi x_1,\dots,\eta+\xi x_n)}$$

where $L(z_1,\dots,z_n;v)$ denotes the likelihood of the first sample and the minimum of the second sample, $\bar{L}(x_1,\dots,x_n) = \int_{-\infty}^{+\infty} d\,v\; L(x_1,\dots,x_n\,;v)$ is the marginal likelihood of the sample and

$$\mu_m(x_i) = \int_{-\infty}^{+\infty} d\,v\;\, v\, L(x_1,\dots,x_n;v)\,/\bar{L}(x_1,\dots,x_n)$$

is the conditional mean. As it is irrelevant, from the expression of $p_m(x_i)$, we can always suppose $\lambda = 0$ and $\delta = 1$. This predictor was shown in that paper to be quasi-linear, i.e., $p_m(\lambda + \delta\, x_i) = \lambda + \delta\, p_m(x_i)$, for every λ and $\delta(>0)$.

Note that, in our case, we have

$\mu_m(x_i) = \mu_{\alpha_{0,m}} = \int_{-\infty}^{+\infty} d\,v\, m^{1/\alpha_0}\, v\, W'_\alpha(m^{1/\alpha_0}\, v) = \Gamma(1+1/\alpha_0)/m^{1/\alpha_0} = \mu_{\alpha_0}, 1/m^{1/\alpha_0}$.

and $\bar{L}(x_i) = L_{\alpha_0}(x_i)$.

As $W'_{\alpha_0}(z) = 0$ if $z \le 0$, the integration takes place in the region $\alpha + \beta\, x_i \ge 0$, that is, $\alpha \ge -\beta\, l(x_i)$ where $l(x_i) = \min(x_1,\dots,x_n)$.

The predictor can be written as

$$p_{\alpha_0}, m(x_i) = \frac{\int_0^{+\infty} d\xi\, \xi^{n-1} \int_{l(x_i)}^{+\infty} d\eta\, (\mu_{\alpha_{0,m}-\eta}) \prod_{i=1}^{n} W'_{\alpha_0}(\eta + \xi\, x_i)}{\int_0^{+\infty} d\xi\, \xi^{n} \int_{l(x_i)}^{+\infty} d\eta \prod_{i=1}^{n} W'_{\alpha_0}(\eta + \xi\, x_i)}.$$

Putting $\zeta = \xi + \eta\, l(x_i)$ we get

$$p_{\alpha_0,m}(x_i) = l(x_i) + \frac{\int_0^{+\infty} d\xi\, \xi^{n-1} \int_0^{+\infty} d\,\zeta (\mu_{\alpha_{0,m}-\zeta}) \prod_{i=1}^{n} W'_{\alpha_0}(\zeta+\xi\,(x_i - l(x_i)))}{\int_0^{+\infty} d\xi\, \xi^{n} \int_0^{+\infty} d\,\zeta \prod_{i=1}^{n} W'_{\alpha_0}(\zeta+\xi\,(x_i - l(x_i)))}.$$

It is evident that computation facilities will not be useful at all, as the expression above implies the computation of the predictor $p_{\alpha_0,m}(x_i)$ for any n (sample size) and for any (practically) observable sample mean, except for special cases. For example, in the case of the exponential distribution $(\alpha_0 = 1)$, we have

$W'_1(z) = 0$ if $z \leq 0$

$= e^{-z}$ if $z \geq 0$,

$\mu_{1,m} = 1/m$, and thus we obtain

$$P1, m(x_i) = l(x_i) + (\frac{1}{m} - \frac{1}{n})[\,\bar{x} - l(x_i)].$$

7.5 Lower bounds for the mean-square error

As the maximum likelihood estimators are not regular or do not exist if $\alpha \leq 2$, it is natural to search for a quasi-linear predictor and thus to try to obtain a lower bound for the mean-square error of any quasi-linear statistic $f(x_i)$ for $\lambda + \mu_{\alpha_0,m}\, \delta$.

The mean-square error of a quasi-linear statistic $f(x_i)$ is

$$M(f(x_i) - W)^2 = \delta^2\, M(f(z_i) - v)^2$$

where $x_i = \lambda + \delta\, z_i$ and $w = \lambda + \delta\, v$ (i.e., z_i and v are the reduced values of the observations). From now on we will deal, thus, with reduced observations $(\lambda = 0, \delta = 1)$ because we are going to evaluate efficiencies and the factor δ^2 cancels out.

The values of $M(f(z_i) - v)^2$ can be split simply by putting $f(z_i) - v = (f(z_i) - \mu_{\alpha_0,m}) - (v - \mu_{\alpha_0,m})$ and so, by independence between the first and second sample, we have

$$M(f(z_i) - v)^2 = M(f(z_i) - \mu_{\alpha_0,m})^2 + M(v - \mu_{\alpha_0,m})^2$$

$$= D^2_{n,m}(f) + \sigma^2_m(\alpha_0) = \text{mean} - \text{square error of } f(z_i) \text{ about } \mu_{\alpha_0,m} + \text{variance of } v.$$

Thus we will study $D^2_{n,m}(f)$. It was shown in Tiago de Oliveira (1968), using the Schwarz inequality conveniently, that, in the general case, if we denote by $\varphi(\lambda, \delta)(-\infty < \lambda < \infty, 0 < \delta < \infty)$

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \mu_m (\lambda + \delta z_1, \dots, \lambda + \delta z_n) \bar{L}(z_1, \dots, z_n) \, dz_1 \dots dz_n$$

and by $w_n(\lambda, \delta)$

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \frac{\bar{L}^2(z_1, \dots, z_n)}{\delta^n \bar{L}(\lambda + \delta z_1, \dots, \lambda + \delta z_n)} \, dz_1 \dots dz_n.$$

then

$$B^2 = \sup_{\lambda,\delta} \frac{(\lambda - \varphi(\lambda, \delta) + \varphi(0,1)\delta)^2}{w_n(\lambda, \delta) - 2\delta + \delta^2}$$

is a lower bound for the mean-square error $D^2_{n,m}(f) = M(f(z_i) - \mu_m(z_i))^2$ of a quasi-linear statistic $f(z_i)$(when $\lambda = 0, \delta = 1$). In the expression of $w_n(\lambda, \delta)$ in the integrals the indeterminate ratio $\frac{0}{0}$ is taken to be equal to zero.

B^2 is indeterminate for $\lambda = 0, \delta = 1$. In our case, owing to independence, we have

$$\varphi(\lambda, \delta) = \mu_{\alpha_0,m} \quad \text{and} \quad w_n(\lambda, \delta) = w^n_{\alpha_0}(\lambda, \delta)$$

where

$$w_{\alpha_0}(\lambda, \delta) = \int_{-\infty}^{+\infty} \frac{W'_{\alpha_0}(z)^2}{\delta \, W'_{\alpha_0}(\lambda + \delta z)} \, dz$$

and thus we have

$$B^2 = \sup_{\lambda,\delta} \frac{(\lambda + \delta \mu_{\alpha_0,m} - \mu_{\alpha_0,m})^2}{w^n_{\alpha_0}(\lambda, \delta) - 2\delta + \delta^2}.$$

As B^2 is almost impossible to obtain, we will put $\lambda = 0$ above and compute

$$\overline{B}^2 = \sup_{\delta} \frac{\mu^2_{\alpha_0,m}(1-\delta)^2}{w^n_{\alpha_0}(0,\delta) - 2\,\delta + \delta^2} \le B^2.$$

and as

$$w_{\alpha_0}(0,\delta) = 1/[\delta^{\alpha_0}(2-\delta^{\alpha_0})] \qquad \text{for} \quad 0 < \delta^{\alpha_0} < 2,$$

we get the (smaller) lower bound

$$\overline{B}^2 = \frac{\mu^2_{\alpha_0,m}}{1 + n\,\alpha_0^2} \sim \frac{\mu^2_{\alpha_0,m}}{n\,\alpha_0^2} = \frac{\mu^2_{\alpha_0,1}}{n\,\alpha_0^2\, m^{2/\alpha_0}}$$

and an (under) evaluation of efficiency is

$$\lim \frac{\overline{B}^2}{D^2_{n,m}(f)} = \frac{\mu^2_{\alpha_0,1}}{\alpha_0^2\, m^{2/\alpha_0} \lim n\, D^2_{n,m}(f)}.$$

For $\alpha_0 = 1, \overline{B}^2 \sim D^2_{n,m}(f)$ when $f(x_i) = l(x_i) + (\frac{1}{m} - \frac{1}{n})\,[\overline{x} - l(x_i)]$ is the best predictor for the exponential distribution, already given.

In case of regularity, as happens for $\alpha_0 > 2$, we can obtain bounds analogous to the Cramér-Rao bounds for estimation. In the last paper referred to it was shown, for the case of independence with the density $g(z)$ for the (reduced) z_i and the mean μ_m for v , to have the expression

$$B'^2 \sim \frac{S_0 \mu_m^2 - 2\,S_1\,\mu_m + S_2 - 1}{n(S_0\,S_2 - S_1^2 - S_0)}$$

with

$$S_0 = \int_{-\infty}^{+\infty} x^p\, \frac{g'(x)^2}{g(x)}\, dx\,.$$

Putting $g(z) = W'_{\alpha_0}(z)$ and $\mu_m = \mu_{\alpha_0,m}$ we get, as

$$S_p = [(\alpha_0 - 1)^2 + p(\alpha_0 + p - 2)]\,\Gamma(1 + \frac{p-2}{\alpha_0})\,,$$

$$B'^2 \sim \frac{\alpha_0^2 - 2\,\alpha_0(\alpha_0-1)\,\Gamma(1-1/\alpha_0)\,\mu_{\alpha_0,m} + (\alpha_0-1)^2\,\Gamma(1-2/\alpha_0)\,\mu^2_{\alpha_0,m}}{n\,\alpha_0^2(\alpha_0-1)^2\,[\Gamma(1-2/\alpha_0) - \Gamma^2(1-1/\alpha_0)]}\,.$$

Asymptotically we have $\overline{B}^2 \le B'^2$ and so the (under-) evaluation of the efficiency is to be taken as

$B'^2/D^2_{n,m}(f)$.

In fact, as the maximum likelihood estimators are the best when regular $(\alpha_0 > 2)$, we are going to obtain a set of simple predictors, chiefly for $\alpha_0 \le 2$, such that they have reasonably good efficiency for some interval in $[0,2]$.

7.6 Some predictors for the shape parameter $\alpha_0 \le 2$

It is desirable to explore different predictors for different ranges of α_0 because of computational difficulties. Let us begin by examining the region $\alpha_0 \le 2$, with a suggestion from the result for the exponential $(\alpha_0 = 1)$.

The simple predictor for the exponential distribution suggests the use of the statistic

$$l(x_i) + a(\alpha_0)\,(\bar{x} - l(x_i))$$

as predictor, where $a(\alpha_0)$ is to be determined so as to minimize the mean-square error, which for reduced values is

$$D^2_{n,m} = M(l(x_i) + a(\alpha_0)\,(\bar{x} - l(x_i)) - \mu_{\alpha_0,m})^2\,.$$

This predictor is suggested also by the fact that the lower limit λ has $l(x_i)$ as a sufficient estimator (but an over-estimator).

The sample average is a quasi-linear statistic converging to $\mu_{\alpha_0,1} = \lambda + \delta\,\Gamma(1+1/\alpha_0)$, with variance $\delta^2\,b(\alpha_0)/n$ with

$$b(\alpha_0) = [\Gamma(1+2/\alpha_0) - \Gamma^2(1+1/\alpha_0)];$$

$l(x_i)$ has the mean value $\mu_{\alpha_0,n} = \lambda + \delta\,\Gamma(1+1/\alpha_0)/n^{1/\alpha_0}$, its variance being $\delta^2\,b(\alpha_0)/n^{2/\alpha_0}$; the correlation coefficient $\rho_{\alpha_0,n} = \rho(\bar{x}, l(x_i))$ converges to zero; see Tiago de Oliveira (1962) and Rosengard (1966).

Thus, the mean-square error $D^2_{n,m}$ for reduced observations as a function of a, can be written as $M[l(x_i) + a(\bar{x} - l(x_i)) - \mu_{\alpha_0,m}]^2 =$

$\mathrm{M}[l(\mathrm{x_i}) - \mu_{\alpha_0,\mathrm{m}}]^2 + 2\,\mathrm{M}\big[\big(l(\mathrm{x_i}) - \mu_{\alpha_0,\mathrm{m}}\big)\big(\bar{\mathrm{x}} - l(\mathrm{x_i})\big)\big]\mathrm{a} + \mathrm{M}[\bar{\mathrm{x}} - l(\mathrm{x_i})]^2\mathrm{a}^2.$
Using the decomposition $l(\mathrm{x_i}) - \mu_{\alpha_0,\mathrm{m}} = l(\mathrm{x_i}) - \mu_{\alpha_0,\mathrm{n}} + \mu_{\alpha_0,\mathrm{n}} - \mu_{\alpha_0,\mathrm{m}}$ and $\bar{\mathrm{x}} - l(\mathrm{x_i}) = \bar{\mathrm{x}} - \mu_{\alpha_0,1} + \mu_{\alpha_0,1} - \mu_{\alpha_0,\mathrm{n}} + \mu_{\alpha_0,\mathrm{n}} - l(\mathrm{x_i})$, we have

$$\mathrm{M}[(l(\mathrm{x_i}) - \mu_{\alpha_0,\mathrm{m}})^2] = \frac{\mathrm{b}(\alpha_0)}{\mathrm{n}^{2/\alpha_0}} + (\mu_{\alpha_0,\mathrm{n}} - \mu_{\alpha_0,\mathrm{m}})^2\,,$$

$$\mathrm{M}\big[\big(l(\mathrm{x_i}) - \mu_{\alpha_0,\mathrm{m}}\big)\,\big(\bar{\mathrm{x}} - l(\mathrm{x_i})\big)\big] = \rho_{\alpha_0,\mathrm{n}}\;\frac{\mathrm{b}(\alpha_0)}{\mathrm{n}^{2/\alpha_0+1/2}} - \frac{\mathrm{b}(\alpha_0)}{\mathrm{n}^{2/\alpha_0}} +$$

$$\big(\mu_{\alpha_0,\mathrm{n}} - \mu_{\alpha_0,\mathrm{m}}\big)\,\big(\mu_{\alpha_0,1} - \mu_{\alpha_0,\mathrm{n}}\big),$$

$$\mathrm{M}[(\bar{\mathrm{x}} - l(\mathrm{x_i}))^2] = \frac{\mathrm{b}(\alpha_0)}{\mathrm{n}} + (\mu_{\alpha_0,1} - \mu_{\alpha_0,\mathrm{n}})^2 + \frac{\mathrm{b}(\alpha_0)}{\mathrm{n}^{2/\alpha_0}} - 2\rho_{\alpha_0,\mathrm{n}}\;\frac{\mathrm{b}(\alpha_0)}{\mathrm{n}^{1/\alpha_0+1/2}}\,.$$

The minimum of the quadratic is attained at some value whose limit is

$$-\mathrm{a}(\alpha_0) = \mu_{\alpha_0,\mathrm{m}}/\mu_{\alpha_0,1} = \mathrm{m}^{-1/\alpha_0},$$

thus the predictor is $l(\mathrm{x_i}) + \mathrm{m}^{-1/\alpha_0}\,(\bar{\mathrm{x}} - l(\mathrm{x_i}))$. The asymptotic values of the mean-square error are then:

$$\mathrm{D}^2_{\mathrm{n,m}} = \frac{\mu^2_{\alpha_0,\mathrm{m}}\;\mathrm{b}(\alpha_0)}{\mu^2_{\alpha_0,1}\;\mathrm{n}} = \frac{\mathrm{b}(\alpha_0)}{\mathrm{n}\,\mathrm{m}^{2/\alpha_0}} \qquad \text{if} \quad \alpha_0 < 2,$$

$$\mathrm{D}^2_{\mathrm{n,m}} \sim \frac{\mathrm{m} - 2\sqrt{\mathrm{m}} + 2 - \pi/4}{\mathrm{m}\,\mathrm{n}} \qquad \text{if} \quad \alpha_0 = 2\,,$$

$$\mathrm{D}^2_{\mathrm{n,m}} \sim \frac{\Gamma\,(1 + 2/\alpha_0)\,(1 - \mathrm{m}^{-1/\alpha_0})^2}{\mathrm{n}^{2/\alpha_0}} \qquad \text{if} \quad \alpha_0 > 2\,.$$

Consequently, the asymptotic efficiency is, since $\overline{\mathrm{B}}^2 \sim \frac{\mu^2_{\alpha_0,\mathrm{m}}}{\mathrm{n}\,\alpha_0^2}$,

$$\mathrm{eff}\,(\alpha_0) = \frac{\Gamma^2\,(1 + 1/\alpha_0)}{\alpha_0^2\;\mathrm{b}(\alpha_0)} \qquad \text{if}\; \alpha_0 < 2\,.$$

$$\mathrm{eff}\,(\alpha_0) = \frac{\pi}{16(\mathrm{m} - 2\sqrt{\mathrm{m}} + 2 - \pi/4)} \qquad \text{if}\; \alpha_0 = 2\,,$$

$$\mathrm{eff}\,(\alpha_0) = 0 \qquad \text{if}\; \alpha_0 > 2\,.$$

Zero efficiency for $\alpha_0 > 2$ results from the slow convergence of $l(\mathrm{x_i})$ to the parameter λ, an effect that in this region overcomes the more rapid convergence of the average $\bar{\mathrm{x}}$. The efficiency for $\alpha_0 > 2$ is independent of m. A short table of the efficiency for the above predictor is, for $\mathrm{m} = 1$.

$$\begin{array}{cccccc} \alpha_0 & & 1/3 & 1/2 & 1 & 2^- \\ \text{eff}\,(\alpha_0) & & 0.5 & 0.8 & 1 & 0.91\,. \end{array}$$

The predictor for $1/2 < \alpha_0 < 2$ is thus reasonably efficient. Outside this range for the shape parameter, another predictor must be found.

Consider now the usage of a sample quantile $Q(x_i)$ and the minimum value $l(x_i)$. It is apparent (by the argument given above), that, since the variance of a quantile is of order $1/n$, a linear combination

$$l(x_i) + a(\alpha_0)\,(Q(x_i) - l(x_i))$$

may be efficient only in the region $\alpha_0 \leq 2$; it will be useful in the region $\alpha_0 < 1/2$. For known $\alpha = \alpha_0$, let us denote by υ the quantile of probability $\omega = W_{\alpha_0}(\upsilon)$. It is well known (Cramér, 1946) that the sample quantile $Q(x_i)$ for probability ω, with reduced observations, is asymptotically normal with expected value υ and asymptotic variance of the order of $\omega(1-\omega)/n\,W'^2_{\alpha_0}(\upsilon)$. As

$$W_{\alpha_0}(z) = 1 - e^{-z^{\alpha_0}}, z \geq 0,$$

the asymptotic variance is of the order of

$$\frac{e(\upsilon^{\alpha_0}) - 1}{n\,\alpha_0^2\,\upsilon^2(\alpha_0 - 1)}\,.$$

The mean-square error, for reduced values, is

$$M[(l(x_i) + a(\alpha_0)\big(Q(x_i) - l(x_i)\big) - \mu_{\alpha_0,m})^2] = M[(l(x_i) - \mu_{\alpha_0,m})^2] +$$

$$2M[\big(l(x_i) - \mu_{\alpha_0,m}\big)\big(Q(x_i) - l(x_i)\big)]\,a(\alpha_0) + M[(Q(x_i) - l(x_i))^2]\,a^2(\alpha_0)$$

with

$$M[(l(x_i) - \mu_{\alpha_0,m})^2\,] = \frac{b(\alpha_0)}{n^{2/\alpha_0}} + (\mu_{\alpha_0,n} - \mu_{\alpha_0,m})^2\,,$$

$$M\big[\big(l(x_i) - \mu_{\alpha_0,m}\big)\big(Q(x_i) - l(x_i)\big)\big] = \rho'_{\alpha_0,n}\,\tau_{\alpha_0,n}\,\frac{b(\alpha_o)^{1/2}}{n^{1/\alpha_0}} - \frac{b(\alpha_0)}{n^{2/\alpha_0}}$$

$$+\big(\mu_{\alpha_0,n} - \mu_{\alpha_0,m}\big)\,\big(\upsilon_{\alpha_0,n} - \mu_{\alpha_0,n}\big),$$

$$M[(Q(x_i) - l(x_i))^2] = \tau^2_{\alpha_0,n} + (\upsilon_{\alpha_0,n} - \mu_{\alpha_0,n})^2$$

$$+\frac{b(\alpha_0)}{n^{2/\alpha_0}} - 2\rho'_{\alpha_0,n}\ \tau_{\alpha_0,n}\ \frac{b(\alpha_0)^{1/2}}{n^{1/\alpha_0}},$$

where we have made a decomposition analogous to the previous one and where $\upsilon_{\alpha_0,n}, \tau^2_{\alpha_0,n}$, and $\rho'_{\alpha_0,n}$ denote the asymptotic mean value and the variance of the sample quantile and the correlation coefficient between the sample quantile and the minimum. As might be expected, $\rho'_{\alpha_0,n} \to 0$ (Rosengard , 1966)*.

Denoting by υ_{α_0} the limiting value of $\upsilon_{\alpha_0,n}$, we see that the limiting value of $a(\alpha_0)$ that minimizes the mean-square error is

$$a(\alpha_0) = \mu_{\alpha_0,m}/\upsilon_{\alpha_0}$$

and the asymptotic mean-square error for $\alpha_0 < 2$ is, for reduced values,

$$D^2_{n,m} = \frac{\mu^2_{\alpha_0,m}\,(e^{\upsilon^{\alpha_0}} - 1)}{n\,\alpha_0^2\ \upsilon^{2/\alpha_0}}.$$

For $\alpha_0 = 2$ we have

$$D^2_{n,m} \sim \{\frac{\pi(e^{\upsilon^2} - 1)}{16\,\upsilon^4} + (1 - \frac{\sqrt{\pi}}{2\sqrt{m}\ \upsilon})^2\}.$$

We will only consider the problem for $\alpha_0 < 2$ .The case for $\alpha_0 = 2$ must be dealt with numerically. The minimum mean-square error is attained for $\upsilon^{\alpha_0} = 1.593624$, and we have approximately

$$D^2_{n,m} \sim \frac{\mu^2_{\alpha_0,m}}{n\,\alpha_0^2}\,(\frac{e^{1.593624} - 1}{(1.593624)^2}) = \frac{1.5441386\ \mu^2_{\alpha_0,m}}{n\,\alpha_0^2}$$

so that the asymptotic efficiency is

$$\frac{1}{1.5441386} = 0.65\,.$$

Taking the approximation $\upsilon^{\alpha_0} = 1.60$, the (constant) efficiency is $\frac{1}{1.5441532} = 0.65$ also. The predictor will then be

---

* In Cramér (1946), p. 368, it is shown implicitly that $\upsilon_{\alpha_0,n} - \upsilon_{\alpha_0} = O(n^{1/2})$.

$$l(x_i) + \mu_{\alpha_0,1}(1.60m)^{-1/\alpha_0}(Q(x_i) - l(x_i))$$

where $Q(x_i) = \chi^*_{.796}$ is then the sample quantile for probability $\omega = 1 - e^{-\upsilon^2} = 1 - e^{-1.60} = 0.796$. This technique, whose asymptotic efficiency is independent of m and α_0, is to be used for values of $\alpha_0 \leq 0.40$ where the efficiency of the predictor $l(x_i) + a(\alpha_0)\ (\bar{x} - l(x_i))$ is smaller than 0.65. For $\alpha_0 = 2$ a better predictor will be considered.

7.7 The maximum likelihood predictor when $\alpha_0 > 2$

The maximum likelihood estimators of λ and δ for known $\alpha_0 (> 2)$ are given by the equations

$$\frac{\sum_1^n (x_i - \hat{\lambda})^{\alpha_0 - 1}}{\sum_1^n (x_i - \hat{\lambda})^{\alpha_0}} = \frac{\alpha_0 - 1}{n\ \alpha_0} \sum_1^n (x_i - \hat{\lambda})^{-1}$$

and

$$\hat{\delta}\alpha_0 = \frac{1}{n} \sum_1^n (x_i - \hat{\lambda})^{)^{\alpha_0}},$$

where the variances and covariance are asymptotically

$$V(\hat{\lambda}) \sim \frac{\alpha_0^2}{n\,\Delta}\ \delta^2, V(\hat{\delta}) \sim \frac{(\alpha_0 - 1)^2\ \Gamma(1 - 2/\alpha_0)}{n\,\Delta}\ \delta^2$$

$$C(\hat{\lambda}, \hat{\delta}) \sim -\frac{\alpha_0(\alpha_0 - 1)\ \Gamma(1 - 1/\alpha_0)}{n\,\Delta}\ \delta^2$$

$$\Delta = \alpha_0^2\ (\alpha_0 - 1)^2\ [\Gamma(1 - 2/\alpha_0) - \Gamma^2(1 - 1/\alpha_0)].$$

The predictor of the quantile $\lambda + \mu_{\alpha_0,m}\,\delta$ to be used must be $\hat{\lambda} + \mu_{\alpha_0,m}\,\hat{\delta}$ with the mean-square error

$$D^2_{n,m}(\hat{\lambda} + \mu_{\alpha_0,m}\,\hat{\delta}) \sim \frac{\alpha_0^2 - 2\alpha_0(\alpha_0 - 1)\Gamma(1 - 1/\alpha_0)\mu_{\alpha_0,m} + (\alpha_0 - 1)^2\Gamma(1 - 2/\alpha_0)\mu^2_{\alpha_0,m}}{n\alpha_0^2(\alpha_0 - 1)^2[\Gamma(1 - 2/\alpha_0) - \Gamma^2(1 - 1/\alpha_0)]}.$$

This is exactly equal to the asymptotic value of B'^2 and so the efficiency is 1. This technique requires the use of computers for the determination of $\hat{\lambda}$ and $\hat{\delta}$. Since, however, it is desirable to have simpler and more practical techniques when computer facilities are not available or, also, to initiate iteration, we consider in the next section some all-range predictors

in this category. Such predictors can be quite useful since the range $\alpha_0 \leq 2$ may be important for practical applications.

7.8 Some all-range predictors

The difficulties found above in respect to computability and efficiency of the proposed predictors can suggest the consideration of the traditional approach, using quantiles and moments, since their variances are of order n^{-1}.

For example, consider the quantile $q(x_i)$ for the probability $\omega = W_{\alpha_0}(\mu_{\alpha_0,m})$, whose asymptotic efficiency for the bound $\overline{B}^2$ is given as

$$\frac{\Gamma^2(1+1/\alpha_0)}{e^{\Gamma^{\alpha_0}(1+1/\alpha_0)/m}-1)\, m^2}$$

which is approximated by $\Gamma^{2-\alpha_0}\,(1+1/\alpha_0)/m$ for large m. For $m = 1$ (one-step predictor) we have the following table of efficiencies:

α_0	0.25	0.5	1	2
$\text{eff}(\alpha_0)$	0.60	0.64	0.58	0.52 .

This short table shows that the sample quantile for probabilities $W_{\alpha_0}(\mu_{\alpha_0,m})$, whose efficiency fades out with m, is not useful in the range $\alpha_0 \leq 2$.

For $\alpha_0 > 2$ the efficiency is given, using the bound B'^2, by

$$\frac{\mu_{\alpha_0,m}^{2(\alpha_0-1)}[\alpha_0^2 - 2\alpha_0(\alpha_0-1)\,\Gamma(1-\frac{1}{\alpha_0})\mu_{\alpha_0,m}] + (\alpha_0-1)^2\,[\Gamma(1-2/\alpha_0)\mu_{\alpha_0,m}^2]}{(e^{\mu_{\alpha_0,m}^{\alpha_0}}-1)(\alpha_0-1)^2\,[\Gamma(1-2/\alpha_0)-\Gamma^2(1-1/\alpha_0)]}\,.$$

The short table of efficiencies for $m = 1$ is:

α_0	2.5	3.0	3.5	4.0	4.5	5.0
$\text{eff}(\alpha_0)$	.53	.55	.57	.58	.59	.60 ,

which shows that the quantile predictor has an almost constant efficiency of about 60%.

Consider now the use of a predictor based on the method of moments $\bar{x} + a(\alpha_0)\, s$, where s denotes the standard deviation of the first sample. The mean-square error for reduced values is

$$D^2_{n,m} = M[(\bar{x} + a(\alpha_0)s - \mu_{\alpha_0,m})^2] = M[(\bar{x} - \mu_{\alpha_0,m})^2] +$$

$$2M\big[(\bar{x} - \mu_{\alpha_0,m})s\big]\, a(\alpha_0) + M(s^2)\, a^2(\alpha_0).$$

The covariance between $\bar{x}$ and s is asymptotic to $[\beta_1(\alpha_0)\, b(\alpha_0)]/2n$ by the δ-method, where $\beta_1(\alpha_0)$ is the skewness coefficient.

Let $\sigma_n(\alpha_0)$ be the mean value of s. We have

$$\sigma_n(\alpha_0) - \sqrt{b(\alpha_0)} \sim -\frac{(\beta_2(\alpha_0) + 3)\sqrt{b(\alpha_0)}}{8n}$$

where $\beta_2(\alpha_0)$ is the kurtosis coefficient (*) and

$$D^2_{n,m} = \frac{b(\alpha_0)}{n} + (\mu_{\alpha_0,1} - \mu_{\alpha_0,m})^2 + 2[C(\bar{x}, s) + (\mu_{\alpha_0,1} - \mu_{\alpha_0,m})\, \sigma_n(\alpha_0)]\, a(\alpha_0) +$$
$$b(\alpha_0)\, a^2(\alpha_0).$$

The value at which the quadratic attains the minimum, as

$$C(\bar{x}, s) \sim \frac{\beta_1(\alpha_0)\, b(\alpha_0)}{2n},$$

(*) A short proof of the asymptotic value of σ_n is given. From Cramér (1946), p. 353, we have the following relation,

$$M(s - \sigma_n(\alpha_0))^2 = \frac{(\beta_2(\alpha_0) - 1)b(\alpha_0)}{4n} + O(n^{-2}).$$

Developing the square and taking mean values, we have

$$\frac{n-1}{n}\, b(\alpha_0) - \sigma_n^2(\alpha_0) = \frac{(\beta_2(\alpha_0) - 1)b(\alpha_0)}{4n} + O(n^{-2}).$$

and so

$$\sigma^2\, n(\alpha_0) - b(\alpha_0) = -\frac{\beta_2(\alpha_0) + 3}{4n}\; b(\alpha_0) + O(n^{-2})$$

from which the desired result follows.

converges to

$$a(\alpha_0) = -\frac{\mu_{\alpha_0,1} - \mu_{\alpha_0,m}}{\sqrt{b(\alpha_0)}}$$

so that the asymptotic value of the mean-square error is

$$D^2_{n,m} = \frac{b(\alpha_0)}{n}\{1 - \frac{\beta_1(\alpha_0)(\mu_{\alpha_0,1}-\mu_{\alpha_0,m})}{\sqrt{b(\alpha_0)}} + \frac{(\beta_2(\alpha_0)+3)(\mu_{\alpha_0,1}-\mu_{\alpha_0,m})^2}{4\, b(\alpha_0)}\}$$

where

$$\beta_1(\alpha_0) = [\Gamma(1+3/\alpha_0) - 3\Gamma(1+2/\alpha_0)\,\Gamma(1+1/\alpha_0) + 2\Gamma^3(1+1/\alpha_0)]/b^{3/2}(\alpha_0)$$

and

$$\beta_2(\alpha_0) = (\Gamma(1+4/\alpha_0) - 4\Gamma(1+3/\alpha_0)\,\Gamma(1+1/\alpha_0) + 6\Gamma(1+2/\alpha_0)\,\Gamma^2(1+1/\alpha_0) - 3\Gamma^4(1+1/\alpha_0)]/b^2(\alpha_0)$$

Then the predictor is

$$\bar{x} - (\frac{\mu_{\alpha_0,1} - \mu_{\alpha_0,m}}{b(\alpha_0)^{1/2}})\, s.$$

The asymptotic efficiency with respect to B^2 is

$$\frac{\mu^2_{\alpha_0,m}}{\alpha_0^2\, b(\alpha_0)\,\{1 - \frac{\beta_1(\alpha_0)(\mu_{\alpha_0,1} - \mu_{\alpha_0,m})}{\sqrt{b(\alpha_0)}} + \frac{(\beta_2(\alpha_0)+3)\,(\mu_{\alpha_0,1} - \mu_{\alpha_0,n})^2}{4\, b(\alpha_0)}\}}$$

and the asymptotic efficiency for $\alpha_0 > 2$, with respect to B'^2, is given by

$$\frac{\alpha_0^2 - 2\alpha_0(\alpha_0-1)\,\Gamma(1-1/\alpha_0)\,\mu_{\alpha_0,m} + (\alpha_0-1)^2\,\Gamma(1-2/\alpha_0)\mu^2_{\alpha_0,m}}{\alpha_0^2(\alpha_0-1)^2\;\; b(\alpha_0)\,[\Gamma(1-\frac{2}{\alpha_0}) - \Gamma^2(1-\frac{1}{\alpha_0})]\,\{1 - \frac{\beta_1(\alpha_0)(\mu_{\alpha_0,1} - \mu_{\alpha_0,m})}{\sqrt{b(\alpha_0)}} + \frac{(\beta_2(\alpha_0)+3)\,(\mu_{\alpha_0,1} - \mu_{\alpha_0,m})^2}{4b(\alpha_0)}\}.}$$

A short table for the first efficiency for $m = 1$ is:

α_0	0.5	1.0	2
$\text{eff}(\alpha_0)$	0.80	1.00	0.91.

7.9 An example

Consider the natural and simple predictor $l(x_i) + m^{-1/\alpha_0}\,(\bar{x} - l(x_i))$.

We will experiment with the predictor generating a sample of n Weibull observations, computing the predictor, generating a new sample of m

under the same conditions and comparing the minimum of the second sample with the value of the predictor.

The experiments are given in Table 7.1 for

Table 7.1

α_0	n	m
1.2:	50	20,40
1.2:	100	50
1.2:	500	50,100,200,300,400
1.5:	50	20,40
1.5:	100	50
1.5:	500	50,100,200,300,400
2.0:	50	20,40
2.0:	100	50
2.0:	500	100,200,300,400.

The values $l_1^{'}, l_2^{'}$, ..., denote the observed smallest values in the second sample of m (generated observations).

Table 7.2

n	m	x_n	x_m	l	$l_1^{'}$	$l_2^{'}$	$l_3^{'}$	$l_4^{'}$	$l_5^{'}$	$l_6^{'}$	$l_7^{'}$
50	20	1.106482	0.618368	0.087202	0.171168	0.013614	0.067171	0.090015	0.170916		
50	40	0.999691	0.864131	0.037770	0.082242	0.012288	0.026883	0.027323	0.057124		
100	50	0.993607	0.870766	0.058743	0.094630	0.009652	0.075735	0.089076			
500	50	0.955362	1.053420	0.017990	0.053974	≥ 1’					
500	100	0.986448	0.811142	0.0005847	0.026973	0.009693	0.021219				
500	200	0.948326	0.995920	0.009625	0.020975	≥ 1’					
500	300	0.995920	0.995920	0.009625	0.016349	≥ 1’					
500	400	0.938055	0.927183	0.000687	0.007048	0.004685					
50	20	0.802585	1.157086	0.041248	0.097344	≥ 1’					
50	40	0.864939	0.919777	0.091047	0.157814	0.020942					
100	50	0.952770	0.885156	0.080250	0.144538	0.041757					
500	50	0.936661	0.987915	0.032320	0.098952	0.083899					
500	100	0.902707	0.901279	0.024205	0.065282	0.044599	0.059561				
500	200	0.931957	0.893814	0.009254	0.036226	0.011021 4					
500	300	0.906875	0.894213	0.003972	0.024120	≥ 1’					
500	400	0.948098	0.958196	0.004700	0.022078	0.001939	0.018320				
50	20	0.936665	0.870119	0.254734	0.407218	0.138128	0.288591	0.386233			
50	40	0.818688	0.914252	0.089966	0.205187	0.018976	0.090137	0.102450			
100	50	0.822177	0.981068	0.100348	0.202430	≥ 1’					
500	100	0.895189	0.837045	0.056219	0.140116	0.021859	0.023695	0.056533	0.063959	0.099370	0.118113
500	200	0.918345	0.862408	0.076011	0.135573	0.010172	0.045231	0.054819	0.064132	0.093357	0.097550
500	300	0.901596	0.894982	0.066431	0.114653	0.065947	0.076226	0.085539	0.095738	0.099727	
500	400	0.897788	0.884023	0.047098	0.089632	0.031186	0.046646	0.055452	0.076625	0.086074	

The table as well a great part of this section was taken from Tiago de Oliveira and Littauer (1976).

7.10 Downpassing probability of level

The technique that follows is the one used for Fréchet distribution. In what follows we will suppose that $\alpha = \alpha_0 (> 2)$ is known and, as said before, we have

$$\mathrm{V}(\hat{\lambda}) \sim \frac{\alpha_0^2}{\mathrm{n}\,\Delta}\,\delta^2,\ \mathrm{V}(\hat{\delta}) \sim \frac{(\alpha_0-1)^2\,\Gamma(1-2/\alpha_0)}{\mathrm{n}\,\Delta}\,\delta^2$$

and

$$\mathrm{C}(\hat{\lambda},\hat{\delta}) \sim -\frac{\alpha_0(\alpha_0-1)\,\Gamma(1-1/\alpha_0)}{\mathrm{n}\,\Delta}\,\delta^2$$

where $\Delta = \alpha_0^2((\alpha_0-1)^2(\Gamma\big(1-2(\alpha_0)-\Gamma^2(1-1/\alpha_0)\big)$; $\alpha_0 > 2$ guarantees the existence and regularity of the maximum likelihood estimators.

Now, we will be interested in the downpassing a of level a. $\mathrm{P}(a) = \mathrm{P}(a\,|\lambda,\delta) = \mathrm{W}_{\alpha_0}(\frac{a-\lambda}{\delta})$ is estimated by $\widehat{\mathrm{P}}(a) = \mathrm{P}\big(\,a\big|\hat{\lambda},\hat{\delta}\big)$.

As before, $\sqrt{\mathrm{n}}\,(\widehat{\mathrm{P}}\,(a) - \mathrm{P}(a))$ is asymptotically equivalent to

$$\sqrt{\mathrm{n}}\,\{\big(\hat{\lambda}-\lambda\big)\frac{\partial\,\mathrm{P}(a)}{\partial\,\lambda} + \big(\hat{\delta}-\delta\big)\frac{\partial\,\mathrm{P}(a)}{\partial\,\delta}\}$$

$$= \sqrt{\mathrm{n}}\,\big(1-\mathrm{P}(a)\big)\log(1-\mathrm{P}(a))\,\alpha_0\,\{\frac{\hat{\lambda}-\lambda}{a-\lambda} + \frac{\hat{\delta}-\delta}{\delta}\}$$

whose asymptotic variance is

$$\mathrm{V}(a) = (1-\mathrm{P}(a\,))^2\,(\log\,(1-\mathrm{P}(a\,)))^2\,\{1+\frac{\alpha_0^2}{\Delta}\,(\frac{\alpha_0-\delta}{a-\lambda} -$$

$(\alpha_0-1)\,\Gamma(1-1/\alpha_0)^2)$ estimated as $\widehat{\mathrm{V}}(a)$, substituting (λ,δ) by $\big(\hat{\lambda},\hat{\delta}\big)$.

As $\sqrt{\mathrm{n}}\,(\frac{\widehat{\mathrm{P}}(a)-\mathrm{P}(a)}{\sqrt{\widehat{\mathrm{V}}(a)}})$ is asymptotically standard normal, the one-sided asymptotic confidence interval with significance level α' is

$$0 \le \mathrm{P}(a) \le \widehat{\mathrm{P}}(a) + \frac{\lambda_{\alpha'}\sqrt{\widehat{\mathrm{V}}(a)}}{\sqrt{\mathrm{n}}}$$

as before.

In the case where the location parameter is known, taken as zero for convenience, we have

$$P(a) = P(a\,|\delta, \alpha) = W_\alpha(a/\delta)$$

and the variance-covariance matrix of $(\hat{\delta}, \hat{\alpha})$ is

$$\frac{1}{n}\begin{bmatrix} \frac{\alpha^2}{\delta^2} & -\frac{1-\gamma}{\delta} \\ -\frac{1-\gamma}{\delta} & \frac{1}{\alpha^2}(\frac{\pi^2}{6} + (1-\gamma)^2) \end{bmatrix}^{-1} = \frac{1}{n}\begin{bmatrix} \frac{\delta^2}{\alpha^2}(1 + \frac{6(1-\gamma)^2}{\pi^2}) & -\frac{6(1-\gamma)}{\pi^2}\delta \\ -\frac{6(1-\gamma)}{\pi^2}\delta & \frac{6}{\pi^2}\alpha^2 \end{bmatrix}$$

The result is analogous to the one for Fréchet distribution (for maxima) with the exchange of $\frac{1-\gamma}{\delta}$ for $-\frac{1-\gamma}{\delta}$ in the both matrices.

Then, evidently $\sqrt{n}(\hat{P}(a) - P(a))$ is asymptotically equivalent to

$$\sqrt{n}((\hat{\delta} - \delta)\frac{\partial P(a)}{\partial \delta} + (\hat{\alpha} - \alpha)\frac{\partial P(a)}{\partial \alpha}) = \sqrt{n}\big(1 - P(a)\big)\log\big(1 - P(a)\big)(\alpha\,\frac{\hat{\delta}-\delta}{\delta} - (\log\frac{a}{\delta})(\hat{\alpha} - \alpha))$$

with asymptotic zero mean value and variance

$$V(a) = (1 - P(a))^2\,(\log(1 - P(a)))^2\,(1 + \frac{6}{\pi^2}\,(1 - \gamma + \alpha\log\frac{a}{\delta})^2)$$

also analogous to the result for the Fréchet distribution. In the same way we see that:

The one-sided asymptotic confidence interval for $P(a)$ is

$$0 \le P(a) \le (\hat{P}(a) + \frac{\lambda_{\alpha'}}{\sqrt{n}}\sqrt{\hat{V}(a)},$$

with significance level α', where $\hat{V}(a)$, as before, is the expression of $V(a)$ with (δ, α) substituted by $(\hat{\delta}, \hat{\alpha})$.

For an application to earthquakes see Tiago de Oliveira (1984'") and references therein.

A worked example

Consider the data given in van Montfort and Otten (1978) of failure times corresponding to a complete sample of $n = 24$ observations, obtained in that order:

Table 7.3

i	t_1	i	t_1
1	6.0	13	109.0
2	8.6	14	118.0
3	17.8	15	119.0
4	18.0	16	138.0
5	27.5	17	141.0
6	33.5	18	144.0
7	50.5	19	146.0
8	51.5	20	150.0
9	69.0	21	151.0
10	74.0	22	153.0
11	74.0	23	153.1
12	89.0	24	153.2

Supposing $\lambda = 0$, we obtain $\hat{\delta} = 101.023905$ and $\hat{\alpha} = 4.31899592$ (with 4 iterations!) from which the usual decision problems can be dealt with.

The Kolmogoroff-Smirnov statistic is $\hat{K}S_n = \sup_x | S_n(x) - \Phi_{\hat{\alpha}}(x/\hat{\delta})| = .14182353$ and as $n = 24$ we have $\sqrt{n}\ \hat{K}S_n = .69479$, accepting thus the Weibull model for minima.

When this is the case — and it seems reasonable in the context of failure times — as $X = -\log T$ has a Gumbel maxima distribution with parameters $\lambda' = -\log \delta$ and $\delta' = 1/\alpha$ and, also, $X = 1/T$ has a Fréchet maxima distribution with parameters $\lambda'' = 0, \delta'' = \delta^{-1}$ and $\alpha'' = \alpha$, as seen before, we can use these transformations.

References

Cohen, Jr A. C. and Wittten, B. J., 1982. Modified maximum likelihood and modified moments estimators for the three - parameter Weibull distribution. *Commun. Statist. - Theor. Meth.*, 11, 2631-2656.

Cramér, H., 1946. *Mathematical Methods of Statistics*, Princeton University Press, New Jersey.

Dubey, D., 1967a. Revised tables for asymptotic efficiencies of the moment estimators for the parameters of the Weibull laws. *Naval Res. Logist. Quat.*, 14, 261-267.

Dubey, D., 1967b. Some percentile estimators for Weibull parameters, *Technometrics*, 9, 112-129.

Dubey, S. D., 1966. Hiper-efficient estimators of the location parameter of Weibull law. *Naval Res. Logist. Quat.*, 13, 253-264.

Dubey, S. D., 1967. On some permissible estimators of the location parameter of the Weibull and certain other distributions. *Technometrics*, 9, 293-307.

Epstein, B. and Sobel, M., 1954. Some theorems relevant to life testing from an exponential distribution. *Ann. Math. Statist.*, 25, 373-381.

Gumbel, E. J., 1954. *Statistical Theory of Extreme Values and Some Practical Applications*. Nat. Bur. Standards, Appl. Math. Series, (33), Washington.

Harter, H. L. and Dubey, S. D., 1967. Theory and tables for tests of hypothesis concerning the mean and the variance of a Weibull population. *Aerospace Res. Labs. Rep*., 67-0059, Wright -Paterson AFB, Ohio.

Jenkinson, A. F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quat. J. Roy. Meteor. Soc.*, 87, 158-171.

Kappenam, R. F., 1981. Estimation of Weibull location. *J. Statist. Comp. Simul.*, 13, 245-254.

Laue, R. V., 1974. Unpublished paper, *Bell Labs*., New Jersey.

Lemon, G. H., 1975. Maximum likelihood estimation of the three parameters Weibull distribution based on censored samples. *Technometrics*, 17, 247-254.

Mann, Nancy .R. and Fertig, Kenneth W., 1973. Tables for obtaining Weibull confidence bounds and tolerance bounds based on best linear invariant estimates of parameters of the extreme-value distribution. *Technometrics*, 15, 87-101.

Mann, Nancy R., 1967. Tables for obtaining the best linear invariant estimates of the parameters of the Weibull distributions. *Technometrics*, 9, 629-645.

Mann, Nancy. R., 1968. Point and interval estimation procedures for the two-parameter Weibull and extreme-value distributions. *Technometrics*, 10, 231-256.

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

Rosengard, A., 1966. Contribution a 1 'etude des liaisons limites entre les differents statistiques dun echantillon, Universite de Paris, These de Doctorat d'Etat.

Tiago De Oliveira J. and Littauer, S. B., 1976. Mean square invariant forecastere for the Weibull distributions. *Naval Res. Logist. Quat*., 23, 487-511.

Tiago De Oliveira J., 1966. Quasi-linearly invariant prediction. *Ann. Math. Statist*., 37, 1684-1687.

Tiago De Oliveira J., 1968. Efficiency evaluation for quasi-linear invariant predictors. *Rev. Belge Statist. Rech. Oper*., 9, 1-9.

Tiago De Oliveira J., 1984. Weibull distributions and large earthquake modeling. in *Probabilistic Methods in the Mechanics of Solids and Structures* (TUTAM Symposium in honour of Dr. W. Weibull), 81-89, S. Eggwertz and N.C. Lind eds., Springer –Verlag, Heidelberg.

Tiago De Oliveira, J., 1962. Structure theory of bivariate extremes extensions. *Estudos, Mat., Estatist. e Econometria, Lisboa*, 7, 165-194.

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.

Weibull, W., 1939. A statistical theory of the strength of materials. *Proc. Roy. Swed. Inst. Eng. Res.*, 151, Stockholm.

Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.*, 73, 812-815.

Woodrofe, M., 1974. Maximum likelihood estimation of a translation parameter of a truncated distribution – II. *Ann. Statist.*, 3, 474-448.

Wyckoff, J., Bain, L. J., and Engelhardt, M., 1980. Some complete and censored sampling results for the three- parameter Weibull Distribution. *J. Statist. Comp. Simul.*, 11, 139-151.

Yegulalp, T. M. and Kuo, J. T., 1974. Statistical prediction of the occurrence of maximum magnitude earthquakes. *Bull. Seismol. Soc. America*, 64, 393-414.

Zanakis, S. M. and Mann, Nancy R., 1982. A good simple percentile estimator of the Weibull shape parameter for use when all three parameters are unknown. *Naval Logist. Res. Quat.*, 29, 419-428.

Zanakis, S. M., 1979. A simulation study of some simple estimators for the three-parameter Weibull distribution. *J. Statist. Comp. Simul.*, 9, 101-116.

*****