# Statistical Modeling of Mechanical Bearing Life Testing

Sébastien Blachère (SKF), Martin Bootsma (Utrecht University), Alessandro Di Bucchianico (Eindhoven University of Technology), Mike Keane (Delft University of Technology), Xinru Li (Leiden University), Andrea Roccaverde (Leiden University), Cristian Spitoni\*(Utrecht University), Dong Yan (Leiden University)

#### Abstract

We investigate reliability test plans under different censoring schemes for estimating performance of bearings with different life characteristics. The test plans, which are based on Weibull distributions, should deliver estimates of performance characteristics with a specified precision. We present results on both a theoretical approach based on Fisher information and a simulation approach.

Keywords: bearing, lifetime, Weibull distribution, censoring, test plan, Fisher information

# 1 Introduction

### 1.1 About SKF

SKF is a global technology provider offering products and services related to bearings and units, seals, mechatronics and lubrication systems. Its headquarters are located in Sweden. The company has around 165 production sites in 28 countries. SKF has several research centres, including one in the Netherlands in Nieuwegein.

Mechanical bearings (see Figure 1) are an important product of SKF. They are mechanical elements that constrain motions to desired motions only, and at the same time reduce friction between moving parts. There is a wide range of applications of bearings, including bicycles, cars, manufacturing machines, trains, wind turbines and airplanes (see Figure 2). Sizes of bearings range from less than 10 mm to 14 m. Since bearings are essential for the proper and safe



Figure 1: Bearings.

functioning of machines and equipment, it is essential for SKF to give customers reliable information on the performance. The performance of mechanical bearings is

<sup>\*</sup>Corresponding author



Figure 2: Examples of applications of mechanical bearings.

expressed through their life, i.e., the amount of time or number of revolutions that a bearing is capable to reach within nominal functioning. Bearing life depends on various parameters like the bearing type or size and the operating conditions (speed, load, lubrication,...). SKF uses an internal calculation tool based on physical models to assess bearing life. However, there is a need for life testing on actual bearings in order to validate these models, evaluate performance of prototypes and obtain insight in effect of design choices. Even identical bearings running under identical operating conditions may experience a wide dispersion in their life. The ratio of the longest to the shortest life may exceed 100 in large samples.



Figure 3: Test rigs used in bearing life tests.

### 1.2 Outline of the problem

Life tests consist of running a group of bearings under identical operating conditions until a stopping criterion is fulfilled. There exist three classical stopping criteria:

- **Type I**. The test is stopped when a preset time has been reached. This is illustrated in the left-hand side of Figure 4.
- **Type II**. The test is stopped when a preset number of failures has been reached. This is illustrated in the right-hand side of Figure 4.
- **Hybrid**. The test is stopped when either a preset time or a preset number of failures has been reached. This is illustrated in Figure 5.

Life times of both failed and not yet failed bearings are recorded since both types of data contain information. Usually most bearings have not yet failed at the end of the life test. As usual in reliability engineering, the data is modelled using a Weibull distribution.



Figure 4: Type I and II stopping criteria.



Figure 5: Hybrid stopping criterion.

The objective of life tests is the estimation of the Weibull parameters (see Subsection 2.1 for details). The precision of such an estimation depends strongly on the test strategy (number of bearings tested, test duration, number of observed failures, replacement policy). Therefore, obtaining precise estimation of the Weibull parameters

may require lengthy (up to several months) and thus costly tests. It is thus necessary to understand the link between the test duration and the precision of the estimation in order to optimize the test time while reaching the target precision. The precision of the estimation is traditionally defined as the length of the confidence intervals for each of the parameters (see Subsection 2.1 for details) on a logarithmic scale:

$$R_{10}(L_{10}) = rac{L_{10,95}}{L_{10,05}}$$
 and  $R_{10}(\beta) = rac{eta_{95}}{eta_{05}}$ 

where  $(L_{10,05}, L_{10,95})$  and  $(\beta_{05}, \beta_{95})$  are the 90% two-sided confidence intervals for  $L_{10}$  and  $\beta$ , respectively.

**Problem statement** Given N life tests (i.e., life tests with bearings of the same type and conducted under identical conditions) to be run within the same fixed test capability, what is the strategy to follow to minimize the total test duration of all N life tests with a given confidence to obtain parameter estimates with a preset precision?

When designing a test strategy, the precision and the available bearings of a certain type are constraints. The degrees of freedom are the sample size, the type of stopping criterion, the replacement policy and the value associated to this stopping criterion (preset time and/or preset number of failures). Some extra degrees of freedom can be added like the replacement of failed bearings by new ones for instance, but they are not treated in this article.

In addition, both the test time and the test precision are depending on the test strategy in a stochastic way. Therefore, the specifications onto the precision and the time need to be expressed in terms of their distributions via the mean, standard deviation or some percentiles.

### 1.3 Approach

This report reflects our first attempt to tackle the general problem. We did so by concentrating ourselves on a simpler problem where we had only 1 type of bearing. For this case, we derived theoretical results based on Fisher information that we will allow us in the future to compare different testing strategies. We complemented these theoretical results by running simulations for different testing strategies based on a fast R code that we developed ourselves.

### 1.4 Outline of this article

This article is organized as follows. In Section 2 we provide details on the Weibull distribution (parametrizations, estimation for the different censoring schemes). Section 3 contains our theoretical approach to test plans based on the Fisher information. The results of our numerical simulations can be found in Section 4. We end our paper with conclusions and recommendations in Section 5.

# 2 Background on the model

### 2.1 Weibull Distribution

The Weibull distribution has been introduced in the setting of material strength by Waloddi Weibull (Weibull (1939)) and was later extended to a wide range of types of experimental data (Weibull (1951)). The Weibull distribution is one of the extreme value distributions and appears if one considers minima of random variables (which is very natural in material strength since material break when the weakest link fails). Another motivation for the use of the Weibull distribution is that the hazard rate is flexible since it is a power function and may thus model both increasing and decreasing failure rates (depending on the sign of the exponent).



Figure 6: Weibull.

The Weibull distribution is extensively used in reliability theory together with its special case, the exponential

distribution. The Weibull distribution possesses two main forms, one with 2 parameters (with domain  $(0, \infty)$ ) and one with 3 parameters (which has an additional location parameter so that the domain need not start at 0).

Here we will describe the background on the Weibull distribution (parameterizations, parameter estimation, censoring,  $\ldots$ ). For a comprehensive overview on the Weibull distribution, we refer to the excellent monograph Rinne (2008).

The two-parameter Weibull distribution has a scale parameter and a shape parameter. The standard representation in term of the cumulative distribution function is

$$F(x|\alpha,\beta) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}} \tag{1}$$

The parameter  $\alpha$  is the scale parameter, while the parameter  $\beta$  is the shape parameter. The exponential distribution is included as the special case  $\beta = 1$ . Note that for  $\beta > 1$ the Weibull distribution has an increasing failure rate. Here we will use another representation of the Weibull distribution which has the same shape parameter  $\beta$ , but a different scale parameter,  $L_{10}$ . In this representation, the cumulative density function has the form:

$$F(x|L_{10},\beta) = 1 - \left(\frac{9}{10}\right)^{\left(\frac{x}{L_{10}}\right)^{\beta}}.$$
(2)

The parameter  $L_{10}$  has a clear interpretation; it is the 10%-quantile of the distribution, i.e. 90% of the bearings survive at least until time  $L_{10}$ . The parameters are linked to each other through the relation

$$L_{10} = \alpha \left( -\log(9/10) \right)^{1/\beta}.$$
 (3)

The Weibull distribution can also be described by the survival function

$$S(x|L_{10},\beta) = \left(\frac{9}{10}\right)^{\left(\frac{x}{L_{10}}\right)^{\beta}}.$$
(4)

In this paper, we will consider  $L_{10}$  as a parameter of the Weibull distribution. However, if one wants to estimate the 10%-quantile of the distribution one could estimate  $L_{10}$  also non-parametrically.

The probability density function of the Weibull distribution in this parameterization is given by

$$f(x|L_{10},\beta) = \frac{1}{L_{10}} \left(\frac{9}{10}\right)^{\left(\frac{x}{L_{10}}\right)^{\beta}} \left(\frac{x}{L_{10}}\right)^{\beta-1} \beta \log\left(\frac{10}{9}\right).$$
(5)

Note that in the remaining part we suppress in the notation for the cumulative distribution function and the probability density function sometimes the dependence on  $L_{10}$  and  $\beta$  to obtain clearer expressions.

### 2.2 Censoring

In real-life testing of bearings, it is impossible to test until all bearings have failed. This is also the case if one uses accelerated life testing, i.e. testing under more severe conditions than normal in such a way that using physical degradation laws one can connect life times under severe conditions to life times under normal conditions (see e.g., Meeker and Escobar (1998)). The data that we obtain from life tests will thus include data on non-failed items. Such observations are called *censored observations*. This terminology is used in engineering contexts (life tests of physical objects) as well as in medical contexts (clinical trials). Note that censored observations in the context of life tests do contain useful information on life times (one usually has a lower bound for the actual lifetimes). In order not to lose information, it is thus important to use statistical techniques that make use of both the censored and uncensored observations. This very much applies to the case of life tests for bearings, since typically the majority of observations is censored.

In order to include censored observations in statistical analyses, it is necessary to model the censoring mechanism at hand. In medical contexts (survival analysis) it is common to model censoring as a random variable independent of the lifetime distribution. In engineering context (reliability theory) it is more common to model censoring in a different way (see e.g., (Rinne, 2008, Section 8.3.1) for a detailed discussion about the different ways of modelling censoring mechanisms). We now consider in more detail censoring types that are common in reliability theory (they also appear in medical contexts under the name administrative censoring). We will restrict ourselves to the case of a single right-censoring. For multiple censoring schemes such as progressive censoring, we refer to the literature (see e.g., Ng et al. (2004) and (Rinne, 2008, Sections 8.3.3 and 8.3.4)).

#### Type I censoring

Suppose we run n bearings on n machines and we stop the experiment at a preset time T. The number r of failed items at time T is then the realisation of a random variable R. We have thus observed r failures at times  $\{x_{(1)}, x_{(2)}, \ldots, x_{(r)}\}$  and we have not observed the failure of the other n - r bearings. With the notation  $x_{(i)}$  we denote the  $i^{\text{th}}$  order statistic, i.e., the  $i^{\text{th}}$  smallest observed failure time. The joint probability density function to observe the failure times is given by:

$$f(x_1, \dots, x_n | L_{10}, \beta) = \frac{n!}{(n-r)!} \prod_{i=1}^r f(x_{(i)} | L_{10}, \beta) (1-F)^{n-r} (T | L_{10}, \beta).$$
(6)

for  $0 \le x_{(1)} \le x_{(2)} \le \ldots \le x_{(r)}$  (where  $r := \max\{i : x_{(i)} \le T\}$ ) and equals 0 otherwise. Note that for this type of censoring (known as administrative censoring in the context of survival analysis) the number of failures r is random and the end time is deterministic.

#### Type II censoring

Suppose we run n bearings on n machines and we stop the experiment once we have  $k \leq n$  failures. At the stopping time we have observed k failures at times  $\{x_{(1)}, x_{(2)}, \ldots, x_{(k)}\}$  and we have not observed the failure of the other n-k bearings. The joint probability density function to observe the failure times is given by:

$$f(x_1, \dots, x_n | L_{10}, \beta) = \frac{n!}{(n-k)!} \prod_{i=1}^k f(x_{(i)} | L_{10}, \beta) (1-F)^{n-k} (x_{(k)} | L_{10}, \beta).$$
(7)

which is defined for  $0 \le x_{(1)} \le x_{(2)} \le \ldots \le x_{(k)}$ . Note that in this type of censoring the number k of failed items is fixed, while the end time is random (it equals the  $k^{\text{th}}$  order statistic).

#### Hybrid censoring

An alternative censoring scheme is possible by combining the stopping criteria of Type I and Type II censoring, i.e. we stop the experiment when either we reach the preset time T or the preset number r of failures. In other words, we stop the experiment at the random time  $\min(T, X_{(k)})$ . This type of hybrid censoring was introduced in Epstein (1954). It is called type-I hybrid censoring in Balakrishnan and Kundu (2013).

Type I censoring is appealing from a practical point of view, since it fixes the duration of the experiment. A mathematical drawback is that it is harder to analyse (one needs to take into account the random time between T and the last failure time before T). Type II is appealing from a mathematical point of view since it is easy to analyse, because the number of failures is deterministic. The practical drawback is that one has no control on the duration of the experiment. A possible drawback of both Type I and hybrid censoring is that if T is chosen too small, that there are may be too few failures and the resulting estimates are poor. In order to overcome this drawback, Childs et al. (2003) introduced an alternative hybrid censoring scheme based on  $\max(T, X_{(k)})$ . However, this has the same practical drawback as type II censoring.

#### 2.3 Parameter estimation

In this subsection we discuss estimation of the  $L_{10}$  and  $\beta$  parameters of the twoparameter Weibull distribution under the three censoring schemes mentioned in the previous subsection. We will assume that failed items are not replaced during a life test. Maximum Likelihood is the preferred way of estimating parameters in reliability theory, because it not only has well-known asymptotic optimality properties but as exemplified by Formulas (6) and (7) it can easily deal with censored data (unlike e.g. the method of moments). The literature mostly deals with estimation for type II censoring, under the assumption that type I censoring can be dealt with in a similar way by conditioning on the number of failures in the interval [0, T] (cf. Remark 26 on page 438 of Rinne (2008)). An exception is Cohen (1965) which treats both types of censoring.

For Type II censoring the standard approach is to take derivatives of the loglikelihood equation with respect to the parameters and to note that after simplification of the equations one obtains  $\hat{\lambda}$  through the following relation (Cohen (1965)):

$$\widehat{\lambda} = \left(\sum_{i=1}^{n} x_i^{\widehat{\beta}}\right)^{\widehat{\beta}}.$$
(8)

The following equation determines  $\hat{\beta}$  in case of type II censoring:

$$\frac{1}{\hat{\beta}} = \frac{\sum_{i=1}^{n} x_i^{\hat{\beta}} \log(x_i) + (n-k) x_{(k)}^{\hat{\beta}} \log(x_{(k)})}{\sum_{i=1}^{n} x_i^{\hat{\beta}} + (n-k) x_{(k)}^{\hat{\beta}}} - \frac{1}{k} \sum_{i=1}^{k} \log(x_{(i)})$$
(9)

A similar relation holds in case of type I censoring. For hybrid censoring, Kundu (2007) describes that one has basically the same procedure where the form of the likelihood depends on which of the stopping criteria applies to the data set at hand. Existence and uniqueness of solutions of (9) are guaranteed unless all observations are equal (see e.g. Farnum and Booth (1997), Pike (1966)). Since the right-hand side of (9) can be proven to be an increasing function of  $\beta$  (see Farnum and Booth (1997)<sup>1</sup>, numerical procedures like Newton-Raphson quickly yield numerical solutions. There are also exist explicit approximate ML estimators based on Taylor expansions of the logarithm of the Weibull distribution (so transforming the Weibull distribution into an extreme value distribution, see e.g. Kundu (2007)) for details).

<sup>&</sup>lt;sup>1</sup>The proofs in Farnum and Booth (1997) are only written down for the uncensored case, but it can easily be shown that a slight adaptation make them work for the censored cases as well

The ML estimators are biased. Corrections are possible by noting that the distribution of  $\hat{\beta}$  is a pivotal quantity for  $\beta$ , i.e. the distribution of  $\hat{\beta}/\beta$  only depends on *n* and depending on the type of censoring *T* or *k* (see e.g., McCool (1970) who evaluates this distribution by Monte Carlo simulation). For a complete discussion of pivotal quantities we refer to Bain and Engelhardt (1991)) and McCool (1970).

In view of the Invariance Principle of Maximum Likelihood estimation (Zehna (1966)), it does not matter which parametrization of the Weibull distribution one chooses when one is only interested in the parameter estimates. However, it does make a difference when computing confidence intervals for the Weibull parameters. This is because the lack of formulas for exact confidence intervals necessitates to use asymptotic intervals for which the width directly depends on the asymptotic standard deviation of the parameter estimator. Explicit asymptotic confidence intervals for the Weibull parameters have been discussed in Meeker and Nelson (1976) and Kahle (1996). Both papers used observed Fisher information, but Meeker and Nelson (1976) do this for the logarithm of the Weibull distribution (so an extreme value distribution) since the asymptotic sampling distribution in that case converges faster. The formulas in both papers involve second derivatives of the incomplete gamma function, which can be expressed in terms of other special functions (see Geddes et al. (1990)). We followed the approach of Meeker and Nelson (1976) but used direct numerical integration to evaluate the integrals directly since there were no convergence problems.

# 3 Theoretical results

If we have a single bearing, we have a continuous decision process as long as the bearing has not failed: do we keep the experiment running or do we stop the experiment and replace the existing bearing with a new one. Which of the two options is most attractive depends on several factors: 1) the amount of information one is expected to get from each choice, 2) how fast this information is obtained and 3) the cost of replacing an existing bearing with a new one. Here we neglect the costs completely and we focus on the first two factors.

### 3.1 Type I censoring

Suppose we test a single bearing and we keep running the experiment till either the bearing fails or a fixed stopping time a has been reached. Suppose the lifetime of the bearing is distributed according to a Weibull distribution with parameters  $L_{10}$  and  $\beta$ . With probability  $S(a|L_{10},\beta)$  the bearing is still functioning at time a. The Fisher information  $I_1(a)$  from this experiment can therefore be calculated as:

$$I_{1}(a) = -S(a|L_{10},\beta))\frac{\partial^{2}}{\partial L_{10}^{2}}\log S(a|L_{10},\beta) - \int_{0}^{a} \left(\frac{\partial^{2}}{\partial L_{10}^{2}}\log f(x|L_{10},\beta)\right)f(x|L_{10},\beta)dx$$
(10)

This can be written as:

$$I_{1}(a) = \left(\frac{9}{10}\right)^{\left(\frac{x}{L_{10}}\right)^{\beta}} \frac{\left(\frac{a}{L_{10}}\right)^{\beta}\beta(1+\beta)\log\frac{10}{9}}{L_{10}^{2}} \\ -\int_{0}^{a} \frac{\beta - \left(\frac{x}{L_{10}}\right)^{\beta}\beta(1+\beta)\log\frac{10}{9}}{L_{10}^{2}} \frac{1}{L_{10}} \left(\frac{9}{10}\right)^{\left(\frac{x}{L_{10}}\right)^{\beta}} \left(\frac{x}{L_{10}}\right)^{\beta-1} \beta\log\left(\frac{10}{9}\right) dx$$
(11)

We do not have a closed-form expression for  $I_1$ , but the numerical evaluation of this integral is fast. If n bearings are tested with type I censoring, the Fisher information  $I_n$  is

$$I_n = nI_1. \tag{12}$$

There are two important time measures for the duration of the experiment, 1) na, i.e., n times the fixed stopping time a and 2) the total time that machines are running, denoted by  $T(L_{10}, \beta, n, a)$ . The first measure is relevant to the situation that the machines on which a bearing fails before time a cannot be used for other purposes or if one has to pay for the time use of these machines even if they are not running. The second time measure is important if machines on which a bearing fails can be used for other experiments.

For the first time measure, the Fisher information per time unit equals  $I_1/a$ . For the second time measure we need to calculate the expected running time of a singlebearing experiment with type-1-censoring.

$$E[T(L_{10},\beta,1,a)] = \int_{0}^{a} xf(x|L_{10},\beta)dx + aS(a|L_{10},\beta)$$
(13)

and the Fisher information per time unit equals  $I_1/E[T(L_{10}, \beta, 1, a)]$ . For the relevant case that  $\beta > 1$ , the larger a, the higher the Fisher information per time unit for the second time measure (see Figure 7). For the first time measure, there is an optimum. There are not many failures before time a if a is small. On the other hand, if ais large, many machines are empty because the bearing on that machine has failed already. The optimum as function of  $\beta$  is plotted in Figure 8.

### 3.2 Type II censoring

The Fisher information I of this experiment with respect to the parameter  $L_{10}$  is given by:

$$I = -E\left(\frac{\partial^2}{\partial L_{10}^2} \log\left(f_{(n,k)}(x_{(1)}, x_{(2)}, \dots, x_{(k)}|L_{10}, \beta)\right)\right).$$
 (14)

The second order derivative with respect to  $L_{10}$  of the logarithm of the probability density function is given by:

$$\frac{\partial^2}{\partial L_{10}{}^2} \log\left(f_{(n,k)}\right) = \frac{k}{\beta} + \frac{\beta(1+\beta)}{L_{10}{}^{2+\beta}} \log\left(\frac{10}{9}\right) \left((n-k)x_{(k)}{}^\beta + \sum_{i=1}^k x_{(i)}{}^\beta\right).$$
(15)



Figure 7: The Fisher information per time unit for type I censoring. In Figure a) the Fisher information is divided by na, in Figure b) the Fisher information is divided by the total running time.



Figure 8: Stopping time for type I censoring that maximizes the Fisher information per time unit  $(I_1(a)/a)$ .

Therefore, the Fisher information is given by:

$$I = -\int_{0}^{\infty} dx_{(1)} \int_{x_{(1)}}^{\infty} dx_{(2)} \dots \int_{x_{(k-1)}}^{\infty} dx_{(k)} f_{(n,k)} \frac{\partial^2}{\partial L_{10}^2} \log\left(f_{(n,k)}\right),$$
(16)

Luckily this k-dimensional integral can also be expressed as a double integral (Park (1996)). Let  $f_{k:n}$  be the density function of the  $k^{\text{th}}$  order statistic in a sample of size n. We have the following expression for  $f_{k:n}$  (see page 224 of Rinne (2008)):

$$f_{k:n}(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{(k-1)} S(x)^{(n-k)} f(x)$$
(17)

If we define  $g(w) := g(L_{10}, \beta, w)$  as

$$g(L_{10},\beta,w) := \int_{w}^{\infty} \left(\frac{\partial}{\partial L_{10}} \log\left(\frac{f(x|L_{10},\beta)}{S(w|L_{10},\beta)}\right)\right)^2 \frac{f(x|L_{10},\beta)}{S(w|L_{10},\beta)} dx$$
(18)

we can write the Fisher Information as:

$$I = n \left(\frac{\beta}{L_{10}}\right)^2 - (n-k) \int_0^\infty g(w) f_{k:n}(w) dw \tag{19}$$

It can be shown, e.g., by performing the substitution  $y = \left(\frac{w}{L_{10}}\right)^{\beta} - \left(\frac{x}{L_{10}}\right)^{\beta}$ , that the function g(w) does not depend on w. We obtain that  $g(w) = \left(\frac{\beta}{L_{10}}\right)^{2}$ . Therefore we obtain as result that the Fisher information for Type II censoring has the following form:

$$I = k \left(\frac{\beta}{L_{10}}\right)^2,\tag{20}$$

i.e., each observed failure time provides the same amount of information, independent of the sample size or the order of the failure.

There are again two important time measures for the duration of the experiment, 1) the time till the  $k^{\text{th}}$  failure, which is given by the  $k^{\text{th}}$  order statistic  $x_{(k)}$  and 2) the total time that machines are running, denoted by  $T(L_{10}, \beta, n, k)$ .

The total running time is given by:

$$T(L_{10},\beta,n,k) = \sum_{i=1}^{k} x_{(i)} + (n-k)x_{(k)}$$
(21)

and the expected total running time can be calculated once we know the expected time of the  $k^{\text{th}}$  order statistic, i.e.,

$$E(T(L_{10},\beta,n,k)) = \sum_{i=1}^{k} E(x_{(i)}) + (n-k)E(x_{(k)})$$
(22)

The probability density function of the  $j^{\text{th}}$  order statistic is given by:

$$f_{j:n}(x) = \frac{n!}{(j-1)!(n-j)!} F(x|L_{10},\beta)^{(j-1)} (1 - F(x|L_{10},\beta))^{n-j} f(x|L_{10},\beta)$$
(23)

and the expectation of the  $j^{\text{th}}$  order statistic is given by (see Formula (5.34) of Rinne (2008)):

$$E(x_{(j)}) = j\binom{n}{j} \frac{\Gamma(1+\frac{1}{\beta})}{(\log\frac{10}{9})^{\frac{1}{\beta}}} L_{10} \sum_{i=0}^{j-1} \frac{(-1)^i \binom{j-1}{i}}{(n-j+i+1)^{1+\frac{1}{\beta}}}.$$
 (24)

We propose as a measure to compare several test strategies the Fisher information per time unit, i.e., the Fisher information divided by the expected total running time of the machines, or the Fisher information divided by the time until the  $k^{\text{th}}$  failure.

If the total running time is relevant, the higher k the more information is obtained (see Figure 9(a)), simply because every failure gives the same amount of information and failures occur more rapidly when the bearing are ageing. If the time until the  $k^{\text{th}}$  failure is relevant, there is an optimal value of k, as can be seen in Figure 9(b). This optimal value of k can also be interpreted as an optimal value of the ratio k/n and this optimal ratio depends on  $\beta$  (see Figure 10).



Figure 9: The Fisher information per time unit for Type II censoring with 30 machines. In Figure a) the Fisher information is divided by the total running time, in Figure b) the Fisher information is divided by the time of the  $k^{\text{th}}$  failure multiplied with the number of machines.



Figure 10: Optimal ratio for the fraction k/n if the time until the  $k^{\text{th}}$  failure is relevant.

# 4 Numerical results

In this section we study by simulation different test strategies in the simplified case when there is only one type of bearing and failed items are not replaced. The test strategies all involve hybrid censoring. In order to have practical relevance, all test strategies should meet the condition that  $R_{10;0.90}^{\beta} < 12$ . The optimality criterion is to minimize to 80%-percentile of the TTT (Total Time on Test) statistic, i.e. the sum of failure times for the failed items and the testing period for the items that did not fail before the end of the test. The failure times were sampled from a Weibull distribution with  $L_{10} = 100$  and shape parameter  $\beta = 1.1$ . In the hybrid censoring testing strategies we varied the number of items on test between 20 and 30, the number of failed items between 4 and 10 and the maximum testing period between 400 and 600 (with steps of size 25). In order to obtain accurate values, we used 30,000 replications for each setting. The irregular shapes in the contour plots are interpolation artifacts caused by the integer values for the number of bearings.



## 4.1 Results Type I Censoring

Figure 11: Type I censoring.

bearings	20	21	22	23	24	25	26	27	28	29	30
trunc	850	725	625	550	475	425	375	325	275	275	250

Table 1: Minimum required number of failures under type I censoring.

For type I censoring we see in Figure 11 that the condition  $R_{10;0.90}^{\beta} < 12$  is met by combinations of stopping times and number of bearings that are to the right of the straight line that goes from 23 bearings and stopping time 475 to 25 bearings and stopping time 425. It follows from Table 1 that the required stopping time decreases linearly with the number of bearings (approximately 50 to 75 hours per bearing). It follows from Figure 11 that the optimal choice is to stay exactly on the straight line to obtain a minimal TTT value.

### 4.2 Results Type II Censoring

For Type II censoring we see artifacts in the plots in spite of the 30,000 replications in the simulation. Therefore we also present the minimum number of failures to ensure  $R_{10;0.90}^{\beta} < 12$  in Table 2. Note the sharp drop when going from 22 to 23 bearings. We note that the TTT value does not change much for a fixed number of failures when we increase the number of bearings. This means that the increase in TTT for a fixed number of bears is fairly constant in the range of bearings that we considered: every extra failure in the stopping criterion causes an increase of 1000 in the TTT value.



Figure 12: Type II censoring.

bearings	20	21	22	23	24	25	26	27	28	29	30
r	11	9	8	4	4	4	3	3	3	3	3

Table 2: Minimum required number of failures under type II censoring

### 4.3 Results Hybrid Censoring

As discussed in Subsection 2.2, the idea of hybrid censoring is to have a beforehand fixed maximum testing time (due to the type I censoring mechanism) which is very important from a practical point of view (scheduling of testing facilities), but at the same time have the option to stop the testing earlier if the test results allow sufficiently precise estimates (due to the type II censoring mechanism). However, this is more complicated than it looks at first sight. For example with type I censoring and 20 bearings we need a stopping time of 850. If we add any corresponding type II censoring stopping criterion, then this means that the value of  $R_{10:0.90}^{\beta}$  will increase since we may stop too early. For example, if we perform hybrid censoring by naively combining the type I stopping time of 850 with the type II criterion for 20 bearings (i.e., stop after 11 failures), then  $R_{10;0.90}^{\beta} = 12.1$ . In order to meet the  $R_{10;0.90}^{\beta} < 12$  condition we need to increase either the type I or the type II criterion. For 20 bearings, this means we could choose the stopping time to be equal to 850 and number of failures to be equal to 12 (with TTT approximately 5900) or choose stopping time 875 and number of failures 11 (with TTT approximately 6150). It is thus better to fix the type I criterion and increase the type II criterion when applying hybrid censoring.



Figure 13: Hybrid censoring - 20 bearings.

# 5 Discussion

In this section we first summarize the conclusions that we may draw from our results. We then present recommendations to SKF for further lines of research.

### 5.1 Conclusions

In this paper we studied a simplified case of the testing problem posed by SKF. The simplification consisted of considering only one type of bearing and no replacement of failed items. We approached the problem through both a theoretical approach based on Fisher information and a numerical approach based on simulation.

A key insight of the approach based on Fisher information is to continue testing as long as there is a positive rate of contributing information. Information per test time leads to optimal values for the stopping time (Type I censoring) or the number of failures (Type II censoring) expressed as a function of  $\beta$ .

Simulations help to minimize tests time along the Pareto front of strategies. For Type I censoring there seems to be a linear dependence of  $R_{10}^{\beta}$  on the truncation time. For Type II censoring we see drastic changes for smaller values of bearings in the number of required failures in order to satisfy a constraint on  $R_{10}^{\beta}$ . The corresponding values for the TTT, however, seem to be fairly constant as a function of the number of bearings. For hybrid censoring one cannot simply combine the criteria of Type I and Type II censoring. In order to get optimal TTT values, one should fix the Type I criterion of a given number of bearings and increase the Type II criterion.



Figure 14: Hybrid censoring - 30 bearings.

### 5.2 Recommendations for Future Research

Future research is needed to extend the results of the current paper to more realistic situations with different types of bearings and more complex replacement strategies (not only replacement of failed items individually, but also in pairs as testing devices usually combines bearings in groups of 2 or 4 bearings).

For the simulations we recommend to use larger simulations or develop variance reduction techniques like importance sampling in order to obtain more stable results. Since the simulations were performed for only one set of values of the Weibull parameters  $L_{10}$  and  $\beta$ , it is recommended to study the influence of  $L_{10}$  and  $\beta$ .

The approach based on Fisher information should be extended to include tests with more than one type of testing as well as hybrid testing. It is also recommended to perform a sensitivity analysis.

A final idea is to explore the idea of approximation the Weibull distribution with an exponential distribution when  $\beta \approx 1$ .

### Acknowledgements

We thank SKF for supplying us with the necessary background information for this interesting problem. S. Blachère wishes to thank Mr. Alexander de Vries, Director SKF Group Product Development, for his kind permission to publish this article.

# References

L. Bain and M. Engelhardt. Statistical Analysis of Reliability and Life Testing Models
 Theory and Methods. Marcel Dekker, New York, 2nd edition, 1991.

- N. Balakrishnan and D. Kundu. Hybrid censoring: Models, inferential results and applications. *Comput. Stat. Data An.*, 57:166–209, 2013.
- A. Childs, B. Chandrasekar, N. Balakrishnan, and D. Kundu. Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution. Ann. Inst. Stat. Math., 55:319–330, 2003.
- A. Cohen. Maximum Likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics*, 7(4):579–588, 1965.
- B. Epstein. Truncated life tests in the exponential case. Ann. Math. Stat., 25:555–564, 1954.
- N. Farnum and P. Booth. Uniqueness of Maximum Likelihood estimators of the 2-parameter Weibull distribution. *IEEE Trans. Rel.*, 46(4):523–525, 1997.
- K. Geddes, M. Glasser, R. Moore, and T. Scott. Evaluation of classes of definite integrals involving elementary functions via differentiation of special functions. *Appli*cable Algebra in Engineering, Communication and Computing, 1(2):149–165, 1990.
- W. Kahle. Estimation of the parameters of the Weibull distribution for censored samples. *Metrika*, 44:27–40, 1996.
- B. Kundu. On hybrid censored Weibull distribution. J. Stat. Plann. Inf., 137:2127– 2142, 2007.
- J. McCool. Inference on Weibull percentiles and shape parameter from Maximum Likelihood estimates. *IEEE Trans. Rel.*, R-19(1):2–9, 1970.
- W. Meeker and L. Escobar. Statistical Methods for Reliability Data. Wiley, New York, 1998.
- W. Meeker and W. Nelson. Weibull percentile estimates and confidence limits from singly censored data by Maximum Likelihood. *IEEE Trans. Rel.*, R-25(1):20–24, 1976.
- H. Ng, P. Chan, and N. Balakrishnan. Optimal progressive censoring plans for the Weibull distribution. *Technometrics*, 46(4):470–481, 2004.
- S. Park. Fisher information in order statistics. J. Amer. Statist. Assoc., 91(433): 385–390, 1996.
- M. Pike. A suggested method of analysis for a certain class of experiments in carcinogenesis. *Biometrics*, 22:142–161, 1966.
- H. Rinne. The Weibull Distribution: A Handbook. CRC Press, 2008.
- W. Weibull. Statistical theory of the strength of materials. Ing. Vetenskaps Akad. Handl., 151:1–47, 1939.

- W. Weibull. A statistical distribution function of wide applicability. J. Appl. Mech., 18(3):293–297, 1951.
- P. Zehna. Invariance of Maximum Likelihood estimators. Ann. Math. Statist., 37(3): 744–744, 1966.