

Time-Dependent Reliability of Random Dynamic Systems Using Time-Series Modeling and Importance Sampling

2011-01-0728

Published
04/12/2011Amandeep Singh and Zissimos Mourelatos
Oakland Univ.Efstratios Nikolaidis
Univ. of Toledo

Copyright © 2011 SAE International

doi:[10.4271/2011-01-0728](https://doi.org/10.4271/2011-01-0728)

ABSTRACT

Reliability is an important engineering requirement for consistently delivering acceptable product performance through time. As time progresses, the product may fail due to time-dependent operating conditions and material properties, component degradation, etc. The reliability degradation with time may increase the lifecycle cost due to potential warranty costs, repairs and loss of market share. Reliability is the probability that the system will perform its intended function successfully for a specified time interval. In this work, we consider the first-passage reliability which accounts for the first time failure of non-repairable systems. Methods are available in the literature, which provide an upper bound to the true reliability which may overestimate the true value considerably. Monte-Carlo simulations are accurate but computationally expensive. A computationally efficient importance sampling technique is presented to calculate the cumulative probability of failure for random dynamic systems excited by a stationary input random process. Time series modeling is used to characterize the input random process from only one sample function of the random process. Examples are presented to demonstrate the accuracy and efficiency of the proposed importance sampling method over the traditional Monte Carlo simulation.

1. INTRODUCTION

In today's technological world, we demand continued functionality of a wide array of complex equipment and processes. We expect our cars and appliances for example, to be functional and reliable. In reality, as time goes on,

degradation in the components (e.g. material properties) and stochastic operating conditions (loading, environmental conditions, etc.), cause these products to deliver inconsistent performance and sometimes to even fail prematurely. As a result, we may have high warranty cost, customer annoyance and lower resale value, or even a catastrophic injury or loss of life. It is important therefore, to maintain good reliability over a long time period in order to maximize the product value over its lifecycle. A desired reliability can be achieved by designing products under the reliability constraint. However, an accurate yet efficient estimation of reliability poses significant challenge, especially, in the presence of random input process. In the current research, we propose an importance sampling (IS) approach based on time-series modeling of a random input process.

Reliability or time-dependent reliability is defined as the probability that the system will perform its intended function successfully for a specified interval of time, under stated operating and environmental conditions. It is therefore, related to product functionality over time which is determined by the so-called "hard" and "soft" failures [1]. In a hard failure the system loses functionality due to a complete breakdown of one or more of its components, while in a soft failure the system is functional but one or more performance measures are out of conformance. The reliability associated with the hard failure is important for non-repairable systems where the replacement or repair of a failed component is not possible. In this case, the failed system is removed permanently from the population. In contrast, repairable systems [2] consist of multiple components which can be

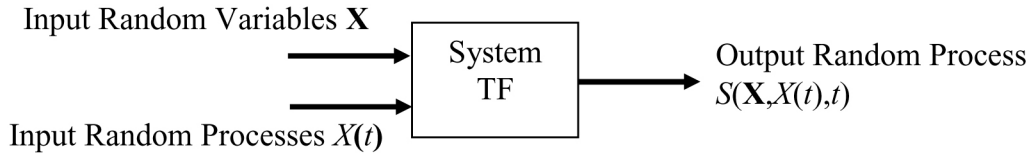


Figure 1. Schematic of input-output relations

repaired or replaced if failed keeping therefore, the original system in the population. In this research, we use time-dependent reliability concepts associated with the so-called first-passage of non-repairable systems. The approach can be used to reduce the lifecycle cost [3, 4] or to predict the reliability in real-time, similarly to the work in [5]. It can also be used to set up a schedule for preventive condition-based-maintenance.

Figure 1 shows our notation for the input and output of a system. The input random vector $\mathbf{Y} = [\mathbf{X}, X(t)]$ consists of a vector of random variables $\mathbf{X}$ and a random processes $X(t)$. An output random process $S(\mathbf{Y}, t)$ is obtained when $\mathbf{X}$ is passed through the system transfer function (TF).

A random input variable is quantified using its probability density distribution (PDF). If the PDF is not available, experimental data can be used to estimate it. A random input process must be however, characterized under certain assumptions (e.g. ergodicity and/or stationarity). In this paper, a time-series modeling approach is used to characterize an input random process using “short” time experimental data. Weak ergodicity is assumed in order to calculate ensemble averages (mean and autocorrelation function) using the temporal averages of only one sample function. Using time-series modeling, we propose in this paper, an importance sampling methodology to calculate the cumulative probability of failure (probability of first passage - upcrossing-) of a dynamic system driven by an ergodic input random process. The basic idea of importance sampling is to concentrate the distribution of sampling points in the region of most importance, which is the area that mainly contributes to the failure probability, instead of spreading the sample points evenly in the whole range of possible values of the basic random variables as is the case in MCS. In importance sampling, a sampling distribution is used to generate sample points which result in many failures and then weights proportional to the likelihood ratio between the actual and sampling distributions are used to estimate the actual probability of failure. A representative example of a quarter car model on a stochastic road is used to demonstrate the accuracy of the proposed method.

2. RELIABILITY ANALYSIS

Reliability analysis estimates the probability of failure due to variability and/or uncertainty. In time-independent (or time-invariant) analysis, the uncertainty is quantified using random

variables. If a time-independent limit state function $g(\mathbf{X})$ is used, where the vector $\mathbf{X}$ represents the input random variables with a joint probability density function $f_{\mathbf{X}}(\mathbf{x})$, the probability of failure is given by

$$P_f = P(g(\mathbf{X}) < 0) = \int \cdots \int_{g(\mathbf{x}) < 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (1)$$

Where upper case denotes a random variable or a random parameter and a lower case denotes a realization of the random quantity. Because the multi-dimensional integral of Eq. (1) is difficult to evaluate, various analytical and simulation-based methods have been developed [6, 7, 8, 9].

Let $S(\mathbf{X}, X(t), t)$ be a random process and $S_{\text{threshold}}$ be an allowable threshold. A limit state is defined as

$$g(\mathbf{X}, X(t), t) = S_{\text{threshold}} - S(\mathbf{X}, X(t), t) \quad (2)$$

where $g(\mathbf{X}, X(t), t) < 0$ indicates failure. Denoting the time interval of interest by $t_L = [t_{\min}, t_{\max}]$, the probability of failure within this time interval is defined as

$$P_f^c(t_{\min}, t_{\max}) = P(\exists t \in [t_{\min}, t_{\max}], g(\mathbf{X}, X(t), t) < 0) \quad (3)$$

where failure occurs if the limit state function becomes negative at least at one instance within the time interval. This is known as *cumulative* probability of failure [3, 10]. The definition of Eq. (3) is different from the *instantaneous* probability of failure

$$p_f^i(t) = P(g(\mathbf{X}, X(t), t) < 0) \quad (4)$$

which is calculated in a quasi-steady sense, by fixing time t and replacing the random process by a random variable. As a result, it does not account for possible failures before time t . The conceptual difference between the cumulative and the instantaneous probabilities of failure is very important.

To estimate the cumulative probability of failure, an outcrossing rate approach is commonly used providing only an upper bound. The outcrossing rate is defined [11] as

$$\nu^+(t) = \lim_{\Delta\tau \rightarrow 0, \Delta\tau > 0} \frac{P[g(\mathbf{X}, X(t), t) > 0 \cap g(\mathbf{X}, X(t + \Delta\tau), t + \Delta\tau) \leq 0]}{\Delta\tau} \quad (5)$$

Ditlevsen and Madsen [12] calculated the mean number of outcrossings as

$$m_1 = E[N^+(t_{\min}, T)] = \int_{t_{\min}}^T \nu^+(t) dt \quad (6)$$

where $N^+(t_{\min}, t)$ is the number of outcrossings in the time interval $[t_{\min}, t]$, and Shinozuka [13] showed that the cumulative probability of failure is bounded as

$$\max_{t_{\min} \leq t \leq t_{\max}} p_f^i(t) \leq p_f^c(t_{\min}, t_{\max}) \leq p_f^i(t_{\min}) + m_1. \quad (7)$$

The outcrossing rate approach was first introduced by Rice [14] followed by extensive studies [15, 16, 17, 18] under the assumption of the outcrossings being statistically independent and Poisson distributed. Hagen and Tvedt [11] suggested a parallel system reliability formulation to compute the outcrossing rate. It uses two successive time-invariant analyses based on FORM, and the binomial cumulative distribution to calculate the probability of the joint event in Eq. (5). This approach was later adopted in the PHI2 method [10]. Methods based on Poisson's distribution and the PHI2 method compute an upper bound of the cumulative probability of failure of Eq. (3). A Monte-Carlo based set theory approach has been also proposed [19] using a similar approach with the PHI2 method. An analytical study in [20] revealed that the PHI2 based approach lacks sufficient accuracy for non-monotonic problems such as vibratory systems. For this reason, analytical approaches were proposed in [20, 21] to accurately estimate the cumulative probability of failure taking into account non-monotonic behavior.

When a dynamic system with random properties is subjected to an input random process, the existing analytical approaches only provide upper bounds that can significantly deviate from the true value. An accurate and efficient MCS based approach was proposed in [22] that estimates time-dependent failure rates over the product lifecycle. The efficiency of this approach is further improved in this paper using a novel importance sampling method.

3. CUMULATIVE PROBABILITY OF FAILURE

3.1. FAILURE RATE FOR FIRST-PASSAGE (NON-REPAIRABLE SYSTEMS)

The probability that an item will fail at time T_f where $t < T_f \leq t + dt$, given that the item is functioning at time t , is given by

$$P(t < T_f \leq t + dt | T_f > t) = \frac{P(t < T_f \leq t + \Delta t)}{P(T_f > t)} = \frac{p_f^c(t + dt) - p_f^c(t)}{R(t)}, \quad (8)$$

where $R(t) = P(T_f > t)$ denotes the reliability at time t . The failure rate function is defined as [23],

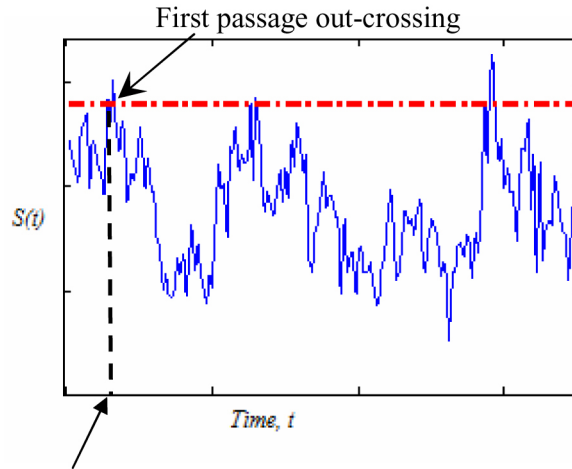
$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P(t < T_f \leq t + dt | T_f > t)}{dt} = \frac{f(t)}{R(t)} \quad (9)$$

Based on Eqs. (8) and (9), the first-passage probability of failure (cumulative probability of failure) $p_f^c(t)$ is given in terms of the failure rate $\lambda(t)$ as

$$p_f^c(t) = 1 - (1 - p_f^i(0)) \exp \int_0^t -\lambda(r) dr. \quad (10)$$

Because it is very challenging to “analytically” calculate $\lambda(t)$, a simulation approach such as MCS can be used. However, MCS can become practically inefficient if the time t is too long or the number of sample functions is large. In [22], MCS efficiency was significantly improved by only calculating the failure rate over a “short” time t_s which represents a small fraction of the system lifecycle.

In the current research, a variance-reduction technique called importance sampling (IS) is described to further improve the computational cost of MCS. The basic idea of importance sampling is to concentrate the distribution of sampling points in the region of most importance, which is the area that mainly contributes to the failure probability, instead of spreading the sample points evenly in the whole range of possible values of the basic random variables as is the case in MCS.



$p(i)$ is incremented by 1. Subsequent out-crossings (up-crossings) of this sample function are not accounted

Figure 2. Schematic of first-passage failure

To use MCS, the time t is discretized using a time step Δt , so that $t = i * \Delta t$ where $i = 0, 1, 2, \dots, I$. The failure rate λ_i at time increment i is then computed as,

$$\lambda_i = \frac{\text{\# of first time failures at } i \text{ with memory of all past events}}{(\text{\# of safe samples at } i-1) * \Delta t} \quad (11)$$

or

$$\lambda_i = \frac{\sum_{n=1}^N p(i)}{N_s(i-1) * \Delta t}, \quad (12)$$

where the counter $p(i)$ is incremented by one if for the n^{th} ($n = 1, 2, \dots, N$) sample function, the event $\{g(k) > 0 \text{ for } k = 1, 2, \dots, i-1 \text{ and } g(i) \leq 0\}$ is true. This event captures the first failure as indicated by the limit state function g . Figure 2 shows this schematically. In Eq. (12), $N_s(i-1)$ denotes the number of sample functions (out of N) for which there is no out-crossing for all time steps $k = 1, 2, \dots, i-1$.

4. RANDOM INPUT PROCESS CHARACTERIZATION USING TIME-SERIES MODELING

It is impractical to get actual data to represent a random input process $X(t)$ over time. The input sample functions are generally available from experimental data for a “short” period of time. In this case, a random process characterization technique such as time-series modeling [24] or spectral representation methods [16] can be utilized to generate synthetic random input process sample functions using the available experimental data.

Time-series modeling techniques have been widely used in the fields of finance, econometrics, and bio-medical engineering, among others, where forecasting under stochastic conditions is needed. It can be used to characterize both stationary and non-stationary processes [24, 25, 26, 27]. The Shinozuka method [28], and the Karhunen-Loeve [16] expansion are commonly used spectral representation approaches. An extension of Karhunen-Loeve method, the Expansion Optimal Linear Estimation (EOLE) method [18] has also been used to characterize a random process. It uses a modal representation of the random process based on a series expansion to represent the random process by a finite set of deterministic functions weighted by random variables. The EOLE method assumes however, stationarity which is not always true. Furthermore for acceptable accuracy, spectral representation based methods generally require a large number of independent harmonics or standard normal random variables which can hinder on their efficiency.

In this study, we use time-series modeling because it does not have the limitations of the EOLE method. There are three broad classes of time-series models which are of practical importance; the Auto-Regressive (AR) models, the Integrated (I) models, and the Moving Average (MA) models [24]. Combinations of these models result in autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models. Here, we use autoregressive models.

Consider a random process $X(t)$. A sample function $x(t)$ is discretized in the time interval $[0, T]$ using a uniform time step Δt so that $x_i = x(t_i)$, where $t_i = i * \Delta t$. For an AR(p) model, the discretized sample function is represented as

$$x_i - \mu = \phi_1(x_{i-1} - \mu) + \phi_2(x_{i-2} - \mu) + \dots + \phi_p(x_{i-p} - \mu) + \varepsilon_i \quad (13)$$

where μ is the temporal mean of the process, $\varepsilon_i \equiv N(0, \sigma_e^2)$ is Gaussian white noise and $\phi_1, \phi_2, \dots, \phi_p$ are feedback parameters to be estimated. Different order AR models can be created to determine the best fit. For an AR(2) model, the variance σ_e^2 of the Gaussian white noise is determined from

$$\gamma(0) = \text{Var}(x_i) = \frac{\sigma_e^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 \dots - \phi_p \rho_p} \quad (14)$$

where $\gamma(0)$ is the variance of the random process, and ρ_p is the value of the autocorrelation function at the time lag $\tau = p \cdot \Delta t$. Similar expressions exist for higher order AR models. Details are provided in [24].

After the feedback parameters are estimated, a residual series $E(t) = x(t) - \hat{x}(t)$ is formed as the difference between the actual $x(t)$ and the estimated $\hat{x}(t)$ processes and statistical tests are performed to make sure that the random variables E_t and $E_{t+\tau}$ are uncorrelated for every τ .

4.1. MODEL IDENTIFICATION

The model type can be roughly identified by visually inspecting the plots of the autocorrelation and the partial sample autocorrelation functions for different lags (multiples of Δt). The autocorrelation provides significant information about the correlation between random variables $X(t_1)$ and $X(t_1 + h)$ where h denotes the lag. For a stationary random process, the autocorrelation depends only on h and not on t_1 . For autoregressive models, the autocorrelation function dies out quickly with increasing h . The sample autocorrelation function $\hat{\rho}(h)$ is defined as

$$\hat{\rho}(h) = \frac{n^{-1} \sum_{i=1}^{n-h} (x_{i+h} - \mu)(x_i - \mu)}{\hat{\sigma}^2} \quad (15)$$

where $\hat{\sigma}$ is the estimated standard deviation of the random process. In Eq. (15), we can have an unbiased estimation of $\rho(h)$ if n^{-1} is replaced by $(n-h)^{-1}$. For convenience however, we used the n^{-1} term similarly to the econometric toolbox of MATLAB.

The partial autocorrelation of lag h represents the autocorrelation between X_i and X_{i+h} with the linear dependence of X_{i+1} through X_{i+h-1} removed. This is equivalent to the autocorrelation between X_i and X_{i-h} that is not accounted for by lags 1 to $h-1$, inclusive. The partial autocorrelation is useful in identifying the order of an

autoregressive model. For an AR(p) model, it is zero for lags greater or equal to $p+1$.

After the order p of the model is identified, the $\hat{\phi}$'s and μ are estimated either by using the Yule-Walker Eqs [24, 29] or equivalently, by minimizing

$$\sum_{i=p+1}^n [\{x_i - \mu\} - \phi_1(x_{i-1} - \mu) - \dots - \phi_p(x_{i-p} - \mu)]^2. \quad (16)$$

Statistical tests are then performed to ensure the goodness of fit as explained in [24].

5. PROPOSED FAILURE RATE ESTIMATION USING IMPORTANCE SAMPLING AND TIME-SERIES MODELING

A comprehensive review of various importance sampling techniques is presented in [30]. An early example of using importance sampling in rare event applications is in estimating the probabilities of failure of nuclear particle penetrating shields [31]. Other areas of application include reliability estimation in the fields of digital communications [32], for example. The majority of IS methods and application thereof are focused on instantaneous probability of failure [33, 34]. Very few applications of IS are available for the first passage time-dependent reliability. Bayer and Bucher [35] presented an IS approach using a spectral representation technique. We have found that the IS approach using spectral representation is limited by the number of random variables. For improved accuracy, the spectral representation method requires a large number of harmonics with frequencies that are treated as random variables. As a result, the approach is usually inaccurate due to a large variance in the likelihood ratio estimation resulting from numerical issues associated with the joint density function. Some studies on importance sampling using the Girsanov transformation and Monte Carlo Markov Chain (MCMC) simulations have also been reported [36].

5.1. IMPORTANCE SAMPLING FUNDAMENTALS

Importance sampling involves two steps. First, samples are drawn from a sampling probability density function (PDF) instead of the original PDF in order to generate more samples in the "important" regions of the sample space which result in more failures. Secondly, the estimation is corrected by averaging the effect of different samples (realizations) using appropriate weights, so that the mean value of the estimated quantity is preserved [30].

Let X be a random variable with a PDF of $f_X(x)$. The expected value of a function $\theta(x)$ of the random variable X where θ is a deterministic function (operator), is given by

$$E[\theta(x)] = \int_{\Omega} \theta(x) f_X(x) dx \quad (17)$$

where Ω is the domain of x . The above equation provides the probability of failure if $\theta(x)$ is an indicator function which is equal to one if failure occurs and zero otherwise; i.e.

$$p_f = E[\theta(x)] = \int_{\Omega} \theta(x) f_X(x) dx. \quad (18)$$

Because the analytical calculation of this integral is very difficult, MCS is used based on sampling from the distribution $f_X(x)$. If N sample points $x_1, x_2, \dots, x_N$ are drawn from $f_X(x)$, the probability of failure is

$$p_f = \frac{\sum_{n=1}^N \theta(x_n)}{N} = \frac{N_f}{N} \quad (19)$$

where N_f is the number of resulting failures from the N sample points.

For rare events however, a very large number N of sample points is required in order to generate a few failures N_f . In importance sampling methods, sampling is performed from a sampling distribution $f_X^s(x)$ in order to generate more failures. The sampling distribution is different from the true distribution $f(x)$ and its choice is arbitrary. [Eq. \(19\)](#) can be then written as

$$p_f = \int_{\Omega} \theta(x) \frac{f_X(x)}{f_X^s(x)} f_X^s(x) dx. \quad (20)$$

If a likelihood ratio is defined as,

$$\omega = \frac{f_X(x)}{f_X^s(x)}, \quad (21)$$

the time-independent probability of failure from [Eq. \(20\)](#) is

$$p_f = \frac{\sum_{n=1}^N \omega(x_n) \theta(x_n)}{N} \quad (22)$$

or

$$p_f = \frac{\sum_{n=1}^{N_f} \omega(x_n)}{N}, \quad (23)$$

where N_f is the number of resulting failures out of the N sample points.

5.2. FAILURE RATE ESTIMATION USING IMPORTANCE SAMPLING

For a random process, the instantaneous probability of failure at time t is defined as

$$p_f(t) = \int_{\Omega} \theta(x;t) f_X(x;t) dx \quad (24)$$

where $f_X(x;t)$ is the PDF of a random variable X obtained by “slicing” the ensemble functions of the random process at time t . The indicator function $\theta(x;t)$ is defined using a conditional failure as

$$\theta(x;t) = \begin{cases} 1 & \text{if } g(x;t) \leq 0 \\ 0 & \text{elsewhere} \end{cases} \quad (25)$$

Using [Eq. \(24\)](#), the instantaneous “conditional” probability of failure,

$$p_f^\lambda(t) = \int_{\Omega} \theta(\mathbf{x};t) f_X(\mathbf{x};t) d\mathbf{x} \quad (26)$$

is defined. We will use it to calculate the failure rate. The instantaneous conditional probability of failure provides the probability of failure at time $t+\Delta t$ given that the system is safe at time t . This definition is different from the instantaneous probability of failure where there is no condition of a safe system at all times before failure. Because the instantaneous conditional probability of failure depends on the past states, it requires the *time-dependent* joint density function $f_X(\mathbf{x};t)$ where $\mathbf{x}$ represents a vector of realizations from the distributions of random variables $X_1 = X(t_1)$, $X_2 = X(t_2)$, ..., $X_k = X(t_k)$ where t_k is the time when failure occurs. In [Eq. \(26\)](#), the conditional indicator function is also time-dependent, expressed as,

$$\theta(\mathbf{x};t) = \begin{cases} 1 & \text{if } g(x, t + \Delta t) \leq 0 \mid T_f > t \\ 0 & \text{elsewhere} \end{cases} \quad (27)$$