

Bachelor's Project

Bayesian Software Reliability Growth Models

A. Di Bucchianico

Keywords software reliability growth models, Bayesian statistics, Markov Chain Monte Carlo.

1 Introduction

In the courses *Mathematical Statistics* and *Regression Analysis and Analysis of Variance* you were introduced to the classical, frequentist approach to statistics. In this approach data is used to select models that fit best to data. Often, these models depend on parameters, in which case the selection boils down to choosing optimal parameters. The parameters are considered as constants. There is another approach to statistics that tries to incorporate prior knowledge by treating the model parameters as random variables, the distributions of which depend themselves on parameters (hyperparameters). Data is then used to update this prior knowledge using Bayes's theorem. This approach is therefore called Bayesian statistics.

Practical use of Bayesian statistics has to deal with two problems:

1. how to extract and model prior knowledge
2. computation of updated probabilities

Using recent advances in Markov Chain Monte Carlo (MCMC) sampling, problem 2) can now be successfully handled in many practical situations. Many algorithms are now known and there is even a free software tool (WINBUGS).

In this bachelor's project you will learn about Bayesian statistics and apply it in a specific applied context. This context is statistical models for software testing. During software testing, it is important to know when to stop testing. Software reliability growth models are a statistical tool to support such decisions. The Probability and Statistics group is developing a software tool (SREtool) to analyse such models. This is a joint project with the Refis company (www.refis.nl) and LaQuSo (www.laquso.com).

2 Project goals

The project focusses around the Littlewood-Verrall model. This is a Bayesian software reliability growth model that allows a frequentist analysis. The main project goals are:

1. describe in detail the Bayesian and frequentist treatments of the Littlewood-Verrall model
2. implement these treatments in R so that they can be included in SREtool (both for grouped and ungrouped data)
3. compare the Bayesian and frequentist treatments of the Littlewood-Verrall model by
 - (a) by applying them to some data sets from practice
 - (b) by simulating scenarios (for this comparison you need to implement this model into WINBUGS or write your own simulation code in R)

3 Project planning

The project should begin with reading the necessary background on software reliability growth models, Bayesian statistics and MCMC. In particular the papers Jeske and Zhang (2005), Kuo and Yang (1996), Littlewood and Verrall (1973), and Xie et al. (2003) should be studied as well as the introductory chapters of Singpurwalla and Wilson (1999). Meanwhile, the student should acquaint him/herself with MCMC implementations

References

- D.R. Jeske and X. Zhang. Some successful approaches to software reliability modeling in industry. *J. Syst. Software*, 74(1):85–99, 2005.
- L. Kuo and T.Y. Yang. Bayesian computation for nonhomogeneous Poisson processes in software reliability. *J. Amer. Statist. Assoc.*, 91(434):763–773, 1996.
- B. Littlewood and J. L. Verrall. A Bayesian reliability growth model for computer software. *J. Roy. Statist. Soc. Ser. C Appl. Statist.*, 22:332–346, 1973.
- N.D. Singpurwalla and S.P. Wilson. *Statistical Methods in Software Engineering*. Springer Series in Statistics. Springer-Verlag, New York, 1999. Reliability and risk.
- M. Xie, G.Y. Hong, and C. Wohlin. Modeling and analysis of software system reliability. In W.R. Blischke and D.N.P. Murthy, editors, *Case Studies in Reliability and Maintenance*, chapter 10, pages 233–249. Wiley, New York, 2003.