ShinyApp for simple Weibull Analysis

Jan Walczak

21. August 2020

1 Introduction

This project revolves around the Weibull distribution

$$F(t) = 1 - \exp\left(-\left(\frac{t - t_0}{T}\right)^b\right),$$

with $t_0 = 0$. The Weibull distribution is used in reliability and survival analysis for answering the question: "How likely is a unit to survive until a certain point in time?" We aim at calculating the parameters b and T using data. With data from experiments we can estimate the underlying quantiles of the Weibull distribution using different methods.

We will mainly use data from reliability tests, where we measure the time to failure for a number of units. The data comes in two forms: right-censored and uncensored data. This basically means that either all units have failed during the test and we know all of their times to failure (uncensored), or we have some survivors that lasted for the time of the experiment (right-censored). The methods differ on what data they use to estimate the quantiles of the underlying distribution. Some only use information about failed units, some will incorporate the information about survivors.

The estimated quantiles are then used to estimate the parameters b and T. Using a proper transformation we can calculate these using a linear regression (see "Methods"for more details).

2 Methods

In general the data has to be sorted by increasing time to failure.

2.1 Weibull Paper

2.2 Median Rank Regression

The Median Rank method uses uncensored data. Hence it is not able to incorporate information about survivors. We use sorting of the data to assign ranks i in an ascending order. The size of the sample is denoted by n. The estimator is then given by

$$F_i = \frac{i - 0.3}{n + 0.4}.$$

2.3 Sudden Death Method

The Sudden Death Method is the only method in this project using multiple samples of sizes n_i and $N = \sum_i n_i$. The idea is to do multiple experiments (of the same design) and end the experiment after the first failure. Hence the data consists of one failed unit and $n_i - 1$ survivors per sample. We sort the time to failure of the failed units. The ranking of the failed units is then given by

$$j_i = j_{i-1} + \Delta j_{i,i-1},$$

where

$$\Delta j_{i,i-1} = \frac{N + 1 - j_{i-1}}{N + 1 - \sum_{k=1}^{i-1} n_k}.$$

The quantile estimator is finally given by

$$F_i = \frac{j_i - 0.3}{n + 0.4}.$$

2.4 Kaplan-Meier Method

The Kaplan-Meier Method can be used in right-censored data. Again, the data is sorted by time to failure. KM can be used when there are multiple events (survivals and fails) at one time. The estimator is given by

$$F_i = 1 - \prod_{t_i \le t} \left(\frac{n_i - d_i}{n_i} \right),$$

where n_i is the number of units in the experiment up until t and d_i is the number of failed units at the time t. If the last event consists of failures only, we should use the adjusted estimator

$$F_i = 1 - \prod_{t_i \le t} \left(\frac{n_i - d_i + 1}{n_i + 1} \right)$$

to avoid a quantile $F_i = 1$. This would lead to an error when transforming the estimation for the linear regression.

2.5 Nelson Method

The Nelson Method can be used on right-censored data. We first assign an inverse Ranking r_i for all of the samples in our data. Then, we compute the failure rate by $\lambda_{\mathrm{Nel},i}(t) = \frac{1}{r_i}$ for all failed units. The cumulative failure rate is given by $H_{\mathrm{Nel},i}(t) = \sum_{j=1}^{i} \lambda_{\mathrm{Nel},i}$. Finally we use the estimator

$$F_i(t) = 1 - e^{-H_{\mathrm{Nel},i}(t)}.$$

Due to an strictly monotonously increasing λ_{Nel}, this method is only suitable for Weibull distributions with $b > 0$. (For a derivation of λ_{Nel} please check other sources.)

2.6 Johnson Method

The Johnson Method also incorporates information about survivors. The ranking is given by the following:

$$j_i = j_{i-1} + \Delta j_{i,i-1},$$

$j_0 = 0$ and

$$\Delta j_{i,i-1} = \left(\frac{N + 1 - j_{i-1}}{N + 1 - K} \right) \cdot n_F,$$

where n_F is the number of failed units at time t and K is the sum of all units that are removed (failures and survivors) from the experiment before time t. The quantile estimator is then given by

$$F_i = \frac{j_i - 0.3}{n + 0.4}.$$

2.7 Linear Regression for Estimation of Weibull Parameters

3 Implementation

4 ShinyApp