3.3 Extreme Laws and Tail Distributions – Application to the probabilistic modelling of material samples for structural reliability analyses

3.3.1 Introduction

The development of probabilistic methods in the mechanics industry is of real interest for design optimisation or optimisation of the maintenance strategy of structures subject to reliability and availability constraints, as well as for the re-qualification of a structure following an incident. One of the main stumbling blocks to the development of probabilistic methods is substantiation of probabilistic models used in the studies. In fact, it is frequently necessary to estimate an extreme value based on a very small sample of existing data. Numerous universities and industrial teams are currently working to resolve those problems connected with industrial concerns and to propose a realistic statistical methodology to substantiate a probability distribution adjusted on a sample, notably in the distribution tail. One of the aims of this paper is to open the debate concerning this issue.

The sections 3.3.2 to 3.3.5 give a reminder of the context of the data statistical treatment and the issues and problems surrounding structural reliability applications. Section 3.3.6 recalls the relationships between extreme laws and the study of tail distribution. Section 3.3.7 describes the problem of the adequacy for extreme value estimation of a probabilistic distribution function (pdf) fitted around the maximum probability area. An example of application is then presented. A small realistic sample of toughness values measured on French nuclear reactor vessel steel is taken as the starting point. This example makes use of new and promising methods developed in the context of joint work conducted by INRIA Rhône-Alpes and Electricité de France Research & Development.

This paper adopts a probabilistic position to model the uncertainty associated with a value or sample dispersion: accordingly, other approaches will not be dealt with in the following. The aim of this section is not to compare the advantages and drawbacks of the various available methods.

3.3.2 Description of uncertainties associated with data

Whether a deterministic or probabilistic approach is implemented, sample or database treatment must be performed. A deterministic approach involves the identification of information like; minimum, maximum values, envelope curves, etc., while a probabilistic vision concentrates on the dispersion or variability of the value, through variation interval or fractile-type data, or a probability distribution. Remember that a fractile or quantile of the order a% is a real number, Xa, satisfying P(X ≤ Xa) = a%. Treatment is compatible with the intended application, as, for example, determining a good distribution representation around a central value or correctly modelling behaviour in a distribution tail, etc.

Tools to describe sample dispersion are taken from the statistic; however, their effectiveness is a function of the sample size. References [1] to [7] describe methods that may be used to adjust a probability distribution on a sample, and then verify the adequacy of this adjusted distribution in the maximum failure probability region. It is obvious that if data is lacking or scarce, these tools are difficult to use. Under such circumstances, it is entirely reasonable to refer to expert opinion in order to model uncertainty associated with a value, and then transcribe said information in the form of a probability distribution. This document does not describe methods available in these circumstances, for example, the maximum entropy principle ([2]), and reference to expert opinion and application of the Bayesian methods ([7], [8]).

* Chapter written by Dr Nicolas Devictor, Dr Michel Marquès and Dr Pieter Van Gelder
The practical approach may be summarised by three scenarios:

Scenario 1. If a lot of experience feedback data is available, the frequentist statistic is generally used. The objectivist or frequentist interpretation associates the probability with the observed frequency of an event. In this interpretation, the confidence interval of a parameter, p, has the property that the actual value of p is within the interval with a confidence level a; this confidence interval is calculated based on measurements.

Scenario 2. If data is not as abundant, expert opinion may be used to obtain modelling hypotheses. The Bayesian analysis is used to correct a priori values established based on expert opinion as a function of observed events. The subjectivist (or Bayesian) interpretation understands probability as a degree of belief in a hypothesis. In this interpretation, the confidence interval is based on a probability distribution representing the analyst's degree of confidence in the possible values of the parameter and reflecting his/her knowledge of the parameter.

Scenario 3. If no data is available on a parameter, its probabilistic representation may be obtained from a model and from the knowledge of the uncertainties on the input parameter of this model. The data to be gathered thus concerns the input parameters. The quality of the probabilistic analysis is a function of the credibility of statistics concerning these input parameters and that of the model. The following may be discerned:

- A structural reliability-type approach if the value sought is a probability,
- An uncertainty propagation-type approach if a statistic around the most probable value is considered.

"Scenario 1", where a large enough sample is available, i.e. the sample allows "characterisation of the relevant distribution with a known and adequate precision", begs the following questions:

Question 1: Is the distribution type selected relevant and justifiable? From the various statistical models available, what would be the optimal distribution choice?

Question 2: Would altering the distribution (all other things being equal) entail a significant difference in the results of the application?

Question 3: How can uncertainty associated with sample representativeness be taken into consideration (sample size, quality, etc)?

Justification is difficult for Scenarios 2 and 3. For example, if the parameters of a density are obtained from the first moments, it must be borne in mind that a precise estimation of the symmetry coefficient requires at least 50 values while kurtosis requires 100 data, except for very specific circumstances. Furthermore, the critical values used by tests to reject or accept a hypothesis are frequently taken from results that are asymptomatic in the sense that sample size tends towards infinity. Thus when sample size is small, the results of conventional tests should be handled with caution.

With respect to question 2, a study examining sensitivity to the probability distribution used provides information. There are two methods available for the sensitivity study:

- Moment identification-type method where it is assumed that the distribution changes while the first moments are preserved (mean, standard-deviation especially).
• Frequentiel or Bayesian approach where the sample is redistributed to establish the reliability data distribution parameters.

Another method is to take the uncertainty associated with some distribution parameters into consideration by replacing the parameters' deterministic value by a random variable. The thesis [9] provides other information in response to questions 2 and 3.

Conventional criticism concerning the statistical modelling of a database concerns:

• Difficulty in interpreting experience feedback for a specific application;
• Database quality, especially if few points are available;
• Substantiation of the probabilistic model built.

The probabilistic modelling procedure should attempt to reply to these questions. If it is not possible to define a correct probabilistic model, it is obvious that, under these circumstances, the quantitative results in absolute value are senseless in the decision process. However, the probabilistic approach always allows results to be used relatively, notably through:

• a comparison of the efficiency of various solutions from the standpoint of reliability, availability, for example;
• or classification of parameters that make the biggest contribution to the uncertainty associated with the response in order to guide R&D works in order to reduce said uncertainty.

This argument concerning the quality of uncertainty probabilistic models also has repercussions on the deterministic approach.

The deterministic approach involves validation of values used and also constitutes a sophisticated problem: it is not easy to prove that a value assumed to be conservative is realistic, especially if the sample is small or is a guarantee that the value is an absolute bound must be provided. Frequently conservative values used are formally associated with small or large fractiles of the orders of the values studied. Whereas the concept of fractile is associated with the probability distribution adjusted on the sample, and even with one of the distribution tails.

The probabilistic approach seems to be even more suitable to deal with the problem. In fact, the probabilistic model reflects the level of knowledge of variables and models, and confidence in said knowledge. By means of sensitivity studies, this approach allows the impact of the probabilistic model choice on risk to be objectively assessed. Furthermore, in the event of new information impugning probabilistic modelling, and consequently the fractiles of a variable, the Bayesian theory, that combines objective and subjective (expertise) data, allows the probabilistic model and results of the probabilistic approach to be updated stringently.

The adjustment of a probability distribution and subsequent testing of the quality of said adjustment around the central section (or maximum failure probability region) of the distribution constitute operations that are relatively simple to implement using the available statistical software packages (SAS, Statgraphics, Splus, Statistica, etc.), for conventional laws in any case. However, the interpretation and verification of results still requires the expertise of a statistician. For example, the following points:

• the results of an adjustment based on a histogram is sensitive to the intervals width;
• the maximum likelihood or moment methods are not suitable for modelling a sample obtained by overlaying phenomena beyond a given limit of an observation variable;
moment methods assume estimations of kurtosis and symmetry coefficients that are only usually specified for large bases (with at least one hundred values for kurtosis);

most statistical tests, specifically, the most frequently used Kolmogorov-Smirnov, Anderson-Darling and Cramer-Von Mises tests, are asymptotic tests;

in the Bayesian approach, the distribution selected a priori influences the result. Furthermore, the debate concerning whether or not the least informative law should be used has not been concluded.

Whereas in industrial studies, the characteristics of databases must be taken into consideration which may render implementation of an adjustment difficult; for example the following frequently encountered scenario:

1. the sample size is small and therefore asymptotic results need to be handled cautiously as well as approximations of more or less valid moments of an order greater than two;
2. sample data values are measured with an uncertainty;
3. sample homogeneity is not verified (mixture of samples taken from different populations, overlaying of phenomena, etc.);
4. if the area of interest is a distribution tail, it may be noted that the statistical theory and above all associated tools are less developed.

3.3.3 Selection of a distribution in practice

In practice, the criteria relevant to selecting a probabilistic model for a random variable would seem to be:

- Use a family compatible with the physical properties (bound value, symmetrical or not, exhibits exponential decay, etc),
- the result of data adjustment,
- distribution not rejected by statistical tests,
- the selection of a distribution that is least "informative" with respect to data or available information (i.e. the introduction of too many hypotheses is avoided).

Statistical tests are used to decide whether or not the adequacy of a selected distribution should be rejected a priori with a confidence level. Whereas, for samples that are not very homogeneous or small in size, several distributions are frequently accepted to represent the sample or, on the contrary, no conventional distribution can be accepted.

3.3.4 Some criticism

To select a specific distribution, the following methods may be applied:

- either rely on selection based on expert opinion or current practice like, for example, a Weibull distribution for reactor vessel steel toughness measurements ([10]) or a logarithmic normal distribution for a constraint;
- or compare confidence levels taken from various statistical tests for each accepted distribution and select the distribution associated with the highest confidence level;
- or further analyse adequacy in the zone of interest and select the distribution most graphically suited in the region of interest (for example a distribution tail for a reliability problem).

Another possibility is to use a resampling method: for example the Bootstrap method or cross-validation method, in order to select "the most relevant" distribution family (ref. for example [11] and [12]).
As regards the non-parametric Bootstrap method, "the most relevant" family is defined as that which is almost always accepted by adjustment tests. This method assumes that sample size is not too small. Experience has also shown that the non-parametric Bootstrap method is rarely suited to the treatment of extreme values.

As regards the parametric Bootstrap method, "the most relevant" family is defined as the family whose confidence intervals on parameters are smallest.

In the event that the graphic adjustment of a sample for various distribution features a "break", the rejection of all conventional distributions may be noted. For example, in Figure 3.3.1, graphically adjustment seems very good in the region of highest probability and bad at the distribution tail. These breaks may be explained by the fact that the sample is a mixture of several homogeneous sub-samples or is taken from a process combining several phenomena. Under such circumstances, if application is a reliability calculation where this distribution tail is, a priori, influential, there is no choice, the model must be rejected. For an uncertainty propagation calculation, this model may be accepted if the number of simulations is not too high (overly frequent simulations in the distribution tail should be avoided).

Practical solutions to obtain a general model include the use of mix models and definition of a distribution through rejoining.

A mix model is assumed to be a weighted sum of random variables, for example, in the case of 2 Gaussian components:

\[ G_{\theta, \gamma} = p \times \Phi_{\mu_1, \sigma_1} + (1-p) \times \Phi_{\mu_2, \sigma_2} \]

where this distribution is a mixture of two normal distributions, means \( m_1 \) and \( m_2 \) respectively, standard deviations \( s_1 \) and \( s_2 \) respectively.

The principle of the rejoining method is to build a distribution of a random variable \( X \) by rejoining two distribution functions, marked \( F \) on the left and \( H \) on the right of a threshold \( u \). Hence, for an observation \( x \) of \( X \), it may be written:

if \( x \leq u \) then \( X \rightarrow F \) and if \( x > u \) then \( X \rightarrow H \)

The two laws \( F \) and \( H \) are therefore truncated. For every value \( a \) of \([0; 1]\), it is possible to construct a continuous rejoining in \( u \) based on the value \( a \) at this point.

\[ G_a(x) = a \frac{F(x)}{F(u)} \quad \text{if} \quad x \leq u \]
Extreme Laws and Tail Distributions

\[ G_a(x) = 1 - (1-a) \frac{1-H(x)}{1-H(u)} \] if \( x > u \)

where \( G_a \) is the distribution resulting from rejoining.

3.3.5 Proposed method to determine adequacy of a distribution (central value)

This proposal is put forward in the context of a Kolmogorov-Smirnov test, however, it may be used for another a priori statistic, like Anderson-Darling, Cramer-von Mises…

Take for example a random sample \( E \), with size \( N \) and order of magnitude \( y \). The \( f(x,c0) \) type distribution (normal, Gumbel, etc.) was adjusted on this sample; \( c0 \), the vector of parameters obtained and KS0 the value of the Kolmogorov-Smirnov test was noted. The following question was raised: what is the probability that a random sample of size \( N \) taken from the distribution \( f(x,c0) \) would allow a value of the Kolmogorov-Smirnov test equal to KS0 to be obtained?

Tables of critical values are available that were obtained for a sample size tending towards infinity. Using these values when \( N \) is small is therefore questionable.

A confidence interval concerning the Kolmogorov-Smirnov test value is constructed as follows:
1. Based on the \( f(x,c0) \) distribution, which has been adjusted, a large number \( p \), of samples of size \( N \) are constructed with the same size as sample \( E \).
2. For each of these samples, the distribution parameter vector is established and the value of the Kolmogorov-Smirnov test, noted \( KSi \) is calculated.
3. This provides a sample with \( p \) values of \( KSi \). As \( p \) is very large, the fractiles of the statistic KS may be calculated with confidence like that the probability of the KS0 value.

Based on the above, a decision may be made whether or not to reject the hypothesis, that according to the Kolmogorov-Smirnov test the distribution \( f(x,c0) \) is an acceptable model of sample \( E \).

3.3.6 Extreme Laws and Tail Distributions

Assume that \( X_1, X_2, ..., X_n \) are independent and identically distributed random variables coming from a parent distribution with a cumulative distribution function \( F(x) \) and probability distribution function \( f(x) \). Define \( H_n=\max(X_1, X_2, ..., X_n) \) then \( H_n \) has a cumulative density function (CDF) given by:

\[ H_n(x)=P(\max(X_1, X_2, ..., X_n)<x)=F^n(x) \] (1)

Notice that the percentiles of \( H_n \) move to the right with increasing \( n \), approaching the upper and lower end points if they are bounded, or going to \( \infty \) if they are unbounded. When \( n \) goes to infinity, we have

\[ \lim_{n \to \infty} H_n(x) = \begin{cases} 1 & \text{if } F(x) = 1 \\ 0 & \text{if } F(x) < 1 \end{cases} \] (2)

that is to say, the limit distribution degenerates to a Dirac function. To avoid this degeneracy, we transform the random variable \( X \) by means of constants \( a_n \) and \( b_n \) such that

\[ \lim_{n \to \infty} H_n(a_n+b_n x) = \lim_{n \to \infty} F^n(a_n+b_n x) = H(x) \] (3)

where \( H(x) \) is a non degenerated CDF.

For instance:
Lifetime Management of Structures

\[
F_{X_{\text{max}}}(x) = (1 - e^{-\lambda x})^n \rightarrow 0 \quad (n \to \infty, \forall x). \quad \text{However if } x \text{ also goes to infinity with } a_n + b_n x \text{ with } a_n = \ln(n)/\lambda \quad \text{and } b_n = 1/\lambda, \text{ then:}
\]

\[
F_{X_{\text{max}}}(x) = F_{X_{\text{max}}}(a_n + b_n x) = F_{X_{\text{max}}}((\ln(n)/\lambda + x/\lambda)) = (1 - e^{-\lambda(\ln(n)/\lambda + x/\lambda)})^n = (1 - e^{-x/n})^n
\]

\[
\to \exp(-e^{-x}) \quad (n \to \infty).
\]

So \(X_{\text{max}}\) converges to Gumbel. \(F_{X_{\text{max}}}(a_n + b_n x)\) is Gumbel distributed with parameters \((0,1)\) for \(n \to \infty\). So \(X_{\text{max}}\) is Gumbel distributed with parameters \((a_n, b_n)\) for \(n \to \infty\).

So:
\[
E(X_{\text{max}}) \to \ln(n)/\lambda + 0.57772/\lambda \quad (n \to \infty) \quad \text{and} \quad \Phi(X_{\text{max}}) \to 1.2825/\lambda \quad (n \to \infty).
\]

Assume the sea levels in Hook of Holland exponentially distributed with parameter \(\lambda\). In Figure 3.3.2 the ten highest sea levels at Hoek of Holland (>180cm) are depicted. In the same Figure the distribution of \(X_{\text{max}}\) (being \((1 - e^{-\lambda x})^n\)) and the limit distribution (being Gumbel with mean 281cm and standard deviation 45cm) are given.

![Figure 3.3.2. Extreme values with n=10 (parent is Exponential)](image)

When \(F(x)\) satisfies the limit \(\lim_{n \to \infty} F_n(a_n + b_n x) = H(x)\), we say that \(F(x)\) belongs to the domain of attraction of \(H(x)\). The surprising result is that there are only 3 possible CDF’s for \(H(x)\) ([20]). They are given by:

Frechet:
\[
H(x) = \exp(-x^{-c}), \quad x \geq 0, \quad c > 0
\]

Weibull for max.:
\[
H(x) = \exp(-(x)^\gamma), \quad x \leq 0, \quad c < 0
\]

Gumbel:
\[
H(x) = \exp(-\exp(-x)), \quad -\infty < x < \infty
\]

The practical importance of this result is that, when we are dealing with extremes, and \(n\) is large enough, the infinite many degrees of freedom that we have with the distribution function for these extremes, are reduced to three parametric families. In other words, whatever the exact parent distribution \(F(x)\) is, one of the above given three distributions can be used for approximating the extremes of \(F(x)\). The problem is now, which distribution (or domain of attraction) is associated with our CDF \(F(x)\)? With the following result from [21], this question can be solved:
A necessary and sufficient condition for \( F(x) \) to belong to the domain of attraction \( H(x) \) is that:

\[
\lim_{\varepsilon \to 0} \left( F^{-1}(1-\varepsilon) - F^{-1}(1-2\varepsilon) \right) \left( F^{-1}(1-2\varepsilon) - F^{-1}(1-4\varepsilon) \right)^{-1} = 2^c
\]

with the following correspondence:

| \( c > 0 \) | \( H \) is Frechet |
| \( c = 0 \) | \( H \) is Gumbel |
| \( c < 0 \) | \( H \) is Weibull (for max.) |

Sometimes physical considerations can be used to reduce the three limiting distributions to two. Parent distributions with a finite endpoint (like wave heights in shallow water) cannot lie in a Frechet-type domain of attraction (because they have their domain for \( x \to \infty \)). Moreover, if we consider that the Gumbel distribution can be approximated as closely as desired by Weibull for maxima or Frechet, we conclude that the limit distribution can be selected solely from physical considerations. If we are dealing with random variables limited in the tail to the right, then a Weibull for maxima distribution is the limiting distribution.

From the assumption that we only have a set of extreme observations for which the parent distribution \( F(x) \) is unknown, we would like to determine the domain of attraction. An estimator for the \( c \)-parameter of Eqn. (5) can be found with the Pickands’ method. [14] shows that this \( c \)-parameter is the same as the one in a Generalized Pareto distribution given by:

\[
GPA(x ; a , c) = 1 - (1 + cx/a)^{-1/c}
\]

Fitting its 2 parameters on the data gives us automatically the domain of attraction. A curvature observation of the data plotted on Gumbel probability paper can also be used:

| Convex curve | \( H \) is Weibull |
| Linear curve | \( H \) is Gumbel |
| Concave curve | \( H \) is Frechet |

If we have concluded for a Weibull for maxima domain of attraction:

\[
H(x) = \exp((-x)^c) \quad x \leq 0, \, c < 0
\]

Then, with \( \lim_{n \to \infty} F^n(a_n + b_n x) = H(x) \), we can derive:

\[
F(x) = H^{1/b}(x-a)/b = \exp(-((\lambda - x)/\delta)^\beta) \quad x \leq \lambda, \, \beta > 0
\]
the general form of a Weibull for maxima distribution in which $\lambda$, $\delta$ and $\beta$ are the unknown parameters, to be fitted to the data. $\lambda$ is the location parameter, $\delta$ is the scale parameter and $\beta$ is the shape parameter. Note that the maximum value that the random variable can adopt is given by $\lambda$.

Caution: there is also another type of Weibull distribution (Weibull Min type III) known in practice:

$$F(x) = 1 - \exp(-(x-\lambda)/\delta)^eta, \quad x \geq \lambda, \quad \beta > 0$$  \hspace{1cm} (9)

This is the type which was suggested by [22]. This is in fact the limiting distribution for minima, but is nowadays used a lot for fitting maximum data as well. It is also commonly used as a lifetime distribution and in corrosion engineering ([23]).

3.3.6.1 Lindley’s argument and the central limit theorem

Lindley [24] has shown that the elicitation and modulation of expert judgment concerning quantities of interest (such as model constants) often involves the use of the normal distribution.

The central limit theorem explains why we might see so many times a normal distribution in practice: a stochastic variable that is influenced by a large number of independent processes will be approximately normally distributed.

Roughly, the central limit theorem says that the sum of a number of (independent) samples taken from any distribution is approximately normally distributed. As we add more terms the approximation becomes better. This does not only apply to the sum but also to the average (which makes sense if one knows that if $X$ has a normal distribution then it follows that also $aX$ has a normal distribution). In mathematical terms, given a set $X_1, X_2, \ldots, X_n$ of identically and independent distributed random variables, suppose that each $X_i$ has mean $E(X_i) = \mu$ and variance $\text{var}(X_i) = \sigma^2 < \infty$. Define $Y_n = \Sigma X_i$

Then:

$$\lim_{n \to \infty} \frac{Y_n - n\mu}{\sigma \sqrt{n}} \text{ has a standard normal } N(0,1) \text{ distribution.}$$  \hspace{1cm} (10)

The central limit theorem for the sum of random variables can easily be applied on the product of random variables by noting that “$\ln \prod = \Sigma \ln$”, and therefore the product of $n$ i.i.d. (à définir) random variables converges to a lognormal distribution.

3.3.6.2 Relationship between PDF’s

Also relations between PDF’s can be derived. The following relations (table 3.3.3) were found by Van Gelder [27]:

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y = \exp(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Shifted Log Normal</td>
</tr>
<tr>
<td>Gumbel</td>
<td>Frechet</td>
</tr>
<tr>
<td>Exponential</td>
<td>Pareto</td>
</tr>
</tbody>
</table>

Table 3.3.3: Relationships between PDFs
The proof for the last two relationships is given by:

\[ F_Y(y) = P(Y \leq y) = P(e^{-y} \leq X = \ln(y)) = e^{-e^{-y}} \]

\[ F_X(\ln(y)) = e^{-e^{-y}} = e^{-e^{-y}} e^{\alpha} = e^{-e^{-y} e^{\alpha}} = e^{-\left(e^{\alpha} - 1\right) \frac{1}{\alpha} e^{-y e^{\alpha}}} \] (11)

This is a Frechet distribution with location 0, scale \( e^\xi \) and shape \( 1/\alpha \).

\[ F_Y(y) = P(Y \leq y) = P(e^{-y} \leq X = \ln(y)) = e^{-e^{-y}} \]

\[ F_X(\ln(y)) = l - e^{-e^{-y}} = l - e^{-e^{-y}} e^{\alpha} = l - e^{-y e^{\alpha}} - l - \left(e^{\alpha} - 1\right) \frac{1}{\alpha} e^{-y e^{\alpha}} \] (12)

This is a Pareto distribution with scale \( e^\xi \) and shape \( 1/\alpha \). Also the following relationship could be proven:

If \( X \) is Weibull then \( Y = \ln(X - \xi) \) is Gumbel for minima distributed:

\[ F_Y(y) = P(Y \leq y) = P(\ln(X - \xi) \leq y) = P(X \leq e^{y + \xi}) \]

\[ = l - \exp\left(-\frac{e^y}{\beta}\right) = l - \exp\left(-\frac{e^y}{\text{ln}(\beta)}\right) = l - \exp\left(-e^{\gamma - \text{ln}(\beta)}\right) = l - e^{-e^{-\gamma - \frac{1}{\alpha}}} \] (13)

This is a Gumbel distribution for minima.

Finally it is possible to examine the tail behaviour of distributions. The following relations (table 3.3.4) were found (properties of the Halphen distributions are described in [25], [26]).

### Table 3.3.4: Tail behaviour as a function of the return interval T

<table>
<thead>
<tr>
<th>X</th>
<th>Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-normal</td>
<td>( \exp(\sqrt{T}) )</td>
</tr>
<tr>
<td>Gamma</td>
<td>( \ln(T) )</td>
</tr>
<tr>
<td>Normal</td>
<td>( \sqrt{T} )</td>
</tr>
<tr>
<td>Halphen</td>
<td>( T^\nu )</td>
</tr>
</tbody>
</table>

#### 3.3.7 Quality of adjustment in the distribution tail

In the structural reliability study, special interest was focused on modelling the distribution tail and hence estimation of extreme quantiles in the distribution tail, that is to say a quantile generally found outside observations. This estimation constitutes an extrapolation problem.

The statistic on distribution tails is characterised mainly by a search in the "dark" (ref. [13]); in fact, points are generally lacking or scarce in this region. In order to be able to deduce information on this zone, a position not too far from the available points must be selected. The following practical questions arise:

**Question No. 1:** Is it possible to have an approximation of the distribution tail distribution?

**Question No. 2:** If yes, with what confidence level? (or, from a practical standpoint, up to what point do available values allow extrapolation in the dark?)

**Question No. 3:** How might it be possible to define complementary tests or complementary calculations to improve information quality?
The problem concerning the statistical treatment of extreme values is intrinsically a non-parametric statistical problem, however, based on tools currently available; it may only be dealt with parametrically. Thus, a problem specific to distribution tails is transformed into an extrapolation problem based on the hypothesis that a model obtained from an analysis in the maximum probability region is valid in "extreme" regions.

References [14] and [15] proposed statistical tests to establish whether or not an extreme quantile may be estimated based on an accepted distribution in the maximum probability domain and especially the ET (Exponential Tail) test, that is adapted for probability distributions exhibiting exponential decay in the Gumbel domain, for example normal, log-normal, exponential, Weibull distributions. One can refer to [15] and other documents of the IS2 project conducted by INRIA Rhône-Alpes concerning this theme for a description of this test (ref. www.inrialpes.fr/is2/).

Another approach, referred to as the Bayesian regularisation procedure, was proposed in [16] to try and improve a probabilistic model that was accepted in the maximum probability region but was rejected in the extreme section by the ET test. The gain in the distribution tail assumes that a degradation in the region of maximum probability is acceptable. A compromise must be reached between minimum degradation in this region and maximum gain in the distribution tail. This approach is applied to the sample studied in the following section and results are presented in [17]. Furthermore, it may be noted that mix or rejoining methods may also be of interest in treating distribution tails.

3.3.8 Illustration by means of an example

Various tests were performed on steel, grade 16MND5, in the context of a research programme; this is the material used to make French nuclear reactor vessels. Twenty-five values of toughness were obtained under similar experimental conditions around –90°C (ref. [18, 19]). Values were (in MPa.m^{1/2}): 133, 160, 112, 171, 146, 59, 82, 150, 112, 130, 78, 112, 100, 80, 92, 118, 155, 142, 139, 73, 41, 50, 116, 76, 130.

These 25 values were between 41 and 171 MPa.m^{1/2}, had a mean value of 110.3 MPa.m^{1/2} and standard deviation of 36 MPa.m^{1/2}. This sample is referred to as "K1c" in the following. Further information concerning this chapter may be found in the paper presented during the λµ13 conference (refer to [17]).

The Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling tests were used to study adjustment of a family of probability distributions on a sample of toughness measurements in a maximum probability region using the software package Statgraphics. These distributions were adjusted using a maximum likelihood method.

It was noted that various probability distributions could be accepted with a confidence level of at least 90%, notably the Weibull distribution (form = 3.60465 and scale = 122.715) and the log-normal distribution (mean = 111.351 and standard deviation = 43.4339). The Kolmogorov-Smirnov, Anderson-Darling and Cramer-von Mises statistic values for these two distributions are provided in table 3.3.5. These would seem to indicate that the Weibull distribution is most suitable; these statistics correspond to various distance measurements between the empirical distribution function of the sample and that of the adjusted distribution. It should be noted that [10] proves that the Weibull distribution is suitable for modelling toughness values in the maximum probability region according to some hypotheses. Figure 3.3.5 compares the two adjusted distributions.
Table 3.3.5: Comparison of statistics

<table>
<thead>
<tr>
<th></th>
<th>Weibull model</th>
<th>Log-normal model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>0.113</td>
<td>0.181</td>
</tr>
<tr>
<td>Cramer-Von Mises</td>
<td>0.051</td>
<td>0.104</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>0.300</td>
<td>0.631</td>
</tr>
</tbody>
</table>

Figure 3.3.3. Comparison of Weibull and Log-normal models on K$_{ic}$ sample

Based on the non-parametric Bootstrap method concept, ten replicas containing 25 values were generated from and then reconstituted into the initial sample. For each replica, the Weibull distribution parameters and normal-log distribution parameters were established first, subsequently their adequacy with respect to the replica was tested. For each of these distribution families, it was noted that the family of distributions tested could only be adopted for five replicas. Furthermore, only 3 replicas allowed the Weibull distribution and log-normal distribution to be simultaneously accepted as a probabilistic model in the maximum probability area. Accepted distribution densities are illustrated in Figure 3.3.4. These may feature significant differences. The parameters of adjusted distributions have a non-negligible variability: 17% on the Weibull distribution form parameter that varies between 3 and 5, and 12% on the standard-deviation of the log-normal distribution. These variabilities have a significant impact on estimation of quantiles in the distribution tail. One of the reasons for this variability is the small sample size.
Weibull distribution Log-normal distribution

Figure 3.3.4. Fitted distributions on replicas

From a reliability standpoint, only the lower distribution tail is of interest, that is to say the domain where the smallest toughness values are found; toughness is a measurement of strength. It was noted that in this distribution tail, values associated with the same quantile could feature significant differences (ref. Table 3.3.6); for example, the Weibull distribution gives a value of 32.4 MPa.m$^{1/2}$ for the fractile at 10$^{-3}$, while the log-normal distribution estimates said fractile at 18.1 MPa.m$^{1/2}$. These differences will have significant consequences during reliability calculations if this variable is influencing. It is therefore important to substantiate the choice between these two adjusted distributions.

Table 3.3.6: Fractiles from the 2 probabilistic models (MPa.m$^{1/2}$)

<table>
<thead>
<tr>
<th>Fractile</th>
<th>Weibull model</th>
<th>Log-normal model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0,00001</td>
<td>20,8</td>
<td>5,0</td>
</tr>
<tr>
<td>0,0001</td>
<td>25,6</td>
<td>9,5</td>
</tr>
<tr>
<td>0,001</td>
<td>32,4</td>
<td>18,1</td>
</tr>
<tr>
<td>0,01</td>
<td>43,2</td>
<td>34,2</td>
</tr>
<tr>
<td>0,1</td>
<td>64,0</td>
<td>65,7</td>
</tr>
</tbody>
</table>

The fact that the log-normal distribution was accepted on the initial sample and transformed sample may be explained by the sample size (25 values).

It is important to know up to what fractile approximation in the distribution tail is acceptable. One way to approximatively establish the limit of this domain is to test the adequacy of the distribution in the distribution tail for increasingly small quantiles (in the case of a lower distribution tail). The approximative limit of the domain is achieved for the first quantile order for which distribution is
rejected. In this example, application of the ET method led to acceptance of the log-normal model for each of the following quantiles 8%, 5%, 3%, 1% and 0.1%. However, it may be noted that the fractile associated with the minimum value of the $K_{1c}^{Inv}$ sample for the log-normal model is equal to 0.7%. Hence, it is possible to extrapolate for this example at least up to the 0.1% fractile; we did not test fractiles less than 0.1%.

A parametric Bootstrap method, under the hypothesis of a log-normal distribution, was implemented to examine the dispersion of estimated fractiles. Based on one hundred replicas, the mean ($m$) and standard deviation ($s$) of ET method fractile estimator values ($q_{ET}$) were estimated. The intervals $[m-s, m+s]$ obtained after the results were retransformed by $1/x$ are presented in Figure 3.3.5.

![Figure 3.3.5. Confidence interval on $q_{ET}$ for the $K_{1c}^{Inv}$ sample – Log-normal model](image)

The application of tools dedicated to distribution tails allowed the Weibull model to be rejected on this sample. This example also illustrates the variability of results obtained based on small samples and hence the necessity for vigilance when using said.

### 3.3.9 Conclusion

One of the most significant criticisms concerning the application of probabilistic methods in the domain of mechanics, as well as in other domains, concerns validation of the probability distribution used to model the various sources of uncertainty. This paper presents various aspects of this issue (criteria for selecting a distribution, extreme laws, distribution tails, sample size, etc.) and presents promising practical guides to provide solutions based around an example. We think that further experiments are necessary to validate some results, notably to allow a comparison with mix or rejoining methods, in order to develop a methodology suitable for industrial requirements.

### 3.3.10 Acknowledgments

The authors would like to extend their sincere thanks to the following persons for discussions concerning this theme: E. Meister and E. Ardillon (Electricité de France), J. Héliot (Framatome), J. Diebolt (Marne-la-Vallée University), M. Garrido (INRIA Rhône-Alpes), M. Baudot and Kadiri (Ligeron SA), M-P. Valeta, M. Eid and P. Lamagnère, S. Boulégue (CEA). Furthermore, we are extremely grateful to both EDF and Framatome for their financial support of part of this work.
3.3.11 References

[13] J-P. Raoult, Tout ce que vous avez voulu savoir…, Proceedings of « IS2 day on extreme values” INRIA Grenoble, 2000 (available on www.inrialpes.fr/is2)
[27] P.H.A.J.M van Gelder, Statistical methods for the risk-based design of civil structures, Publisher TU Delft, Series: Communications on Hydraulic and Geotechnical Engineering, 2000