Bayesian methods for evaluation fatigue tests and their application for a glider airplane Góbé

Krisztián Kovács

1 General introduction

The typical ways of damage of the vehicle structures are the fatigue crack propagation and fracture. The fatigue fracture of the important structural components during the service induces high risk of severe accidents. This is why the distribution of the service lives up to fracture should be known. In case of many structural components, the permitted service life is determined on the basis of the probability of fatigue fracture. In some cases the permissible probability of fracture is only about $10^{-3}$ ... $10^{-6}$. When investigating in this region of extremely low probabilities, deep difficulties are arising:

- The probability region investigated is not accessible directly by fatigue test. To reach this region $\approx 10^{-3}$ ... $0^6$ specimens would be needed, depending on the probability level.

- The relatively small samples can be acceptably described by many known distribution types (shapes), but the behaviour of these distributions are significantly different in the region of extremely low probabilities. The exact shape of the distribution describing the sample is unknown.

- The service load acting on the investigated structure is only particularly known. Therefore the validity of the loads used in fatigue test is always questionable. Because of the uncertainty of loads, the distribution of real service lives is not identical to the distribution of test fatigue lives.

Practically the task cannot be solved: probability extrapolation should be done based on small sample, knowing that the questioned distribution is different from the distribution of the sample, and the types of both distributions are unknown.

The most important properties of the task to be solved are the uncertainty and the lack of information. We are in instant need of usable information!

2 The sources of information about Góbé

2.1 Fatigue Tests

The Department of Mechanics of the Faculty of Transportation Engineering BME and the Steel-industrial Research Institute in the years 1975-76 performed fatigue tests using five Góbé
airplane taken out from service. The test load was determined based on theoretical and experimental results. A block program was composed to represent the service circumstances of this aircraft type. The fatigue lives measured in the experiments were transformed to equivalent flight hours (the individual flight hours performed in real service were taken into consideration also). Fractures arose in three regions of the wing structure:

- in the joint of the main wing beam (the notation of this region: $B$)
- close to the joint of the main and the diagonal beam (the notation of this region: $F$)
- in the rear joint of the wing (the notation of this region: $H$)

The airplane is almost perfectly symmetric, it contains all three kinds of critical region duplicated. A fracture occurring in the regions $B$ and $F$ makes the wing unable to bear the load... A fracture occurring in the region $H$ leaves chance to land safely (and the rebuilding of this region is relatively easy).

### 2.2 Survived service lives

In 1976 2000 equivalent flight hours were permitted, later this was increased to 2800 equivalent flight hours. Almost all Góbés reached or approached these service lives. The planes fell out from service of various reasons and the aircrafts being in use performed also remarkable service times. Exact actual information is not available about every aircraft, but Table 2 describes the survived service lives with an acceptable accuracy.

#### Tab. 2. The survived service lives without fatigue fracture

<table>
<thead>
<tr>
<th>Survived service life (equivalent flight hours)</th>
<th>1600</th>
<th>2000</th>
<th>2800</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of survived aircrafts</td>
<td>55</td>
<td>6</td>
<td>75</td>
</tr>
</tbody>
</table>

### 3 The sources of information about the distribution of lifetimes

There is prior information about the shape of the distribution and about the standard deviation of the logarithms of the life times.

#### 3.1 Theoretical Considerations

The lognormal, the gamma, the Birnbaum-Saunders and the Weibull distribution types can be justified theoretically using more or less approximate assumptions. Taking into consideration every known theoretical justifications it can be rendered likely that the fatigue lives can be described as a sum of the lifetimes of “chains” built from link of lognormal lifetime distribution (chain: sequential elements from aspect of failure; see part[5]). This composite distribution can be called lognorm-chain sum distribution. This distribution contains all the lognormal, the gamma, the Birnbaum-Saunders and the Weibull distribution types as extreme cases [6]. The fatigue fracture due to the real service loading is a very complex process. It is not surprising, that the distribution of fatigue lives can not be described at an acceptable accuracy by a simple distribution type with 2...3 parameters.

It is likely that the failure rate of the fatigue failures is monotonic increasing function of the service time performed [3]. This assumption gives a condition for the shape of lifetime distribution (for its type and/or parameters).

The existence of a lifetime $T_0$ of 0% fracture is a controversial question. Gedeon refers to the researches of Gillemot, who stated that the rupture of material bindings needs some work in any case, and the accumulation of this work needs some load cycle [4]. Therefore there should be a surely fracture free lifetime. This may be valid in laboratory. But it seems possible that a structure manufactured in poor quality meets extreme rough service loads and after a short service life fatigue fracture occurs. Saunders is definitely against the use of parameter $T_0$ [8]. However, if $T_0$ exist its reliable statistical determination can be considered impossible. Additionally, the parameter $T_0$ is not needed for the decision making: because of the numerous hazards always being present we are forced to take risks anyway in our every action.

### 3.2 Fatigue lives of different airplane structural components

Based on real service fatigue events occurring in the structures of airliners published [8] the approximate identification of the shape for the distribution function of the aluminium components is possible [6]. The shape rendered likely is plotted on Fig.1 (named empiric) using Weibull probability paper. On the horizontal axis the logarithm of the lifetime is standardized to zero mean and unit standard deviation (see part[4,2]). For comparison, the standardized two-parameter Weibull distribution is also shown on Fig.1(W2).

Fortunately there are published data about large number of laboratory fatigue tests [2]. The specimens of the investigated samples were real airplane structures, structural components or similar to them. The standard deviation of the logarithm of the fatigue lives let be indicated with $\sigma$ and its estimated value $S(\log t)$. The $S(\log t)$ value was computed for every sample. The hypothesis of the constant $\sigma$ seems not likely (but can not be rejected on pure statistical basis). The $\sigma$ value of a sample de-
The logarithmic standard deviation of different uncertainty. The characteristic measure of variation of the effects of the individual load processes are approximately equal—increase significantly the deviation of lifetimes, the random effects that the different components are approximately equal to the value of Boeing Scientific Research Laboratories comparable.

Saunders Sam C. The thicknesses, rivet sizes are the level of components, there is no significant difference between the Góbé and the airliners: the thicknesses, rivet sizes are shape and size, and these all were treated together by the basis of the data available). An objective hypothesis (more exact orientation was not possible on the applied distribution of the standard deviations is only a subset of the values. The values used in our investigations are shown in Table 3. In the practice, the expected value is comparable. Therefore, in spite of the different load spectrum, the application of distribution shape and deviation properties of airliner components seems acceptable even in the case of Góbé.

4 The elements of the estimation method applied
4.1 Distribution shape given in tabular form
For statistical estimations usually those distribution types are used which are given in a closed form depending on several parameters. In our study the shape of the distribution is described by a multi-parametric composite type (lognormal-chain sum distribution, see part [3]), the number of parameters investigated is 6. Estimation of all the parameters based on a small sample is obviously meaningless, and the distribution function cannot be written in closed form. Therefore the shape of the distribution is not given for the estimator method by parameters but in a tabular form. (Due to the tabular handling, the direct use of an empirical distribution based on an extremely large sample is also possible.)

The experiments show that the distributions of fatigue lives more or less become approximately a straight line on Weibull probability paper (the distributions are not far from the two-parameter Weibull distribution, for example see Fig. 1). The values of the tabular given distribution $F_t$ can be determined accurately enough even using a simple linear interpolation over a wide range of argument $\log t$.

4.2 The transformation of the given shape of distribution
The distribution given in tabular form is fixed, it has no parameter and there is nothing to estimate. For the fitting of the given distribution shape to the sample investigated parameters must be introduced. The logarithmic expected values and deviations of the individual samples can be different. A simple two-parameter transformation of the argumentum $\log t$ answers the purpose: shifting and stretching are used on the basis of the following simple equation:

$$\log t^* = S^* \cdot \log t + M^*$$

where $S^*$ and $M^*$ are the stretch and the shift parameters of the transformation.

Using the parameterised transformation (1) a two-parameter distribution can be introduced:

$$F(\log t, S^*, M^*) = F_t(S^* \cdot \log t + M^*)$$

After this, theoretically all fitting methods can be used which are used for fitting of the distributions given by closed form.

<table>
<thead>
<tr>
<th>Table 3. The equally probable values of the standard deviation of samples of logarithmic fatigue lives for modelling the distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.060</td>
</tr>
<tr>
<td>0.127</td>
</tr>
<tr>
<td>0.189</td>
</tr>
</tbody>
</table>

Bayesian methods for evaluation fatigue tests and their application for a glider airplane Góbé 2009 37 1-2
The question arises whether the transformation (1) deforms the shape of the distribution. A distribution type is not deformed when the transformed distribution can be generated by the proper selection of its natural parameters. In this case the transformation (1) can be replaced with the proper transformation of the parameters. The lognormal and the two-parameter Weibull types are non-deformable. The three-parameter Weibull distribution (W3) is a deformable type, but in our application its deformation can be neglected, as our investigations show. In strict sense the lognorm-chain sum distribution (section 3.1) is also deformable, but its deformation is less than that of W3, and can be considered practically non-deformable.

4.3 The likelihood function

The sample to be evaluated might have come from different distributions. The probability is investigated that the sample came from a distribution $F$. This probability can be rendered to the distribution $F$ or for its parameters, in the case of fixed distribution type.

Let us assume that the information about the sample $x_i$ element is only that it is in the interval $(x_{ai}, x_{bi})$. The probability $P_{li}$ of this event can be written:

$$P_{li} = F(x_{bi}) - F(x_{ai})$$

If in the given interval contains not only $l$ but $m_j$ independent sample element then the probability of this event can be computed by simple multiplication (do to the assumption of independence):

$$P_{mi} = (F(x_{bi}) - F(x_{ai}))^{m_j}$$

If $n$ intervals are known containing totally $\sum m_i$ independent sample element coming from the same distribution $F$ then the most general form of the likelihood function can be written:

$$L(F) = \prod_{i=1}^{n} (F(x_{bi}) - F(x_{ai}))^{m_j}$$

Because of all information associated to the given sample is fixed the likelihood function is the function of $F$ only. When computing, the type of the distribution is fixed and the likelihood function is considered as the function of the parameters of the distribution.

When evaluating fatigue lives, it is possible that some specimens do not brake. In this case $x_{bi} = \infty$ and $F(x_{bi}) = F(\infty) = 1$. It is usual that every fatigue life is known accurately:

$$x_{ai} \approx x_i \approx x_{bi}$$

In this case the finite difference of distribution $F$ can be approximated using the probability density function $f$:

$$F(x_{bi}) - F(x_{ai}) \approx f(x_i)(x_{bi} - x_{ai})$$

Using the approximations

$$L(f) \approx \prod_{i=1}^{n} f(x_i) \prod_{i=1}^{n} (x_{bi} - x_{ai})$$

the second product does not depend on the distribution. It influences the function $L$ with a constant multiplying factor only which is indifferent for us. Therefore the second product can be omitted. In the task investigated some lifetime are known accurately and there are survived lifetimes also. The most useful form likelihood function in this case is the following:

$$L(F) = \prod_{i=1}^{n} \frac{dF(x_i)}{dx} \cdot \prod_{j=1}^{n_s} (1 - F(x_{aj}))^{m_j}$$

where $n_s$ stands for the number of lifetimes up to fracture, $n_t$ stands for the number of survived lifetimes and $m_j$ is the number of specimens survived the time $x_{aj}$.

The likelihood function may appear in different forms. The likelihood function is proportional to the probability of the distribution $F$ (or its parameter). The reciprocal factor of the proportionality is the integral of the likelihood function over the region of the distributions (or parameters) coming into question. (In case of different distribution types or discrete parameters instead of integration a finite summation can be used.) If a continuous parameter is investigated, then the likelihood function is its probability density function multiplied by a constant. After integration this function determines the probability distribution function also.

4.4 The Bayesian approach

The method of taking into account the prior additional information is the Bayesian method. For the lifetime distributions $F$ coming into question a probability is ordered which gives the prior probability of the event that the sample investigated came from the distribution $F$. This prior information can be described by a function $P_{prior}(F)$. The occurrence of every data in the sample gives a condition which modifies the prior expectation of $F$. Using the knowledge of the sample the probability $P_{post}(F)$ of the distribution $F$ can be computed as follows:

$$P_{post}(F) \sim P_{prior}(F) \cdot L(F)$$

The Bayes method is known since the 18th century but it is the subject of intensive mathematical researches in the last decades. Using the so-called “noninformative” prior distributions the “objective Bayesianism” was defined which is applicable when there is no prior information. Against the traditional “frequentist” approach the Bayesian approach offers a new paradigm in statistics which has numerous advantages. The publication 1 gives an inspiring insight into this new topic and into the classical Bayesian approach.

4.5 The direct estimation of the probabilities of fracture

In the task to be solved the probabilities of the fracture are needed therefore it is obvious that directly these probabilities should be estimated. Even so, traditionally the estimation of parameters is done at first and after that the estimated parameter values are applied in the formula of the distribution. The
danger of this traditional method will be presented in this paper. The estimated values of the parameters of a hypothesised distribution type are perfectly uninteresting. In the practical realization of the Bayesian approach the parameters are used for the selection of the distribution $F$ to be evaluated using Eq. (3). The computed probability $P_{\text{post}}$ can be interpreted as the probability of the probabilities of fracture given by distribution $F$ at every lifetime investigated. Therefore the expected value of the probabilities of fracture can be determined by evaluation of all distribution (types and parameter values) coming into question. (In the practice the probability of fracture is estimated in a finite number of the lifetime values only, but theoretically there exists a continuous estimated distribution function.)

4.6 The estimation of the expected value instead of the most likely value

The likelihood function or Eq. (3) gives the probability density and distribution functions of the parameter to be estimated (distribution parameter or probability of fracture). With the knowledge of the functions mentioned three possible choices offer themselves for the selection of the estimated value:

- the most likely value
- the median
- the expected value.

The selection of the most likely value is widespread for parameter estimation (maximum likelihood method). A significant reason for doing this is the computational convenience: in many cases closed forms are known for computing the maximum likelihood estimation without a lot of computations. But the numerical integration needed for computing the other two values can be performed quickly using a computer. Nowadays none of the choices causes practical problem. In the task to be solved the probabilities of the fracture are parameter to be estimated. When using the expected values of the probabilities, the estimated distribution appearing differs from the distribution type used in the estimation method. The distribution estimated in this way, the distributions appointed by the expected or the most likely parameter values are all different. In the decision making under uncertainty the expected value of the loss of the fatigue fracture has role. This expected value is determined by the expected value of the fracture. Therefore in the task to be solved the proper method is the estimation of the expected values of the probability of fatigue fracture.

5 On the chain property

5.1 The chain model

If load bearing elements are connected to each other like a chain then the lifetime of this chain structure is determined by the element (link) of smallest lifetime. If the chain contains $r$ number of nominally identical, in probability sense independent elements of lifetime distribution $F_i(x)$ then the resultant probability distribution function of the lifetime of the chain $F_r(x)$ can be computed. The relation between $F_i(x)$ and $F_r(x)$ determines the relation between the related probability density functions $f_i(x)$ and $f_r(x)$ also. The relations mentioned are as follows:

$$F_r(x) = 1 - (1 - F_i(x))^r$$
$$f_r(x) = r \cdot f_i(x) \cdot (1 - F_i(x))^{r-1}$$

(4)

It is proved that in case of $r \to \infty$ the distribution $F_r(x)$ became Weibull distribution, independently from the shape of $F_i(x)$. The distribution of a chain built from arbitrary number of elements of identical Weibull distribution remains a Weibull distribution with the same location- and shape parameters, the scale parameter changes only.

5.2 The invariance of the likelihood function in the chain property

Let be investigated a sample of nominally identical chains each containing $r$ elements. After a series of fatigue tests $x_i$ lifetimes up to fracture and $x_{aj}$ are known (the latter with occurrence frequencies $m_j$). When estimating the lifetime distribution of chain $F_r(x)$ two trains of thought are possible:

1 The object investigated is regarded as a simple specimen, and its chain property is neglected. The distribution $F_r(x)$ is estimated directly from data.

2 The chain property is taken into consideration. As first step the distribution of chain element $F_i(x)$ is estimated. As second step the distribution $F_r(x)$ is computed using the relation (4).

It is important that when applying method number 2 a fracture lifetime of a chain element $x_i$ is a survived lifetime for the other $r - 1$ elements. And a survived lifetime $x_{aj}$ for the whole chain with multiplicity $m_j$ means at element level a survived lifetime with multiplicity $r \cdot m_j$. The likelihood function can be formulated on the basis of both train of thought, whether as a function of distribution $F_i$ or $F_r$. Some computation proves that the likelihood functions $L_1$ and $L_2$ determined by the two trains of thought are identical in both forms:

$$L_1(F_1) = L_2(F_1)$$
$$L_1(F_r) = L_2(F_r)$$

The invariance of the likelihood function means that every estimation which is based on the likelihood function gives the same result using both of the 1st and 2nd trains of thought in every case. The validity of this invariance does not depend on the shape of the distribution or the value of the parameter $r$. In the engineering practice the chain property of the objects investigated is not obvious, the value of chain parameter $r$ is uncertain or unknown.

If an estimation method is not invariant then the result depends on our thoughts about the object. Every estimation
method is subjective in this sense which is not invariant. Therefore every estimation method is subjective which can not use the survived lifetimes. The method of moments, the probability plotting methods (using median ranks) and the linear estimations are subjective. It is strongly recommended to avoid these methods.

5.3 The dependence of the shape of the distribution on the chain property

The Weibull distribution type is invariant in the chain property. In case of other types the distribution of the chain has more or less different shape than the distribution of the elements. Therefore the distribution shape plotted on Fig. 1 may be applied for the components similar to the aircraft components investigated. For example, the mentioned shape may not be applied directly for a whole structure containing numerous components. Before using fatigue data of an other object the chain property of the objects must be investigated and compared. (The estimation of a relative chain parameter is easier than an absolute measure.)

5.4 The appearance of the chain property in the case of Góbé

The wing structure of the Góbé is symmetrical at the structure level. The main beams and the joint components are symmetrical at the component level also. Therefore there is 4 regions of type “F”, and 2 regions of type “B”. If these two types of regions are considered equivalent in the resistance against fatigue then the total number of the equivalent regions are \( r = 6 \). In sense of reliability these regions can be considered as the links of a chain.

5.5 The uncertainties related to the chain property

The independence of the components of a given aircraft is a question. The components may be manufactured from the same material portion at same circumstances. The environmental circumstances and the main load process are the same for them. They share a similar “destiny” in the aircraft. In extreme case, if the correlation between the components of a given aircraft were very strong then the chain model would not be needed. It would be the ideally friendly case. The clear chain model is the other extreme case, the worst. The reality is between these extreme cases, probably not far from the worst. Other question is that the distribution plotted on Fig. 1 is valid for symmetrical components or not (Saunders gives no information about this question). If yes then only the half of the number of links should be used in the computations. Fortunately our investigations show that a factor 2 in the parameter \( r \) practically does not influence the results. The uncertainties related to the chain property have small significance.

6 On the three-parameter Weibull distribution

6.1 Problem with the location parameter

Due to the location parameter \( T_0 \) introduced the domain of the three-parameter Weibull (W3) distribution is parameter dependent. As a consequence the Cramér-Rao relation can not be applied to the W3 distribution \([9]\). A more important problem arises also: in the practical cases the maximum likelihood estimation (MLE) often gives an obviously wrong result. The estimated value of \( T_0 \) is often the smallest element of the sample while the estimated shape parameter \( A < 1 \). In this case \( f(T_0) = \infty \) and the value of likelihood function is also infinity. In the case of Weibull distribution the monotic property of the fatigue failure rate (mentioned in section 3) gives the condition \( A \geq 1 \). (The exponential type, with its “ageless” property, is a special case of the Weibull distribution when \( A = 1 \).) The Fig. 2 shows an unreal, degenerated W3 distribution given by MLE (curve notation is W3 Fmax).

The MLE method of great theoretical significance has a clear background. The above problem can be considered an imperfection of the W3 distribution not that of the MLE method. Additionally, the theoretical and experimental arguments behind W3 are not persuasive \([9]\). The data published by Saunders and by Butler and W3 seems incongruent \([4]\). For description of the samples of size \( n < 100 \) the W3 distribution is usable but this can confirm the reliability of W3 in the probability region \( 1\%...99\% \) only.

6.2 The application of the expected value principle

An estimator method for the W3 distribution is outlined in this part, which method is based on the likelihood function and is near to the MLE method. If the value of parameter \( T_0 \) is fixed, then the MLE method determines the other two parameters without problems. The value of the likelihood function \( L(T_0,A,B) = L(T_0) \) can be determined, which is yet the function of \( T_0 \) only. On the basis of function \( L(T_0) \) the expected value of \( T_0 \) can be computed. As the part 4.5 shows, the expected value principle can be applied to the probabilities of fracture. Every distribution related to the \( T_0 \) investigated is weighted with the value of \( L(T_0) \) when computing the estimated distribution. The result of this estimation is not a W3 distribution yet.

7 Results

In this paper the contracted evaluation of fracture regions “F” and “B” is presented. In this case the Table 1 gives 3 fractures and 2 survived lifetimes. The survived lifetimes of Table 2 are taken into consideration also, and the chain parameter used is \( r = 6 \). The Bayesian estimation is performed using the distribution plotted on Fig. 1 and the \( \sigma \) values of the Table 3. The estimator software used is implemented on the basis of principles outlined in part 4. The results of the W3 based estimations (see part 6.2) are also given. In case of both distribution shapes (Fig. 1 and W3) three estimated distribution are plotted on Fig. 2 the distributions appointed by the expected parameter.
values ($F_{xpp}$), by the most likely parameter values ($F_{max}$) and
the distribution based on the expected values of fracture proba-
bilities. For comparison, the maximum likelihood estimation of
$W_2$ distribution is also shown ($W_2$).

After computing the results the sensitivity analysis should
be performed for every input data. (In some cases the Bayes
method can be sensitive to the prior information [1]). As a con-
trol, in the case of the Góbé the region „F” should be evaluated
separately and the contracted evaluation of all fracture regions
should be performed using all of the methods presented. The
estimated values of the fracture probabilities give the basis of
the determination of the permitted service life. This decision
making needs the basic principle of game theory: the expected
value of the win should be maximized. But it is not enough to
make the “best” decision under uncertainties. The reliability of
the decision should be investigated properly. If the probability
level of avoiding a “wrong” decision is not high enough then
new information must be acquired unavoidably.

8 Conclusions

The estimation of the extremely low probabilities of frac-
tures is a very difficult task. Therefore all usable information
is needed, beyond the direct data of the sample investigated. For
the reliable solving of the task the systematic acquisition of data
of real service fatigue failures and laboratory tests is necessar-
ily needed in every field of application. The Bayesian approach
is proper for taking into consideration of the prior information
acquired. Instead of the maximum likelihood principle the ex-
pected value principle is to be applied. Instead of estimating of
the distribution parameters, the probabilities of fracture are to
be estimated directly. The estimated distribution in this way in-
volves the uncertainties of the estimation also, not only the mod-
elled natural uncertainties of the fatigue lives. There is no real
hope that a simple theoretical distribution type can be proper
for describing of the service lives up to fracture in the whole
range of probability. When using the three-parameter Weibull
distribution ($W_3$) the low probabilities are significantly under-
estimated. The error can be reduced when the expected values
of the probabilities of fracture are estimated. This application
of the $W_3$ distribution can be acceptable only. The two-parameter
Weibull ($W_2$) distribution itself overestimates the low probabili-
ties. But in the example of Góbé the simple maximum likelihood
estimation (MLE) of $W_2$ distribution gives results very close to
the more advanced Bayes estimation. Therefore when there is
no prior information, the $W_2$ distribution and the MLE method
are suggested, as first orientation even in the case when the esti-
mated $W_2$ distribution seems not fit to the sample in the medial
probability range ($P>5\%$).

For more information or for a free trial copy of the estimator
software developed please contact the author per e-mail.

References

1 Bernardo J M. Objective Bayesianism, 2008, available at www.uv.es/
bernardo/publications.html
2 Butler J P. Reliability Analysis in the Estimation of Transport-Type Air-
craft Fatigue Performance. International Conference on Structural Safety and
pp. 181-211.
3 Castillo E, Fernández-Canteli A, Hadi A S. On fitting a fatigue