Using Bayesian Theory for Estimating Dependability Benchmark Measures

Michel Cukier and Carol S. Smidts

Center for Reliability Engineering
Department of Materials and Nuclear Engineering
University of Maryland at College Park
{mcukier, csmidts}@eng.umd.edu

1. Introduction

Assessing the quality of service of a computer system is a difficult task. A lot of work has been conducted on evaluating quality of service attributes like performance, robustness, and dependability. Two approaches used for evaluating performance and robustness are modeling and benchmarking. For evaluating dependability, modeling can be used either alone or combined with fault injection [Sie92, Kan91]. However, less work has been conducted on building dependability benchmarks. A dependability benchmark can be defined as "a way to evaluate the behavior of components and computer systems in the presence of upsets, allowing the quantification of dependability attributes or the characterization of the systems into well defined dependability classes" [Mad01].

This paper focuses on the quantification part of the definition. The goal of this paper is to propose the use of Bayesian estimation methods for quantifying dependability attributes. We first will give a brief overview of two estimation Schools in Section 2. We will then illustrate our proposal by focusing on a key parameter for fault-tolerant systems, the coverage factor. We will introduce the coverage factor in Section 3 and present coverage factor estimations in Section 4.

2. How does the Bayesian theory differ from the frequentist theory?

The frequentist School and the Bayesian School are two important estimation branches in statistics. We now briefly compare the applicability of the two theories. After having conducted some experiments, there are different ways for processing the obtained results in order to get an estimation.

An estimation obtained using the frequentist theory is based only on the results collected during the experiments. The distribution associated with the experiment is often introduced in order to obtain more accurate estimations.

When applying the Bayesian theory, an estimation is based on the results collected during the experiments and a prior knowledge of the estimation. This prior knowledge could be based on previous experimental results or on expert knowledge. As for the frequentist case, a distribution is often associated with the experiment. In the Bayesian case, another distribution is introduced to include the prior knowledge, called a *prior distribution*. The combination of these two distributions leads to the *posterior distribution*. The posterior distribution is then used to calculate the estimation.

This combination of two sources of information often has the advantage that, if the experimental results confirm the prior knowledge, a smaller number of experimental results compared to the frequentist approach will be needed to obtain the same estimation.

One of the purposes of a benchmark is to obtain an estimation without having to gather too many experimental results. However, in that case, the estimation might then not be very accurate. The use of Bayesian methods might increase the accuracy since, besides experimental results, a prior knowledge is also used to calculate an estimation. The following sections will present some examples showing the advantages of the Bayesian method for an important parameter of fault-tolerant systems: the coverage factor.

3. Coverage factor

Let us first formally define the coverage factor. The reaction of a fault-tolerant system will depend on two different inputs: the inserted *upsets* and the submitted *workload*. The upsets are the inputs specific to the fault tolerance mechanisms. The workload represents the environment. The overall input space of a fault-tolerant system is then the Cartesian product $G = U \times W$, where U is the upset space and W is the workload space. Let us define H as a variable characterizing the handling of a particular upset. The coverage (factor) of a fault-tolerant system can then be defined as:

$$c = \Pr\{H = 1 | g \in G\}$$

i.e., the conditional probability of correct upset handling, given the occurrence of an upset/workload pair g.

4. Frequentist and Bayesian estimations

From now, we assume a representative sample, i.e., where the selection probability is equal to the relative probability of the occurrence of a given upset/workload pair.

Since fault-tolerant systems will, most of the time, correctly handle an upset, we will focus on the non-coverage as a measure ($\bar{c} = 1 - c$), and more specifically on the upper confidence limit of the non-coverage. Since the number of upsets not correctly handled follows a binomial distribution with parameters n (number of inserted upsets) and $\bar{c}$ (non-coverage), the 100γ upper confidence limit is given by:

- Frequentist theory:

$$\bar{c}_\gamma = \frac{(x+1)F_{2(x+1),2(n-x),\gamma}}{(n-x)+(x+1)F_{2(x+1),2(n-x),\gamma}}$$

- Bayesian theory:

$$\overline{c}_\gamma = \frac{(x+k)F_{2(x+k),2(n-x+l),\gamma}}{(n-x+l)+(x+k)F_{2(x+k),2(n-x+l),\gamma}}$$

where n is the number of inserted upsets, x is the number of upsets incorrectly handled, γ is the confidence level, and $F_{\upsilon 1,\upsilon 2,\gamma}$ is 100γ percentile points of an F distribution with $\upsilon 1, \upsilon 2$ degrees of freedom [Joh69, p.59]. Moreover, in the Bayesian case, we have selected a prior distribution following a beta distribution with parameters l and k: beta(k, l). The reasons for this choice are: first, that one special type of a beta distribution is the uniform distribution (when k=l=1), which effectively expresses the fact that there is *no* prior knowledge, and second, that since the experiment follows a binomial distribution, the posterior distribution will also be a binomial distribution with parameters x+k and n-x+l: beta(x+k, n-x+l).

Let us now compare the estimations obtained with the frequentist approach with the ones from the Bayesian. We will consider three prior distributions: the uniform distribution (k=l=1) where no assumption is made on the fault-tolerant system behavior, a distribution with a slight confidence that the fault tolerance system will behave correctly (k=1, l=3), and a distribution with a very high confidence that the system will correctly handle upsets (k=1, l=100). The density function ($f_Z(z)$) for these three prior distributions with a beta(1, 2) for completeness are shown on Figure 1.

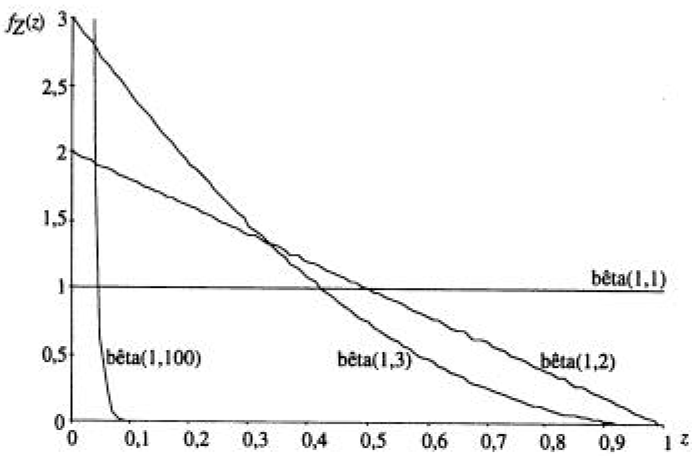

Figure 1 Prior distribution density functions

Since fault-tolerant systems will correctly handle most of the upsets, we will compare the frequentist and Bayesian estimations when 0 or 1 upset is not correctly handled. We compare these estimations when 10, 100 and 1000 upsets are inserted. The obtained estimations are shown in Table 1.

From the table, we observe that the confidence limit is smaller when using the Bayesian approach, leading to a smaller confidence interval and thus a more accurate estimation. When applying a prior distribution with high confidence in the correct upset handling, the obtained estimations become more accurate. The reason is that the small number of incorrectly handled upsets (0 or1) is in accordance with the prior knowledge that almost no upset will be incorrectly handled. An important observation is that for a small number of inserted upsets, an order of magnitude might be gained in the estimation accuracy when using the Bayesian approach and a strong confidence of upsets being correctly handled by the system.

Up-sets	Incor.	Frequentist	Bayesian		
N	X		k=1, l=1	k=1, l=3	k=1, l=100
10	0	0.369	0.342	0.298	0.0410
	1	0.504	0.470	0.413	0.0588
100	0	0.0450	0.0446	0.0437	0.0227
	1	0.0645	0.0639	0.0627	0.0327
1000	0	0.00459	0.00459	0.00458	0.00418
	1	0.00662	0.00661	0.00660	0.00602

Table 1 Comparison of frequentist and Bayesian estimations

5. Conclusion

We have shown in this paper that Bayesian theory might be worth considering when quantifying dependability attributes. We have illustrated the relevance of the Bayesian theory by presenting a simple example: the estimation of the coverage factor of a fault-tolerant system.

Acknowledgments

The authors would like to thank David Powell and Jean Arlat for their comments and for providing the example used in this paper.

References

[Cuk99] M. Cukier, D. Powell, J. Arlat, *Coverage Estimation Methods for Stratified Fault-Injection,* in IEEE Trans. On Computers, vol. 48, no. 7, pp.707-723, July 1999.
[Joh69] N.L. Johnson, S. Kotz, *Distributions in Statistics – Discrete Distributions*, New York, John Wiley & Sons, 1969.
[Kan91] K. Kanoun, J. Arlat, L. Burrill, Y. Crouzet, S. Graf, E. Martins, A. MacInnes, D. Powell, J.-L. Richier, J. Voiron, "Validation", in *Delta-4: A Generic Architecture for Dependable Distributed Computing*, (D. Powell, Ed.), pp. 371-406, Berlin, Germany, Springer-Verlag, 1991.
[Mad01] H. Madeira, P. Koopman, *Dependability Benchmarking: making choices in an n-dimensional problem space*, in First Workshop on Evaluating and Architecting System Dependability (EASY), joint organized with IEEE/ACM 28th International Symposium on Computer Architecture (ISCA) and the IEEE International Conference on Dependable Systems and Networks (DSN), Gothemburg, Sweden, July 2001.
[Sie92] D.P. Siewiorek, R.S. Swarz, *Reliable Computer Systems – Design and Evaluation*, Bedford, MA, USA, Digital Press, 1992.