Reliability Engineering and System Safety ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Deriving a frequentist conservative confidence bound for probability of failure per demand for systems with different operational and test profiles

Peter Bishop *, Andrey Povyakalo

City University, London, United Kingdom

ARTICLE INFO

Article history:
Received 15 January 2016
Received in revised form
19 July 2016
Accepted 23 August 2016

Keywords:
Statistical testing
Confidence bounds
Operational profile
Software reliability

ABSTRACT

Reliability testing is typically used in demand-based systems (such as protection systems) to derive a confidence bound for a specific operational profile. To be realistic, the number of tests for each class of demand should be proportional to the demand frequency of the class. In practice, however, the actual operational profile may differ from that used during testing. This paper provides a means for estimating the confidence bound when the test profile differs from the profile used in actual operation. Based on this analysis the paper examines what bound can be claimed for different types of profile uncertainty and options for dealing with this uncertainty. We also show that the same conservative bound estimation equations can be applied to cases where different measures of software test coverage and operational profile are used.

1. Introduction

Nuclear protection systems are designed to protect against a range of safety-related plant incidents (known as postulated initiating events or PIE). A PIE can affect one or more plant parameters (such as temperature, pressure and neutron flux). These plant parameters are monitored by the protection system and the reactor is tripped if the plant parameters go outside the safe operational envelope.

In the UK, a probabilistic safety assessment (PSA) is required to justify the safety of nuclear plant. As part of this process, the performance of the protection system must be quantified in terms of probability of failure on demand, *pfd*, where the demand can be any of the PIE events. There are accepted means for estimating the *pfd* arising from hardware failures, but we also need to include an estimate for the *pfd* of the software if the protection system is computer-based. Statistical reliability testing [1,2] is one means of estimating the software *pfd* of a demand-based system to some confidence bound, and it is recommended in functional safety standards such as IEC 61508 [3]. For example, reliability testing was performed as part of the independent confidence building programme required by the UK Office for Nuclear Regulation (ONR) for the computer-based primary protection system (PPS) at Sizewell B nuclear power station [4]. The PPS was subjected to 5000 simulated demands to support a *pfd* claim of 10^{-3}. Reliability testing is also planned for new nuclear power stations to be installed in the UK [5].

The confidence bound derived from statistical reliability testing is based on a number of modelling assumptions. The stated assumptions in IEC 61508 [3] for the low demand rate case are:

1. The test data distribution is equal to the distribution of demands during on-line operation.
2. Test runs are statistically independent from each other, with respect to the cause of a failure.
3. An adequate mechanism exists to detect any failures which may occur.
4. Number of test cases $N > 100$.
5. No failure occurs during the N test cases.

The second assumption can be met in the protection system context as the protection system is normally reset after a reactor trip (so the software always starts from the same initial state).

The third assumption requires a perfect "oracle" that determines if a failure has occurred. The required response is relatively easy to determine for PIE events in a nuclear plant since each simulated PIE is expected to result in a reactor trip.

The last two assumptions will also be met in a nuclear protection context as many thousands of tests are needed for the required *pfd* and the software has to be corrected and retested from scratch if a failure is observed.

To satisfy assumption 1, the number of tests for each class of

* Corresponding author.
E-mail address: pgb@csr.city.ac.uk (P. Bishop).

http://dx.doi.org/10.1016/j.ress.2016.08.019
0951-8320/

demand (i.e. for each PIE) should be proportional to the demand frequency of that class during operation, so the confidence bound estimate cannot be used if the test and operational profiles differ.

This paper presents a means for estimating the confidence bound when the test profile differs from the profile used in actual operation. Based on this analysis, the paper examines what bound can be claimed for different types of profile uncertainty and the options for dealing with this uncertainty.

We also show that the same conservative bound equations can be applied in contexts where the software reliability bound and input profile are characterised in different ways.

2. Problem statement

If a system is subjected to N test demands without failure [1], we can follow the approach suggested by Neyman and Pearson [6], Neyman [7], and Clopper and Pearson [8] as it is presented by Wang [9] and identify an upper confidence bound, q, on the probability of failure on demand Q to a confidence $1-\alpha$ as the largest value such that the hypothesis "$H0$: $Q = q$" is not rejected against the alternative "$H1$: $Q < q$" at the significance level α.

Thus, q must satisfy the following equation:

$$(1-q)^N = \alpha \tag{1}$$

However, it is often the case that the system handles different *classes* of demand, e.g. a protection system that protects against different PIE events. These demand classes are assumed to be disjoint, i.e. only a single demand can occur at any point in time.

Testing over a series of classes can be characterised by a test plan vector:

$$\mathbf{n} = \{n_1, n_2, \ldots, n_m\} \tag{2}$$

where m is a number of demand classes, n_i is the number of tests for demand class i, and the total number of tests is:

$$N = \sum_{i=1}^{m} n_i \tag{3}$$

The distribution of tests over the demand classes can be characterised by a test distribution profile vector:

$$\hat{\mathbf{p}} = \{\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_m\} \tag{4}$$

where $\hat{p}_i = n_i/N$, $i = 1 \ldots m$

When this multiple demand class system is used in operation it will be subject to an operational profile:

$$\mathbf{p} = \{p_1, p_2, \ldots, p_m\} \tag{5}$$

Ideally the operational and test profile distributions will match so that $\mathbf{p} = \hat{\mathbf{p}}$. However, in practice the operational profile $\mathbf{p}$ will vary if the system is used in different environments or there is uncertainty in the likelihood of different external events. So we need some means to determine a bound q_s to some confidence $(1-\alpha)$ for a different operational profile $\mathbf{p}$ given a prior set of tests $\mathbf{n}$.

3. Problem formulation

For some (unknown) vector of demand class *pfds*

$$\mathbf{q} = \{q_1, q_2, \ldots, q_m\} \tag{6}$$

the likelihood of observing no failures with test plan $\mathbf{n}$ is:

$$P(\mathbf{q}, \mathbf{n}) = \prod_{i=1}^{m} (1-q_i)^{n_i}, (0 \le q_i \le 1) \tag{7}$$

The $(1-\alpha)$ confidence area for all possible *pfd* vectors, $\mathbf{q}'$, is

$$D(\mathbf{n}, \alpha) = \{\mathbf{q}' : P(\mathbf{q}', \mathbf{n}) \ge \alpha\} \tag{8}$$

For an arbitrary vector of demand class *pfds* $\mathbf{q}$ and operational profile $\mathbf{p}$, the system *pfd*, Q_S, is simply the weighted average of the vector of $\mathbf{q}$ values, i.e.

$$Q_S(\mathbf{q}, \mathbf{p}) = \mathbf{q}\cdot\mathbf{p} = \sum_{i=1}^{m} q_i p_i \tag{9}$$

The confidence area (8) constrains the set of permissable $\mathbf{q}$ vectors and induces a confidence interval for Q_S with the upper bound:

$$q_s = \max_{\mathbf{q} \in D(\mathbf{n}, \alpha)} Q_s(\mathbf{q}, \mathbf{p}) \tag{10}$$

We therefore need a method for solving (10) for an arbitrary demand profile $\mathbf{p}$.

It is straightforward to solve (10) numerically for any profile $\mathbf{p}$ and test vector $\mathbf{n}$. However a numerical analysis does not permit any general conclusions to be drawn about the impact of changes in the operational profile $\mathbf{p}$.

With an analytic derivation of the confidence bound, we can model the impact of a mismatch between the test profile and the actual demand profile and identify general strategies for designing test profiles that reduce the sensitivity of the bound to uncertainties in the operational profile.

The next section describes the approach we developed to derive an analytic solution for the confidence bound.

4. Solution approach

In Appendix A we use Lagrangian multipliers to identify the stationary points that represent the potential solutions to (10) but the solution space is complex. There are $2^m - 1$ stationary points and the optimal point depends on the specific values used in $\mathbf{p}$ and $\mathbf{n}$. As a result, there is no simple analytic solution that can be applied to all operational profiles. So we developed an alternative approach for obtaining an analytic solution by deriving a conservative approximation for (10) that makes the problem easier to solve.

In this reformulation, the likelihood (7) is approximated as:

$$\tilde{P}(\mathbf{q}, \mathbf{n}) = \prod_{i=1}^{m} \exp(-q_i n_i), \quad 0 \le q_i \le 1 \tag{11}$$

It is a standard result [10] that

$$\exp(-q_i n_i) \ge (1-q_i)^{n_i} \tag{12}$$

Thus

$$P(\mathbf{q}, \mathbf{n}) \ge \alpha \tag{13}$$

implies,

$$\tilde{P}(\mathbf{q}, \mathbf{n}) \ge \alpha \tag{14}$$

Therefore, the approximated confidence area

$$\tilde{D}(\mathbf{n}, \alpha) = \{\mathbf{q}' : \tilde{P}(\mathbf{q}', \mathbf{n}) \ge \alpha\} \tag{15}$$

is a superset of the exact confidence area, i.e.

$$\tilde{D}(\mathbf{n}, \alpha) \supseteq D(\mathbf{n}, \alpha) \tag{16}$$

As a result, the approximate solution will always be conservative relative to the exact solution, i.e. for a given α, $\mathbf{n}$, $\mathbf{p}$

$$\tilde{q}_s \ge q_s \tag{17}$$

where

Please cite this article as: Bishop P, Povyakalo A. Deriving a frequentist conservative confidence bound for probability of failure per demand for systems with different operational and test profiles. Reliability Engineering and System Safety (2016), http://dx.doi.org/10.1016/j.ress.2016.08.019