

*INJURY BIOMECHANICS RESEARCH*  
*Proceedings of the Twelfth International Workshop*

FITTING INJURY VERSUS EXPOSURE DATA INTO A RISK FUNCTION.

Arne Ran, Magnus Koch and Hugo Mellander  
 Departments 02372, 56355 and 56540 at  
 Volvo-Data and Volvo Car Corporation  
 S-40508, Göteborg, Sweden

1. Introduction and Background

A problem in safety research is how injury data from accidents, and damage data from laboratory experiments, should be concisely reduced into a single descriptive formula. Well known is the uncertainty that comes from variability of a statistical nature in biological subjects and in the physical course of events producing injuries and damages. A less known source of uncertainty in the violence and loading data is that they most often are what statisticians call "censored". This complicates the establishment of a relation between cause and effect. The purpose of this report is to describe what censoring is and how censored data can be effectively analyzed. A few examples and possible developments are presented. The presentation is somewhat lengthy because we want to make the reader understand how the mathematics work. We also want to illustrate the types of reasoning that are possible with the aid of the presented model. We will employ the biomechanical terminology proposed in reference <1><sup>1</sup>.

Censored data are data that are biased in one direction or another. The sign of the bias is known but not the magnitude. This complicates the application of conventional correlation and regression techniques since these methods assume data to be free from bias.

The occurrence of censored data is very common in biomechanics. The reason is the need to classify the injuries and damages in a few recognized levels in the AIS or ADS scales<sup>2</sup>.

Consider as an example an experimental test series where a group of human substitutes are exposed to a range of dynamic loadings, which are measured as mechanical peak forces. The test specimens are carefully examined for bone fractures. The actual force in each particular case will probably not be of such a magnitude that it exactly would cause just a barely observable fracture. All force data have to be assigned to either of two groups. One group consists of the experiments where the forces have been too low to cause a fracture. The other group contains those experiments where fractures are detected. In the latter group the forces have been sufficient, and probably even excessively so, for the causation of fractures. Therefore each group now consists of data which are censored in either direction.

Our attention to the problem with censored data has been raised by discussions within the ISO/TC22/SC12/Working Group 6. The task for that group is related to road vehicle crash testing and to "Performance Criteria Expressed in Biomechanical Terms". Several approaches to the treatment and analysis of censored biomechanical data have been proposed by biomechanical and traffic safety researchers <3> <5> <7> <8> <9> <13> <14>. Nobody seems, however, to have presented an unambiguously satisfactory method. Knowing that censored data are advantageously handled by the Maximum Likelihood Method - first presented by Sir Ronald Fisher <4> - we have prepared this report to describe the method to the biomechanical community. Workers in other technical areas, such as explosives testing or toxicology or reliability engineering have also faced this problem and have worked with solutions similar to what is presented here. A readable short survey is reference <11>. An additional source containing some worksheets and formulas is chapter 10 in reference <10>.

<sup>1</sup> Numbers within < > refer to a list of references at the end of the report.  
<sup>2</sup> ADS = Abbreviated Damage Severity, see reference <1>.

## 2. The Maximum Likelihood Method

We will present a known statistical method for the treatment of censored data. The method should be generally applicable to a wide range of biomechanical interests and studies. For convenience we will describe the method in terms of the hypothetical laboratory experiment in the above introduction. But the method might equally well be applied to accident data, or any other cause-effect analysis involving censored data. The interested reader can easily transform the text to fit his or her particular analysis.

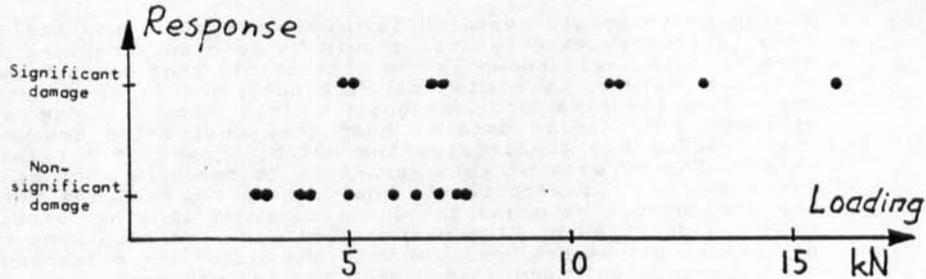


Figure 1. Experimental data

Figure 1 is a plot of a series of typical experimental data. For low loadings all experiments gave a non-significant response, while very high loadings always gave significant damage response. In the mid-range there is obviously some randomness involved. Here it would be difficult to predict the outcome of any single experiment, but a coarse statement would say that there is about a fifty per cent risk for a fracture to occur.

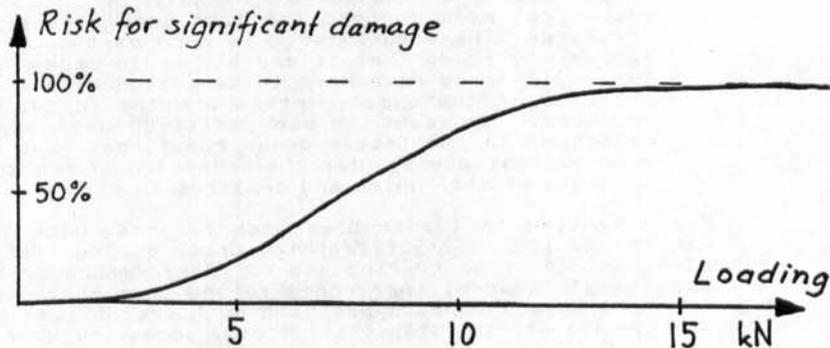


Figure 2. Risk vs. loading

Figure 2 is a useful and concise graphic presentation of this inferred risk for damage. It rises from zero to one hundred per cent over the range of loadings. Mathematically the curve is a "cumulative frequency distribution". Our statistical method will employ all the recorded data to arrive at the best possible representation of the experimental test series. If one dares to generalize from the observations, one might use the curve as a probability curve and as a prediction tool.

**2.1 Assumptions:** The following assumptions are made concerning the experiment and the data.

- A population of possible laboratory test specimens is assumed to be definable.

- Each specimen is exposed to a certain loading from external physical forces during the course of a laboratory experiment. All loading forces are measured and expressed in the same way for the whole population, and all observations are assigned the same weight. Depending upon the circumstances, the magnitudes of the loading exposures may be more or less controllable in advance by an experimenter.
- Each specimen is examined for damages resulting from the exposure to the recorded loading. The examination results in an accurate record of minor and major damages. Here, we reduce the contents of that record to a single bi-level response classification: significant or non-significant damage. The criterion for significance is defined by the analyst, as required by the goal of the investigation.
- The vulnerability of the specimens is such that the risk for significant damages is higher when the loading is higher. The actual loading in each particular case is, however, seldom of such precise magnitude that it exactly causes a barely significant damage. On the contrary, in most cases the loading exposure is either too small or too excessive for the causation of a significant damage. Consequently all loading measures are censored.

**2.2 Notations:** Mathematical notations for the above assumptions and for some useful aids are as follows:

- $z$  is the loading variable with some special values:  $v$  is the recorded loading to each specimen,  $V$  is the set of all the recorded loadings, and  $x$  is the "threshold" loading that just barely would result in a damage classified as significant to each specimen.
- The non-significant damages are indexed  $i = 1 \ 2 \ \dots \ n$
- The significant damages are indexed  $j = 1 \ 2 \ \dots \ m$
- $K$  is a parameter, or a set of parameters
- $P(\text{event})$  is the probability for a particular event to occur
- $L$  is the likelihood function. It is the product of all the independent individual probabilities in a complex arrangement.
- $F(z;K)$ , or shorter:  $F(z)$ , is an arbitrary cumulative frequency distribution. Since  $z$  here shall represent the loading to a test specimen,  $z$  can only assume values greater than zero.  $K$  is a set of parameters that affect the shape and location of the distribution. In this case  $K$  will have such values that  $F(z;K)$  is defined within the  $z$  interval from zero to plus infinity.
- $W(z;K)$ , or shorter:  $W(z)$ , is the notation for a special case of  $F(z;K)$ . It is the Weibull cumulative frequency distribution with one variable and three parameters:

$$W(z;K) = W(z; \alpha, \beta, \gamma) = 1 - e^{-\left(\frac{z-\gamma}{\alpha}\right)^\beta}$$

A summary of the properties of the Weibull function is placed in an Appendix at the end of this report.

- $Q(z;K)$  is the hazard intensity function, defined as:

$$Q(z;K) = \frac{\frac{d}{dz} W(z;K)}{1 - W(z;K)}$$

It can be shown <6> that a constant value for  $Q$ , independent of  $z$ , corresponds to a  $\beta$  value equal to 1. If  $Q(z;K)$  increases with  $z$ , then  $\beta$  is greater than 1.

**2.3 The Probability of the Observation Set:** As mentioned above, the conclusion from each single experiment can only be that the recorded loading has been too excessive or too small for the causation of a significant damage. However, each exposure of a specimen to a loading can be regarded as one statistical experiment with a response (i.e. damage classification) that is determined by a probability mechanism. A collective treatment of all the probabilistic responses will give the best overall description of that damage vs. loading mechanism.

Tentatively we denote the mechanism simply with the aid of the arbitrary function  $F(z;K)$ , depicted in Figure 2 on page 2. With  $P$ ,  $K$ ,  $v$  and  $x$  defined as above, the probability for a significant damage response of each experiment is written:

$$P(\text{significant damage, given } v) = F(v;K)$$

or, since  $x \leq v$  implies a significant damage:

$$P(x \leq v) = F(v;K)$$

The data for the observed significant damage response experiments give:

$$P(x_j \leq v_j) = F(v_j;K); \quad m \text{ probability expressions}$$

Correspondingly, for each of the non-significant damage experiments with their complementary responses, we get

$$P(x_i > v_i) = 1 - F(v_i;K); \quad n \text{ probability expressions}$$

The  $m + n$  probabilities, containing all the censored data, are now treated according to the Maximum Likelihood Method. The likelihood function  $L$  for the entire test series is:

$$L(v_1 v_2 v_3 \dots v_j v_i \dots v_m v_n) = L(V) = \prod P(v)$$

$$L(V) = L(V;K) = \prod F(v_j;K) * \prod (1-F(v_i;K))$$

The likelihood function  $L$  is thus dependent on the recorded data set  $V$  and on the yet undetermined distribution function  $F(z;K)$ . Now our next concern is not the absolute magnitude of  $L$ , but the conditions for a suitable set of values for  $K$  to provide a maximum for  $L$ .

Symbolically, this maximum is analyzed as follows. The first derivative  $dL/dK$  and its relation to a maximum is studied in the usual manner. Treating  $V$  as constants and  $K$  as variables, the equation

$$0 = \frac{d}{dK} L(V;K)$$

will give values for  $K$  which depend on the recorded data  $V$ . That set of roots to the equation which yield a maximum  $L$  is denoted  $\bar{K}$ . It will give the largest possible credibility to the experimental data. The probability mechanism which best fits the observations is then written:

$$P(\text{significant damage, given a loading } z) = F(z;\bar{K})$$

2.4 A Practical Solution: What remains now is to find a suitable expression for  $F$ , which can cover many possible modes of cumulative distributions. One good choice is the Weibull distribution which has these possibilities and also some other interesting features (see the Appendix). Alternate distribution choices will be discussed later. Replacing  $F(z;K)$  by  $W(z;K)$  and using the short notation  $W(z)$  gives:

$$L(V;K) = L(V;\alpha,\beta,\gamma) = \prod W(v_j) * \prod (1-W(v_i))$$

The calculations become somewhat simpler by taking the natural logarithm of both members of the equation. The likelihood  $L$  has its maximum simultaneously with  $\ln(L)$ .

$$\ln(L(V;\alpha,\beta,\gamma)) = \sum \ln W(v_j) + \sum \ln(1-W(v_i))$$

This equation contains the set of recorded loadings  $V$  and it also contains the three undetermined parameters  $\alpha$ ,  $\beta$  and  $\gamma$ , which we for the moment treat as variables. Instead of proceeding by taking the traditional derivatives and solving the resulting three simultaneous equations for the parameter set  $K$ , a convenient alternate way is possible. Computer programme packages are nowadays available for the solving of maximum problems. We have employed the NAG subroutine package <12>, which directly seeks a numerical maximum for  $\ln(L)$  and reports the values for  $\alpha$ ,  $\beta$  and  $\gamma$  which are associated to that maximum.

The roots - corresponding to  $\bar{K}$  in the previous section - are denoted by  $\bar{\alpha}$ ,  $\bar{\beta}$  and  $\bar{\gamma}$  and inserted in the Weibull function, which now is a good and concise representation of the experimental data:

$$P(\text{significant damage, given a loading } z) = W(z; \bar{\alpha}, \bar{\beta}, \bar{\gamma})$$

An experimenter needs of course not go through all the formal steps above in his routine analysis. The formulas are conveniently programmed into a nearby computer, together with a suitable statistical programme package.

### 3. Mathematical and Statistical Comments

This section will present and discuss some technical comments and developments to the previously derived cumulative function. The first subsections address particularly the Weibull distribution and its features. The later ones are of a more general nature. Some further comments are also included in the Applications section. Some comments are straightforward, others are more tentative and intended to indicate what a deeper analysis might give if this method is developed and tested with more actual biomechanical data.

3.1 Alternate Distributions: The presented solution is limited to what is possible to model with a Weibull function. But since its three parameters provide a wide flexibility, we will get a solution not far from the ideal one. Other distributions than the Weibull might give still somewhat higher values for the likelihood function  $L$ . The shape of the Weibull distribution may serve as an initial indication for a selection of alternate distributions (e.g. normal, lognormal, gamma, extreme value) to replace  $F(z;K)$  in the search for a maximized  $L$ . Nelson<sup>11</sup> mentions the possibility to determine a non-parametric distribution. The possibility of fitting the sum of two distributions might also be considered if there are physical indications for this.

3.2 Constraints: In order to avoid physically ambiguous solutions, some constraints are necessary during the search for the maximum  $L$ . Obvious are the trivial requirements that  $\alpha$  and  $\beta$  shall be greater than zero, and that  $\gamma$  shall be equal to or greater than zero. However, some physical reasoning can lead us further.

A reasonable assumption is that the risk for a significant damage is relatively greater in an interval from e.g. 10 to 11 loading units than in an interval from 5 to 6 loading units. This is equivalent to the statement that the hazard intensity  $Q(z)$  is increasing, which also implies  $\beta > 1$ . This is also a suitable constraint to include in the search for a maximum for  $L$ . If a lower bound than 1 for  $\beta$  is contemplated, two remarks are appropriate. A  $\beta < 1$  would violate the initial assumptions on the increasing hazard and the vulnerability of the specimens. Additionally, some literature indicates that certain Maximum Likelihood computer algorithms may show poor convergence for  $\beta < 1$ .

3.3 Treatment of Unlikely Low Observations: If there are physical reasons for a possible "always endurable loading", which must be exceeded before significant damages can be caused<sup>3</sup>, this corresponds to a lower bound for  $\bar{\gamma}$  equal to that "endurable loading". Technically this corresponds to the removal of a few of the lowest non-significant damages from the input data. Such a removal must be based on a thorough understanding of the physics of the event.

A mathematical indication for the removal of some of the lowest data points is as follows. Assume that the maximizing of  $L$  gives a  $\bar{\gamma}$  equal to the lowest  $v_i$  value (the lowest recorded non-significant damage). Then  $W(z;K)$  is equal to zero at  $z = v_1 = \bar{\gamma}$ . This is equal to a 100 per cent probability for non-significant damage at the loading level  $v_1$ . The value of the likelihood function  $L$  is not affected by multiplication by 1, and consequently this lowest  $v_1$  value may be removed from the input data to the maximizing programme. A new search is then initiated and may result in a better fit and a higher  $L$ . Several low  $v_i$  values may be consecutively removed until the highest  $L$  value is found. A Maximum Likelihood computer programme will need the experimenter's decisions how to treat the input data in this aspect.

It must, however, again be pointed out that a possible assignment of zero damage probability to some data must be based on an understanding of the physics. The mathematics will only suggest how the data could be modified to give a better fit

<sup>3</sup> Example: Vehicle retardation must exceed a certain magnitude before a restrained occupant hits the dashboard.<sup>4</sup>

to a chosen distribution model! The situation has philosophical aspects, see for instance chapter 17, "Rejection of outlying observations" in reference <10>.

3.4 Reliability of the Solution: The risk distribution which has been determined from the data is representative only for this particular sample of specimens. Probability inferences about other possible samples or individuals from the same population can be made if confidence limits are calculated for the Weibull parameters and distribution. The confidence limits available from median rank tables are conservative. Narrower limits are available by more advanced methods, see for instance references <2> and <11>.

#### 4. Applications

We will illustrate the use of the Maximum Likelihood method and the Weibull distribution for three data sets. The examples give opportunity for special comments to the experimental and analysis procedures. The data sets are taken as exercise examples and we have no intentions here to go into biomechanical details or judgements on the sampled specimens behind the data.

4.1 Accident Data: A set of injury data from a traffic accident data file has been selected for analysis. The injuries are AIS data for the legs of 725 drivers who have been involved in severe car accidents. The accidents were all classified as frontal collisions, and the released energy in the collisions is represented by a Relative Deformation Index (RDI) for the cars. This type of data has many uncertainties in the deterministic relation between the loading measure and the injury, but it is often the only data that are available. The figures below show the number of recorded cases for each AIS and RDI category.

RDI	AIS=0	AIS=1	AIS=2	AIS=3	AIS=0-3
00	2	0	0	0	2
05	40	4	0	0	44
10	110	4	1	0	115
15	140	20	2	0	162
20	116	22	1	0	139
25	73	19	3	1	96
30	54	9	4	0	67
35	29	5	5	5	44
40	16	6	4	0	26
45	4	1	2	1	8
50	1	3	0	2	6
55	2	1	2	1	6
60	0	0	0	1	1
65	0	0	2	1	3
70	0	1	0	0	1
75	0	0	0	0	0
80	0	2	0	0	2
85	0	0	1	0	1
90	0	0	0	1	1
95	0	1	0	0	1

The analysis is performed for three groupings of the data: AIS=3 AIS=2,3 AIS=1-3. The resulting risk distributions are shown in Figure 3 on page 7.

The curves show the risk for for injury vs. RDI number for the different AIS groupings. One can see how the risk curve is displaced rightwards as the criterion for significant injury is increased. The number of observations for the RDI intervals 0 - 20 - 40 - 60 - 80 - 100 are noted in the boxes below the curves. The curves show three different features as follows. In the RDI interval 0 - 60 the observations are plenty and the credibility can be assumed to be high. In the interval 60 - 100 the observations are much fewer and the curves must then be of a more approximate nature. Above RDI 100 no observations exist and the shape of the curve does not have much meaning. This type of credibility for different segments of a risk curve should be possible to illustrate by some kind of confidence band. A suitable way to calculate such a band remains to be developed.

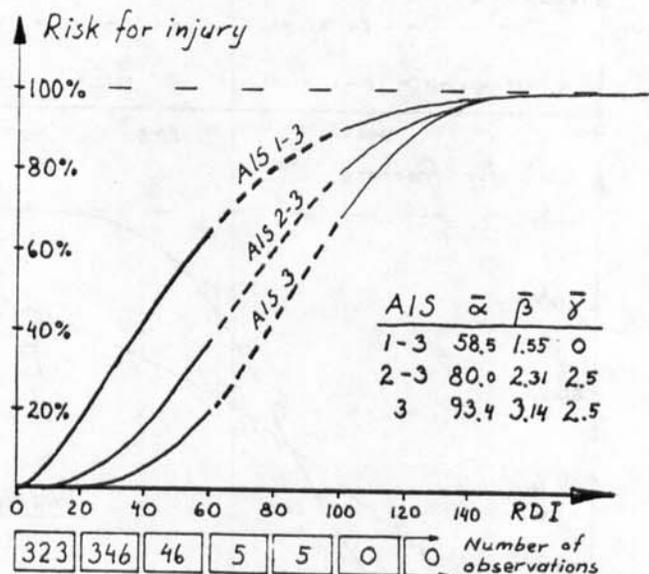


Figure 3. Risk distributions for leg injuries

**4.2 HIC Data:** The document that initiated our work was the WG6 "U.S. Position Paper" <13>. It included an analysis of 54 cadaver head impact data from three research laboratories. The paper analyzed cranial fractures vs. HIC numbers for the heads with the aid of the Mertz-Weber method <9>. That method assumes a Normal distribution to be applicable and such a distribution is fitted between the end points of an "overlap range", i.e. from the lowest damage point to the highest non-damage point. The risk values for the end points of the range are taken from a median rank table for 43 observations (the number of data in the overlap range). That distribution is plotted with a solid line in Figure 4 on page 8. One can see how that fit shows a risk for fracture exceeding 95 per cent at the two non-fractures recorded at 2138 and 2351 HIC units.

We have fitted two Weibull distributions to these data, also plotted in Figure 4. The fit denoted "A" employs all the recorded data from the lowest 175 to the highest 3400 HIC units giving the risks 0 and 88 per cent, respectively. The Maximum Likelihood fit shows a continuous increase in risk vs. HIC number in contrast to the suggested Normal distribution, which has a more "sudden" transition from zero to one hundred per cent in the range from 1000 to 2000 HIC units.

The median at 50 per cent risk corresponds to 1400 and 1470 HIC units for the Normal and the Weibull distributions, respectively. In order to establish the median, an experimenter usually tries to select his experimental conditions so that every second specimen receives significant damages. In a way, this is similar to Dixon's Up-and-Down Staircase design, see chapter 10 in reference <10>.

During research work towards the establishment of Performance Criteria Limits <1>, the goal of an investigation might be to determine a loading level that has a risk of, say, 25 per cent to give significant damages. The experimenter should then strive at a yield of 25 per cent significant damages among his test specimens. He has then no interest in the right tail of the risk distribution.

We have made a simulation of such an investigation by removing all observations with HIC  $\geq$  1500 from the cranial fracture data. This left us 42 observations to analyze, and curve "B" in Figure 4 was the result. Above 1500 HIC units there is a marked difference between curves "A" and "B", which of course is due to the lack of data for "B" in this range. However, below 1500 the curves follow each other within plus minus 1.5 per cent. This indicates that the twelve specimens above 1500 HIC could have been spared - or exposed to lower loadings - if loadings above 1500 HIC units had been believed to be unnecessary for the establishment of the 25 per cent risk level.

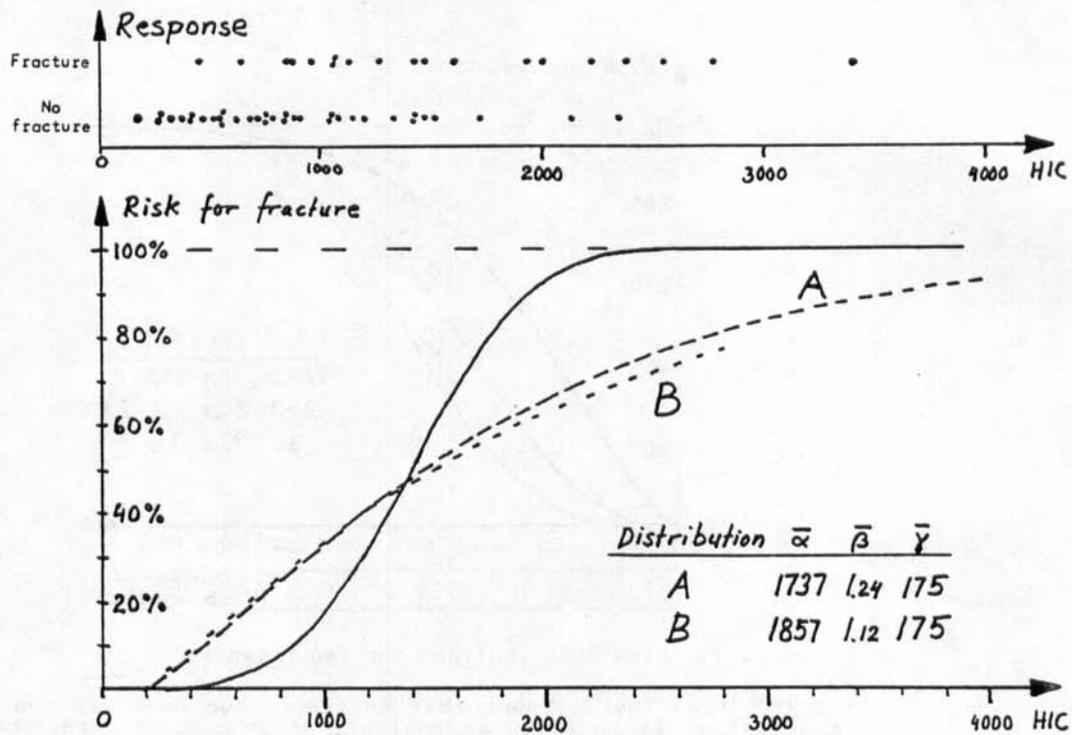


Figure 4. Risk distributions for head damages

Thus it seems possible to use the Maximum Likelihood method in the analysis of one tail of a risk distribution, and in the planning of experimental schemes.

A  $\bar{\beta}$  close to 1 indicates that this particular Weibull distribution is close to an exponential distribution. Such a distribution is often interpreted to represent a random failure mechanism with no relation between accumulated loading level and risk for failure<sup>4</sup>. Would this indicate that the recorded HIC loading values for these particular cadavers bear little deterministic relation to the occurrence of fractures? If so, something may be missing in today's supposed HIC vs. injury determinism. Another possible reason for the low  $\bar{\beta}$  might, however, be the experimental scheme itself. Perhaps ample observations at the low end and few at the high end of a loading range will tend to give low  $\bar{\beta}$  values? A third possibility is that the Weibull distribution is not flexible enough and that alternate distribution models should be sought. More experimental data sets need to be analyzed before a possible use of  $\bar{\beta}$  as a diagnostic parameter can be employed in biomechanics.

**4.3 Airbag Data:** A third example is taken from reference <9>, where Mertz and Weber introduced the above "overlap range" fitting of a Normal distribution. In a research programme a number of pigs, baboons and anthropomorphic dummies were exposed to deploying airbags. The chosen example is a test series where the loading was measured as maximum rate of chest compression, in km/h. The response was classified in a six-interval Threat-to-Life scale, based upon clinical examination of the damages.

<sup>4</sup> C.f. the chance for your telephone to ring next minute. It is generally independent of how long time you have been within hearing range. See also the Appendix for some possible interpretations of  $\bar{\beta}$ .

The authors fitted a Normal distribution to their data as described in the previous section. Their result is plotted as a solid line in Figure 5 on page 9. Their distribution assigns a risk level of 95 per cent at 35 km/h, in spite of their two low data points at this velocity.

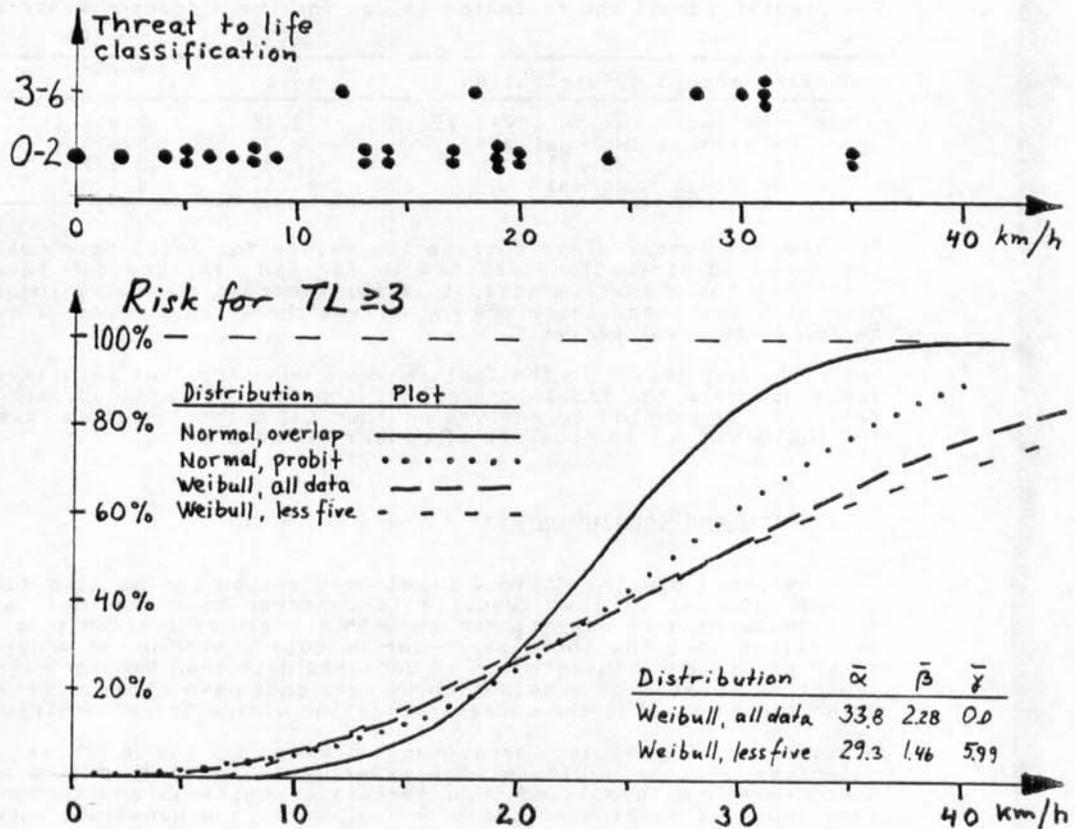


Figure 5. Risk distributions for chest damages

The Weibull parameters for these data have been determined to:  $\bar{\alpha} = 33.8$  km/h,  $\bar{\beta} = 2.28$ ,  $\bar{\gamma} = 0.0$  km/h. The resultant Weibull cumulative function is plotted with long dashes in Figure 5.

One might be willing to assign zero damage probability to some of the lowest data, for some possible physical or medical reasons<sup>5</sup>. The maximizing procedure then suggests the following Weibull parameters:  $\bar{\alpha} = 29.3$  km/h,  $\bar{\beta} = 1.46$ ,  $\bar{\gamma} = 5.99$  km/h. That is, the highest likelihood value  $L$  is available when five of the lowest values are discarded. This distribution is more skewed to the right and is plotted with short dashes in Figure 5. A numerical comparison between the two Weibull distributions gives a difference of less than 3 per cent between them over the range of observations. Both Weibull distributions give higher risk than the Normal distribution up to 20 km/h. Above this velocity the Normals are higher, and the Weibull risks increase at a slower pace.

During the final preparation of this paper we became aware of some investigators <3> <7> <8> who had proposed the "Probit Method" for biomechanical analyses, see chapter 10 in reference <10>. That method does take account to censored data. However, the data are collected into groups, thereby losing some precision. There must also be a statistically sufficient number of data and groups. Also,

<sup>5</sup> At least one possible experimental artefact might be a TL-2 damage at 0 km/h.

the symmetric Normal distribution is often assumed without possibilities to search for a shape parameter. Lowne <8> has analyzed the data in <9> and for the series here he has found the distribution plotted with dots in Figure 5. It comes closer to our Weibull distributions up to 25 km/h. Then, due to its inherent symmetry it has to increase more rapidly towards the hundred per cent level.

The Likelihood has the following values for the different distribution models.

Fitting method	Distribution	ln(L)	$e^{\ln(L)/(m+n)}$
Max Likelihood	Weibull $\bar{y}=5.99$	-12.14	0.676
Max Likelihood	Weibull $\bar{y}=0$	-12.22	0.674
Probit	Normal	-12.34	0.672
Overlap Range	Normal	-14.42	0.628

For the two latter distributions the values for ln(L) have been calculated from the Normal distributions plotted in <8> and <9>. One can see that our method, which has taken maximum account of the recorded numerical information gives the highest likelihood. Therefore we believe the Weibull model to be a good representation of the test series.

The rightmost column in the table shows a quantity that intuitively can be used to judge how well the model interpretes the data. It will approach unity for a perfect fit. It remains to analyze whether the quantity can be developed into a useful indicator for the quality of models and data sets.

### 5. Summary and Conclusions

We have shown how the Maximum Likelihood method can be used for the fitting of a Weibull cumulative risk function to censored biomechanical data. We have found that the method is easy to handle with the aid of available computer programmes. We believe that the three parameter Weibull distribution provides a result that gives a better representation of censored data than earlier methods used in biomechanics. Several of a selection of data sets have shown a marked right skewness, which speaks against the usual application of the Normal distribution.

Reasons for unexpected behaviour of data may be a physical mechanism, the selection of test subjects, the experimental scheme, or the analysis model. We have therefore formulated some tentative questions and comments concerning the behaviour and meaning of biomechanical data. The questions should be possible to analyze when the method is developed more and applied to more biomechanical data sets. We encourage other interested researchers to analyze their data with this method and see which conclusions that are possible to draw. The successful application of statistical models will come when there is a close cooperation between the experimentalist and the statistician. This would benefit injury research as well as the development of proper protection measures.

### 6. Appendix on the Weibull Distribution

The Weibull cumulative frequency distribution with one variable and three parameters is defined as:

$$W(z;K) = W(z;\alpha,\beta,y) = 1 - e^{-\left(\frac{z-y}{\alpha}\right)^\beta}$$

where

z is the independent variable, defined from y to infinity  
 $\alpha$  is the scale parameter, always greater than zero  
 $\beta$  is the shape parameter, always greater than zero  
y is the location parameter

The three parameters permit a wide flexibility when modeling distributions for various purposes as shown in Figure 6 on page 11.

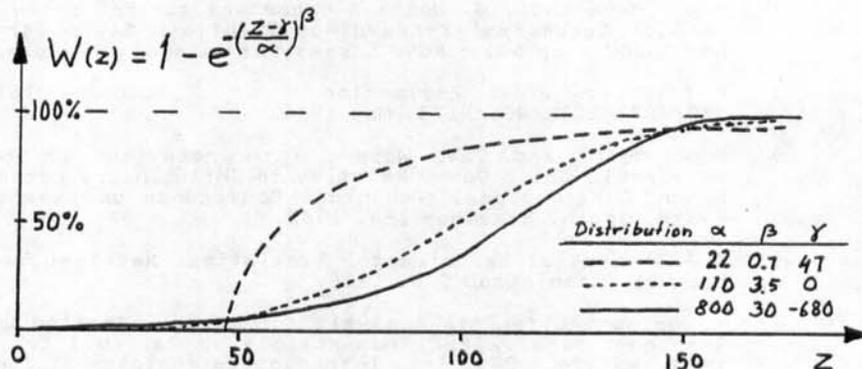


Figure 6. Various forms of the Weibull distribution

The Weibull distribution has found wide use in reliability engineering <6>, where failures or other significant events occurring in technical systems and components are studied. The variable  $z$  then denotes a suitable accumulating load variable like: age, distance traveled, number of operating cycles, time to next call in a telephone switchboard, time to perform a maintenance task, etc.

The parameter  $\beta$  is of special interest, because it has been possible to associate it with different types of physical failure mechanisms. In reliability engineering the following age related failure types have been identified:

- $\beta < 1$  Decreasing hazard intensity, inherent physical deficiency and weakness leading to highest risk in early life, sometimes called "infant mortality".
- $\beta = 1$  Constant hazard intensity, exponentially random failures. No relation between risk and accumulated life. Often several independent failure mechanisms.
- $\beta > 1$  Increasing hazard intensity, wearout failures at higher loads. Preventive measures (inspection and maintenance) can be scheduled before high load (age) failures occur.
- $\beta 1-3$  Distribution skewed right.
- $\beta 3-4$  One dominant wearout mechanism centers an approximately symmetric (Normal) distribution on average failure age.
- $\beta > 4$  Distribution skewed left.

#### 7. List of References

- 1 B. Aldman, H. Mellander and M. Mackay: The Structure of European Research into the Biomechanics of Impacts. Proceedings Twenty-Seventh Stapp Car Crash Conference, paper no. 831610, pp 129 - 136, Society of Automotive Engineers Inc., 1983.
- 2 R.C.H. Cheng and T.C. Iles: Confidence bands for Cumulative Distribution Functions of Continuous Random Variables. Technometrics, Vol.25, No.1, February 1983, pp 77-86.
- 3 R.H. Eppinger, J.H. Marcus and R.M. Morgan: Development of Dummy and Injury Index for NHTSA's Thoracic Side Impact Protection Research Program. Government/Industry Meeting and Exposition, Washington D.C., May 21-24, 1984, paper no. 840885, Society of Automotive Engineers Inc., 1984.
- 4 R.A. Fisher: Statistical Methods for Research Workers. First edition Edinburgh 1925, many later editions.
- 5 G. Grime: Probabilities of Injury to Car Occupants in Accidents with a Practical Example. Proceedings of the Institution of Mechanical Engineers, Automobile Division, No. 35, Volume 196, 1982.
- 6 K.C. Kapur and L.R. Lamberson: Reliability in Engineering Design. John Wiley and Sons, 1977.

- 7 R.W. Lowne and J.G. Wall: A Procedure for Estimating Injury Tolerance Levels for Car Occupants. Proceedings Twentieth Stapp Car Crash Conference, paper no. 760820, pp 641 - 664, Society of Automotive Engineers Inc., 1976.
- 8 R.W. Lowne: Estimation of Human Tolerance Distribution. ISO/TC22/SC12/WG6/N173, May 1984.
- 9 H.J. Mertz and D.A. Weber: Interpretations of the Impact Responses of a 3-Year-Old Child Dummy Relative to Child Injury Potential. Proceedings of the Ninth International Technical Conference on Experimental Safety Vehicles. Kyoto, Japan, November 1-4, 1982.
- 10 M.G. Natrella: Experimental Statistics, National Bureau of Statistics Handbook 91. Washington D.C. 1966.
- 11 W. Nelson: Life Data Analysis for Units Inspected Once for Failure (Quantal Response Data). IEEE Transactions on Reliability, Vol.R-27, No.4, October 1978, pp 274 - 279. This introduction contains 21 further references on analysis methods.
- 12 Numerical Algorithms Group Ltd. NAG Central Office, 7 Banbury Road, Oxford OX2 GNN, England, 1983.
- 13 Position Paper on HIC Levels, Document N-146 prepared by the U.S. delegation to ISO/TC22/SC12/WG6, June 1983.
- 14 J.A. Searle, J. Bethell and G. Baggaley: The Variation of Human Tolerance to Impact and Its Effect on the Design and Testing of Automotive Impact Performance. Proceedings Twenty-Second Stapp Car Crash Conference, paper no. 780885, pp 3 - 32, Society of Automotive Engineers Inc., 1978.