Probabilistic Engineering Mechanics 38 (2014) 54–69

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

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

Reliability of dynamic systems in random environment by extreme value theory

M. Grigoriu *, G. Samorodnitsky

Cornell University, Ithaca NY 14853-3501, USA

ARTICLE INFO

Article history:
Received 6 January 2014
Received in revised form
18 August 2014
Accepted 18 August 2014
Available online 10 September 2014

Keywords:
Generalized extreme value distributions
Highly reliable dynamic systems
Monte Carlo simulation
Random vibration
Reliability

ABSTRACT

A practical method is developed for estimating the performance of highly reliable dynamic systems in random environment. The method uses concepts of univariate extreme value theory and a relatively small set of simulated samples of system states. Generalized extreme value distributions are fitted to state observations and used to extrapolate Monte Carlo estimates of reliability and failure probability beyond data. There is no need to postulate functional forms of extreme value distributions since they are selected by the estimation procedure. Our approach can be viewed as an alternative implementation of the method in [7] for estimating system reliability. Numerical examples involving Gaussian and non-Gaussian system states are used to illustrate the implementation of the proposed method and assess its accuracy.

1. Introduction

The probability $p_s(\tau)=P(X(t)\in D, 0\le t\le \tau)$ that a system state $X(t)$ does not leave a safe set D during a time interval $[0,\tau]$, and the probability $p_f(\tau)=1-p_s(\tau)$ that $X(t)$ exits D at least once in $[0,\tau]$, referred to as reliability and failure probability, are essential quantities of interest in dynamics. They provide useful metrics for the development of economical and safe designs.

Monte Carlo simulation is the only method that can be used to find the probabilities $p_s(\tau)$ and $p_f(\tau)$ irrespective of a system size and complexity. System reliability is estimated by $\hat{p}_s(\tau)=n^{-1}\sum_{i=1}^{n} 1(x_i(t)\in D, 0\le t\le\tau)$, where $\{x_i(t)\}$ are n independent samples of $X(t)$. However, the method is not feasible in applications involving large-dimensional $(d \gg 1)$, highly reliable $(p_f(\tau)\simeq 0)$ dynamic systems because of the computational demand that can be excessive. For example, approximately 10^6 independent samples of $X(t)$ are needed to estimate failure probabilities of order 10^{-5} and the generation of these samples would require about 10^5 h if a single sample can be obtained in 10 min.

To reduce the computation time for estimating $p_s(\tau)$ and $p_f(\tau)$, it has been proposed to change the measure of the state vector such that failure and survival events occur approximately in equal proportion [1]. If such measure change can be constructed, accurate estimates of $p_s(\tau)$ result from a relatively small number of samples of $X(t)$ even for highly reliable systems. Although the Girsanov theorem provides the theoretical framework for measure change when dealing with dynamic system driven by Gaussian noise, the construction of measures providing efficient reliability estimates poses significant difficulties in applications [5, Section 5.4].

Analytical solutions for $p_s(\tau)$ and $p_f(\tau)$ are available in special cases of limited practical interest. Generally, numerical methods need to be employed to calculate these probabilities. The theory of random vibration provides exact and approximate methods for calculating the reliability of dynamic systems. The exact methods involve solutions of partial differential equations, e.g., the Fokker–Plack equation and partial differential equations for the characteristic function of $X(t)$ with appropriate boundary conditions. They are practical only for dynamic systems with small state vectors [5, Section 7.3]. The approximate methods, e.g., the stochastic averaging, stochastic linearization, moment closure, perturbation, and crossing of random processes, have been used with mixed success to solve a broad range of applications [5, Section 7.3.1.5]. It is difficult to assess the accuracy of these methods in a general setting since they are based on heuristic assumptions.

Recently, an alternative method has been proposed for estimating the performance of highly reliable dynamic systems [7,8]. The method uses a relatively small number of independent samples of $X(t)$ to estimate the mean rate ν_D at which this process exits D, referred to as mean D-outcrossing rate, and approximate $p_s(\tau)$ from ν_D under the assumption that the D-outcrossings of $X(t)$ are Poisson events with intensity ν_D. Since the sample size is relatively small, Monte Carlo estimates of $p_s(\tau)$ based solely on samples of $X(t)$ can only be obtained for at most moderately reliable systems. Monte Carlo estimates of mean D-outcrossing rates and concepts

* Corresponding author.
E-mail address: mdg12@cornell.edu (M. Grigoriu).

http://dx.doi.org/10.1016/j.probengmech.2014.08.005
0266-8920/

of the extreme value theory are employed in [7] to construct approximations of $p_s(\tau)$ beyond data that can be used to assess the performance of highly reliable systems.

This paper presents an alternative implementation of the main idea in [7,8]. Like in these studies, our objective is to find the probabilities $p_s(\tau)$ and $p_f(\tau)$ for highly reliable dynamic systems from a relatively small set of samples of $X(t)$. In contrast to these studies, we use exclusively the theory of univariate extreme value distribution to estimate system reliability. Our estimates of $p_s(\tau)$ and $p_f(\tau)$ are derived from generalized extreme value (GEV) and generalized Pareto (GP) distributions fitted to samples of $X(t)$. The type of the extreme value distribution used to construct our approximations for $p_s(\tau)$ and $p_f(\tau)$ does not have to be postulated. It is selected by the estimation procedure.

The proposed GEV and GP estimates of $p_s(\tau)$ and $p_f(\tau)$ are satisfactorily in all numerical examples presented in the paper. GEV estimates are more attractive since they are conceptually simple, apply to both stationary and non-stationary states, and have low storage demand. On the other hand, GP estimates involve some technicalities, apply only to stationary states in the form considered in our discussion, and may require significant storage. Our preference for GEV estimates is at variance with current estimates of floods, high wind speeds, excessive ozone levels, and other environmental extreme events that are frequently based on GP distributions [3,10]. In this setting, GP estimates are often preferred since they have the potential of extracting more information from single records. For example, excesses of daily wind speed maxima over a specified threshold can be used to fit GP distributions and estimate wind speeds of specified return periods. Depending on the threshold, the number of these excesses can be larger or smaller than the number of years of daily wind speeds. The sample size for corresponding GEV estimates is equal to the number of years in the wind speed record.

Our results are limited to stationary state processes. Following a review of essentials of extreme value theory, the probabilities $p_s(\tau)$ and $p_f(\tau)$ are estimated for independent/dependent, Gaussian/non-Gaussian discrete-time state processes. Time series rather than stochastic processes are used to represent samples of $X(t)$ since observations of system states are recorded at discrete times.

2. Problem definition

Let $X(t)$ be an $\mathbb{R}^d$−valued stationary stochastic process defining the state of a dynamic system in the steady-state regime. Let $D_z = \{x \in \mathbb{R}^d : g(x) \le z\}$ denote a safe set, where $g : \mathbb{R}^d \to (0, \infty)$ is a specified smooth function. The safe set D used previously to define system reliability $p_s(\tau)$ coincides with D_z for a particular value of $z > 0$. We use D_z rather than D in our further considerations since this safe set can be expanded and contracted while preserving its shape so that it can accommodate low- to high-reliable systems. This definition of D_z is used to calculate the reliability of a broad range of structural/mechanical systems [2, Chapter 6].

In this setting, our objective is to estimate the probability

$$p_s(z;\tau) = P(X(t) \in D_z,\ 0 \le t \le \tau) \tag{1}$$

that $X(t)$ does not leave D_z during the time interval $[0, \tau]$, referred to as system reliability. The complement $p_f(z;\tau) = 1 - p_s(z;\tau)$ of this probability is the system probability of failure.

It is assumed that (1) the system state $X(t)$ is a stationary process, (2) the information on $X(t)$ consists of samples of this process, and (3) the time interval $[0, \tau]$ is sufficiently long in a sense defined later in the paper. We develop estimates for $p_s(\tau;z)$ and $p_f(\tau;z)$ under these assumptions and construct confidence intervals on these estimates. The accuracy of the estimates depends essentially on properties of $X(t)$, available samples of these processes, and the length of the reference time τ.

An alternative formulation of the reliability problem in Eq. (1) is

$$p_s(z;\tau) = P(Z(t) \le z, 0 \le t \le \tau) = P(Z_\tau \le z), \tag{2}$$

where

$$Z(t) = g(X(t)), \quad t \ge 0, \tag{3}$$

and $Z_\tau = \max_{0 \le t \le \tau}\{Z(t)\}$. The latter formulation is adequate for our objective and is used exclusively in the paper. The construction of estimates of $p_s(z;\tau)$ based on Eq. (2) is simpler since it involves concepts of the univariate extreme value theory. In contrast, estimates of this probability based on the formulation in Eq. (1) involve elements of the multivariate extreme value theory.

3. Estimates of system reliability

Generalized extreme value distributions are fitted to data and used to approximate the law of Z_τ and the probabilities $p_s(z;\tau)$ and $p_f(z;\tau) = 1 - p_s(z;\tau)$. The type of extreme value distribution used for the law of Z_τ does not have to be specified. It is selected by the estimation procedure.

Samples of system state are generated from the defining equation of $X(t)$ by Monte Carlo simulation. If the law of the stationary system state is known, stationary samples result by assuming that the initial state $X(0)$ follows the stationary distribution of $X(t)$. Otherwise, stationary samples of $X(t)$ are produced in two steps. First, samples of $X(t)$ are generated in a time interval $[0, \tau + \tau']$ starting from an arbitrary initial condition, where $\tau' > 0$ is such that transients do not extend beyond τ'. Second, sections of these samples during the time interval $[\tau', \tau + \tau']$ are kept and viewed as stationary samples of $X(t)$.

Monte Carlo estimates of $p_s(z;\tau)$ and $p_f(z;\tau)$ are not feasible when dealing with realistic, highly reliable dynamic systems since they require large numbers of samples of $Z(t)$ and the generation of a single sample of this process is likely to be computationally demanding. Moreover, these estimates are only available in the data range. In contrast to Monte Carlo estimates, the estimates of $p_s(z;\tau)$ and $p_f(z;\tau)$ by the proposed method require relatively small numbers of samples of $Z(t)$ and extend beyond data as illustrated by the numerical examples in Section 4. Block maxima and threshold models are used to implement the proposed estimates of $p_s(z;\tau)$ and $p_f(z;\tau)$.

Exact, GEV and GP estimates, and asymptotic approximations of $p_f(z;m)$ are reported whenever available, e.g., exact failure probabilities are reported only for iid series. GEV and GP estimates of $p_f(z;m)$ are derived from GEV and GP distributions fitted to observations, respectively. Asymptotic approximations are theoretical distributions of extremes of $Z(t)$ corresponding to infinite reference times.

3.1. Block maxima model

Estimates of the distribution of Z_τ are constructed from independent samples $z_{\tau,i} = \max_{0 \le t \le \tau}\{z_i(t)\}$, $i = 1, \dots, n_b$, of the random variable $Z_\tau = \max_{0 \le t \le \tau}\{Z(t)\}$. Resulting estimates are used to approximate system reliability and failure probability.

As previously stated, it is assumed that the exposure time τ is sufficiently long such that Z_τ can be assumed to follow a generalized extreme value distribution approximately. Otherwise, approximations of $p_s(z;\tau)$ based on the assumption that Z_τ follows a GEV distribution will be biased irrespective of the size of n_b. Since the samples of $Z(t)$ are recorded at discrete times, the available information consists of independent samples of time