Coverage Estimation Methods for Stratified Fault-Injection

Michel Cukier, *Member*, *IEEE*, David Powell, *Member*, *IEEE*, and Jean Arlat, *Member*, *IEEE*

Abstract—This paper addresses the problem of estimating fault tolerance coverage through statistical processing of observations collected in fault-injection experiments. In an earlier paper, various estimators based on simple sampling in the complete fault/activity input space and stratified sampling in a partitioned space were studied; frequentist confidence limits were derived based on a normal approximation. In this paper, the validity of this approximation is analyzed. The theory of confidence regions is introduced to estimate coverage without approximation when stratification is used. Three statistics are considered for defining confidence regions. It is shown that one—a vectorial statistic—is often more conservative than the other two. However, only the vectorial statistic is computationally tractable. We then consider Bayesian estimation methods for stratified sampling. Two methods are presented to obtain an approximation of the posterior distribution of the coverage by calculating its moments. The moments are then used to identify the type of the distribution in the Pearson distribution system, to estimate its parameters, and to obtain the coverage confidence limit. Three hypothetical example systems are used to compare the validity and the conservatism of the frequentist and Bayesian estimations.

Index Terms—Fault tolerance coverage, coverage estimation, fault-injection, stratified sampling, confidence regions, confidence limits, frequentist estimation, Bayesian estimation.

1 Introduction

Computer systems have to be designed to be fault-tolerant if they are to meet the stringent dependability requirements imposed by critical applications. The degree of dependability that can be achieved heavily relies on the efficiency of the fault tolerance mechanisms—the fault tolerance coverage—that is usually defined as *the probability of system recovery given that a fault exists* [4].

Fault tolerance mechanisms are commonly assessed by carrying out fault-injection experiments [3], [6], [7], [10], [14], [15], [18], [20]. A single fault-injection experiment consists of injecting a fault condition into a simulation or a prototype of a fault-tolerant system and observing the behavior of the system to determine whether or not the injected fault has been properly handled by the system's fault tolerance mechanisms. In practical systems, it is not possible to inject all fault conditions that could possibly occur during the system's operational life, so coverage evaluation is usually achieved by statistical estimation.

In this paper, we address the problem of estimating fault tolerance coverage by processing the observations collected in a set of fault injection experiments. Since the effect of a fault is dependent on system activity at the moment the fault is sensitized, we consider a sample space consisting of the combination of the fault space and the set of system "activities" at the moment of fault sensitization. We consider both frequentist and Bayesian estimation methods.

Many papers have been dedicated to frequentist estimation methods for sampling in the complete "fault/activity" sample space (simple sampling), but only a few papers have studied estimation methods when the sample space is partitioned into *classes* or *strata* (*stratified sampling*). These methods were first applied to fault-injection in [16].

Stratification offers several practical advantages. It can be used to reinterpret the results of a fault-injection campaign should it be necessary, for example, to use test data for coverage estimation or should new knowledge be acquired about the actual distribution of real faults. Stratified sampling allows, under certain conditions, for the composition of results of fault-injection experiments. Sometimes, stratification simplifies the fault-injection process, for example, for physical fault injection.

Wang et al. [21] used stratified sampling to introduce *equivalence classes*, defined as a set of faults that can be determined to have the same effect on system behavior. The aim of Wang et al. was to reduce the number of fault-injection experiments necessary to obtain a statistically significant estimate of coverage while only injecting one fault in each equivalence class. However, it was concluded that no benefit could be obtained from this approach in practical systems. In this paper, we do not assume that strata define equivalence classes. Here, the strata result simply from a "convenient" partition of the sample space (for example, each stratum may represent an individual hardware board of a fault-tolerant system).

In [16], several coverage estimation techniques based on simple sampling and on stratified sampling were considered. It was concluded that stratified sampling techniques were of particular interest. The coverage confidence limit estimations presented in that paper were based on a normal distribution approximation obtained when applying the central limit theorem. This approximation leads to the well-known result that the estimator variance reduction

- *M. Cukier was at LAAS-CNRS. He is now with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL 61801.*
  E-mail: cukier@crhc.uiuc.edu.
- *D. Powell and J. Arlat are with LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse Cedex 4, France. E-mail: {dpowell, arlat}@laas.fr.*

For information on obtaining reprints of this article, please send e-mail to:

0018-9340/99/$10.00 © 1999 IEEE

provided by stratification improves the precision of the estimation.

It is shown in this paper that the approximation leading to a normal distribution is not usually valid for coverage confidence limit estimations. First, most fault tolerance mechanisms are characterized by very high coverage factor values. Second, the number of faults injected during a campaign is often relatively low (in the order of thousands). Both aspects restrict the validity of the normal distribution approximation. The confidence region theory is therefore applied for stratified sampling. Confidence regions obtained by two point estimators and a vectorial statistic are introduced. For more than three strata, the coverage confidence limit estimations using the point estimators become computationally intractable so, most of the time, only the vectorial statistic can be applied. When applying the confidence region theory, the result that stratification improves the precision of the estimation is no longer correct. In fact, with the vectorial statistic, stratification leads to very conservative upper confidence limits on noncoverage.

We then investigate whether coverage estimation methods from the Bayesian School can be used with stratified sampling to obtain less conservative, yet tractable, upper confidence limits. In Bayesian estimation, the coverage confidence limit depends directly on the overall noncoverage posterior distribution. Since the analytical expression of this posterior distribution can be easily obtained only for partitions into a small number of classes, we will consider methods giving an approximation of the posterior distribution for any number of classes. Two methods are presented to obtain the first moments of the distribution. The first method introduces independence assumptions between the coverages of the various classes. The second method is based on the *moment generating function* of the posterior distribution. The moment values are then used to identify, by using the Pearson distribution system, the type of the posterior distribution, to estimate its parameters, and, finally, to obtain the coverage confidence limit.

The most appropriate way to assess the validity and the precision of an estimation method is to apply it to experimental data obtained from a simulation of a system with known characteristics. First, simulation allows the generation of data sets that push an estimation method toward its limits. Second, it is possible to compare the estimates obtained from the simulated data with the parameters of the distributions used to obtain the data. We follow this approach here to compare the various frequentist and Bayesian estimation techniques.

The remainder of the paper is organized as follows. Section 2 reiterates some notions from [16] concerning basic definitions and the two categories of considered sampling techniques: simple sampling and stratified sampling. Section 3 then recapitulates the frequentist coverage confidence limit estimators obtained with stratified sampling when applying the normal distribution approximation. The problem of exact sampling distributions for stratified sampling is then investigated using the theory of confidence regions. Section 4 is dedicated to the Bayesian estimation methods and describes, in particular, how to obtain an approximation of the posterior distribution of the overall noncoverage by its moments and how to identify the type of the approximated distribution in the Pearson distribution system. Section 5 compares the validity and the precision of the frequentist and Bayesian estimations for three hypothetical examples with both stratified and simple sampling. Finally, Section 6 concludes the paper. A set of appendices groups most of the mathematical developments that are necessary for an in-depth study of the estimation methods. When the background mathematics can be found in classic textbooks, an appropriate reference is given.

2 Definitions

The definitions presented in this section recapitulate the concepts introduced in [16].

2.1 Coverage factor

The effect of a given fault is dependent on system activity at the moment of, and following, the sensitization of the fault, so the complete input space of a fault-tolerance mechanism consists not only of the considered *fault space F* but also the *activity space A* due to the system's functional inputs (a single "activity" is a trajectory in the system's state space). We can thus formally define the *coverage factor* of a fault tolerance mechanism in terms of the complete input space defined as the Cartesian product $G = F \times A$. Let H be a variable characterizing the handling of a particular fault, such that $H = 1$ if the mechanism correctly handles the fault and $H = 0$ otherwise. The coverage factor is defined as:

$$c = \Pr\{H = 1 \mid g \in G\}, \tag{1}$$

i.e., the conditional probability of correct fault handling, given the occurrence of a fault/activity pair $g \in G$.

H is a discrete random variable that can take the values 0 or 1 for each element of the fault/activity space G. Let $h(g)$ denote the value of H for a given point $g \in G$ and let $p(g)$ be the *relative probability of occurrence* of g. Expression (1) can then be rewritten as:

$$c = \sum_{g \in G} h(g)p(g). \tag{2}$$

The coverage factor c can be viewed as $E\{H\}$, the expected value of H. It should be stressed that the distribution $p(g)$ is an inherent part of the very definition of coverage as a conditional probability parameter. Without knowledge about $p(g)$, one cannot make any meaningful statements about the coverage factor of a system. This point was also clearly stressed in [17] when analyzing the results of fault injection experiments on a system decomposed into subsystems. The best that can be done without knowing the distribution $p(g)$ is to use the *coverage proportion*, $\tilde{c} = \frac{1}{|G|}\sum_{g \in G} h(g)$, to describe the effectiveness of a given fault tolerance mechanism. However, such a measure of effectiveness can only be used as a branching probability parameter for predicting system dependability (e.g., in stochastic Markov chain models) if it can be assumed that all fault/activity pairs in G are equally probable, i.e., $p(g) = 1/|G|$.

The most accurate way to determine c would be to submit the system to all $g \in G$ and to observe all outcomes $h(g)$. However, such exhaustive testing is rarely possible. For this reason, coverage evaluation is, in practice, carried out by submitting the system to a subset of fault/activity occurrences $G^* \subset G$ obtained by random sampling in the space G and then using statistics to estimate c.

In practice, random sampling in the combined fault/activity space G is decomposed into two concurrent sampling processes: a fault in space F and an activity in space A. Whereas the fault space sampling process is explicit, the activity sampling process is achieved implicitly: The target system executes its operational workload and the selected fault is injected asynchronously, at some random point in the workload execution [16].

We now recall the theory of coverage factor estimation by sampling first in a nonpartitioned sample space and then in a partitioned sample space.

2.2 Sampling in a Nonpartitioned Space

We consider sampling with replacement of a subset G^* of n fault/activity pairs in G. To each element of G is assigned a *selection probability*, $t(g)$, such that $\forall g \in G,\ t(g) > 0$ and $\sum_{g \in G} t(g) = 1$. Note that the experiments are Bernoulli trials with outcome $H = 1$ with probability θ and $H = 0$ with probability $1 - \theta$ where $\theta = \sum_{g \in G} h(g)t(g)$. By setting $t(g)$ equal to:

- $1/|G|$, we obtain a *uniform sample* and $\theta = \tilde{c}$, the coverage *proportion*,
- $p(g)$, we obtain a *representative sample* and $\theta = c$, the coverage *factor*.

Let Γ_i denote the ith fault/activity pair added to the sample. Let $h(\Gamma_i)$ be the random variable representing the outcome of the experiment for Γ_i (described by the predicate H). An unbiased point estimator for c is given by [16]:

$$\hat{c}(\Gamma_1, \ldots, \Gamma_n) = \frac{1}{n}\sum_{i=1}^{n} h(\Gamma_i)\frac{p(\Gamma_i)}{t(\Gamma_i)}, \tag{3}$$

where $p(\Gamma_i)$ and $t(\Gamma_i)$, respectively, denote the random variables giving the occurrence probability and the selection probability corresponding to Γ_i.

With representative sampling, we have $\forall i \in \{1..n\}$, $p(\Gamma_i) = t(\Gamma_i)$, and (3) may be rewritten as:

$$\hat{c}(\Gamma_1, \ldots, \Gamma_n) = \frac{1}{n}\sum_{i=1}^{n} h(\Gamma_i).$$

Equivalently, an unbiased estimator of the system noncoverage $\bar{c}\ (= 1 - c)$ is given by:

$$\hat{\bar{c}}(X) = \frac{X}{n}, \tag{4}$$

where $X = n - \sum_{i=1}^{n} h(\Gamma_i)$ is the number of fault tolerance deficiencies observed for n injected fault/activity pairs. In the sequel, we refer to the random variable X as the *deficiency number* and an observation of X is noted x.

2.3 Sampling in a Partitioned Space

For the sampling techniques that follow, the sample space G is considered as partitioned into M classes or strata:

$$G = \bigcup_{i=1}^{M} G_i \text{ such that } \forall i, j, i \neq j, G_i \cap G_j = \emptyset.$$

We can rewrite the coverage factor definition (2) as follows:

$$\begin{aligned} c &= \sum_{i=1}^{M}\sum_{g \in G_i} h(g)p(g) \\ &= \sum_{i=1}^{M} p(g \in G_i)\sum_{g \in G_i} h(g)p(g \mid g \in G_i) = \sum_{i=1}^{M} p_i c_i, \end{aligned}$$

where $p(g \in G_i) \equiv p_i$ is the relative probability of fault/activity occurrences in class G_i and c_i is the corresponding coverage factor:

$$c_i = \sum_{g \in G_i} h(g)p(g \mid g \in G_i).$$

A target system can then be characterized by:

- $\mathbf{p} = [p_1, \ldots, p_M]$, the vector of fault/activity occurrence probabilities in each class of the fault/activity space G,
- $\bar{\mathbf{c}} = [\bar{c}_1, \ldots, \bar{c}_M]$, the (unknown) vector of class noncoverages,

such that the overall noncoverage is given by the scalar product:

$$\bar{c} = \sum_{i=1}^{M} p_i \bar{c}_i = \mathbf{p}.\bar{\mathbf{c}}^{\mathbf{T}}.$$

In a stratified fault-injection campaign, a fixed number of experiments are carried out in each class (using representative sampling within each class, cf. Section 2.2). The random variables characterizing the deficiency number X_i for each class together constitute a random *deficiency vector* $\mathbf{X}$. A stratified fault-injection campaign is therefore defined by:

- $\mathbf{n} = [n_1, \ldots, n_M]$, the number of fault/activity pairs injected per class;
- $\mathbf{x} = [x_1, \ldots, x_M]$, the observed value of the deficiency vector $\mathbf{X}$.

We consider two stratified allocations of n experiments corresponding, respectively, to a *representative* and a *homogeneous* allocation:

$$\mathbf{n}_R = [p_1 n, \ldots, p_M n] \qquad \mathbf{n}_H = \left[\frac{n}{M}, \ldots, \frac{n}{M}\right].$$

With a representative allocation, the number of fault/activity pairs injected in each class is proportional to the fault/activity occurrence probability in that class, whereas a homogeneous allocation results in the injection of an identical number of fault/activity pairs in each class.

Sections 3 and 4 are dedicated to the frequentist and Bayesian confidence limit estimations with stratified sampling.

3 Frequentist Confidence Limits

In frequentist theory, the upper 100γ percent confidence limit estimator for $\bar{c}$ is defined by:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{X}) : \Pr\Big[\bar{c} \leq \bar{c}^{\uparrow}_{\gamma}(\mathbf{X}) \mid \bar{c}\Big] = \gamma, \tag{5}$$

i.e., if a large number of individual values $\mathbf{x}$ of the deficiency vector $\mathbf{X}$ are considered, $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ will be greater than the unknown value $\bar{c}$, 100γ percent of the time.

We first consider two estimators defined, respectively, by the *arithmetic* and *weighted* averages of the elements of the deficiency vector $\mathbf{X}$:

$$Y_A(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{M} X_i \qquad Y_W(\mathbf{X}) = \sum_{i=1}^{M} \frac{p_i}{n_i} X_i.$$

When each statistic is considered with each allocation, three point estimation techniques can be considered:

$$\begin{aligned}
&Y_A(\mathbf{X}) \text{ or } Y_W(\mathbf{X}) \text{ with } \mathbf{n_R} \Rightarrow && \hat{\hat{c}}_{*R}(\mathbf{X}) = \tfrac{1}{n}\textstyle\sum_{i=1}^{M} X_i = \tfrac{X}{n}\\
&Y_W(\mathbf{X}) \text{ with } \mathbf{n}_H \Rightarrow && \hat{\hat{c}}_{WH}(\mathbf{X}) = \tfrac{M}{n}\textstyle\sum_{i=1}^{M} p_i X_i\\
&Y_A(\mathbf{X}) \text{ with } \mathbf{n}_H \Rightarrow && \hat{\hat{c}}_{AH}(\mathbf{X}) = \tfrac{1}{n}\textstyle\sum_{i=1}^{M} X_i = \tfrac{X}{n}.
\end{aligned} \tag{6}$$

The latter technique is called naive stratification since the estimate it provides is biased (its expected value is, in fact, the average class noncoverage). By extension, we call $Y_A(\mathbf{X})$ the naive estimator and $Y_W(\mathbf{X})$ the stratified estimator since only the latter includes the fault/activity occurrence distribution and is therefore always unbiased.

In [16], it is shown that the variance of the stratified estimator with a representative allocation, $\hat{\hat{c}}_{*R}(\mathbf{X})$, is never greater than that of the estimator for simple sampling, $\hat{\bar{c}}(X)$.

3.1 Confidence Limits Using Approximations

A classic way of defining a frequentist confidence limit is to assume that, due to the central limit theorem, the noncoverage estimators are assumed to be normally distributed around their expected values, which leads to the well-known form for the confidence limit estimator:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{X}) = \hat{\hat{c}}_{@}(\mathbf{X}) + z_{\gamma}\sqrt{V\{\hat{\hat{c}}_{@}(\mathbf{X})\}}, \tag{7}$$

where @ is $^{*}R$, WH, or AH, as defined in (6), $V\{\hat{\hat{c}}_{@}(X)\}$ is the variance of the noncoverage estimator and z_{γ} is the 100γth standard normal percentile.

For the point estimation technique $\hat{\hat{c}}_{*R}(\mathbf{X})$, by replacing $V\{\hat{\hat{c}}_{@}(X)\}$ by its corresponding estimator [16], (7) becomes:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{X}) = \sum_{i=1}^{M} X_i + z_{\gamma}\sqrt{\sum_{i=1}^{M}\left(\frac{p_i}{n_i}\right)^2\left(\frac{X_i(n_i - X_i)}{n_i - 1}\right)}. \tag{8}$$

3.2 Confidence Limits without Approximation

It can be easily demonstrated that, even for modest coverage values, the previous confidence limits based on the normal approximation are only valid for large total sample sizes. Moreover, when coverage is very high, it is very likely that one or more classes will not reveal any deficiencies, i.e., for some classes we will have $x_i = 0$. Expression (8) would appear to imply that such classes do not contribute *any* uncertainty to the overall estimate. This is clearly not a very satisfactory implication. These problems can be avoided if we use, for both sampling strategies, the true sampling distribution instead of an approximation. For a partitioned sample space, we therefore need to apply the theory of confidence regions. We apply the theory first to estimators which are positive functions of $\mathbf{X}$ and then to a special vectorial statistic. The well-known result for simple sampling will also be recapitulated as a special degenerate case of stratified sampling in which there is only one class.

A *confidence region* for the noncoverage parameter vector $\bar{\mathbf{c}} = [\bar{c}_1, \ldots, \bar{c}_M]$ is a function $I_{\gamma}(\mathbf{X})$ of the deficiency vector $\mathbf{X}$ such that for any given value of the vector $\bar{\mathbf{c}} : \Pr[I_{\gamma}(\mathbf{X}) \ni \bar{\mathbf{c}} \mid \bar{\mathbf{c}}] = \gamma$, i.e., if a large number of individual values $\mathbf{x}$ of the random vector $\mathbf{X}$ are considered, $I_{\gamma}(\mathbf{x})$ will contain the unknown value of the parameter vector $\bar{\mathbf{c}}$ 100γ percent of the time. The frontier values of a confidence region are *confidence limits*.

Since the *overall* noncoverage is defined by the scalar product $\bar{c} = \mathbf{p}.\bar{\mathbf{c}}^T$, an upper 100γ percent confidence limit estimator on $\bar{c}$ is given by:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{X}) = \max_{\bar{\mathbf{c}} \in I_{\gamma}(\mathbf{X})}\big(\mathbf{p}.\bar{\mathbf{c}}^T\big). \tag{9}$$

Appropriate functions $I_{\gamma}(\mathbf{X})$ can be defined in a number of different ways. Here, we consider: 1) the use of point estimators that are positive functions of $\mathbf{X}$, and 2) a vectorial statistic. The corresponding mathematics is given in Appendices A and B. In this section, we only present the resulting expressions.

3.2.1 *Point Estimators*

The point estimators $Y_A(\mathbf{X})$ and $Y_W(\mathbf{X})$ are positive functions of the X_i. For any such positive function of the X_i, noted $Y(\mathbf{X})$, the upper 100γ percent confidence limit $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ can be expressed as the following maximization problem (see Appendix A):

Maximization of $\bar{c} = \mathbf{p}.\bar{\mathbf{c}}^T$ under the constraints:

- given by the confidence region frontier $\bar{\mathbf{c}}_{\gamma}(Y(\mathbf{x}))$ for $\bar{\mathbf{c}}$:

$$\bar{\mathbf{c}} : \sum_{\mathbf{x}' : Y(\mathbf{x}') \leq Y(\mathbf{x})} \prod_{i=1}^{M} \binom{n_i}{x'_i} \bar{c}_i^{\,x'_i} (1 - \bar{c}_i)^{n_i - x'_i} = 1 - \gamma$$

- given by the limits of the parameter space:

$$\forall i \in \{1..M\}, \bar{c}_i \in [0, 1].$$

3.2.2 *Vectorial Statistic*

We now consider an alternative formulation in which each element of the deficiency vector is used individually to define an upper confidence limit on the noncoverage of its corresponding class.

For each class, the confidence limit is estimated based on a confidence level such that the product of the confidence levels of the different classes is equal to the global confidence level γ. Since we will obtain a confidence interval for each class, the confidence region will be a hypercube in the parameter space. The overall confidence limit is defined by the apex of the hypercube. The least conservative upper 100γ percent confidence limit $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ is given by the combination of the confidence levels of the different classes such that the corresponding apex is the closest to the origin. The vectorial statistic leads thus to the following minimization problem (see Appendix B):

Minimization of $\bar{c} = \mathbf{p}.\bar{\mathbf{c}}^T$ under the constraints:

- given by the global confidence:

$$\bar{\mathbf{c}} : \prod_{i=1}^{M} \left(1 - \sum_{x'_i=0}^{x_i} \binom{n_i}{x'_i} \bar{c}_i^{x'_i} \left(1 - \bar{c}_i\right)^{n_i - x'_i} \right) = \gamma$$

- given by the limits of the parameter space:

$$\forall i \in \{1..M\}, \bar{c}_i \in [0, 1].$$

3.3 Discussion

In case of simple (nonstratified) sampling, the expression obtained for the upper 100γ percent confidence limit $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ for the estimators which are positive functions of $\mathbf{X}$ and for the vectorial statistic are identical. Expression (5) can be rewritten in its well-known form:

$$\bar{c}^{\uparrow}_{\gamma}(X) : \sum_{j=0}^{x} \binom{n}{j} \left(\bar{c}^{\uparrow}_{\gamma}(X)\right)^{j} \left(1 - \bar{c}^{\uparrow}_{\gamma}(X)\right)^{n-j} = 1 - \gamma$$

This equation can be solved analytically. By introducing 100γ percent percentile points, $F_{\nu_1,\nu_2,\gamma}$, of an F distribution with ν_1, ν_2 degrees of freedom [11, p. 59], we obtain:

$$\bar{c}^{\uparrow}_{\gamma}(X) = \frac{(X+1)F_{2(X+1),2(n-X),\gamma}}{(n-X) + (X+1)F_{2(X+1),2(n-X),\gamma}}. \tag{10}$$

A general demonstration concerning the comparison of the estimations obtained from both estimators and the vectorial statistic is beyond the scope of this paper. However, some interesting insights can be gained by analyzing an example (see [8]):

- The limit given by $Y_W(\mathbf{X})$ is less conservative than, or identical to, that given by $Y_A(\mathbf{X})$,
- The least conservative confidence limit may be given by the vectorial statistic *or* by $Y_w(\mathbf{X})$, depending on the fault/activity occurrence distribution $p(g)$.

It has been shown in [8] that the vectorial statistic provides a numerical solution for a large range of fault-injection experiments, which is not the case for the estimators $Y_A(\mathbf{X})$ and $Y_W(\mathbf{X})$. The computational time necessary to solve the maximization problem for the point estimators is also higher than that for the vectorial statistic (solved numerically by using the `e04vcf` routine of the NAG library) since the constraint must be approximated. Therefore, the comparisons given later in the paper only consider the vectorial statistic.

4 Bayesian Confidence Limits

In the Bayesian theory, the overall noncoverage $\bar{c}$ and the class noncoverages $\bar{c}_i$ are no longer considered as parameters but as random variables. Accordingly, they will be noted, respectively $\bar{C}$ and $\bar{C}_i$.

When applying the Bayesian theory, an upper 100γ percent confidence limit is directly defined by the distribution of the random variable $\bar{C}$:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{x}) : \Pr\left[\bar{C} \leq \bar{c}^{\uparrow}_{\gamma}(\mathbf{X}) \mid \mathbf{X} = \mathbf{x}\right] = \gamma, \tag{11}$$

i.e., the value $\bar{c}_{\gamma}(x)$ such that, for the observed value $\mathbf{x}$ of the deficiency vector $\mathbf{X}$, the probability that $\bar{C}$ is less than or equal to $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ is equal to γ.

4.1 Outline of the Method

To calculate the confidence limit defined by (11), we must first obtain the posterior distribution of $\bar{C}$, noted $f_{\bar{C}}(\bar{c}|\mathbf{X} = \mathbf{x})$. With stratified sampling, we seek to obtain this distribution from the posterior distributions of the class noncoverages $\bar{C}_i$, noted $f_{\bar{C}_i}(\bar{c}_i|X_i = x_i)$. The posterior distributions combine the prior distributions, which contain the belief about the distribution of $\bar{C}_i$ and the distributions associated to the experiments (in our case, a binomial distribution) for the observed number of deficiencies x_i. It is of course essential to choose the priori distributions for the class noncoverages $\bar{C}_i$ very carefully.

We advocate the use of a beta distribution as a prior distribution for the $\bar{C}_i$ for several reasons. First, since the deficiency number of each class X_i has a binomial distribution, a beta prior distribution belongs to a family of conjugate prior distributions, which ensures that the prior and the posterior distributions are both from the same family. The posterior distribution will then also have an analytical expression. Second, with both its parameters equal to one, the beta distribution corresponds to a uniform distribution over the interval $[0, 1]$, which means that all values $\bar{c}_i$ have the same weight (for this reason, such a prior distribution is sometimes called an "ignorance" prior).

Since $\bar{C} = \sum_{i=1}^{M} p_i \bar{C}_i$, we can look for the posterior distribution of $\bar{C}$ by combining the posterior distributions for the various $\bar{C}_i$. It has been shown in [8] that obtaining an analytical expression of the posterior distribution for $\bar{C}$ is far too complicated for partitions into more than about three classes. Therefore, in general, the posterior distribution for $\bar{C}$ can only be obtained by applying approximate methods.

When a distribution cannot be calculated exactly, then, for most ordinary purposes, knowledge of the moments is equivalent to knowledge of the distribution function, in the sense that it should be possible *theoretically* to exhibit all the properties of the distribution in terms of the moments [19, pp. 108-109]. It is further stressed in [19, pp. 109-110] that distributions that have a finite number of the lower moments in common will, in a sense, be approximations of one another. In practice, approximations of this kind often turn out to be remarkably good, even when only the first two, three, or four moments are equated. Section 4.2 is dedicated to the calculation of the first four moments of the posterior distribution for $\bar{C}$.

The Pearson distribution system will be used to identify the type of the posterior distribution for $\bar{C}$ based on the value of its moments. When the type of the posterior distribution is identified, the upper 100γ percent confidence limit can be obtained by applying (11) for the considered posterior distribution. Section 4.3 presents the estimation method based on the Pearson distribution system and the implementation issues.

4.2 Moment Calculation

This section is dedicated to two different methods for calculating the first four moments of the posterior distribution for $\bar{C}$. The first method introduces independence assumptions between the coverages of the various classes. The second method is based on the moment generating function of $\bar{C}$.

4.2.1 Independence Assumptions

The first moment of $\bar{C}$ is directly obtained by a linear combination of the first moments of the $\bar{C}_i$. By making assumptions on powers of the noncoverages $\bar{C}_i$, we can obtain linear expressions between the second, third, and fourth central moments of $\bar{C}$ and, respectively, the second, third, and fourth central moments of the $\bar{C}_i$. The latter can be easily obtained using: 1) the well-known relationship between central moments and moments about zero, and 2) the fact that the $\bar{C}_i$ have beta distributions.

Given that the deficiency number of each class X_i has a binomial distribution with parameters n_i and $\bar{c}_i$, it follows from [2, p. 28] that if the prior distribution for $\bar{C}_i$ is a beta distribution with parameters k_i and l_i, then the posterior distribution for $\bar{C}_i$ is a beta distribution with parameters $k'_i = x_i + k_i$ and $l'_i = n_i - x_i + l_i$. To minimize the subjectivity of the Bayesian approach, we assume a uniform prior distribution for the $\bar{C}_i$. The posterior distributions for the $\bar{C}_i$ are then beta distributions with parameters k'_i and l'_i, where:

$$\left.\begin{array}{l} k'_i = x_i + 1 \\ l'_i = n_i - x_i + 1 \end{array}\right\} \tag{12}$$

The rth moment μ'_{r_i} $(r = 2, 3, 4)$ of $\bar{C}_i$ is thus given by:

$$\mu'_{r_i} = \frac{B(k'_i + r, l'_i)}{B(k'_i, l'_i)} = \frac{k'_i(k'_i + 1)\dots(k'_i + r - 1)}{(k'_i + l'_i)(k'_i + l'_i + 1)\dots(k'_i + l'_i + r - 1)},$$

where $B(k'_i, l'_i)$ is a Beta function with parameters k'_i and l'_i [1, p. 258].

For each class, the three central moments μ_{r_i} $(r = 2, 3, 4)$ can be obtained from the moments μ'_{r_i} by the expressions [11, p. 18, eq. 100]:

$$\begin{aligned} \mu_{2_i} &= \mu'_{2_i} - \mu'^2_{1_i}; \\ \mu_{3_i} &= \mu'_{3_i} - 3\mu'_{2_i}\mu'_{1_i} + 2\mu'^3_{1_i}; \\ \mu_{4_i} &= \mu'_{4_i} - 4\mu'_{3_i}\mu'_{1_i} + 6\mu'_{2_i}\mu'^2_{1_i} - 3\mu'^4_{1_i}. \end{aligned} \tag{13}$$

It can be shown that the second, third, and fourth central moments of a weighted sum of independent random variables Z_i are equal to a weighted sum of the central moments of these independent random variables if:

$$\begin{array}{l} \forall i, j \text{ and } i \neq j : Z_i \text{ and } Z_j;\ Z_i^2 \text{ and } Z_j; Z_j^2 \text{ and } Z_j^2; \\ Z_i^3 \text{ and } Z_j \text{ are independent.} \end{array} \tag{14}$$

Applied to the overall noncoverage $\bar{C} = \sum_{i=1}^{M} p_i\bar{C}_i$, if μ_{r_i} is the rth central moment of $\bar{C}_i$ and if (14) is satisfied for the $\bar{C}_i$ taken pair by pair, then the second, third, and fourth central moments of the overall noncoverage $\bar{C}$ can be simply expressed in function of central moments of the noncoverages $\bar{C}_i$:

$$\mu_2 = \sum_{i=1}^{M} p_i^2 \mu_{2_i}; \quad \mu_3 = \sum_{i=1}^{M} p_i^3 \mu_{3_i}; \quad \mu_4 = \sum_{i=1}^{M} p_i^4 \mu_{4_i}.$$

4.2.2 Moment Generating Function

We develop here another technique for calculating the moments of $\bar{C}$ which, unlike that of the previous section, does not rely on a simplifying hypothesis.

Since the moment generating functions of the $\bar{C}_i$ are well-known for beta distributions, the moment generating function of $\bar{C}$ can easily be obtained by combining the moment generating functions of the $\bar{C}_i$. The derivatives of the moment generating function of $\bar{C}$ define the moments (about zero) of $\bar{C}$. The second, third, and fourth central moments of $\bar{C}$ are then obtained by applying the classic expressions that relate the central moments to the moments about zero.

As in the previous section, we assume a uniform prior distribution for the $\bar{C}_i$, so the posterior distributions of the $\bar{C}_i$ are beta distributions with parameters k'_i and l'_i given by (12).

Now, the moment generating function [11, p. 20] of a random variable which has a beta distribution can be expressed as a confluent hypergeometric function [1, chapter 13], [13, p. 40]. So, the moment generating function of $\bar{C}_i$ is given by:

$$\phi_{\bar{C}_i}(t) = {}_1F_1(k'_i; k'_i + l'_i; t), \tag{15}$$

where ${}_1F_1(k'_i; k'_i + l'_i; t)$ is the confluent hypergeometric function.

The moment generating function of $p_i\bar{C}_i$ is equal to: $\phi_{p_i\bar{C}_i}(t) = {}_1F_1(k'_i; k'_i + l'_i; p_i t)$. Since the moment generating function of a sum of random variables is equal to the product of the moment generating functions of the various random variables [11, p. 21], the moment generating function of $\bar{C} = \sum_{i=1}^{M} p_i\bar{C}_i$ is given by:

$$\phi_{\bar{C}}(t) = \prod_{i=1}^{M} {}_1F_1(k'_i; k'_i + l'_i; p_i t). \tag{16}$$

Since the rth derivative of the moment generating function of $\bar{C}$ for $t = 0$ defines the rth moment μ'_r of $\bar{C}$, the various moments can easily be obtained.

The first derivative of $\phi_{\bar{C}}(t)$ for $t = 0$ leads to:

$$\mu_1' = \frac{d\phi_{\bar{C}}(t)}{dt}\,|_{t=0}$$
$$= \left(\sum_{i=1}^{M} \frac{d_1F_1(k_i'; k_i' + l_i'; p_i t)}{dt} \prod_{\substack{j=1 \\ j \neq i}}^{M} {}_1F_1(k_j'; k_j' + l_j'; p_j t \right) \Bigg|_{t=0}.$$

By the properties of the confluent hypergeometric function [1, pp. 504-515], we have:

$$\begin{aligned} {}_1F_1(a; b; 0) &= 1 \\ \frac{d_1F_1(a; b; z)}{dz} &= \frac{a}{b} {}_1F_1(a+1; b+1; z). \end{aligned}$$

So, μ_1' becomes finally:

$$\mu_1' = \sum_{i=1}^{M} \frac{p_i k_i'}{k_i' + l_i'}.$$

The expressions of the first four moments are presented in Fig. 1.

Using the same relationships as in (13), but this time applied to the moments of $\bar{C}$, we obtain:

$$\begin{aligned} &\mu_2 = \mu_2' - \mu_1'^2; \\ &\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3; \mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4. \end{aligned} \tag{17}$$

4.3 Confidence Limit Calculation

The confidence limit estimation $\bar{c}_\gamma^{\uparrow}(\mathbf{x})$ corresponds, for the Bayesian theory, to the inverse function of the posterior distribution of $\bar{C}$ (11). For some classic distributions, the expression of the confidence limit estimation can be found in statistic textbooks. Otherwise, this estimation is obtained numerically either by using a routine from a mathematical library (for the classic distributions) or by successive approximations (for the less common distributions).

The moments obtained by either of the methods of Sections 4.1 and 4.2 enable the identification of an approximate posterior distribution of $\bar{C}$ from which the confidence limit can be calculated. In this paper, we identify a distribution from the Pearson system of distributions.

The distributions of the Pearson system all have a probability density function $f(z)$ that satisfies a differential equation of the form [12, p. 9]:

$$\frac{1}{f}\frac{df}{dz} = -\frac{a + z}{b_0 + b_1 z + b_2 z^2}. \tag{18}$$

The shape of the distribution depends on the values of the four parameters a, b_0, b_1, and b_2. These parameters are related to the four first moments of distribution $f(z)$ by the expressions [12, p. 13]:

$$\left.\begin{aligned} b_0 &= \tfrac{\mu_2(4\mu_2\mu_4 - 3\mu_3^2)}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3} \\ a = b_1 &= \tfrac{\mu_3(\mu_4 + 3\mu_2^2)}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3} \\ b_2 &= \tfrac{2\mu_2\mu_4 - 3\mu_3^2 - 6\mu_2^3}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3} \end{aligned}\right\} \tag{19}$$

The Pearson system contains seven distributions that can be represented in a plane (β_1, β_2), where β_1 is the *skewness* coefficient ($\beta_1 = \frac{\mu_3^2}{\mu_2^3}$) and β_2 is the *kurtosis* coefficient ($\beta_2 = \frac{\mu_4}{\mu_2^2}$) (Fig. 2).

The seven types of distribution include:

- the beta distribution (type I),
- the gamma distribution (type III),
- the inverse Gaussian (Wald) distribution (type V),
- the t distribution (type VII).

Bowman and Shenton [5] introduced a rational fraction approximation for any percentile y_γ $(F(y_\gamma) = \gamma)$ of a standardized distribution $(\mu_1 = 0, \mu_2 = 1)$ of the Pearson system. This approximation uses a 19-point formula:

$$y_\gamma = \frac{\pi_{\gamma,1}\big(\sqrt{\beta_1}, \beta_2\big)}{\pi_{\gamma,2}\big(\sqrt{\beta_1}, \beta_2\big)}, \tag{20}$$

where for $i = 1, 2$:

$$\pi_{\gamma,i}\Big(\sqrt{\beta_1}, \beta_2\Big) = \sum_{0 \le r+s \le 3} \sum a_{\gamma,r,s}^{(i)} \Big(\sqrt{\beta_1}\Big)^r \beta_2^s$$

with $a_{\gamma,0,0}^{(2)} = 1$. The values for a 99 percent level are presented in Fig. 3 (values for other levels are given in [5]). The approximation is valid for $0 \le \beta_1 \le 4$. In the case of a 99 percent level, the numerical error due to the approximation is less than 0.4 percent. When $\mu_3 \ge 0$, the γ-percentile of the nonstandardized distribution is given by: $z_\gamma = \mu_1 + \sqrt{\mu_2} y_\gamma$. When $\mu_3 < 0$, $z_\gamma = \mu_1 - \sqrt{\mu_2} y_{(1-\gamma)}$ [9]. The confidence limit estimation $\bar{c}_\gamma^{\uparrow}(\mathbf{x})$ is thus given by:

$$\begin{cases} \bar{c}_\gamma^{\uparrow}(\mathbf{x}) = \mu_1 + \sqrt{\mu_2} y_\gamma, & \mu_3 \ge 0 \\ \bar{c}_\gamma^{\uparrow}(\mathbf{x}) = \mu_1 - \sqrt{\mu_2} y_{(1-\gamma)}, & \mu_3 < 0. \end{cases} \tag{21}$$

5 Results and Comparisons

To assess the validity and the precision of an estimation method, it is necessary to apply it to data sets obtained from a system with known characteristics. The most appropriate way to obtain such controlled data sets is to use simulation. In this way, it is possible to compare the estimates obtained from the simulated data with the parameters of the distributions used to obtain the data. Therefore, to compare the Bayesian estimates with the frequentist ones, we will use three hypothetical systems whose characteristics are purposely chosen to aggressively test the developed statistical methods. Finally, these estimations are compared to those obtained when simple sampling is used.

5.1 Definition of the Systems

The characteristics of the three example systems are defined in Figs. 4a, 4b, and 4c. Each system is partitioned into $M = 50$ classes. The figures show the noncoverage $\bar{c}_i$ and the fault/activity occurrence probability p_i of each class, as well as the values of the system noncoverage, $\bar{c}$, and the *mean* class noncoverage, $\tilde{\bar{c}} = \frac{1}{M}\sum_{i=1}^{M} \bar{c}_i$. Each system is also characterized by the correlation factor ρ between the p_i and $\bar{c}_i$ [16]. The characteristics of the three example systems are quite different and have been chosen to *stress* the

$$\mu_1' = \sum_{i=1}^{M} \frac{p_i k_i'}{k_i' + l_i'}$$

$$\mu_2' = \sum_{i=1}^{M} \frac{p_i^2 k_i'(k_i'+1)}{(k_i'+l_i')(k_i'+l_i'+1)} + \sum_{i=1}^{M} \frac{p_i k_i'}{k_i'+l_i'} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j k_j'}{k_j'+l_j'}$$

$$\mu_3' = \sum_{i=1}^{M} \frac{p_i^3 k_i'(k_i'+1)(k_i'+2)}{(k_i'+l_i')(k_i'+l_i'+1)(k_i'+l_i'+2)} + 3\sum_{i=1}^{M} \frac{p_i^2 k_i'(k_i'+1)}{(k_i'+l_i')(k_i'+l_i'+1)} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j k_j'}{k_j'+l_j'}$$
$$+ \sum_{i=1}^{M} \frac{p_i k_i'}{k_i'+l_i'} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j k_j'}{k_j'+l_j'} \sum_{\substack{u=1\\ u\neq j\\ u\neq i}}^{M} \frac{p_u k_u'}{k_u'+l_u'}$$

$$\mu_4' = \sum_{i=1}^{M} \frac{p_i^4 k_i'(k_i'+1)(k_i'+2)(k_i'+3)}{(k_i'+l_i')(k_i'+l_i'+1)(k_i'+l_i'+2)(k_i'+l_i'+3)} + 4\sum_{i=1}^{M} \frac{p_i^3 k_i'(k_i'+1)(k_i'+2)}{(k_i'+l_i')(k_i'+l_i'+1)(k_i'+l_i'+2)} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j k_j'}{k_j'+l_j'}$$
$$+ 3\sum_{i=1}^{M} \frac{p_i^2 k_i'(k_i'+1)}{(k_i'+l_i')(k_i'+l_i'+1)} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j^2 k_j'(k_j'+1)}{(k_j'+l_j')(k_j'+l_j'+1)} + 6\sum_{i=1}^{M} \frac{p_i^2 k_i'(k_i'+1)}{(k_i'+l_i')(k_i'+l_i'+1)} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j k_j'}{k_j'+l_j'} \sum_{\substack{u=1\\ u\neq j\\ u\neq i}}^{M} \frac{p_u k_u'}{k_u'+l_u'}$$
$$+ \sum_{i=1}^{M} \frac{p_i k_i'}{k_i'+l_i'} \sum_{\substack{j=1\\ j\neq i}}^{M} \frac{p_j k_j'}{k_j'+l_j'} \sum_{\substack{u=1\\ u\neq j\\ u\neq i}}^{M} \frac{p_u k_u'}{k_u'+l_u'} \sum_{\substack{v=1\\ v\neq k\\ v\neq j\\ v\neq i}}^{M} \frac{p_v k_v'}{k_v'+l_v'}$$

Fig. 1. The first four moments.

coverage estimation methods by applying them in very extreme situations. System A (Fig. 4a) features a relative homogeneity among the classes regarding the noncoverage, a relative low variability of the fault/activity occurrence probabilities p_i and a slight (positive) correlation factor ($\approx$ 14 percent). System B (Fig. 4b), is a very atypical system with a large variability of the noncoverage and occurrence probabilities in each class. Furthermore, the correlation factor is negative and less than -40 percent. System C (Fig. 4c) is a near-perfect system with a very low noncoverage. It has quite a high variability over the classes and a large positive correlation between the p_i and c_i ($\approx$ 90 percent).

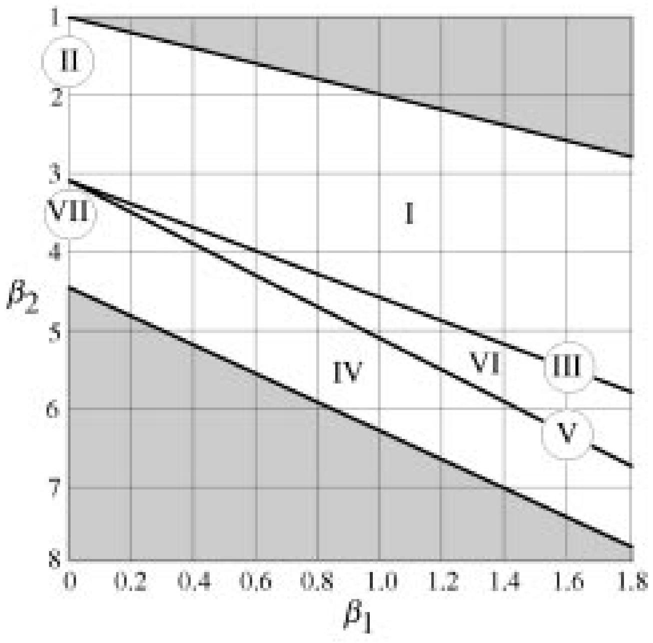
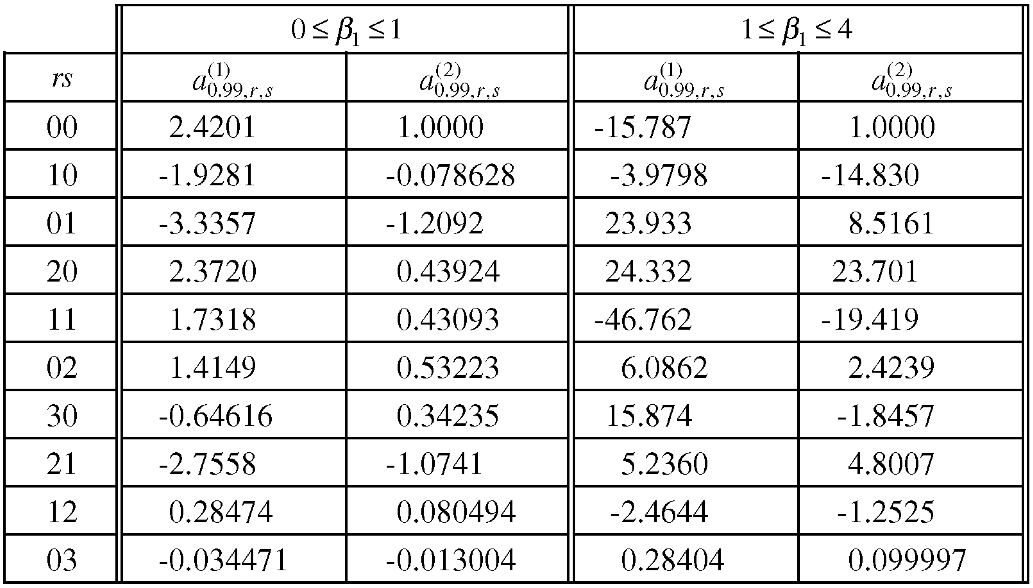


Fig. 2. The Pearson distribution system [12, p. 13].

5.2 Comparison Method

We compare the various confidence limit estimation methods by simulating a large number (25,000) of fault-injection campaigns carried out on each of the three systems, with a total sample size (the total number of injected faults in each campaign) taking six values ranging from 500 to 50,000,000. For each simulated campaign, we draw a random deficiency vector using the known class coverages of our hypothetical systems. We then compute the 99 percent confidence limits that would be obtained with each considered method.

Since we have total knowledge about our three hypothetical systems, we can verify whether the confidence we claim for the various methods is justified by calculating the proportion of simulated fault-injection campaigns for which the confidence limit statement $\bar{c} < \bar{c}_{\gamma}^{\dagger}(\mathbf{x})$ is true. We call this



	$0 \le \beta_1 \le 1$		$1 \le \beta_1 \le 4$	
rs	$a^{(1)}_{0.99,r,s}$	$a^{(2)}_{0.99,r,s}$	$a^{(1)}_{0.99,r,s}$	$a^{(2)}_{0.99,r,s}$
00	2.4201	1.0000	-15.787	1.0000
10	-1.9281	-0.078628	-3.9798	-14.830
01	-3.3357	-1.2092	23.933	8.5161
20	2.3720	0.43924	24.332	23.701
11	1.7318	0.43093	-46.762	-19.419
02	1.4149	0.53223	6.0862	2.4239
30	-0.64616	0.34235	15.874	-1.8457
21	-2.7558	-1.0741	5.2360	4.8007
12	0.28474	0.080494	-2.4644	-1.2525
03	-0.034471	-0.013004	0.28404	0.099997

Fig. 3. Coefficients $a^{(i)}_{\gamma,r,s}$ in the Bowman and Shenton approximation for $\gamma = 0.99$.

proportion the *success proportion*. With the prescribed 99 percent confidence level, the success proportion should also be about 99 percent. If this is not so for a given system and sample size, then the corresponding method is inadequate for that situation and should not have been used.

To assess the conservatism of the confidence limits for each method, we would also like to evaluate the expected values of the confidence limits for each total sample size. For all frequentist confidence limits derived from the normal approximation, the expected value of the limit can easily be calculated analytically by taking expectations of the corresponding expression (7) and replacing $E\{X_i\}$, respectively, by the known $\bar{c}_i$. The other expected confidence limits can be calculated by summing the corresponding confidence limit expressions over the complete sample space weighted by the (known) sample distribution. However, this summation would be much too time-consuming, so we resort to the simulated fault-injection campaigns and use the average of all the simulated confidence limits as an approximation of the expected value. The average is finally divided by the value of the system noncoverage. We call this measure the *normalized expected 99 percent upper confidence limit*. The nearer this normalized limit is to one, the less conservative is the corresponding estimation method.

5.3 Presentation of the Results

Each estimation method will be identified by a mnemonic using the notation defined in the table of Fig. 5. As explained in Section 3.3, for the frequentist estimations without the normal distribution approximation, only estimations obtained by the vectorial statistic are presented.

It happens sometimes that it is not possible to obtain a confidence limit estimation $\bar{c}^{\dagger}_{\gamma}(\mathbf{x})$ for a given fault-injection campaign. When this occurred in more than 10 percent of the simulated campaigns, we choose not to give any result at all. When the confidence limit estimations could be calculated for more than 90 percent of all simulated campaigns, the presented results are normalized by the number of simulated fault-injection campaigns leading to a result.

5.4 Frequentist Estimations

For systems B and C, some results are not presented for a representative stratification with small total sample sizes since this combination leads to a null allocation ($n_i = 0$) for some classes and is therefore meaningless.

Since, for system B, whatever the total sample size is, and for systems A and C, for high total sample sizes, the optimization routine failed to find a solution for more than 10 percent of the simulated campaigns, no result will be presented for the exact sample distribution.

5.4.1 Approximated Sample Distribution

Fig. 6 shows the success proportions (Fig. 6a) and the normalized expected 99 percent upper confidence limits (Fig. 6b) for all three systems.

The biased estimation technique F/AH, based on $\hat{\bar{c}}_{AH}(\mathbf{X})$ (cf. (6)), shows that the success proportion increases or decreases when the sample size increases depending on the sign of the correlation factor ρ between the p_i and $\bar{c}_i$. When the correlation factor is negative (system B), the success proportion increases with the sample size and tends to 100 percent, even with small sample sizes. When the correlation factor is positive (systems A and C), the success proportion tends to 0 percent.

When the sample size increases, the normalized expected confidence limit will also tend to a value function of the sign of ρ. When the correlation factor is negative, the expected confidence limit tends to a conservative value (normalized estimation greater than one). When the correlation factor is positive, the expected confidence limit tends to an optimistic value (normalized estimation less than one).

Contrary to F/AH, the success proportion of the nonbiased estimation techniques F/WH and F/*R first increases when the sample size increases and then tends to

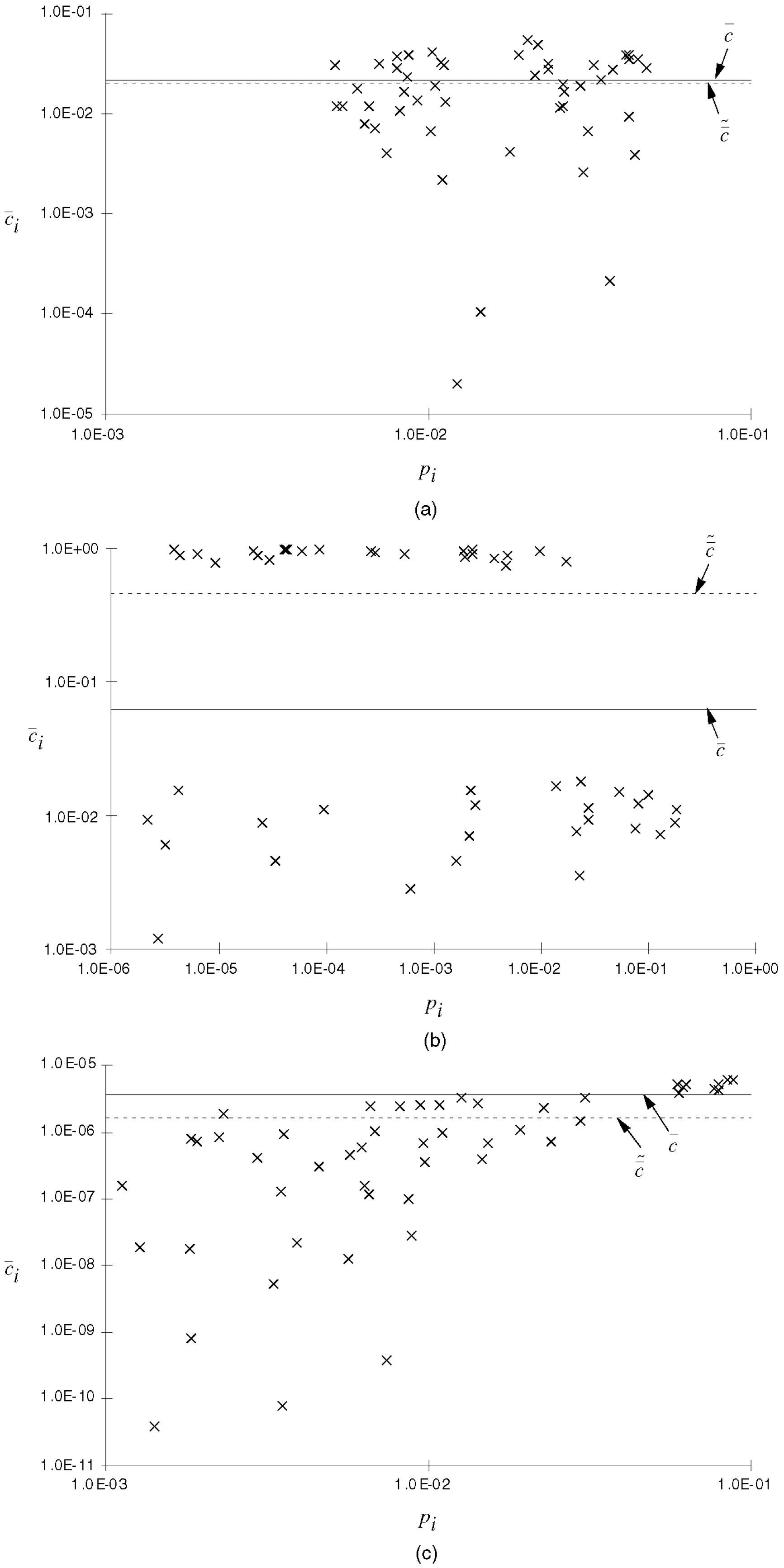
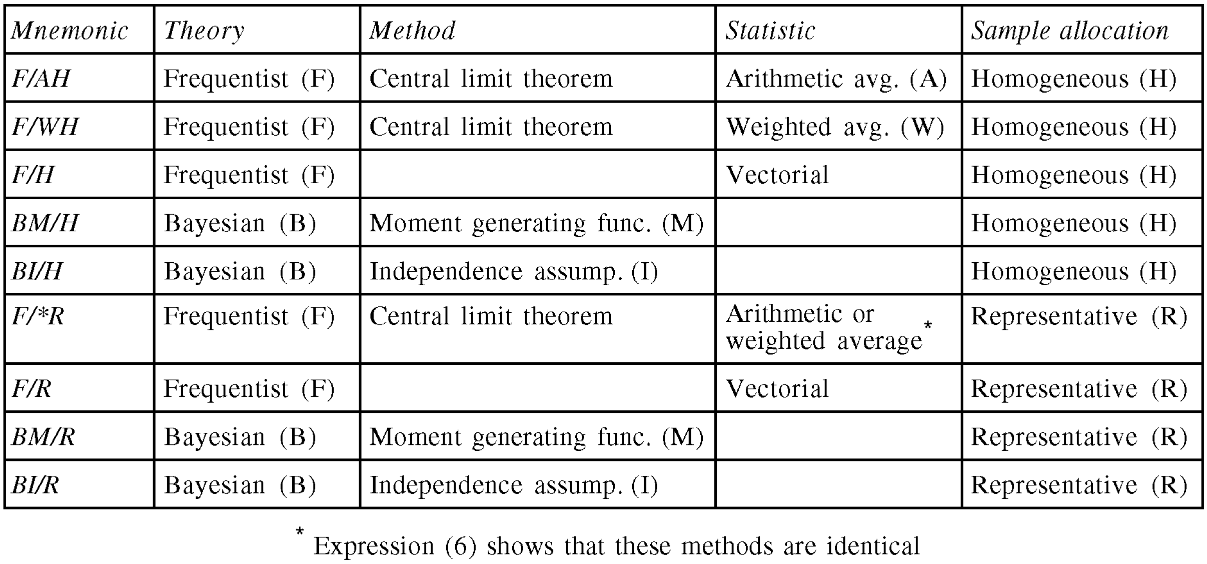


Fig. 4. System characteristics. (a) System A ($\bar{c}$ = 2.20E-02, ρ = 13.5%). (b) System B ($\bar{c}$ =6.21E-02, ρ = −41.5%). (c) System C ($\bar{c}$ = 3.72E-06, ρ = 89.1%).



Mnemonic	*Theory*	*Method*	*Statistic*	*Sample allocation*
F/AH	Frequentist (F)	Central limit theorem	Arithmetic avg. (A)	Homogeneous (H)
F/WH	Frequentist (F)	Central limit theorem	Weighted avg. (W)	Homogeneous (H)
F/H	Frequentist (F)		Vectorial	Homogeneous (H)
BM/H	Bayesian (B)	Moment generating func. (M)		Homogeneous (H)
BI/H	Bayesian (B)	Independence assump. (I)		Homogeneous (H)
*F/*R*	Frequentist (F)	Central limit theorem	Arithmetic or weighted average*	Representative (R)
F/R	Frequentist (F)		Vectorial	Representative (R)
BM/R	Bayesian (B)	Moment generating func. (M)		Representative (R)
BI/R	Bayesian (B)	Independence assump. (I)		Representative (R)

* Expression (6) shows that these methods are identical

Fig. 5. Notation.

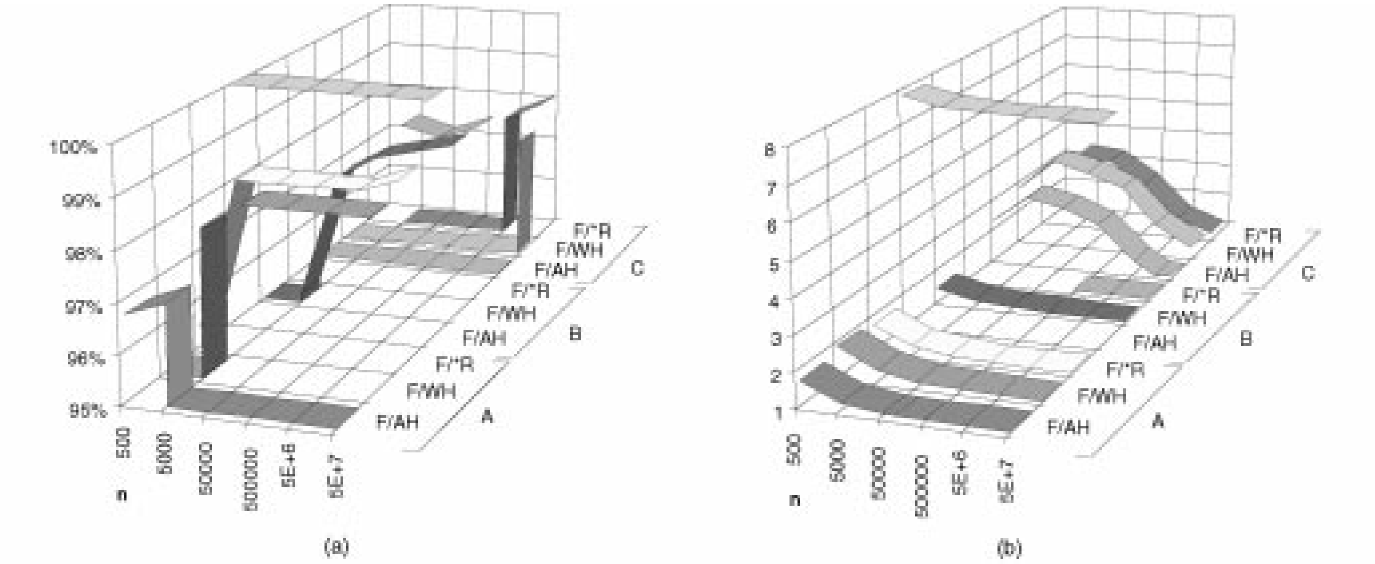


Fig. 6. Results with the approximated sample distribution. (a) Success proportion. (b) Normalized expected 99 percent upper confidence limits.

the requested value of 99 percent of the success proportion. Because of the very high coverage of system C, the requested value of the success proportion will only be reached when injecting a higher number of faults than shown in Fig. 6b.

For systems A and B, the expected confidence limit decreases with the sample size whatever the sign of the correlation factor. For system C, the expected confidence limit first increases and then decreases. For all three systems, the asymptotic normalized confidence limit is one.

For systems A and B, the estimation method based on the central limit theorem is valid for sample sizes greater than about 5.10^5. On the contrary, for system C, the requested value of 99 percent for the success proportion is not reached. The approximation based on the central limit theorem is thus not valid for sample sizes lower than 5.10^7.

For fault-tolerant systems with modest coverage, the estimation method based on the central limit theorem is valid only if many faults are injected in each class. For fault-tolerant systems with high coverage, the approximation given by the central limit theorem is not valid even if many faults are injected in each class. Therefore, for ultra-reliable systems, an estimation method without approximation is necessary.

5.4.2 Exact Sample Distribution

Fig. 7 shows the success proportions (Fig. 7a) and the normalized expected 99 percent upper confidence limits (Fig. 7b) for systems A and C.

In Fig. 7a, the success proportion for the available frequentist estimations is always equal to 100 percent. This indicates that these estimations are conservative. However, Fig. 7b shows that the degree of conservatism does decrease rapidly as the sample size increases. The conservatism depends also on the coverage of the fault-tolerant system. The confidence limits corresponding to system C (a system

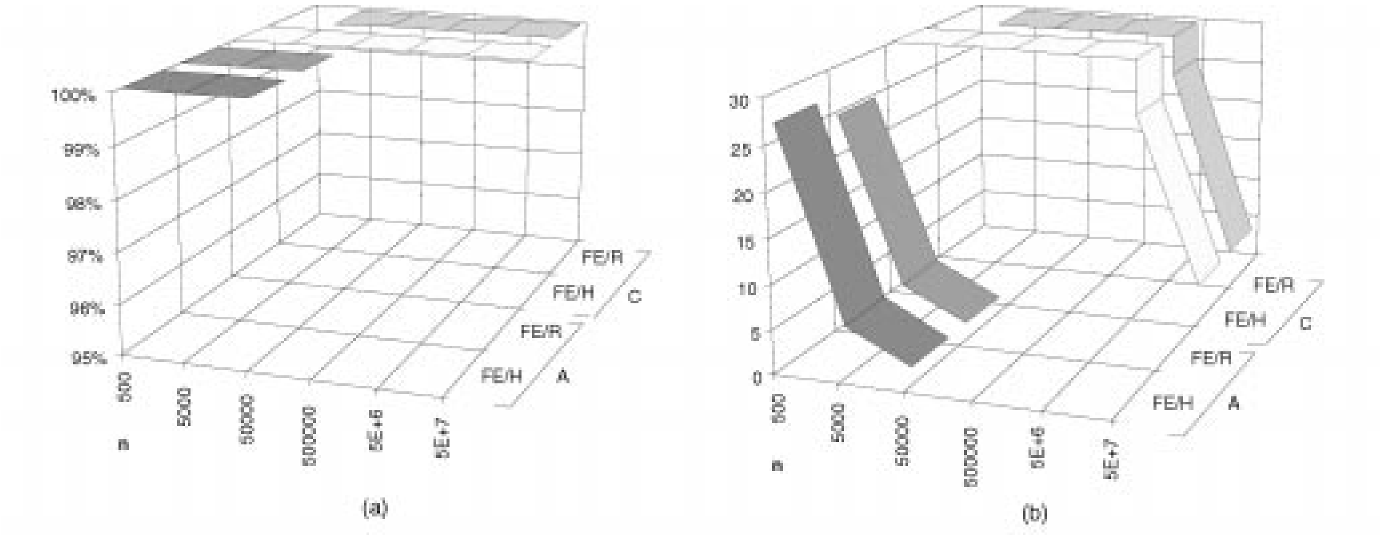


Fig. 7. Frequentist results using the exact sample distribution. (a) Success proportion. (b) Normalized expected 99 percent upper confidence limits.

with high coverage) are significantly more conservative than those for system A (a system with modest coverage).

It can also be noted, for both systems, that there is no great difference between the expected confidence limits corresponding to a homogeneous allocation and a representative one. Moreover, it is interesting to observe that a representative allocation does not necessarily lead to a less conservative confidence limit. As can be seen for system C in Fig. 7b, for example, when $n = 5.10^6$, this classic result for stratified sampling when using the normal distribution approximation no longer holds when considering the exact sample distribution with the vectorial statistic.

5.5 Bayesian Estimations

For system B, the estimation of the confidence limit required a much longer computation time than the other two. Due to this, and since we could not calculate frequentist limits for this system, the comparison between the frequentist and the Bayesian estimations will only concern systems A and C. The results for systems A and C are shown in Figs. 8 and 9, respectively.

For system C, results are not presented for a representative stratification with small total sample sizes ($n = 500$) since this combination leads to a null allocation ($n_i = 0$) for some classes and is therefore meaningless.

For system A (a system with modest coverage), the success proportion for the Bayesian estimation methods decreases from 100 percent to a value of about 99 percent when the number of faults injected increases (curves BM/H, BI/H, BM/R, and BI/R in Fig. 8a). It should be noted that, when more than 5.10^6 faults are injected, the success proportion is slightly under 99 percent (about 98.7 percent). This slight error occurs systematically and is probably due to the imprecision introduced by the Bowman and Shenton approximation (cf. Section 4.3). Since the success proportion is about 99 percent, all the Bayesian estimation methods are valid when applied to this system.

Fig. 8b shows that the frequentist estimations are always more conservative than the Bayesian ones (the normalized expected confidence limits are higher for F/H and F/R than for the Bayesian methods). However, this conservatism decreases as the number of faults injected increases. Finally, for the Bayesian methods, no significant difference can be noticed between the homogeneous and representative allocations.

For system C (a system with very high coverage), the success proportion for both Bayesian estimation methods (curves BM/H, BI/H, BM/R, and BI/R in Fig. 9a) decreases from 100 percent to a value close to 100 percent (about 99.8 percent) when the number of faults injected increases. Since the success proportion is higher than 99 percent, both Bayesian methods are also valid when applied to this system. The high success proportion indicates that these estimations will be conservative.

Fig. 9b shows that the frequentist estimations are again more conservative than the Bayesian estimations and that this conservatism decreases as the number of faults injected increases. In most cases, the Bayesian estimations with a representative allocation are very slightly less conservative than the ones with a homogeneous allocation.

Both Bayesian estimation methods prove to be valid when applied to these two quite different systems. This means that, in practice, we can always use one of these two methods since their validity does not depend on the value of the (unknown) system coverage. Furthermore, the corresponding estimations are less conservative than the frequentist ones.

5.6 Stratified Sampling vs. Simple Sampling

Since, for stratified sampling, the Bayesian estimations are less conservative than the frequentist ones, it is interesting to study how well they fare compared to the estimations obtained when simple sampling is used. For simple sampling, frequentist and Bayesian estimations are quite identical for the considered sample sizes [8]. For stratified sampling, since the Bayesian methods lead to upper confidence limits of about the same value, we show in this comparison only the results for the Bayesian method based on the moment generating function for a representative allocation.

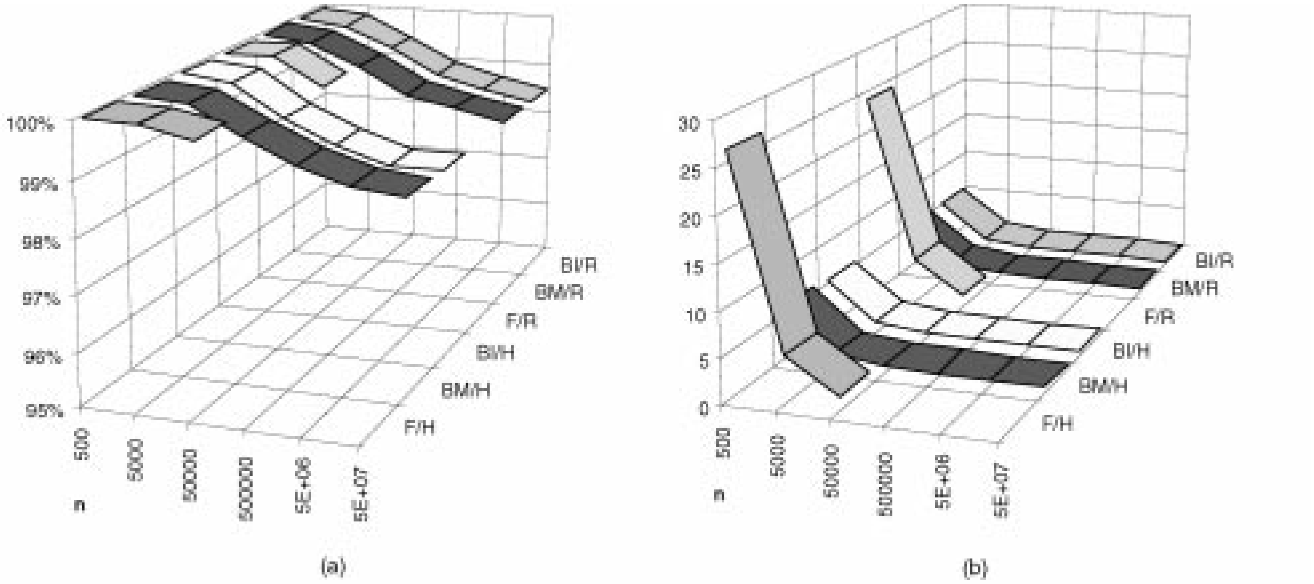


Fig. 8. Frequentist vs. Bayesian results (System A). (a) Success proportion. (b) Normalized expected 99 percent upper confidence limits.

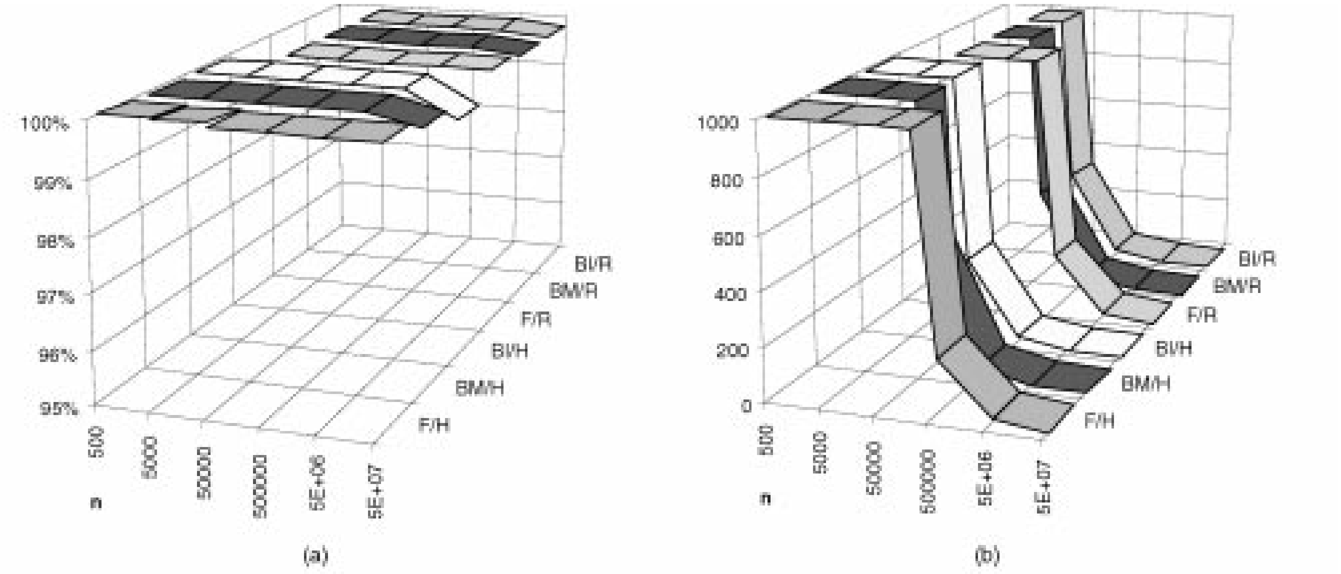


Fig. 9. Frequentist vs. Bayesian results (System C). (a) Success proportion. (b) Normalized expected 99 percent upper confidence limits.

Figs. 10 and 11 present the normalized expected 99 percent upper confidence limits, respectively, for systems A and C. The simple sampling estimations are noted "NS" (Non-Stratified) on these figures.

The figures show that the Bayesian estimations based on the moment generating function (curve BM/R) are more conservative than the estimations when simple sampling is used. However, the Bayesian estimations are much less conservative than the frequentist estimations with the vectorial statistic (curve F/R).

6 Summary and Conclusion

This paper has presented frequentist and Bayesian estimation methods applied to fault-injection for assessing coverage figures of fault-tolerant systems. In particular, we have focused on the statistical methods that can be applied to the results of a stratified fault-injection campaign on a system with high coverage.

Most previous work on frequentist estimation methods for stratified sampling has been based on the normal approximation of the distribution of a coverage point estimator. However, we have shown that the approximation given by the central limit theorem and leading to a normal distribution is not usually valid for coverage confidence limit estimations. First, most fault tolerance mechanisms are characterized by very high coverage factor values. Second, the number of faults injected during a campaign is often relatively low. Both contribute to the erroneousness of the normal distribution approximation.

If the confidence limit estimation cannot be based on a normal distribution, the confidence region theory must be used. Confidence regions obtained by two point estimators and a vectorial statistic have been introduced in this paper.

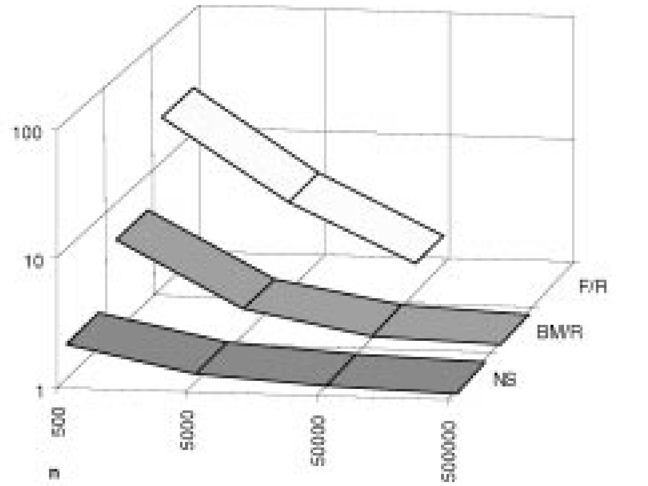


Fig. 10. Normalized expected limits for stratified and nonstratified sampling (System A).

In many cases, the coverage upper confidence limit estimations obtained by using the point estimators are less conservative than those obtained by the vectorial statistic. However, for more than three strata, the coverage confidence limit estimations using the point estimators become computationally intractable. So, most of the time, only the vectorial statistic can be applied.

For the Bayesian theory, the coverage confidence limits directly depend on the posterior distribution of the overall noncoverage. For partitions into more than three classes, the analytical expression of the posterior distribution becomes too complicated, so two methods have been presented to approximate the posterior distribution by calculating its first four moments. The first method uses assumptions on the independence between the coverages of the classes, while the second method is based on the moment generating function of the distribution. If the posterior distribution is identified in the Pearson system as a beta distribution (which is usually the case), the noncoverage upper confidence limit estimation can be obtained without approximation by calculating the inverse function [8]. In this paper, we used the Bowman and Shenton approximation to derive the confidence limit estimations without having to identify the particular distribution of the Pearson system. The Bayesian estimations prove to be considerably less conservative than the frequentist estimations obtained when using the vectorial statistic.

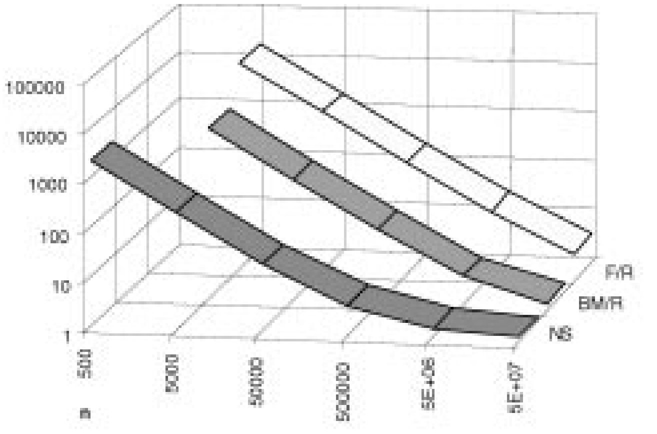


Fig. 11. Normalized expected limits for stratified and nonstratified sampling (System C).

Compared to simple sampling estimations, the stratified sampling frequentist and Bayesian estimations are more conservative, which means that stratification can degrade the coverage confidence limits. Therefore, when the practical advantages of stratification for fault injection are not paramount, simple sampling should be used. Indeed, the stratified estimation methods that do not rely on the normal approximation based on the central limit theorem lead to more conservative results than with simple sampling. For simple sampling, the frequentist and Bayesian estimations are very close to each other [8], so either of them can be chosen. Since both are easy to implement (e.g., see (10) for the frequentist estimator), there is no need to resort to the normal approximation.

When stratification is used and the normal approximation due to the central limit theorem is valid in each class, the frequentist estimation method based on this approximation should be applied. When the normal approximation is not valid in each class, the number of classes needs to be considered when choosing an estimation method. When the sample space is partitioned into a small number of classes (two or three), the frequentist estimations obtained by the point estimators and the Bayesian estimations based on the calculation of an analytical expression of the posterior distribution for $\bar{C}$ are very slightly less conservative than those obtained with the Bayesian methods using the moment calculations. However, their implementation is quite complicated. In practice, the Bayesian methods based on moment calculation are easier to implement and anyway become necessary when the number of classes is greater than three. Since, most of the time, the estimations obtained with a representative allocation are slightly less conservative, this allocation is preferable (but not mandatory). Finally, despite the complexity of the formulas given in Fig. 1 for the Bayesian method based on the moment generating function, this method is preferable since it avoids the introduction of unnecessary assumptions.

Appendix A

Confidence Region Theory for Point Estimators

We will note $Y(\mathbf{X})$ a general estimator which is a positive function of X_i and let $y_\gamma(\bar{\mathbf{c}})$ be the value of $Y(\mathbf{X})$ that satisfies:

$$F_{Y(\mathbf{X})}\big(y_\gamma(\bar{\mathbf{c}}) \mid \bar{\mathbf{c}}\big) = 1 - \gamma. \qquad \text{(A.1)}$$

Since sampling experiments in each class are Bernoulli trials, the deficiency number of each class X_i is distributed according to the binomial distribution $B(n_i, \bar{c}_i)$. The cumulative distribution of each X_i is thus a decreasing function of $\bar{c}_i$. Since $Y(\mathbf{X})$ is a positive function of the X_i $(i = 1, \ldots, M)$, its cumulative distribution $F_{Y(\mathbf{X})}(y|\bar{\mathbf{c}})$ is also a decreasing function of the $\bar{c}_i$ $(i = 1, \ldots, M)$.

Furthermore, by the very definition of a cumulative distribution, $F_{Y(\mathbf{X})}(y|\bar{\mathbf{c}})$ is an increasing function[1] of y. This means that, for $F_{Y(\mathbf{X})}(y|\bar{\mathbf{c}})$ to remain constant, any positive variation of a $\bar{c}_i$ must be compensated by a positive variation of y. Therefore, $y_\gamma(\bar{\mathbf{c}})$ defined by the solution of (13) is an increasing function of $\bar{\mathbf{c}}$.

We will now successively consider the special cases $M = 1$ and $M = 2$.

Case $M = 1$: This special case corresponds to simple (nonstratified) sampling. We have $\bar{\mathbf{c}} = \bar{c}$ and equation (A.1) becomes:

$$y_\gamma(\bar{c}) : F_{Y(X)}\big(y_\gamma(\bar{c}) \mid \bar{c}\big) = 1 - \gamma. \tag{A.2}$$

The solution of this equation, $y_\gamma(\bar{c})$, is typically of the form illustrated in Fig. 12.

We will now show that an upper 100γ percent confidence limit on $\bar{c}$ is given by the inverse of function $y_\gamma(\bar{c})$, that is:

$$\bar{c}_\gamma(Y(x)) = \bar{c} : F_{Y(X)}(Y(x) \mid \bar{c}) = 1 - \gamma. \tag{A.3}$$

For a given value $\bar{c}^*$ of $\bar{c}$, Fig. 12 shows that the estimate $Y(x)$ will be greater than $y_\gamma(\bar{c}^*)$ for 100γ percent of the time, and less than $y_\gamma(\bar{c}^*)$ for $100(1-\gamma)$ percent of the time. Consequently, the inverse function $\bar{c}_\gamma(Y(x))$ will lead to a value on the $\bar{c}$ axis that falls to the right of $\bar{c}^*$ 100γ percent of the time and to the left $100(1-\gamma)$ percent of the time.

We observe that the random variable $\bar{c}_\gamma(X)$ that is so defined satisfies the requirement for an upper 100γ percent confidence limit given by (5). So, for $M = 1$, $\bar{c}^\uparrow_\gamma(X) = \bar{c}_\gamma(X)$.

For this special case, (5) can be rewritten in its well-known form:

$$\bar{c}^\uparrow_\gamma(X) : \sum_{j=0}^{x} \binom{n}{j} \left(\bar{c}^\uparrow_\gamma(\mathbf{x})\right)^j \left(1 - \bar{c}^\uparrow_\gamma(X)\right)^{n-j} = 1 - \gamma. \tag{A.4}$$

This equation can be solved analytically. By introducing 100γ percent percentile points, $F_{\nu_1,\nu_2,\gamma}$, of an F distribution with ν_1, ν_2 degrees of freedom [11, p. 59], we obtain:

$$\bar{c}^\uparrow_\gamma(X) = \frac{(X+1)F_{2(X+1),2(n-X),\gamma}}{(n-X) + (X+1)F_{2(X+1),2(n-\mathbf{X}),\gamma}}. \tag{A.5}$$

Case $M = 2$: The solution $y_\gamma(\bar{\mathbf{c}})$ of (A.1) in the space $(y, \bar{c}_1, \bar{c}_2)$ defines a surface around the $Y(\mathbf{x})$ axis, typically of the form shown in Fig. 13.

The contours on this surface are solutions of expression (A.1), noted $y_\gamma(\bar{\mathbf{c}})$, corresponding to different values of $Y(\mathbf{X})$. These contours can be projected onto the plane $(\bar{c}_1, \bar{c}_2)$ as shown in Fig. 13. They represent the inverse of function $y_\gamma(\bar{\mathbf{c}})$, noted $\bar{\mathbf{c}}_\gamma(\mathbf{x})$, defined by:

$$\bar{\mathbf{c}}_\gamma(Y(\mathbf{x})) = \bar{\mathbf{c}} : F_{Y(\mathbf{X})}(Y(\mathbf{x}) \mid \bar{\mathbf{c}}) = 1 - \gamma. \tag{A.6}$$

For a given value $\bar{\mathbf{c}}^*$ of $\bar{\mathbf{c}}$, the region delimited by $\bar{\mathbf{c}}_\gamma(Y(\mathbf{x}))$ and the $(\bar{c}_1, \bar{c}_2)$ axes, which can be denoted $I_\gamma(\mathbf{x})$, will include $\bar{\mathbf{c}}^*$ 100γ percent of the time, and exclude it $100(1-\gamma)$ percent of the time. $I_\gamma(\mathbf{X})$ is thus a 100γ percent confidence region estimator for the vector $\bar{\mathbf{c}}$. An upper 100γ percent confidence limit estimator on $\bar{c}$ is obtained by (9). Since $\bar{c}$ is a positive function of $\bar{c}_i$, the value of $\bar{\mathbf{c}}$ which maximizes $\mathbf{p}.\bar{\mathbf{c}}^T$ is on the contour $\bar{\mathbf{c}}_\gamma(Y(\mathbf{x}))$.

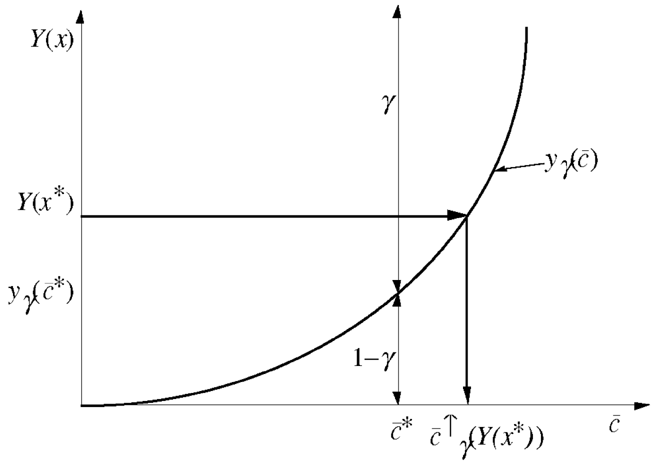


Fig. 12. Noncoverage upper confidence limit.

General case: For any $M > 0$, the solution $y_\gamma(\bar{\mathbf{c}})$ of (A.1) defines a hypersurface around the $Y(\mathbf{x})$ axis and the inverse function $\bar{\mathbf{c}}_\gamma(Y(\mathbf{x}))$ defines the limit of a confidence region around the origin:

$$\bar{\mathbf{c}}_\gamma(Y(\mathbf{x})) = \bar{\mathbf{c}} : F_{Y(\mathbf{X})}(Y(\mathbf{x}) \mid \bar{\mathbf{c}}) = 1 - \gamma. \tag{A.7}$$

As stated before, an upper 100γ percent confidence limit estimator on $\bar{c}$ is obtained for a vector $\bar{\mathbf{c}}$ on the frontier $\bar{\mathbf{c}}_\gamma(Y(\mathbf{x}))$. Since the elements X_i of $\mathbf{X}$ are independently and binomially distributed with parameters $\bar{c}_i$ and n_i, (A.7) can be rewritten as:

$$\bar{\mathbf{c}} : F_{\mathbf{X}}(Y(\mathbf{x}) \mid \bar{\mathbf{c}}) = \sum_{\mathbf{x}' : Y(\mathbf{x}') \le Y(\mathbf{x})} f_{\mathbf{X}}(\mathbf{x}' \mid \bar{\mathbf{c}}) = 1 - \gamma \tag{A.8}$$

with

$$f_{\mathbf{X}}(\mathbf{x}' \mid \bar{\mathbf{c}}) = \prod_{i=1}^{M} \binom{n_i}{x'_i} \bar{c}_i^{x'_i} (1 - \bar{c}_i)^{\mu_i = x'_i}.$$

The computation of $\bar{c}^\uparrow_\gamma(\mathbf{x})$ can thus be expressed as the following maximization problem:

Maximization of $\bar{c} = \mathbf{p}.\bar{\mathbf{c}}^T$ under the constraints:

- given by the confidence region frontier $\bar{\mathbf{c}}_\gamma(Y(\mathbf{x}))$ for $\bar{\mathbf{c}}$:

$$\bar{\mathbf{c}} : \sum_{\mathbf{x}' : Y(\mathbf{x}') \le Y(\mathbf{x})} \prod_{i=1}^{M} \binom{n_i}{x'_i} \bar{c}_i^{x'_i} (1 - \bar{c}_i)^{n_i - x'_i} = 1 - \gamma \tag{A.9}$$

- given by the limits of the parameter space:

$$\forall i \in \{1..M\}, \bar{c}_i \in [0, 1]$$

[1] Since the X_i and, therefore, $Y(\mathbf{X})$ are discrete, the distribution $F_{Y(\mathbf{X})}(y|\bar{\mathbf{c}})$ is, in fact, a staircase function. However, for simplicity of explanation, we will use a continuous representation of this function.

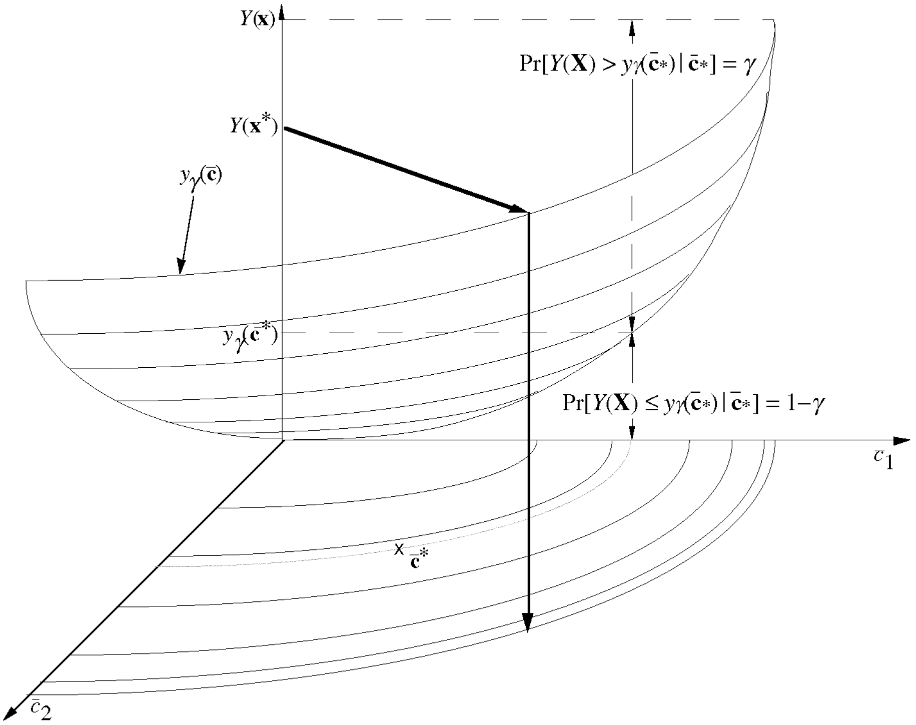

$Y(\mathbf{x})$
$\Pr[Y(\mathbf{X}) > y_\gamma(\bar{\mathbf{c}}*) \mid \bar{\mathbf{c}}*] = \gamma$
$Y(\mathbf{x}^*)$
$y_\gamma(\bar{\mathbf{c}})$
$y_\gamma(\bar{\mathbf{c}}^*)$
$\Pr[Y(\mathbf{X}) \le y_\gamma(\bar{\mathbf{c}}*) \mid \bar{\mathbf{c}}*] = 1-\gamma$
c_1
x $\bar{\mathbf{c}}^*$
$\bar{c}_2$

Fig. 13. Noncoverage upper confidence frontiers and projections for $M = 2$.

Appendix B

Confidence Region Theory for the Vectorial Statistic

For each class i, we can use (A.4) to define an upper $100\gamma_i$ percent confidence limit estimator:

$$\bar{c}^{\uparrow}_{i\gamma_i}(X_i) : \sum_{j=0}^{x_i} \binom{n_i}{j} \left(\bar{c}^{\uparrow}_{i\gamma_i}(X_i)\right)^j \left(1 - \bar{c}^{\uparrow}_{i\gamma_i}(X_i)\right)^{n_i - j} = 1 - \gamma_i \quad \forall i = 1, \ldots, M. \tag{B.1}$$

If we choose the M values of γ_i such that:

$$\prod_{i=1}^{M} \gamma_i = \gamma, \tag{B.2}$$

then the hypercube in the parameter space formed by the axes and the $\bar{c}^{\uparrow}_{i\gamma_i}(x_i)$ is a 100γ percent confidence region for which the corresponding upper 100γ percent confidence limit on $\bar{c}$ is given by:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})|_{(\gamma_1, \gamma_2, \ldots, \gamma_M)} = \sum_{i=1}^{M} p_i \bar{c}^{\uparrow}_{i\gamma_i}(x_i), \tag{B.3}$$

corresponding to the value of $\bar{\mathbf{c}}$ at the apex of the confidence region hypercube.

Among the infinite number of ways of choosing the γ_i such that (B.2) is satisfied, we are looking for the one leading to the lowest value of $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})|_{(\gamma_1, \gamma_2, \ldots, \gamma_M)}$. Thus:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{x}) = \min_{(\gamma_1, \gamma_2, \ldots, \gamma_M)} \bar{c}^{\uparrow}_{\gamma}(\mathbf{x})|_{(\gamma_1, \gamma_2, \ldots, \gamma_M)} = \min_{(\gamma_1, \gamma_2, \ldots, \gamma_M)} \sum_{i=1}^{M} p_i \bar{c}^{\uparrow}_{i\gamma_i}(x_i). \tag{B.4}$$

By combining (B.1) and (B.4), the confidence limit $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ is given by:

$$\bar{c}^{\uparrow}_{\gamma}(\mathbf{x}) = \min_{(\gamma_1, \gamma_2, \ldots, \gamma_M)} \sum_{i=1}^{M} p_i \bar{c}_i$$

under the constraint:

$$\sum_{j=0}^{x_i} \binom{n_i}{j} \bar{c}_i^j (1 - \bar{c}_i)^{n_i - j} = 1 - \gamma_i \quad \forall i = 1, \ldots, M.$$

From (B.2), the upper 100γ percent confidence limit $\bar{c}^{\uparrow}_{\gamma}(\mathbf{x})$ is therefore given by the following minimization problem:

Minimization of $\bar{c} = \mathbf{p}.\bar{\mathbf{c}}^T$ under the constraints:

- given by the global confidence:

$$\bar{\mathbf{c}} : \prod_{i=1}^{M} \left(1 - \sum_{x'_i=0}^{x_i} \binom{n_i}{x'_i} \bar{c}_i^{x'_i} (1 - \bar{c}_i)^{n_i - x'_i}\right) = \gamma \tag{B.5}$$

- given by the limits of the parameter space:

$$\forall i \in \{1..M\}, \bar{c}_i \in [0, 1].$$

Acknowledgments

This work was partially supported by the *DEVA* project (ESPRIT no. 20072).

REFERENCES

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* New York: Dover, 1972.

[2] J. Aitchison and I.R. Dunsmore, *Statistical Prediction Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1975.

[3] J. Arlat, A. Costes, Y. Crouzet, J.-C. Laprie, and D. Powell, "Fault Injection and Dependability Evaluation of Fault-Tolerant Systems," *IEEE Trans. Computers,* vol. 42, no. 8, pp. 913-923, Aug. 1993.

[4] W.G. Bouricius, W.C. Carter, D.C. Jessep, P.R. Schneider, and A.B. Wadia, "Reliability Modeling for Fault-Tolerant Computers," *IEEE Trans. Computers,* vol. 20, no. 11, pp. 1,306-1,311, Nov. 1971.

[5] K.O. Bowman and L.R. Shenton, "Approximate Percentage Points for Pearson Distributions," *Biometrika,* vol. 66, no. 1, pp. 147-155, 1979.

[6] R. Chillarege and N.S. Bowen, "Understanding Large System Failures—A Fault Injection Experiment," *Proc. 19th Int'l Symp. Fault-Tolerant Computing (FTCS-19),* pp. 356-363, Chicago, June 1989.

[7] G.S. Choi, R.K. Iyer, R. Saleh, and V. Carreno, "A Fault Behavior Model for an Avionic Microprocessor: a Case Study," *Dependable Computing for Critical Applications,* A. Avizienis and J.-C. Laprie, eds., pp. 171-195, 1991.

[8] M. Cukier, "Estimation of the Coverage of Fault-Tolerant Systems," doctoral dissertation, Nat'l Polytechnic Inst. Toulouse, France, July 1996 (in French).

[9] C.S. Davis and M.A. Stephens, "Approximate Percentage Points using Pearson Curves, Algorithm AS192," *Applied Statistics,* vol. 32, pp. 322-327, 1983.

[10] M.-C. Hsueh, T.K. Tsai, and R.K. Iyer, "Fault Injection Techniques and Tools," *Computer,* vol. 40, no. 4, pp. 75-82, Apr. 1997.

[11] N.L. Johnson and S. Kotz, *Distributions in Statistics—Discrete Distributions.* New York: John Wiley & Sons, 1969.

[12] N.L. Johnson and S. Kotz, *Distributions in Statistics—Continuous Univariate Distributions-1.* New York: John Wiley & Sons, 1970.

[13] N.L. Johnson and S. Kotz, *Distributions in Statistics—Continuous Univariate Distributions-2.* New York: John Wiley & Sons, 1970.

[14] J. Karlsson, P. Lidén, P. Dahlgren, R. Johansson, and U. Gunneflo, "Using Heavy-Ion Radiation to Validate Fault-Handling Mechanisms," *IEEE Micro,* vol. 14, no. 1, pp. 8-23, Feb. 1994.

[15] G.A. Kanawati, N.A. Kanawati, and J.A. Abraham, "FERRARI: A Flexible Software-Based Fault and Error Injection System" *IEEE Trans. Computers,* vol. 44, no. 2, pp. 248-260, Feb. 1995.

[16] D. Powell, E. Martins, J. Arlat, and Y. Crouzet, "Estimators for Fault Tolerance Coverage Evaluation," *Proc. 23rd Int'l Symp. Fault-Tolerant Computing (FTCS-23),* pp. 228-237, Toulouse, France, 1993 (extended version in *IEEE Trans. Computers,* vol. 44, no. 2, pp. 347-366, Feb. 1995).

[17] D.A. Rennels and A. Avizienis, "RMS: A Reliability Modeling System for Self-Repairing Computers," *Proc. Third Int'l Symp. Fault-Tolerant Computing (FTCS-3),* pp. 131-135, Palo Alto, Calif., 1973.

[18] Z. Segall, D. Vrsalovic, D. Siewiorek, D. Yaskin, J. Kownacki, J. Barton, D. Rancey, A. Robinson, and T. Lin, "FIAT—Fault Injection Based Automated Testing Environment," *Proc. 18th Int'l Symp. Fault-Tolerant Computing (FTCS-18),* pp. 102-107, Tokyo, June 1988.

[19] A. Stuart and J.K. Ord, *Distribution Theory,* Kendall's Advanced Theory of Statistics, 1. London: Edward Arnold, 1987.

[20] C.J. Walter, "Evaluation and Design of an Ultra-Reliable Distributed Architecture for Fault Tolerance," *IEEE Trans. Reliability,* vol. 39, no. 4, pp. 492-499, Oct. 1990.

[21] W. Wang, K.S. Trivedi, B.V. Shah, and J.A. Profeta III, "The Impact of Fault Expansion on the Interval Estimate for Fault Detection Coverage," *Proc. 24th Int'l Symp. Fault-Tolerant Computing (FTCS-24),* pp. 330-337, Austin, Tex., June 1994.

Michel Cukier (M'99) received the Engineer degree from the Free University of Brussels, Belgium, in 1991 and the Doctor of Engineering degree from the National Polytechnic Institute of Toulouse, France, in 1996. During 1991-1992, he was a teaching assistant at the Free University of Brussels. From 1992 to 1996, he was at LAAS-CNRS, Toulouse, France, for his doctoral work on coverage estimation of fault-tolerant systems. He is currently a visiting research asistant professor in the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. Dr. Cukier's research interests include adaptive fault tolerance in distributed systems, the evaluation of fault-tolerant systems combining modeling and fault injection, and the estimation of fault tolerance coverage. As part of this work, he is a co-developer of the AQuA Architecture, an architecture that provides dependable distributed objects. He is a member of the IEEE.

David Powell (M'92) received his Bachelor of Science degree in electronic engineering from the University of Southampton, England, in 1972, a Specialty Doctarate degree from the Toulouse Paul Sabatier University in 1975, and his Docteur ès-Sciences degree from the Toulouse National Polytechnic Institute in 1981. He is currently "Directeur de Recherche" within CNRS, the French National Organization for Scientific Research and leads the research group on Dependable Computing and Fault Tolerance at LAAS-CNRS. He has managed several national and European research contracts. He was the scientific director of the six-year Delta-4 Esprit project on open dependable distributed computing. He is currently the chief architect of the GUARDS Esprit project, which aims to design a generic fault-tolerant architecture for embedded, ultra-dependable real-time systems for space, railway, and nuclear submarine applications. Dr. Powell's current research work concerns the design and validation of fault-tolerant distributed computing systems. Particular interests are distributed algorithms for software-implemented fault tolerance, stochastic Petri net modeling for dependability evaluation, and the use of fault injection for estimating fault tolerance coverage. He has written more than 90 papers for international and national journals and conferences, is coauthor of two books on dependable computing, and holds a patent for a fault- and damage-tolerant network for data trnasmission. He was program chair of the first European Dependable Computing Conference (Berlin, 1994) and program co-chair of the 26th IEEE Symposium on Fault-Tolerant Computing (Sendai, Japan, 1996). He was guest editor of a special section of *Communication of the ACM* devoted to group communication (April 1996). Dr. Powell is a member of the ACM, the IEEE, and the SE Working Group on Dependable Computing.

Jean Arlat (M'80) received the Engineer degree from the National Institute of Applied Sciences of Toulouse in 1976 and the Doctor in Engineering degree and the Docteur ès-Sciences degree from the National Polytechnic Institute of Toulouse in 1979 and 1990, respectively. He is currently "Directeur de Recherche" within CNRS, the French National Organization for Scientific Research and a member of the group on Dependable Computing and Fault Tolerance at LAAS-CNRS. Since 1997, he has been director of the Laboratory for Dependable Engineering (LIS: Laboratoire d'Ingénierie de la Sûreté de fonctionnnement) hosted by LAAS. Dr. Arlat's research interests focus on the evaluation of hardware-and-software fault-tolerant systems, including both analytical modeling and experimental fault injection approaches, subjects on which he authored or coauthored more than 70 papers for international and national journals and conferences. Dr. Arlat chaired the IEEE Computer Society's Technical Committee on Fault-Tolerant Computing (TC-FTC) in 1994-1995. In 1998, he served as the program co-chair for the 28th International Symposium on Fault-Tolerant Computing. Since January 1999, he has been chairman of the IFIP Working Group 10.4 on Dependable Computing and Fault Tolerance. He is a member of the ACM, the IEEE, and the SEE Working Group on Dependable Computing.