

Internship proposal – 2014/2015

GOODNESS-OF-FIT TESTS FOR THE WEIBULL DISTRIBUTION WITH CENSORED DATA

CONTEXT:

In the field of industrial reliability, the Weibull law is a very common distribution used to model lifetimes of non repairable components. Indeed, this two-parameter distribution is very flexible and allows to model equipment with increasing (wear-out) or decreasing (rejuvenation) failure rate. Even though the Weibull distribution is popular, the checking of its relevance for a given data set is often carried out with elementary tools, such as the Weibull-plot. Yet there exist more advanced techniques, the statistical goodness-of-fit (GoF) tests, which aim to determine if a distribution is adapted to a given data sample.

Many GoF tests for the Weibull distribution have been developed and a review of them is proposed in the doctoral dissertation of Meryam Krit, entitled "Goodness-of-fit tests in reliability: Weibull distribution and imperfect maintenance models". This PhD work has been jointly supervised by EDF R&D and the Jean Kuntzmann Laboratory (LJK) from Grenoble University. In particular, this thesis gives GoF tests adapted to censored data, that is to say to data that are partially observed, as is the case of components that never failed and for which only a lower bound of the lifetime is known (right censoring).

OBJECTIVE:

The objective of the internship is to continue the work introduced by Meryam Krit, by carrying out a power study of the GoF tests using Monte Carlo simulations. This analysis will allow to identify the most efficient tests for censored data. New GoF tests can be developed, for instance by using non parametric techniques (Kaplan-Meier estimator of the survival function) or by looking for sufficient statistics.

ACTION PLAN:

1. In the first step, the student will have to understand the industrial context and take note of the GoF tests described in Meryam Krit's dissertation.
2. Then the student will carry out the power study to compare the tests. For different alternative distributions, different sample sizes, different types of censoring and different censoring rates, this analysis will consist in simulating a large number of data sets and applying the GoF tests on these samples in order to estimate the powers of the tests. A detailed analysis of the results will enable to identify which GoF tests are the most efficient and under which operation conditions. The R package EWGoF can be used since it implements all the GoF tests listed in Meryam Krit's dissertation.
3. Next the student will consider the development of new GoF tests for the Weibull distribution with censored data.
4. Finally a report presenting all the work carried out during the internship will be written.

DESIRED PROFILE:

Students in Master 2 or from an engineering school in statistics or applied mathematics

REQUIRED SKILLS:

Probabilistic and statistical methods

Computer science

IT ENVIRONMENT:

Microsoft Office

R software

REPRESENTATIVES:

emmanuel.remy@edf.fr

olivier.gaudoin@imag.fr

LOCATION:

Depending on the arbitration between EDF R&D and LJK:

- either EDF R&D – Groupe "Maintenance, Optimisation, Décisions d'Investissements, Fiabilité" - Lab Chatou – 06 quai Watier, 78401 Chatou Cedex
- or LJK – Campus de Saint Martin d'Hères – 51, rue des Mathématiques, BP 53, 38041 Grenoble Cedex 09

FEES:

Between 700 and 850 € per month