Importance Sampling for the Estimation of the Failure Probability of the Degradation Process*

Oleg Lukashenko[1,2] and Alexandra Borodina[1,2]

[1] Institute of Applied Mathematical Research of the Karelian Research Centre of RAS, Petrozavodsk, Russia
[2] Petrozavodsk State University, Petrozavodsk, Russia
{lukashenko,borodina}@krc.karelia.ru

Abstract. The estimation of the failure probability is an important and hard problem, arising in the optimal control of unreliable systems. The required performance measure is usually not analytically available. Thus, simulation becomes the most adequate technique to evaluate a rare event failure probability. We develop a variance reduction technique to estimate a small failure probability in a special unreliable system with repairs. A few variants of the importance sampling method are adapted to estimate the quantity of interest.

Keywords: Importance Sampling, Variance Reduction, Failure Probability, Reliability Analysis, Degradation Process

1 Introduction

Currently, research in the field of the technical system reliability is associated with rather complex models of degradation processes, for which it is difficult to obtain analytical or heuristic solutions. An important feature of assessing the parameters of the degradation process or the probability of system failure is the varying rate of degradation and failures with a natural or predetermined threshold level, which is usually a random variable [8,15]. The key point in the analysis of complex shock models is the transition to simulation methods and *naive Monte Carlo method* is still popular in a large number of modern works (see [5,6,11,16,17]), despite its widely known inefficiency for the study of highly reliable systems. Speed-up simulation methods and variance reduction techniques provide a more efficient way of investigating models associated with rare events than the naive Monte Carlo method.

In case of applying variance reduction techniques for the analysis of the degradation models the non-standard formulation of the rare events estimation problem to be considered, when the level of system failure is a random variable.

* The study was carried out under state order to the Karelian Research Centre of the Russian Academy of Sciences (Institute of Applied Mathematical Research KarRC RAS) and supported by the Russian Foundation for Basic Research, projects 18-07-00187, 19-07-00303

In particular, earlier in [3] the variance reduction technique based on Asmussen-Kroese method [1] has been extended to estimate the failure probability that a random sum exceeds a random threshold. This approach provides a decreasing of the relative error for the heavy-tailed degradation stages.

In this work, we describe the different strategies of applying the importance sampling (IS) method for the considered problem of estimation of the rare-event failure probabilities related to a degradation process containing a few steps, in which a maintenance repair is used to prevent a failure. The rest of the paper is organized as follows. In the second section, we recall the basic definitions of the degradation process. Then, in the next third section, the problem of rare-event estimation is formulated. Section 4 provides the description of the importance sampling (IS) estimators adapted to the considered failure probabilities. Finally, numerical results are presented in Section 5.

2 Model Description

Consider the degradation process $X := \{X(t),\, t \geq 0\}$ with a complete failure state F and state space

$$E = \{0, 1, \ldots, L, \ldots, M, \ldots, K; F\},$$

that represents *degradation stages* of the system (see Fig. 1). The states (stages) L, K, M play a specific role and assumed to be fixed. With the help of this model, the dynamics of the state of the anti-corrosion coating was previously studied, and in some cases, an analytical solution for the characteristics of the process is known (see [4]). Nevertheless, this model allows us to develop and test a number of methods for variance reduction and further extend these techniques to more general degradation processes with independent increments.

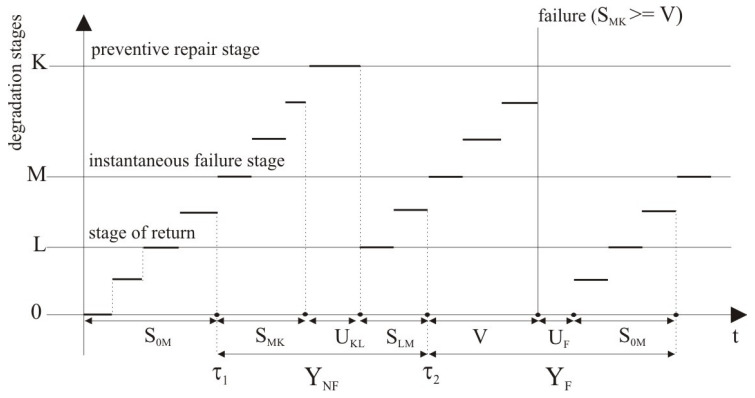

Fig. 1. Realization of the degradation process with two types of cycle

In what follows, we will assume that the two-level policy for managing preventive maintenance and restoration of the system is fixed and the values of (K, L) are known, where K stands for preventive repair stage, after which the process returns to stage L. Consider the possible states of the system, illustrated in Fig. 1. The process starts in state $X(0) = 0$ and then passes $K-1$ intermediate degradation stages.

Two types of failures are possible in the system: *gradual* failure, which corresponds to the next stage transition, and *instantaneous* failure, when the process enters the final state F. Tracking of instantaneous system failures is a key feature for highly reliable systems modeling.

Let T_i be a random length of the ith stage of the degradation, with a given probability density function (pdf) f_{T_i}. (Note that in general, distribution of T_i may depend on the stage number i.)

An instantaneous failure becomes possible starting from stage M, when the system does not have time to go to stage K and, accordingly, the time

$$S_{M,K} = \sum_{i=M}^{K-1} T_i$$

of transition from stage M to stage K exceeds the time to failure V, where V is a random value (r.v.) with given distribution function F_V. As a result, the process X jumps to the state F. Following the transition to the state F, the system again is ready for work from the state 0 after a random repair time U_F with a known distribution. If, on the contrary, the process X enters the state K, then repair is performed during random time $U_{K,L}$, and after that, the process returns to stage L.

Degradation process X is regenerative with two types of regeneration cycles. Denote by Y_F the length of the regeneration cycle with a failure, and let Y_{NF} be the length of the cycle without failure. Thus (unconditional) regeneration cycle length Y can be written as

$$Y = Y_F \cdot I_{\{V \le S_{M,K}\}} + Y_{NF} \cdot I_{\{S_{M,K} < V\}}.$$

where $I_{\{\cdot\}}$ denotes indicator function and cycle lengths are determined as follows:

$$\begin{aligned} Y_F &= V + U_F + S_{0,M} \\ Y_{NF} &= S_{M,K} + U_{K,L} + S_{L,M}, \end{aligned}$$

where r.v. V, U_F, $S_{0,M} = \sum_{j=0}^{M-1} T_j$ are independent as well as r.v. $S_{M,K}$, $U_{K,L}$, $S_{L,M} = \sum_{j=L}^{M-1} T_j$.

Since for reliability models the greatest risk comes from the instantaneous failure, then the main target is to find the probability of instantaneous failure within the regeneration cycle, that is

$$\begin{aligned} l &= \mathbb{P}(S_{M,K} \ge V) = \mathbb{E}_{S_{M,K}} \left[\mathbb{E}[I(S_{M,K} \ge V) | S_{M,K}] \right] \\ &= \mathbb{E}_{S_{M,K}} [F_V(S_{M,K})], \end{aligned}$$

where it was taken into account independence of the r. v.'s V and $S_{M,K}$.

Another critically important reliability descriptor of the system quality is a *reliability function* which is defined as

$$R(t) = \mathbb{P}[T > t | X(0) = 0],\ t \geq 0,$$

where T stands for the lifetime of the system. In our model, the mean lifetime can be written as

$$\mathbb{E}[T] = \mathbb{E}[Y_{NF}](\mathbb{E}[N] - 1) + \mathbb{E}[V | V \leq S_{M,K}],$$

where N is the number of cycles completed before the total failure happens. It is easy to see that the mean number of such cycles equals $\mathbb{E}[N] = 1/\ell$.

For the exponential reliability model, the classical evaluation of the function $R(t)$ based on the solution of the Kolmogorov differential equations for the state probabilities, thus the results are obtained in terms of the Laplace transform. However, this approach is not acceptable in the rare failure case in general. Nevertheless, for rare failures, the reliability function asymptotic can be used (see [4]):

$$R(t) \sim e^{-t\frac{\ell}{\mathbb{E}[Y_{NF}]}} \text{ as } \ell \to 0,$$

and the mean lifetime becomes (approximately)

$$\mathbb{E}[T] = \int_0^\infty R(t)dt \approx \frac{\mathbb{E}[Y_{NF}]}{\ell}.$$

If ℓ is close to zero, then the *naive Monte-Carlo* fails, thus the variance reduction methods must be applied.

3 Rare event simulation of the failure probability

In this section, we describe a general problem arising in the estimation of a small probability (the rare event simulation problem). Suppose that $\theta > 0$ is the so-called rarity parameter, namely

$$\ell_\theta := \mathbb{P}(S_{M,K} \geq V_\theta) \to 0, \quad \text{as} \quad \theta \to \infty.$$

The rarity parameter is associated with the parametric set of the distributions V_θ, in particular, one can consider $\theta = \mathbb{E}V$.

Denote by Z_θ an estimator of ℓ_θ, that means $\mathbb{E}Z_\theta = \ell_\theta$. To estimate ℓ_θ by Monte Carlo (MC) simulation, one has to generate the independent and identically distributed (i.i.d) copies of the r. v. Z_θ and calculate the sample mean

$$\hat{\ell}_\theta \ := \ \frac{1}{N} \sum_{n=1}^{N} Z_\theta^{(n)}. \tag{1}$$

The typical measure of the goodness of the estimator is expressed by the *relative error* (RE) defined as follows:

$$\mathrm{RE}\left[\hat{\ell}_\theta\right] := \frac{\sqrt{\mathbb{V}\mathrm{ar}\left[\hat{\ell}_\theta\right]}}{\mathbb{E}\left[\hat{\ell}_\theta\right]}. \tag{2}$$

The simple example is indicator function of the target event, i.e.

$$Z^{\mathrm{MC}} = I(S_{M,K} \geq V),$$

which corresponds to the naive Monte Carlo. It is straightforward to show that

$$\mathrm{RE}\left[\hat{\ell}_\theta^{\mathrm{MC}}\right] \sim \frac{1}{\sqrt{\ell_\theta N}} \quad \text{as} \quad \ell_\theta \to 0,$$

where $a \sim b$ means that $a/b \to 1$.

From the equation above it is clear that the RE of the naive MC estimator diverges for small values of the target probability and the sample size required to get a suitable RE is inversely proportional to $\sqrt{\ell_\theta}$.

There are many rare event simulation techniques [12,9] which aim at modifying the estimator (1) in order to reduce its variance, hence requiring less sample size for the desired accuracy. The best-desired case is the vanishing RE, that is

$$\mathrm{RE}\left[\hat{\ell}_\theta\right] \to 0, \text{ as } \theta \to \infty.$$

In the latter case, the accuracy of the estimation increases in the rarity of the event. However such property is less implementable in practice, and in many situations, the RE grows but slower than in naive MC. Note that for the tail probabilities of the sum of independent subexponential r. v.'s such a property takes place under some additional conditions on the distributions of the r.v. T, when V is a constant. In other words, in this case, $V = \theta$ is a rarity parameter. (For more detail see [7].)

4 Importance Sampling for the Failure Probability Estimation

Importance sampling is a popular method for variance reduction aimed at changing the probability measure, so that the target rare event becomes more likely to occur.

We assume that T_i are i.i.d. with a common pdf f_T, then the pdf of the random vector $(T_M, ..., T_K)$ has the following product form

$$f(x) = f(x_M, .., x_{K-1}) = \prod_{i=M}^{K-1} f_T(x_i).$$

Let further denote

$$h_\theta(x) = F_{V_\theta}\left(\sum_{i=M}^{K-1} x_i\right)$$

and $g(x) = g(x_M, ..., x_{K-1})$ be some proposal density function, then the target probability

$$\ell_\theta = \int h_\theta(x)\frac{f(x)}{g(x)}g(x)dx = \mathbb{E}_g\left[h_\theta(X)\frac{f(X)}{g(X)}\right],$$

$$\widehat{\ell}_\theta^{\,\mathrm{IS}} = \frac{1}{N}\sum_{i=1}^{N} h_\theta(X^i)\frac{f(X^i)}{g(X^i)}, \tag{3}$$

where $X^1, ..., X^N$ is the sample of the random vector with the density g.

4.1 Exponential Twisting

In this subsection, we describe the classical algorithm based on the importance sampling with a *exponential change of measure*. This approach can be applied to a restricted class of problems when the moment generating function

$$Q(\gamma) = \mathbb{E}[e^{\gamma T_i}]$$

exists, that is for the *light-tailed* stages T_i. The classical approach consists in selecting an importance sampling distribution from a family of pdf's

$$f_T^\gamma(x) = \frac{e^{\gamma x}}{Q(\gamma)}f_T(x). \tag{4}$$

Note that in general sampling from the twisting pdf could be technically difficult, but it is possible to use acceptance-rejection method for this aim.

The target probability can be expressed as

$$\ell_\theta = \mathbb{E}_\gamma\left[F_V(S_{M,K})W_\gamma(S_{M,K})\right],$$

where $\mathbb{E}_\gamma$ is the expectation under pdf (4), and then the corresponding likelihood ratio is defined as

$$W_\gamma(S_{M,K}) = \exp\Big\{-\gamma S_{M,K} + (K-M)\ln Q(\gamma)\Big\}.$$

So, given the two sequences of the samples

$$\{T^{(i)} = (T_M^{(i)}, ..., T_{K-1}^{(i)}),\ i = 1, ..., N\},\ \{V^{(i)},\ i = 1, ..., N\},$$

where elements of the random vector $T^{(i)}$ are independent realizations of the r.v. with the twisted pdf f_T^γ defined in (4), one can estimate the failure probability in the following way

$$\hat{\ell}_\theta^{\,\mathrm{IS}} := \frac{1}{N}\sum_{i=1}^{N} F_V(S_{M,K}^{(i)})W_\gamma(S_{M,K}^{(i)}),$$

where $S_{M,K}^{(i)}$ are independent copies of $S_{M,K}$.

To select the twisting parameter γ, adaptive IS is used. Let consider the second moment of the IS estimator:

$$\mathbb{E}\left[I(S_{M,K} \geq V) W_\gamma^2(S_{M,K})\right] \leq (Q(\gamma))^{2(K-M)} \mathbb{E}[e^{-2\gamma V}]$$

It is reasonable to choose γ^* which minimizes given above upper bound. For example, when $T_i \sim \exp(\lambda)$ and $V = const$, one can obtain

$$\gamma^* = \frac{V\lambda - K + M}{V}.$$

In general, such a one-dimensional optimization problem can be solved numerically.

4.2 Laplace approximation of the zero-variance distribution

For the positive function h consider the following distribution

$$g_*(x) = \frac{h_\theta(x) f(x)}{\int h_\theta(x) f(x) dx} = \frac{h_\theta(x) f(x)}{\ell}.$$

It is quite straightforward to show that g_* is the optimal density which provides the zero variance of the estimator.

But in practice, it is not implementable because it requires knowledge of the target quantity ℓ. Nevertheless, given above expression provides an insight into the general form of a good proposal distribution. Following [2] rewrite the intractable optimal proposal:

$$g_*(x) = \frac{1}{\ell} h_\theta(x) f(x) = \frac{1}{\ell} \exp(-E(x)),$$

where

$$E(x) = -\log h_\theta(x) - \log f(x).$$

The Laplace method gives a Gaussian approximation of $g_*(x)$ based on a local expansion around a mode

$$x^* = \operatorname*{argmin}_x E(x).$$

In general, the mode can be calculated by numerical methods. Then a Taylor expansion up to second order around the mode gives

$$E(x) \approx E(x^*) + (x - x^*) \nabla E\big|_{x=x^*} + \frac{1}{2}(x - x^*)^T H (x - x^*).$$

where ∇E is the gradient vector and H is the Hessian matrix evaluated at the mode. This leads to the following Gaussian approximation of the zero-variance distribution:

$$\widetilde{g}_*(x) = \frac{1}{\ell^*} \exp\left(-\frac{1}{2}(x - x^*)^T H (x - x^*)\right) = N(x \,|\, x^*, H^{-1}).$$

Then, having the sample $X^1, ..., X^N$ of the multivariate normal distribution, the corresponding estimator $\widehat{\ell}^{\mathrm{LA}}$ is calculated by formula (3) with

$$g(\cdot) = N(\cdot \,|\, x^*, H^{-1}).$$

4.3 Cross-Entropy Method

The aim is to find the proposal distribution g close to the desired zero-variance distribution g_* in the sense of the Kullback-Leibler divergence

$$\begin{aligned}\mathcal{D}(g_*, g) &= \mathbb{E}_{g_*}\left[\log \frac{g_*(x)}{g(x)}\right] \\ &= \int g_*(x) \log g_*(x) dx - \int g_*(x) \log g(x) dx.\end{aligned}$$

Additionally assume that the nominal pdf f_T is parameterized by a finite dimensional vector $\mathbf{u}$: $f_T(x) = f_T(x; \mathbf{u})$. Let the proposal pdf is $f(\cdot, \mathbf{v})$ for some parameter $\mathbf{v}$ (i. e. belongs to the same parametric class). The CE algorithm is based on finding an optimal parameter $\mathbf{v}^*$:

$$\begin{aligned}\mathbf{v}^* &= \operatorname*{argmin}_{\mathbf{v}} \mathcal{D}(g_*, f(\cdot, \mathbf{v})) \\ &= \operatorname*{argmax}_{v} \mathbb{E}_{\mathbf{w}} \left[h_\theta(X) W(X; \mathbf{u}, \mathbf{w}) \log f(X; \mathbf{v})\right],\end{aligned}$$

where

$$W(X; \mathbf{u}, \mathbf{w}) = \frac{f(X; \mathbf{u})}{f(X; \mathbf{w})}.$$

The given above stochastic optimization problem is usually replaced by its stochastic counterpart:

$$\mathbf{v}^* = \operatorname*{argmax}_{\mathbf{v}} \frac{1}{N_1} \sum_{k=1}^{N_1} h_\theta(X^k) W(X_k, \mathbf{u}, \mathbf{w}) \log f(X_k; \mathbf{v}), \tag{5}$$

where $X^1, .., X^{N_1}$ is the sample from the distribution $f(\cdot; \mathbf{w})$. In some cases the solution of this optimization problem is available in closed form.

The main problem when dealing with the rare-event simulation is that most values of the $h_\theta(X^k)$ are zero. In this case the so-called multi-level CE procedure [10,13,14] can be applied. This procedure constructs the sequence $(\theta_i, \mathbf{v}_i)$ of both rarity and reference parameters, where $\mathbf{v}_i$ is a solution of the problem (5) with previous values of the parameters: $\mathbf{w} = \mathbf{v}_{i-1}$; $h_\theta = h_{\theta_{i-1}}$.

5 Numerical Examples

In this section, we present the simulation analysis of the considered IS estimators. We have used $N = 10000$ replications for all experiments. We compare proposed IS estimators with the naive MC estimator in terms of the relative error (RE) denoted by the formula (2) and the relative experimental error (RER) defined as

$$\mathrm{RER} = \frac{|\widehat{\ell} - \ell|}{\ell} \cdot 100\%,$$

where ℓ is the true value (when available).

5.1 Exponential Twisting

We assume that the stages T_i have exponential distribution with a rate parameter λ. Performance analysis of the IS with exponential twisting estimator is presented in Tables 1, 2. In the first example, we assume that V is uniformly distributed on the interval $(\theta-\delta, \theta+\delta)$, with $\theta > \delta > 0$, in which case θ is the rarity parameter. The following numerical parameters are used: $K - M = 10$; $\lambda = 1$; $\delta = 5$. The second example, where V is exponentially distributed with the rate parameter $1/\theta$, is used as a benchmark for which the closed-form expression is available [4]

$$\ell = 1 - \left(\frac{\lambda}{\lambda + 1/\theta}\right)^{K-M}.$$

The same set of parameters was used as for the first experiment.

Table 1. Performance of the exponential twisting estimator: $T_i \sim exp(\lambda)$, $V \sim \mathcal{U}(\theta - \delta, \theta + \delta)$

θ	$\widehat{\ell}^{\mathrm{MC}}$	$\widehat{\ell}^{\mathrm{IS}}$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{MC}})$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{IS}})$
10	0.4804	0.4827	0.0104	0.0057
15	0.1224	0.1239	0.0267	0.0065
20	0.0151	0.0136	0.0807	0.0108
25	9e-04	8.1e-04	0.3332	0.0144
30	–	3.2e-05	–	0.0172
35	–	9.9e-07	–	0.0196
40	–	2.4e-08	–	0.0221
45	–	4.9e-10	–	0.0244

Numerical results indicate a significant decrease of the RE. Moreover, simulation results are very close to the theoretical ones. However, as it was pointed out above, in general (non-exponential) case the sampling from the twisting distribution could be technically difficult.

Table 2. Performance of the exponential twisting estimator: $T_i \sim exp(\lambda)$, $V \sim exp(1/\theta)$

θ	ℓ	$\widehat{\ell}^{\mathrm{MC}}$	$\widehat{\ell}^{\mathrm{IS}}$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{MC}})$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{IS}})$	$\mathrm{RER}(\widehat{\ell}^{\mathrm{IS}})$
10^3	9.94e-03	1.16e-02	9.92e-03	0.0939	0.0015	0.1898
10^4	9.99e-04	5e-04	1.58e-03	0.4471	0.0015	0.1425
10^5	9.99e-05	–	9.97e-05	–	0.0016	0.2629
10^6	9.99e-06	–	1e-05	–	0.0016	0.0929
10^7	9.99e-07	–	9.98e-07	–	0.0015	0.1689
10^8	9.99e-08	–	1e-07	–	0.0016	0.3619

Table 3. Performance of the estimators: Weibull stages T_i, $V \sim \mathcal{U}(\theta - \delta, \theta + \delta)$

γ	$\widehat{\ell}^{\mathrm{MC}}$	$\widehat{\ell}^{\mathrm{LA}}$	$\widehat{\ell}^{\mathrm{CE}}$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{MC}})$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{LA}})$	$\mathrm{RE}(\widehat{\ell}^{\mathrm{CE}})$
20	2.3e-04	2.8e-04	2.5e-04	0.2493	0.0480	0.0293
21	9.5e-05	5.9e-05	6.4e-05	0.3831	0.0298	0.0231
22	2.7e-05	1.5e-05	1.5e-05	0.6246	0.0436	0.0261
23	–	3.4e-06	3.5e-06	–	0.0448	0.0289
24	–	6.1e-07	7.2e-07	–	0.0304	0.0397
25	–	1.2e-07	1.3e-07	–	0.0355	0.0344
26	–	2.0e-08	2.3e-08	–	0.0256	0.0397
27	–	3.4e-09	3.5e-09	–	0.0332	0.0455
28	–	5.3e-10	5.3e-10	–	0.0428	0.0763
29	–	7.1e-11	7.3e-11	–	0.0363	0.0852
30	–	9.5e-12	9.2e-12	–	0.0305	0.0718

5.2 Approximation of the Zero-Variance Distribution

In this subsection, two strategies of the approximation of zero-variance distribution are compared, namely the Laplace approximation (LA) and the cross-entropy (CE) algorithm.

We assume that the stage T has Weibull distribution,

$$F_T(t) = 1 - \exp(-(t/b)^a), \quad t \geq 0,$$

where $a > 0$ is a shape parameter and $b > 0$ is a scale parameter. In the experiments we put $b = 1$.

Whereas the implementation of the Laplace approximation is quite straightforward, the CE procedure requires additional specification of the parameters, namely the reference parameter $\mathbf{u}$ of the nominal pdf $f(\cdot; \mathbf{u})$ and the number of replications N_1 in the stochastic counterpart optimization problem (5). It is assumed that the nominal pdf is parametrized by the set of the scale parameters, thus

$$\mathbf{u} = (u_M, ..., u_{K-1}) = (1, ..., 1).$$

The CE algorithm attempts to find the new set of the scale parameters

$$\mathbf{v} = (v_M, ..., v_{K-1})$$

which provides the optimum to (5). Under such a parametrization the closed form solution of (5) is available (see [13] for more details).

The simulation experiments were performed with the following values of the parameters: $K - M = 10$; $\delta = 5$; $N_1 = 100$. The numerical results presented in the Table 3 show that both LA and CE algorithms provide a significant decrease of the RE. It seems that the LA procedure slightly outperforms the CE for sufficiently small probabilities but we also note that the RE is also estimated from generated samples, thus is a r. v. itself. Furthermore, increasing the number of samples N_1 leads to further decreasing of the RE.

6 Conclusion

In this paper, we consider a variance reduction technique based on the importance sampling with the different strategies of choosing the proposal distribution, to estimate the failure probability of the degradation process. A few numerical examples were presented to demonstrate the properties and accuracy of the estimators. It is assumed in future research, to study the performance of the proposed estimator by both numerical and analytic methods and combine it with the advanced acceleration techniques based on the splitting of the trajectories. As a possible extension one can consider the so-called heterogeneous case when the stages are not identically distributed.

References

1. Asmussen, S., Kroese, D.P.: Improved algorithms for rare event simulation with heavy tails. Advances in Applied Probability **38**, 545–558 (2006)
2. Barber, D.: Bayesian Reasoning and Machine Learning. Cambridge University Press (2012)
3. Borodina, A., Lukashenko, O., Morozov, E.: On conditional monte carlo for the failure probability estimation. In: Proceedings of 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems (ICUMT 2018). pp. 202–207. IEEE (2018)
4. Borodina, A.V., Efrosinin, D.V., Morozov, E.V.: Application of splitting to failure estimation in controllable degradation system. In: Communications in Computer and Information Science. vol. 700, pp. 217–230. Springer International Publishing (2017). https://doi.org/10.1007/978-3-319-66836-9, `http://www.springer.com/gb/book/9783319668352`
5. Dong, Q., Cui, L.: A study on stochastic degradation process models under different types of failure thresholds. Reliability Engineering & System Safety **181**, 202–212 (2019)
6. Gao, H., Cui, L., Kong, D.: Reliability analysis for a wiener degradation process model under changing failure thresholds. Reliability Engineering & System Safety **171**, 1–8 (2018)
7. Hartinger, J., Kortschak, D.: On the efficiency of the Asmussen–Kroese-estimator and its application to stop-loss transforms. Blätter der DGVFM **30**, 363–377 (2009)
8. Jensen, H., Papadimitriou, C.: Reliability analysis of dynamical systems. Substructure Coupling for Dynamic Analysis pp. 69–111 (2019)

9. Kroese, D.P., Taimre, T., Botev, Z.I.: Handbook of Monte Carlo Methods. John Wiley & Sons (2011)
10. Kroese, D.P., Rubinstein, R.Y., Glynn, P.W.: Chapter 2 - the cross-entropy method for estimation. In: Rao, C., Govindaraju, V. (eds.) Handbook of Statistics, Handbook of Statistics, vol. 31, pp. 19 – 34. Elsevier (2013). https://doi.org/https://doi.org/10.1016/B978-0-444-53859-8.00002-3
11. Lin, Y.H., Li, Y.F., Zio, E.: A comparison between monte carlo simulation and finite-volume scheme for reliability assessment of multi-state physics systems. Reliability Engineering & System Safety **174**, 1–11 (2018)
12. Ross, S.M.: Simulation. Elsevier (2006)
13. Rubinstein, R.Y., Kroese, D.P.: The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte Carlo Simulation and Machine Learning. Springer-Verlag, New York (2004)
14. Rubinstein, R.Y., Kroese, D.P.: Simulation and the Monte Carlo method. John Wiley & Sons, Inc., New Jersey (2017)
15. Shahraki, A., Yadav, O.P., Liao, H.: A review on degradation modelling and its engineering applications. International Journal of Performability Engineering **13**(3), 299–314 (2017)
16. Sun, Q., Ye, Z.S., Zhu, X.: Managing component degradation in series systems for balancing degradation through reallocation and maintenance. IISE Transactions **52**(7), 797–810 (2020)
17. Yousefi, N., Coit, D.W., Song, S., Feng, Q.: Optimization of on-condition thresholds for a system of degrading components with competing dependent failure processes. Reliability Engineering & System Safety **192**, 106547 (2019)