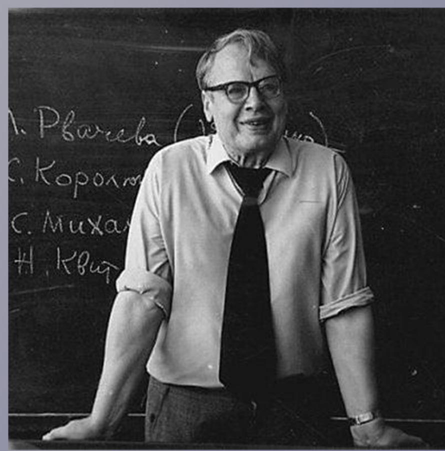
ELECTRONIC JOURNAL
OF INTERNATIONAL GROUP
ON RELIABILITY

Gnedenko Forum Publications



JOURNAL IS REGISTERED
IN THE LIBRARY
OF THE U.S. CONGRESS

•

ISSN 1932-2321

•

VOL.8 NO.1 (28)
MARCH, 2013

RELIABILITY:
THEORY&APPLICATIONS

RELIABILITY
RISK ANALYSIS
MAINTENANCE
SAFETY

Vol.8
No.1 (28)
•
March
2013

San Diego

ISSN 1932-2321

http://www.gnedenko-forum.org/Journal/index.htm

RELIABILITY:
THEORY & APPLICATIONS

Vol.8 No.1 (28),
March, 2013

San Diego
2013

Send your paper

e-Journal ***Reliability: Theory & Applications*** publishes papers, reviews, memoirs, and bibliographical materials on Reliability, Quality Control, Safety, Survivability and Maintenance.

Theoretical papers have to contain new problems, finger practical applications and should not be overloaded with clumsy formal solutions.

Priority is given to descriptions of case studies.

General requirements for presented papers

1. Papers have to be presented in English in MSWord format. (Times New Roman, 12 pt , 1.5 intervals).
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper has to be spell-checked.
4. For those whose language is not English, we kindly recommend to use professional linguistic proofs before sending a paper to the journal.

* * *

The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Publication in this e-Journal is equal to publication in other International scientific journals.

Papers directed by Members of the Editorial Boards are accepted without referring.

The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Send your papers to

the Editor-in-Chief ,
Igor Ushakov
igusha22@gmail.com

or

the Deputy Editor,
Alexander Bochkov
a.bochkov@gmail.com

Table of Contents

Block and Basu (1974) proposed an absolutely continuous bivariate exponential distribution whose marginals are weighted average of exponentials. Chandrasekar and Sajesh (2010) considered location, scale and location-scale families arising out of absolutely continuous bivariate exponential (ACBVE) distribution with equal marginals and derived the minimum risk equivariant estimators of location, scale and location-scale parameters. In this paper we consider a two-component system when failure times follow location-scale ACBVE distribution with equal marginals. We obtain the reliability performance measures of two-component parallel, series and standby systems. Also we provide the UMVUE, the MLE and the MREE of these reliability performance measures.

The paper considers using renewable wind energy for electricity generating. The system is characterized by high reliability and ecological purity. The authors briefly present the main methodological principles of choosing the parameters of the considered wind-hydro power system.

In Accelerated life testing if the accelerated test stress level is not high enough then many of the test items will not fail during the available time and one has to be prepared to handle a lot of censured data. To avoid such type of problems, a better way is step-stress ALT. In Step-stress ALT all test items are first tested at a specified constant stress for a specified period of time and then Items which are not failed will be tested at next higher level of stress for another specified time and so on until all items have failed or the test stops for other reasons. In this paper simple step stress pattern of ALT assuming that the lifetime of a product at any constant level of stress follow a two parameter Pareto distribution is considered. The maximum likelihood and asymptotic confidence interval estimate of the parameters are obtained. Optimal step stress ALT plan is proposed by minimizing the asymptotic variance of the MLE of the 100 Pth percentile of the lifetime distribution at normal stress condition. A simulation study is also performed to analyse the performance of parameter estimates.

Method of Universal Generating Function (UGF) was introduced in many papers. Here we give an example how method of UGF can be implemented to solution problems of optimal redundancy for systems consisting of multi-level units.

Mean residual life criteria of first passage time of semi-Markov process is considered. Properties of transition probability functions when using scaled Total Time on Test (TTT) transform for some criteria of mean residual life are discussed. Application to Multistate reliability system is also addressed.

Smagin V.A. the brief review of a history of introduction of delta - function on a complex plane. The proof of the mathematical form of complex delta - function is given. The example of application of complex delta - function for a presence of stationary value alternation casual process of accumulation with the information income and the charge is given.

The paper discusses the processes of inertial reacting and self-regulation of the environment impacted by hazards of railway accidents involving dangerous goods and the queuing system Markovian model is proposed to determine the probable consequences of such accidents development.

Inliers in a data set are subset of observations not necessarily all zeroes, which appears to be inconsistent with the remaining data set. They are either the resultant of instantaneous or early failures usually encountered in life testing, financial, clinical trial and many other studies. We study the estimation of inliers in Normal distribution. The masking effect problem for correctly identifying the inliers is also discussed. An illustration and a real life example is presented with detailed discussions.

This paper aims at bringing out the usefulness of Chebyshev- and Markov- type inequalities in structural engineering design decision making. By examining whether the bounds arising from Chebyshev - type inequality (associated with these are weak upper bound probabilities) enclose the respective experimental values for deflections of six ferrocement I-beams and web shear fatigue life of a steel plate girder it is inferred that the bounds and the associated probabilities estimated are realistic and hence can be used in structural engineering design decision making. The paper also presents some recent developments in application of Markov type inequalities (which are due to Steliga and Szynal (2010)) for estimation of bounds on probability of an event sought. The importance of such bounds in structural engineering applications is brought out. It is shown from the results of Monte Carlo simulation that the bounds on probability of an event, estimated using the method presented by Steliga and Szynal, are sharp. One of the important advantages of the bounds presented by Steliga and Szynal (2010) is that the original (hidden/internal) random variable need not have well defined moments. Possible engineering applications are also pointed out.

RELIABILITY MEASURES OF SYSTEMS WITH LOCATION-SCALE ACBVE COMPONENTS

B. Chandrasekar[1] and T.A. Sajesh[2]

•

[1]Loyola College, Chennai, India
[2]St. Thomas' College, Thrissur, India

e-mail: bchandrasekar2003@yahoo.co.in,
sajesh.abraham@yahoo.com

ABSTRACT

Block and Basu (1974) proposed an absolutely continuous bivariate exponential distribution whose marginals are weighted average of exponentials. Chandrasekar and Sajesh (2010) considered location, scale and location-scale families arising out of absolutely continuous bivariate exponential (ACBVE) distribution with equal marginals and derived the minimum risk equivariant estimators of location, scale and location-scale parameters. In this paper we consider a two-component system when failure times follow location-scale ACBVE distribution with equal marginals. We obtain the reliability performance measures of two-component parallel, series and standby systems. Also we provide the UMVUE, the MLE and the MREE of these reliability performance measures.

Keywords and phrases: Bivariate exponential, location – scale, reliability measures, two-component system.

1 Introduction

Evaluating performance measures associated with systems having dependent component failure times is rare in the literature. In this paper we consider a two-component system when failure times follow location – scale absolutely continuous bivariate exponential (ACBVE) distribution with equal marginals. Chandrasekar and Sajesh (2010) discussed about the equivariant estimation for parameters of location-scale exponential models. We obtain the reliability performance measures of the two component systems. Also we discuss the estimation of these reliability performance measures.

The plan of the paper is as follows: Section 2 provides some definitions and notations required in this paper. Some distributional results are discussed in Section 3. In Section 4 we obtain the reliability performance measures of two-component parallel, series and standby systems when the component failure times follow location – scale ACBVE with equal marginals. Section 5 provides the UMVUE, the MLE and the MREE of the reliability performance measures.

2 Preliminaries

Consider a two component system with failure times T_1 and T_2 respectively. Assume that (T_1, T_2) follows location – scale ACBVE $(\alpha, \beta, \xi, \tau)$ with pdf

$$f(t_1, t_2; \alpha, \beta, \xi, \tau) = \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2} \exp\left[-\frac{1}{\tau}\{\alpha(t_1+t_2)+\beta(t_1 \vee t_2)-(2\alpha+\beta)\xi\}\right],$$
$$\alpha, \beta \text{ fixed}, \xi \in R, \tau > 0; t_1 \wedge t_2 > \xi. \quad (1)$$

It is assumed that (α, β) is known.

Suppose we observe n identical systems with observations (t_{1i}, t_{2i}), i = 1, 2, 3,…, n. Then the joint pdf of the sample is

$$p_{\xi,\tau}(t_1, t_2) = \left\{\frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\right\}^n \exp\left[-\frac{1}{\tau}\sum_i \{\alpha(t_{1i}+t_{2i})+\beta(t_{1i} \vee t_{2i})-(2\alpha+\beta)\xi\}\right],$$
$$\underset{i}{\mathrm{Min}}(t_{1i} \wedge t_{2i}) > \xi; \ \xi \in R, \tau > 0. \quad (2)$$

It is easy to observe that the maximum likelihood estimator (MLE) of (ξ, τ) is given by $\left(\hat{\xi}, \hat{\tau}\right)$, where

$\hat{\xi} = \underset{i}{\mathrm{Min}}(t_{1i} \wedge t_{2i})$ and $\hat{\tau} = \frac{1}{n}\sum_i \left\{\alpha(t_{1i} + t_{2i}) + \beta(t_{1i} \vee t_{2i}) - (2\alpha+\beta)\hat{\xi}\right\}$.

3 Distributional results

Theorem 1: Let (T_1, T_2) follow location – scale ACBVE (α, β, ξ, τ) with pdf given in (1). Then

$(T_1 \wedge T_2) \sim E\left(\xi, \frac{\tau^2}{2\alpha+\beta}\right)$. Here E (a, b) refers to a random variable with pdf $\left(\frac{1}{b}\right)\exp\left\{-\frac{1}{b}(x-a)\right\}$, $x > a;\ a \in R,\ b > 0$.

Proof: For $u > \xi$,

$$P_\xi[T_1 \wedge T_2 > u] = P_\xi[T_1 > u, T_2 > u]$$
$$= \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\int_u^\infty\int_u^\infty \exp\left[-\left(\frac{1}{\tau}\right)\{\alpha(x+y) + \beta(x \vee y) - (2\alpha+\beta)\xi\}\right] dt_1\, dt_2 \qquad (3)$$

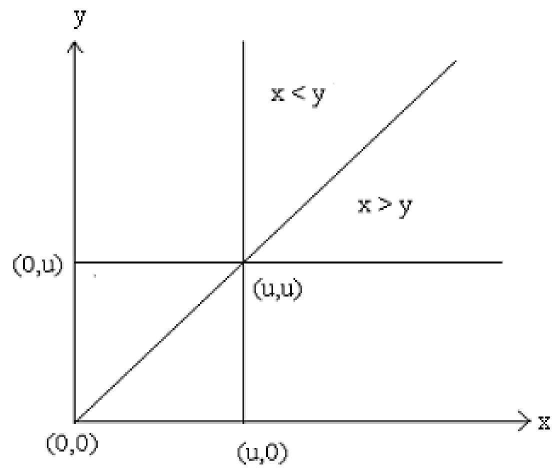


Figure 1

Using Figure 1, equation (3) can be written as,

$$P_\xi[T_1 > u, T_2 > u] = \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\int_u^\infty\left\{\int_u^{t_1}\exp\left[-\left(\frac{1}{\tau}\right)\{\alpha(t_1+t_2) + \beta t_1 - (2\alpha+\beta)\xi\}\right]dt_2\right\}dt_1$$
$$+ \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\int_u^\infty\left\{\int_{t_1}^\infty\exp\left[-\left(\frac{1}{\tau}\right)\{\alpha(t_1+t_2) + \beta t_2 - (2\alpha+\beta)\xi\}\right]dt_2\right\}dt_1$$
$$= \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\left\{\frac{\tau}{\alpha}\int_u^\infty\exp\left[-\left(\frac{1}{\tau}\right)\{(\alpha+\beta)t_1 + \alpha u - (2\alpha+\beta)\xi\}\right]dt_1\right\}$$
$$-\frac{\tau}{\alpha}\int_u^\infty\exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(t_1-\xi)\right\}dt_1 + \frac{\tau}{(\alpha+\beta)}\int_u^\infty\exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(t_1-\xi)\right\}dt_1\Bigg]$$
$$= \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\left[\frac{\tau^2}{\alpha(\alpha+\beta)}\exp\{-(2\alpha+\beta)(u-\xi)\} - \frac{\tau^2}{\alpha(2\alpha+\beta)}\exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(u-\xi)\right\}\right.$$
$$\left. + \frac{\tau^2}{(\alpha+\beta)(2\alpha+\beta)}\exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(u-\xi)\right\}\right]$$

$$= \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\left[\exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(u-\xi)\right\}\right]\left[\frac{\tau^2\{(2\alpha+\beta)-(\alpha+\beta)+\alpha\}}{\alpha(\alpha+\beta)(2\alpha+\beta)}\right]$$

$$= \exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(u-\xi)\right\}.$$

Therefore, $P_\xi\,[T_1 \wedge T_2 > u] = \exp\left\{-\left(\frac{1}{\tau}\right)(2\alpha+\beta)(u-\xi)\right\},\ u > \xi.$

Hence $(T_1 \wedge T_2) \sim E\left(\xi, \frac{\tau}{2\alpha+\beta}\right)$.

Theorem 2: Let (T_1, T_2) follow location – scale ACBVE $(\alpha, \beta, \xi, \tau)$ with pdf given in (1). Then

(i) $T_1 + T_2 - 2\xi \overset{d}{=} U_1 + U_2$,

(ii) $T_1 \vee T_2 - \xi \overset{d}{=} U_1 + U_3$ and

(iii) $T_1 \wedge T_2 - \xi \overset{d}{=} U_3$,

where $U_1 \sim E\left(0, \frac{\tau}{\alpha+\beta}\right)$, $U_2 \sim E\left(0, \frac{2\tau}{2\alpha+\beta}\right)$ and $U_3 \sim E\left(0, \frac{\tau}{2\alpha+\beta}\right)$.

Proof: MGF of $(T_1 + T_2, T_1 \vee T_2)$ at (u_1, u_2) is

$m(u_1, u_2) =$

$$\frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\int_\xi^\infty\int_\xi^\infty \exp\left[-\frac{1}{\tau}\{(\alpha-u_1\tau)(t_1+t_2)+(\beta-u_2\tau)(t_1\vee t_2)-(2\alpha+\beta)\xi\}\right]dt_1\,dt_2$$

$$= \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2}\,\frac{2\tau^2\exp\left[\left\{\left(\frac{2\alpha+\beta}{\tau}\right)\xi - \frac{2(\alpha-u_1\tau)+(\beta-u_2\tau)}{\tau}\xi\right\}\Big/\tau\right]}{\{(\alpha-u_1\tau)+(\beta-u_2\tau)\}\{2(\alpha-u_1\tau)+(\beta-u_2\tau)\}}$$

$$= \left(1-\frac{u_1+u_2}{\alpha+\beta}\tau\right)^{-1}\left(1-\frac{2u_1+u_2}{2\alpha+\beta}\tau\right)^{-1}\exp\{(2u_1+u_2)\xi\}.$$

(i) $m(u_1, 0) = \left(1-\frac{u_1}{\alpha+\beta}\tau\right)^{-1}\left(1-\frac{2u_1}{2\alpha+\beta}\tau\right)^{-1}\exp(2u_1\xi)$.

This implies that $T_1 + T_2 - 2\xi \overset{d}{=} U_1 + U_2$, where $U_1 \sim E\left(0, \frac{\tau}{\alpha+\beta}\right)$,

$U_2 \sim E\left(0, \frac{2\tau}{2\alpha+\beta}\right)$ and $U_1 \coprod U_2$.

(ii) $m(0, u_2) = \left(1-\frac{u_2}{\alpha+\beta}\tau\right)^{-1}\left(1-\frac{u_2}{2\alpha+\beta}\tau\right)^{-1}\exp(u_2\xi)$.

This implies that $T_1 \vee T_2 - \xi \overset{d}{=} U_1 + U_3$, where $U_1 \sim E\left(0, \frac{\tau}{\alpha+\beta}\right)$, $U_3 \sim E\left(0, \frac{\tau}{2\alpha+\beta}\right)$

and $U_1 \coprod U_3$.

(iii) From Theorem 1, we have,

$T_1 \wedge T_2 - \xi \overset{d}{=} U_3$, where $U_3 \sim E\left(0, \frac{\tau}{2\alpha+\beta}\right)$.

4 Reliability performance measures

4.1 Parallel system

MTBF: Consider a two-unit parallel system with component failure times T_1 and T_2 respectively. Then the system failure time is

$$T = T_1 \vee T_2.$$

Assume that $(T_1, T_2) \sim$ location – scale ACBVE $(\alpha, \beta, \xi, \tau)$. Then from Theorem 2, we have

$$T - \xi \overset{d}{=} U_1 + U_3.$$

Therefore,
$$\begin{aligned} \text{MTBF} &= E(T) \\ &= E(U_1 + U_3) + \xi \\ &= \frac{\tau}{\alpha+\beta} + \frac{\tau}{2\alpha+\beta} + \xi \\ &= \frac{3\alpha+2\beta}{(\alpha+\beta)(2\alpha+\beta)}\tau + \xi. \end{aligned} \tag{4}$$

Reliability function: Consider, $P(T - \xi > t),\ t > 0.$
$= P(U_1 + U_3 > t),\ t > 0.$

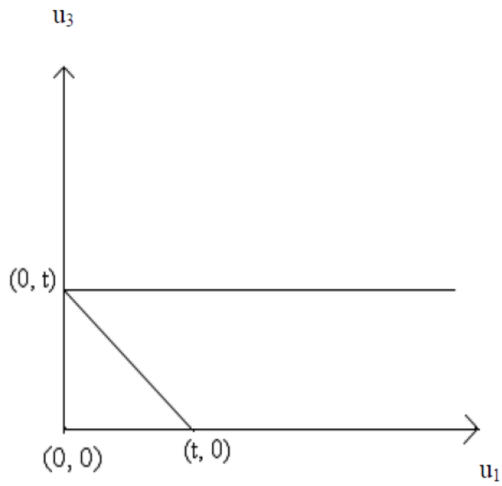


Figure 2

From figure 2 we have,

$$P(T - \xi > t) = \int_0^t \int_{t-u_3}^{\infty} \frac{(\alpha+\beta)(2\alpha+\beta)}{\tau^2} \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)u_1 - \left(\frac{2\alpha+\beta}{\tau}\right)u_3\right\} du_1 du_3$$

$$+ \int_t^{\infty}\int_0^{\infty} \frac{(\alpha+\beta)(2\alpha+\beta)}{\tau^2} \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)u_1 - \left(\frac{2\alpha+\beta}{\tau}\right)u_3\right\} du_1 du_3$$

$$= \int_0^t \frac{(2\alpha+\beta)}{\tau} \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)(t-u_3) - \left(\frac{2\alpha+\beta}{\tau}\right)u_3\right\} du_3\, u_1$$

$$+ \int_t^{\infty} \frac{(2\alpha+\beta)}{\tau} \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)u_3\right\} du_3$$

$$= \left(\frac{2\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)t\right\}\left\{1 - \exp\left(-\frac{\alpha}{\tau}t\right)\right\} + \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)t\right\}$$

$$= \left(\frac{2\alpha+\beta}{\alpha}\right)\left[\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)t\right\} - \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)t\right\}\right] + \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)t\right\}$$

$$= \left(\frac{2\alpha+2\beta}{\alpha}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)t\right\} - \left(\frac{\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)t\right\},\ t > 0.$$

Therefore,

$$P\,(T - \xi > t - \xi) = \left(\frac{2\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)(t-\xi)\right\} - \left(\frac{\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)(t-\xi)\right\},\ t > \xi.$$

Then, the reliability function is given by

$$R\,(t) = P(T > t) = \left(\frac{2\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)(t-\xi)\right\} - \left(\frac{\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)(t-\xi)\right\},\ t > \xi. \qquad (5)$$

4.2 Series system

MTBF: Consider a two unit series system with component failure times T_1 and T_2 respectively. Then the system failure time is

$$T = T_1 \wedge T_2.$$

Assume that $(T_1, T_2) \sim$ location – scale ACBVE $(\alpha, \beta, \xi, \tau)$. Then from Theorem 2, we have

$$T - \xi \overset{d}{=} U_3.$$

Therefore, MTBF = E (T)

$$= E\,(U_3) + \xi$$

$$= \frac{\tau}{2\alpha+\beta} + \xi. \qquad (6)$$

Reliability function: Consider, $P\,(T - \xi > t),\ t > 0$

$$= P\,(U_3 > t),\ t > 0.$$

$$= \int_t^\infty \left(\frac{2\alpha+\beta}{\tau}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)u_3\right\}du_3$$

$$= \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)t\right\},\ t > 0.$$

Therefore,

$$P\,(T - \xi > t - \xi) = \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)(t-\xi)\right\},\ t > \xi.$$

Then, the reliability function is given by

$$R\,(t) = P\,(T > t) = \exp\left\{-\left(\frac{2\alpha+\beta}{\tau}\right)(t-\xi)\right\},\ t > \xi. \qquad (7)$$

4.3 Standby system

MTBF: Consider a two unit standby system with component failure times T_1 and T_2 respectively. Then the system failure time is

$$T = T_1 + T_2.$$

Assume that $(T_1, T_2) \sim$ location – scale ACBVE $(\alpha, \beta, \xi, \tau)$. Then from Theorem 3.2, we have

$$T - 2\xi \overset{d}{=} U_1 + U_2$$

Therefore, MTBF = E (T)

$$= E(U_1 + U_2) + 2\xi$$
$$= \frac{\tau}{\alpha+\beta} + \frac{2\tau}{2\alpha+\beta} + 2\xi$$
$$= \frac{4\alpha+3\beta}{(\alpha+\beta)(2\alpha+\beta)}\tau + 2\xi. \qquad (8)$$

Reliability function: Consider, $P(T - 2\xi > t),\ t > 0.$
$= P(U_1 + U_2 > t),\ t > 0.$

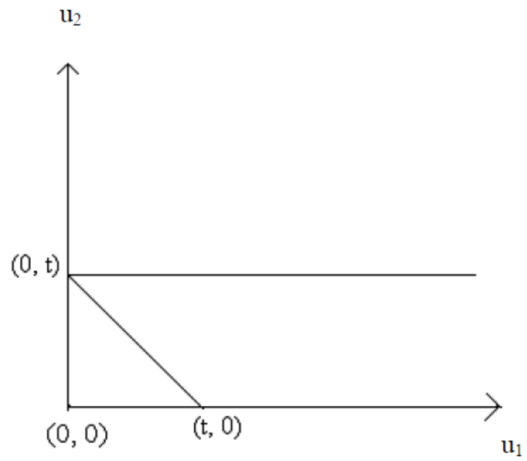


Figure 3

From figure 3 we have,

$$P(T - 2\xi > t) = \int_0^t \int_{t-u_2}^{\infty} \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2} \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)u_1 - \left(\frac{2\alpha+\beta}{2\tau}\right)u_2\right\} du_1 du_2$$
$$+ \int_t^{\infty}\int_0^{\infty} \frac{(\alpha+\beta)(2\alpha+\beta)}{2\tau^2} \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)u_1 - \left(\frac{2\alpha+\beta}{2\tau}\right)u_2\right\} du_1 du_2$$
$$= \int_0^t \frac{(2\alpha+\beta)}{2\tau} \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)(t-u_2) - \left(\frac{2\alpha+\beta}{2\tau}\right)u_2\right\} du_2$$
$$+ \int_t^{\infty} \frac{(2\alpha+\beta)}{2\tau} \exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)u_2\right\} du_2$$
$$= \left(\frac{2\alpha+\beta}{\beta}\right) \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)t\right\}\left\{\exp\left(\frac{\beta}{2\tau}t\right) - 1\right\} + \exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)t\right\}$$
$$= \left(\frac{2\alpha+\beta}{\beta}\right)\left[\exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)t\right\} - \exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)t\right\}\right] + \exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)t\right\}$$
$$= \left(\frac{2\alpha+2\beta}{\beta}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)t\right\} - \left(\frac{2\alpha+\beta}{\beta}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)t\right\},\ t > 0.$$

Therefore,

$$P(T - 2\xi > t - 2\xi) = \left(\frac{2\alpha+2\beta}{\beta}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)(t-2\xi)\right\} - \left(\frac{2\alpha+\beta}{\beta}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)(t-2\xi)\right\},\ t > 2\xi.$$

Then, the reliability function is given by

$$R(t) = P(T > t) = \left(\frac{2\alpha+2\beta}{\beta}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{2\tau}\right)(t-2\xi)\right\} - \left(\frac{2\alpha+\beta}{\beta}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\tau}\right)(t-2\xi)\right\},\ t > 2\xi. \quad (9)$$

5 Estimation of reliability measures

Sajesh (2007) considered location, scale and location-scale families arising out of ACBVE distribution proposed by Block and Basu (1974) with equal marginals and derived MREE, UMVUE and MLE of location, scale and location-scale parameters. In this section we discuss optimal estimation of reliability measures for parallel, series and standby systems.

5.1 Parallel system

UMVUE: The MTBF is $M = \frac{3\alpha+2\beta}{(\alpha+\beta)(2\alpha+\beta)}\tau + \xi$.

From Sajesh [3], the UMVUEs of ξ and τ are $\left\{T_1^* - \frac{T_2^*}{n(2n-1)(2\alpha+\beta)}\right\}$ and $\frac{T_2^*}{(2n-1)}$ respectively, where $T_1^* \sim E\left(\xi, \frac{\tau}{n(2\alpha+\beta)}\right)$ and $T_2^* \sim G\left(\frac{1}{\tau}, 2n-1\right)$.

Hence the UMVUE of M is

$$M^* = \left\{\frac{3\alpha+2\beta}{(\alpha+\beta)(2\alpha+\beta)}\right\}\left(\frac{T_2^*}{2n-1}\right) + \left\{T_1^* - \frac{T_2^*}{n(2n-1)(2\alpha+\beta)}\right\}.$$

$$= T_1^* + \left\{\frac{(3n-1)\alpha+(2n-1)\beta}{n(2n-1)(\alpha+\beta)(2\alpha+\beta)}\right\}T_2^*. \quad (10)$$

MREE: From Chandrasekar and Sajesh (2010), the MREE of $c\xi + d\tau$ is

$$M^{**} = cT_1^* + \frac{1}{2n}\left\{d - \frac{c}{n(2\alpha+\beta)}\right\}T_2^*.$$

Here $c = 1$, $d = \left\{\frac{3\alpha+2\beta}{(\alpha+\beta)(2\alpha+\beta)}\right\}$.

Then $M^{**} = T_1^* + \left\{\frac{(3n-1)\alpha+(2n-1)\beta}{2n^2(\alpha+\beta)(2\alpha+\beta)}\right\}T_2^*. \quad (11)$

MLE: The MLE of (ξ, τ) is $\hat{\xi} = T_1^*$ and $\hat{\tau} = \frac{1}{2n}T_2^*$.

Then the MLE of MTBF is $\hat{M} = \frac{3\alpha+2\beta}{(\alpha+\beta)(2\alpha+\beta)}\hat{\tau} + \hat{\xi}. \quad (12)$

The MLE of the reliability function is

$$\hat{R}(t) = \left(\frac{2\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{\alpha+\beta}{\hat{\tau}}\right)(t-\hat{\xi})\right\} - \left(\frac{\alpha+\beta}{\alpha}\right)\exp\left\{-\left(\frac{2\alpha+\beta}{\hat{\tau}}\right)(t-\hat{\xi})\right\},\ t > \hat{\xi}. \quad (13)$$

5.2 Series system

UMVUE: The MTBF is $M = \frac{\tau}{(2\alpha+\beta)} + \xi$.

From Sajesh (2007), the UMVUEs of ξ and τ are $\left\{T_1^* - \frac{T_2^*}{n(2n-1)(2\alpha+\beta)}\right\}$ and $\frac{T_2^*}{(2n-1)}$ respectively.

Hence the UMVUE of M is

$$M^* = \left\{\frac{1}{(2\alpha+\beta)}\right\}\left(\frac{T_2^*}{2n-1}\right) + \left\{T_1^* - \frac{T_2^*}{n(2n-1)(2\alpha+\beta)}\right\}.$$

$$= T_1^* + \left\{\frac{(n-1)}{(2n-1)(2\alpha+\beta)}\right\}T_2^*. \quad (14)$$

MREE: From Chandrasekar and Sajesh (2010), the MREE of $c\xi + d\tau$ is

$$M^{**} = cT_1^* + \frac{1}{2n}\left\{d - \frac{c}{n(2\alpha+\beta)}\right\}T_2^*.$$

Here c =1, d = $\frac{1}{(2\alpha+\beta)}$.

Then $M^{**} = T_1^* + \left\{\frac{(n-1)}{2n^2(2\alpha+\beta)}\right\}T_2^*. \quad (15)$

MLE: The MLE of (ξ, τ) is $\hat{\xi} = T_1^*$ and $\hat{\tau} = \frac{1}{2n}T_2^* T_2$.

Then the MLE of MTBF is $\hat{M} = \frac{\hat{\tau}}{(2\alpha+\beta)} + \hat{\xi}$. (16)

The MLE of the reliability function is $\hat{R}(t) = \exp\left\{-\left(\frac{2\alpha+\beta}{\hat{\tau}}\right)(t-\hat{\xi})\right\}$, $t > \hat{\xi}$. (17)

5.3 Standby system

UMVUE: The MTBF is $M = \frac{4\alpha+3\beta}{(\alpha+\beta)(2\alpha+\beta)}\tau + 2\xi$.

From Sajesh (2007), the UMVUEs of ξ and τ are $\left\{T_1^* - \frac{T_2^*}{n(2n-1)(2\alpha+\beta)}\right\}$ and $\frac{T_2^*}{(2n-1)}$ respectively.

Hence the UMVUE of M is

$$M^* = \left\{\frac{4\alpha+3\beta}{(\alpha+\beta)(2\alpha+\beta)}\right\}\left(\frac{T_2^*}{2n-1}\right) + 2\left\{T_1^* - \frac{T_2^*}{n(2n-1)(2\alpha+\beta)}\right\}.$$

$$= 2T_1^* + \left\{\frac{(4n-2)\alpha + (3n-2)\beta}{n(2n-1)(\alpha+\beta)(2\alpha+\beta)}\right\}T_2^*. \quad (18)$$

MREE: From Chandrasekar and Sajesh(2010), the MREE of $c\xi + d\tau$ is

$$M^{**} = cT_1^* + \frac{1}{2n}\left\{d - \frac{c}{n(2\alpha+\beta)}\right\}T_2^*.$$

Here c =2, d = $\frac{4\alpha+3\beta}{(\alpha+\beta)(2\alpha+\beta)}$.

Then $$M^{**} = 2T_1^* + \left\{ \frac{(4n-2)\alpha + (3n-2)\beta}{2n^2(\alpha+\beta)(2\alpha+\beta)} \right\} T_2^*. \tag{19}$$

MLE: The MLE of (ξ, τ) is $\hat{\xi} = T_1^*$ and $\hat{\tau} = \frac{1}{2n} T_2^*$.

Then the MLE of MTBF is $$\hat{M} = \frac{4\alpha + 3\beta}{(\alpha+\beta)(2\alpha+\beta)} \hat{\tau} + 2\hat{\xi}. \tag{20}$$

The MLE of the reliability function is

$$\hat{R}(t) = \left(\frac{2\alpha + 2\beta}{\beta} \right) \exp\left\{ -\left(\frac{2\alpha+\beta}{2\hat{\tau}} \right)(t - 2\hat{\xi}) \right\} - \left(\frac{2\alpha+\beta}{\beta} \right) \exp\left\{ -\left(\frac{\alpha+\beta}{\hat{\tau}} \right)(t - 2\hat{\xi}) \right\}, \; t > 2\hat{\xi}. \tag{21}$$

References

Block, H.W. and Basu, A.P. (1974). A continuous bivariate exponential extension. *Journal of American Statistical Association*, 69, 1031 - 1037.

Chandrasekar, B. and Sajesh, T.A. (2010), Equivariant estimation for parameters of location-scale bivariate exponential models. *Journal of Applied Statistical Science*, 17, 541-548.

Sajesh, T.A. (2007). Absolutely continuous bivariate exponential location-scale models. *M.Phil Dissertation submitted to Loyola College (Autonomous), Chennai-600 034, University of Madras, India.*

WIND-HYDRO POWER SYSTEM AS AN EXAMPLE OF DIVERSIFICATION OF DISTRIBUTED GENERATION

G. F. Kovalev, M. A. Rychkov

•

(Institution of the Russian Academy of Sciences Melentiev
Energy Systems Institute of the Siberian Branch of the RAS, Irkutsk, Russia)

e-mail: kovalev@isem.sei.irk.ru

ABSTRACT

The paper considers using renewable wind energy for electricity generating. The system is characterized by high reliability and ecological purity. The authors briefly present the main methodological principles of choosing the parameters of the considered wind-hydro power system.

Key words: distribution generation, wind plant, reservoir, hydropower plant, electricity consumption, electricity supply, reliability, efficiency.

Introduction

Nowadays renewable energy sources attract attention because the depletion of conventional nonrenewable energy (coal, gas, oil, etc.) is getting increasingly obvious. Wind energy is characterized by a considerable potential among the renewable resources.

Human civilizations have been using wind for a long time. In the Ancient times wind was used to propel boats. It is known that even 3000 years BC the citizens of Alexadria had used “wind wheels”. In the 16^{th} century the Netherlands had more than ten thousand wind-driven plants that were used to dry lakes for cultivation area. In 1888 the USA constructed a large wind power plant for electricity production. The multi-blade wind motors invented by the engineer Davydov appeared at the Russian Exhibition in Nizhny Novgorod in 1896 [1]. Wind mills found wide application. In the USSR the first 100 kV wind power plant was built in the Crimea in 1931 and was in operation until World War II.

Currently wind energy is widely used in more than 60 countries of the world. Today 10 leading countries account for about 86% of all wind power capacities installed in the world, of which more than 38% are situated in China and the USA. In Europe wind energy is mostly used in Germany, Denmark, Spain, Portugal, and France. The total installed capacity in the world reached 194 GW [2] in 2011 and continues to soar.

When used as distributed generation, modern wind power plants along with advantages (free primary energy) have some drawbacks:

- lack of regularity and constancy in electricity generation due to variability of wind parameters;
- relatively high cost and low reliability;
- complexity of automated control of wind power plants both in case of their autonomous operation and in case of their operation within a grid;
- environmental problems (noise and allocation of large territories).

Elimination of these drawbacks is associated with additional costs of creating storage devices to replace generation capacities, sophisticated distributed automation of control system of parallel operation of a large number of wind generators “virtual power plant”; and removal of wind power plants from populated settlements to the uninhabited areas.

In this paper the authors take account of all the above circumstances and consider the advantages and disadvantages of a wind-hydro power complex (WHPC) that consists of wind - driven pumps, a storage capacity (water reservoir) and a hydropower plant.

An advantage of the system is its principal simplicity versus other designs of wind power plants and consists mainly in a simple scheme of converting power generated by the wind power plant.

The research aims to find out the conditions to make this system more efficient as compared to the other types of wind power plants. The authors suggest a technique for feasibility study on the efficiency of the wind hydropower system. Moreover, special attention is paid to reliability of power supply to consumers connected to such systems.

The main stages of the technique for calculation of parameters and estimation of the WHPC efficiency include:

1. Study on electricity consumption, load curves and requirements for electricity supply to the existing consumers.
2. Analysis of database on wind conditions (wind speeds and duration) in the studied area.
3. In the case of sufficient wind conditions – study and choice of water sources the most appropriate for the considered local conditions to be used to fill the hydropower plant reservoir with the aid of wind-driven pumps (available nearby water source: sea, lake, river, underground sources, etc.).
4. Determination of a required installed capacity of hydropower plant, characteristics of the main equipment and construction of the hydropower plant, taking account of electricity demand, load curves and reliability.
5. Collection of information on nomenclature and parameters of commercially manufactured hydropower units. Choice of an effective number of units and their rated capacity for concrete conditions.
6. Determination of a required capacity of reservoir and its main characteristics on the basis of local topographic and weather climatic conditions. The reservoir capacity can be increased depending on other economic needs of the region. Calculations of structures and hydro constructions of the reservoir.
7. Determination of the required installed capacity of wind-driven pumps and their characteristics, on the basis of requirements for reservoir filling within a calculation period determined by the wind speeds in this area and reliability requirements.
8. Acquisition of information about nomenclature and parameters of commercially manufactured wind-driven pumps. Choice of an effective number and delivery of the pumps to meet specific conditions, reliable and sufficient to fill the reservoir to the required level. A special order can be placed to manufacture exclusive pumps.
9. Preparation of technical and economic data for comparative estimation of the suggested and alternative variants, including the case of receiving electricity from power grid, and traditional wind power plants with capacities to backup them, etc.
10. Choice of the final variant of electricity supply in the region on the basis of feasibility study of the variants.
11. Solving the other problems related to the construction of WHPS. For example, consideration of possibility of WHPS operation using the reservoir in the low head of hydropower plant (HPP) to pump water from it to the upper reservoir with the aid of wind-driven pumps (a closed cycle scheme), etc.

Brief characteristic of the complex

The flowchart of the WHPC is presented in Figure 1. Its main modules are:

1. Reservoir filling module – wind-driven pumps.
2. Energy storage module – reservoir (water reservoir).
3. Generation module – HPP.

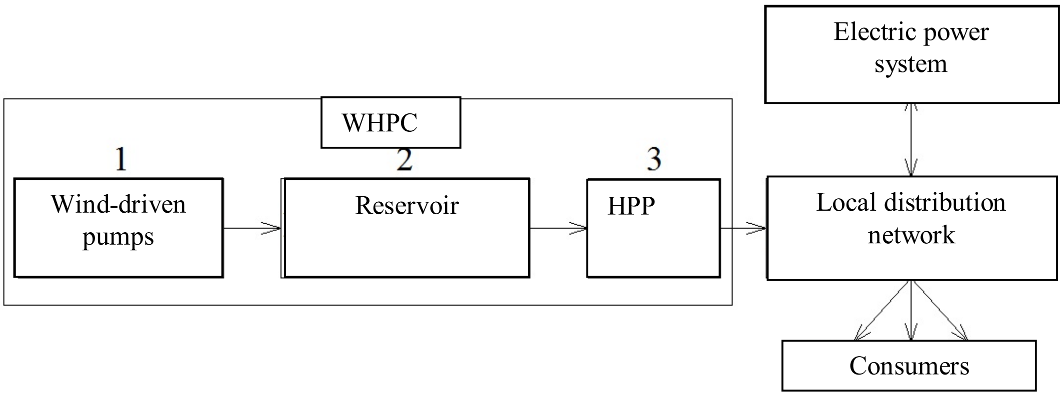


Figure 1. Flowchart of WHPC and its interfaces with the system of electricity supply to the area

As Figure 1 shows the idea of WHPC is not original. However, the reason why we address this issue is related to the circumstances that are growing urgent nowadays.

The essence of the suggested plant lies in the fact that wind being primary energy resource for WHPC operation is harnessed to fill reservoir which represents the energy storage stage in the cycle of electricity production. The major advantage of the WHPC is the combination of wind power plant advantages and the idea of pumped water storage. This excludes the main flaw of wind power generation, i.e. a mismatch between unpredictable variations in wind speed and electricity consumption schedule. This distinguishes WPHC from similar plants in which wind energy is used directly to generate electricity to meet the demand and to charge storage battery. When there is no wind and energy storage systems such plants are backed up by diesel power plants (DPP) or gas-turbine power plants (GTPP).

The wind-hydro power complex is in many parameters similar to the pumped storage power plants (PSPP) [3]. However, the main difference is the use of free wind power to fill the reservoir. The hydropower plant included in the system covers not only peak loads but the entire local load. Moreover, for WHPC the connection (at least a weak one) with power grid is desirable but not mandatory. With this connection WHPC can perform the functions imposed on distributed generation.

Compared to other wind power plants this system has the following advantages:

1. The WHPS designed according to such a flowchart makes it possible to separate and consider individually two random non-correlated processes:
- the use of wind power for electricity supply under any wind conditions;
- reliable (continuous) high quality electricity supply to consumers, irrespective of wind conditions at any time moment.
2. The possibility of applying relatively simple (and hence cheap and reliable wind-driven plants including wind-driven pumps with mechanical transmission of wind power to hydro pump (piston or centrifugal).

3. Application of reservoir as an energy storage which is environmentally friendlier and simpler than storage batteries of the same capacity, pressed air, hydrogen etc. as well as backup diesel units with diesel fuel stocks, or gas turbine plants.
4. Use of reservoir for the purposes other than storage, i.e. as a reservoir for tap water supply to the nearest populated settlements and productions, as a drinking place, for fish breeding, poultry farming, irrigation of agricultural lands, recreation needs, etc.

Application of such a system is efficient first of all for remote populated settlements because their full and reliable electricity supply from centralized power grid can be difficult or expensive. Although for the reasons of reliability and cost effectiveness WHPC can be constructed when consumers are supplied with electricity from grid. In this case the connection to the power grid can:

- increase the electricity supply reliability;
- decrease the reservoir capacity, particularly, if there are consumers of all categories. Then it is sufficient to use hydropower under emergency conditions to meet only the demand of consumers of the first category and partially of the second category;
- considerably improve power quality if transmission lines connected to the grid are very long and have a low rated voltage (110 kV and lower, down to 10-6 kV at lengths of 50-150 km and longer). Owing to HPP there will be local surplus active and reactive power for continuous electricity supply and voltage control;
- enable transmission of surplus power to the grid.

An additional advantage of the suggested system is the fact that to fill the reservoir it is not necessary to use high-speed wind machines, on the contrary it is more expedient to apply slow-speed wind-driven pumps. This increases the period of wind use and does not require such high aerodynamic characteristics as those necessary to use wind turbines to directly supply an electric load. The possibility of using natural relief of an area in order to construct a reservoir should also be considered as a benefit of the WHPS. The authors consider the possibility of applying such a system for electricity and water supply particularly in the arid regions.

However, along with the advantages the system has some flaws. First of all this is the impact of climatic conditions on water storage in the reservoir. With allocation of the plant in severe climatic conditions there appears a danger of reservoir and water supply system freezing and as a consequence the impossibility of their further use in the winter period. Elimination of this drawback requires additional investment. The relief of the territory may not always be suitable for the reservoir construction. Then it can be necessary to create an artificial water reservoir because of flat ground or insufficient ground strength which can lead to additional investment.

Taking into account the known electricity consumption variability over time electricity generation from WHPC should provide a reliable power supply to meet the demand.

Bearing in mind the advantages and disadvantages of the suggested system to clearly understand its efficiency as applied to specific conditions as well as to make a specific design it is necessary to develop an efficiency estimation technique as a calculation tool for solving the problem of electricity supply in specific conditions on the basis of renewable energy sources.

Thus, the technique for determining the WHPC parameters in general should include the following steps:

1. Calculation of HPP parameters on the basis of electricity consumption forecast and the need to provide reliable electricity supply to consumers;
2. Calculation of reservoir characteristics on the basis of calculated HPP parameters.
3. Determination of parameters for wind power plants on the basis of calculation results for p.2 and, bearing in mind wind characteristics of the area in which the plant is situated and reliability of wind-driven pumps.
4. Solving the other problems related to the construction of WHPC.

The HPP parameters, reservoir characteristics and parameters of wind power units are calculated using the known techniques but taking into account specific operating features of these facilities within the WHPC and estimation of the power supply reliability. Below the authors present the specific features of selecting the parameters and characteristics for the indicated components of WHPC.

Determination of HPP parameters

Determination of HPP parameters in fact implies selection of a rated capacity and type of hydropower units with respect to selected head, and their number bearing in mind the required level of reliability of electricity supply.

The algorithm for calculation is as follows. After determining a regular annual load peak from the forecast of social and economic development of the region $N^L_{reg.peak}$ for the respective time period we find an irregular annual peak $N^L_{ir.peak}$ by the expression:

$$N^L_{ir.peak} = (1 + 3\sigma_L)N^L_{reg.peak}. \qquad (1)$$

In (1) σ_L is standard deviation of load from a regular value of capacity, per unit. It is known [4] that these deviations follow the normal distribution.

The optimal reliability of HPP is calculated according to the Bernoulli formula [5].

The method based on the Bernoulli formula makes it possible to estimate the required reserve and probability of shortage-free operation of the facilities consisting of n components according to their rated parameters. In order to assess reliability of HPP these parameters will be represented by rated capacity N_r, number of hydropower units n and probability of failure-free operation p. With the assumed N_r we determine the required number of units which will provide supply of the required load with a specified (rated) probability of shortage-free operation under the minimum capacity reserve.

The calculations are made according to the formula of binomial distribution

$$\begin{aligned}(p[N_r]+q[0])^n &= p^n[nN_r] + C_n^1 p^{n-1} q[(n-1)N_r] + C_n^2 p^{n-2} q^2[(n-2)N_r] + \ldots + \\ &+ C_n^i p^{n-i} q^i [(n-i)N_r] + \ldots + q^n[0] = 1\end{aligned}, \qquad (2)$$

where p – probability of operable state of hydropower unit (taken from the data of manufacturer or according to the emergency rate statistics for HPP equipment); $q = 1 - p$ – probability of emergency downtime of the hydropower unit; n – the number of units to be installed at HPP; $i = \overline{1, n}$ - the number of units that can be in an inoperable state; C_n^i – number of combinations from n units with respect to i; expressions in square brackets characterize the values of the HPP available capacity in respective calculated states.

The presented binomial expansion is a full group of events with different possible states of the HPP components. In this case this is a combination of operable (n - i) and inoperable i components from their total number n.

From (2) we find the probability of shortage-free load supply:

$$P = \sum_{i=0}^{I} C_n^i p^{n-i} q^i [(n-i)N_r]$$

for all i, for which

$$(n \quad i)N_r \geq N^L_{ir.peak}. \tag{3}$$

In this case we choose n for which

$$P \geq P_{stnd}. \tag{4}$$

In (4) P_{stnd} is a standard value of probability of shortage-free electricity supply to consumers. In Russia P_{stnd} is assumed at the level of 0.996 [4], in Western Europe – 0.9996.

If (4) is not provided then N_r and/or n, at which (4) is met, varies.

Normally N_r is taken equal to the capacity given in the catalogues of plant manufacturers of hydropower equipment and the number of units n of the required rated capacity is specified.

Thus, from the above calculation we determine the main electric parameters of HPP: rated capacity of units N_r^0, number of units n^0 and installed capacity of HPP

$$N^{inst}_{HPP} = n^0 N_r^0 \text{ kW}.$$

Calculation of reservoir characteristics

At the second stage of calculation of the WHPC parameters we determine the required reservoir capacity.

Since wind conditions vary water supply to water reservoir varies too. The main objective here is to provide such volume and conditions for reservoir filling as to have sufficient water to meet the demand for electricity in a required amount and at a required time throughout the entire calculation period T, that depends on the wind conditions.

The wind conditions are estimated on the basis of data from climatologic reference books on wind for the area where the WHPS is going to be constructed. These data are used to determine the duration of periods with wind speed insufficient for operation of wind turbines and duration of energy inefficient wind speed. The two parameter Weibull distribution [1] is used in calculations to determine the repetition of wind speeds. The obtained information is then used to determine the duration of period with a wind speed that ensures useful work of wind-driven pumps.

In this case the calculation period T should be considered as

$$T = T_w + T_{i/w}, \tag{5}$$

where T_w – time of sufficient wind conditions, day; $T_{i/w}$ – time of insufficient wind conditions, day.

Energy storage is used to solve the following problems:

- reciprocal matching of energy production and consumption schedules in order to provide uninterrupted electricity supply to consumers;
- increase in the efficiency of wind energy utilization through complete use of the total output of wind turbines.

When resolving the issues related to storage of energy produced by wind power plants we should take into account the following characteristics:

- relative sizes;
- duration of energy storage;
- admissible amount of energy to be stored;
- complexity of energy transformations (rectification, inversion, frequency transformation, etc.);
- simplicity and safety of maintenance, etc.

The main criterion for determining the reservoir capacity is the need to provide the required water flow rate Q by operating hydropower units. Flow rate decreases with an increase in the water head H. This condition is taken into account to design the reservoir in terms of the area relief.

The calculated water head is assumed according to the possibilities of reservoir construction in a specified area. With the assumed calculated head on the basis of manufacturer's data for the chosen type of hydropower unit we determine a specific water flow rate Q_0 (m^3/kWh). The required available reservoir capacity is determined by the formula

$$V_{avlb} = W_{req} \cdot Q_0 \text{ m}^3, \quad (6)$$

where W_{req} –required HPP output determined by the load curve for a respective calculation period T:

$$W_{req} = \int_0^T N(t)dt \text{ kWh},$$

where $N(t)$ – required power of electricity consumption at hour t of the load curve.

Knowing the reservoir surface area F_s (m^2), we estimate the depth of water layer of the available reservoir capacity:

$$\Delta H = V_{avlb} / F_s \text{ m}.$$

Based on the known H_{min} (from manufacturer's data) we determine the maximum head water level:

$$H_{max.} = H_{min} + \Delta H \text{ m}.$$

To estimate the reservoir surface area and depth of the available reservoir capacity we should seek to reduce the surface area (which decreases the alienation of land surface for reservoir, evaporation surface etc.), and the depth of periodic reservoir drawdown since large variations in the water level have a negative impact on the flora and fauna of the reservoir itself and its seashore. Generally, the reservoir surface area can be regulated by diking the reservoir of small sizes. The regulation of the reservoir surface area makes it possible to choose the depth of drawdown and vice versa, depending on the specific circumstances.

The selection of the hydropower units for WHPC implements the key principle of determining the reservoir capacity, i.e. makes it possible to provide the minimum possible flow rate Q in order to minimize the reservoir capacity and as a result decrease investment in its construction and operation.

Since an HPP supplies electricity to consumers, covering the whole of the load curve, the available reservoir capacity should satisfy the water flow rate by hydropower units for the assumed calculation period T of power supply. The calculation period can be taken equal to the time interval from a day (daily storage) to a year (yearly storage). The calculation period is chosen based on specific wind parameters – the more frequent is the wind, the shorter is the calculation period and hence, the smaller is the water reservoir and the lower is the investment in its construction.

In addition to the available reservoir capacity V_{avlb} to meet the water flow rate by hydropower units, account should be taken of losses caused by water evaporation V_{evap} from reservoir surface, by filtration V_f through ground and by ice formation V_i for the areas of cold climate [6, 7].

The methods for determination of water flow rate to compensate for these losses are empirical and applied depending on every specific case.

The total storage capacity will be

$$V_{total} = V_{avlb} + V_{evap} + V_f + V_i + V_{dead} = V_{av} + V_{loss} + V_{dead} \text{ m}^3,$$

where $V_{loss} = V_{evap} + V_f + V_i$; V_{dead} – dead storage capacity.

Since construction of a purely man-made lake is an expensive measure, it is more expedient to arrange it on the basis of natural relief roughness with minimum involvement of materials and labor inputs in construction of a storage reservoir. The possibilities for provision of V_{total} at the site are evaluated by calculating the storage capacity of a prospective reservoir through the sequential summation of capacities ΔV_i of individual layers between two adjacent contour lines on the topographic maps. This is done for determination of reservoir surface F_s by way of their planimetering based on the knowledge of topographic characteristics of the site.

In general the storage capacity V_{total} can be increased in case of the need to solve other economic problems in the area of WHPC construction that were mentioned above. This is, however, a separate problem.

The method of reservoir arrangement and its type by its design features are determined based on technical and economic indices of one or another variant for specific conditions. However, preference should be given to reservoir arrangement, taking advantage of natural relief as much as possible [6, 7].

At the initial stage of reservoir filling the wind-driven pumps will have to fill the total storage capacity V_{total}, which will require some time. Then after filling of the dead and available capacities, the wind-driven pumps will have to fill only the capacity $V_{avlb} + V_{loss}$ to be emptied during the calculation period T. And the capacity filled in advance is emptied at the current period and simultaneously the new volume $V_{avlb} + V_{loss}$ is stored for HPP operation at the next period.

Reservoir dislocation can be chosen based on a great number of variants: gorge, ravine, notch, depression on the upland. Apart from natural conditions, it is possible to consider creation of a man-made diked lake, a lake with consolidation of its bed with impermeable materials, etc.

Calculation of wind-driven pump parameters

In operation of any plant using wind energy, including WHPC, the wind parameters are of prime importance as a source of energy production.

Wind depends on many complex geophysical and climatic factors. Its variability, therefore, can be predicted only with some probability that is determined as a result of statistical processing of the results of wind speed observations in the considered area for a long-term period.

Wind speed is the most important energy characteristic that estimates its kinetic energy. Under the impact of some meteorological factors (atmosphere perturbations, changes in solar activity and amount of heat energy arriving from the space to the Earth, etc.), and also the relief of the site, the wind speed changes in rate and direction. The powerful winds favorable for operation of wind power plants alternate thereby with calms.

The wind-driven pump capacity depends on the wind speed and the surface area swept by the wind wheel and is calculated by the formula:

$$N_w = \frac{\upsilon^3 D^2}{7000} \text{ kW}, \qquad (7)$$

where υ – wind speed, m/s; D – wind wheel diameter, m.

The wind-driven pump converts part of this capacity into effective capacity that is estimated by the wind energy utilization factor ζ:

$$N_{w\,avlb} = \zeta \cdot N_w. \qquad (8)$$

In general it is not expedient to use wind power plants for direct covering of electric loads without additional expensive smoothing and replacing facilities and also automatic controllers by virtue of essential distinctions between the consumer load curve and the curve of wind speed variation as random functions of time.

As was noted, the wind-driven pump (module for reservoir filling) capacity can be determined on the basis of the necessary storage capacity $V_{avlb} + V_{loss}$ and the reservoir filling time T_r to be known, during the calculation period T. It is apparent that the dead storage capacity is filled once before the beginning of WHPS operation, and installation of additional wind-driven pumps for this purpose will be inexpedient.

The total pumping capacity of wind-driven pumps $Q_{w\Sigma}$ is calculated by the formula:

$$Q_{w\Sigma} = \frac{V_{avlb} + V_{loss}}{T_w} \text{ m}^3\text{/s},$$

where T_w is in seconds.

The total delivery of wind-driven pumps is determined by the expression:

$$N_{w\Sigma} = \frac{9.81 Q_{w\Sigma} \mathrm{H}}{\eta} \text{ kW}, \qquad (9)$$

where η – pump efficiency; H – height of water, m.

The time of wind-driven pump operation T_w is determined from the formula in [1]:

$$T_w = \frac{f(\upsilon \geq \upsilon_0) \cdot T}{100}, \qquad (10)$$

where $f(\upsilon \geq \upsilon_0)$ – probability that the initial speed of the wind-driven pump will be exceeded, %; υ_0 – initial speed of the wind wheel, m/s. In calculations υ_0 is taken equal to 3 m/s. The multi-blade wind-driven pumps that are targeted for use in WHPS start to operate at this speed.

The values of $f(\upsilon \geq \upsilon_0)$ as a function of the wind parameters $\upsilon_0 / \bar{\upsilon}$ and c_υ are determined as tabular data in accordance with the Weibull distribution [1].

The down time of the wind-driven pumps $T_{i/w}$ is determined as:

$$T_{i/w} = T - T_w \text{ h}. \qquad (11)$$

The relation between T_w and $T_{i/w}$ may be arbitrary, and the time of sufficient wind speed T_w may be both longer and shorter than $T_{i/w}$. The wind-driven pump capacity depends on the relation between T_w and $T_{i/w}$. The longer is T_w, the lower is the capacity $N_{w\Sigma}$.

The rated capacity of the wind-driven pump $N_{w\,rat}$ is chosen based on the machine industry capabilities. It is obvious that for these purposes the choice should be made of maximum possible capacity in terms of specific design conditions.

Then the minimum needed number of wind-driven pumps is calculated as

$$n_w = N_{w\Sigma} / N_{w\,rat}. \qquad (12)$$

The number of wind-driven pumps and their rated capacity considering reliability of wind-driven pumps can be evaluated more accurately on the basis of their emergency rate q_w to be determined and formula (2). The standard reliability of all the wind-driven pumps is taken as a function of wind conditions in the considered region, however, not lower than the probability of shortage-free power supply (for RF – 0.996).

The described stages of WHPC calculation are basic for designing the considered plant. Technical and economic problems dealing with power supply from WHPC are solved at the next stage.

From the *technical* standpoint they include:

- generation of an electric circuit of hydropower plant, choice of voltages of generator, auxiliaries, master switchgear;
- generation of an electric circuit of local distribution network;
- assurance of reliability of power supply to consumers and required power quality in accordance with the standards of electric installation code, maintenance rules, and other documents;
- consideration of specific features of natural-climatic and geographical conditions for WHPS construction when being designed;
- implementation of capabilities for additional utilization of the man-made reservoir in socio-economic development of the region;
- assessment of WHPC security.

From the *economic* standpoint it is necessary to carry out a feasibility study on the effectiveness of the proposed system in comparison with other alternative options of power supply:

1) wind power plants in combination with replacing power sources (diesel power plants, geothermal power plants, etc.).
2) wind power plants in combination with storage facilities of other types (thermal, chemical, mechanical, etc.).

In addition one should bear in mind advantages of the suggested complex over the mentioned ones:

- simplicity of design;
- simplicity of meeting the main requirements to reliability and quality of power supply to consumers;
- wider range of economic usage (not only power generation).

Thus, the optimal decision for all these options can be chosen only based on the specific feasibility analysis that considers secondary advantages such as environmental.

Assessment of technical and economic effectiveness

Technical and economic characteristics of the suggested WHPS should be determined to compare them with similar characteristics of other alternative power generation sources in the considered region and to estimate their dependence on local conditions.

The technical and economic characteristics of power supply options should be determined on the basis of their functional comparability: full satisfaction of demand, power supply reliability and power quality. Besides, one should bear in mind such advantages of WHPCs over other options as absence of fuel costs, simplicity, low cost and high reliability in comparison with wind power plants, substantially simplified control of WHPC, possibility for solving other (in addition to power supply) socio-economic problems in the considered region, higher environmental compatibility of WHPC, etc.

Effectiveness can be assessed based on the information about expenditures for both WHPS and alternative power supply options such as renewable and non-renewable energy sources, i.e. expenditures for construction of replacing diesel or gas-fired power plants, fuel cost, cost of diverse storage facilities and their practical limits on capacity.

For the most general case the expression for the simplified technical and economic evaluation of WHPS can be written in the following form:

$$aN_{WP} + bV_{total} + cN_{HPP} < dN_{WPP} + eN_{SF} + fN_{DPP} + gB_fW\frac{T_{i/w}}{T}n. \tag{13}$$

In this expression:

a – unit cost of wind-driven pump (WP), RUR/kW;
N_{WP} – installed capacity of WP, kW;
b – unit cost of reservoir construction, RUR/m^3;
V_{total} – reservoir storage capacity, m^3;
c – unit cost of hydropower plant (HPP) construction (without reservoir cost), RUR/kW;
N_{HPP} – installed capacity of HPP, kW;
d – unit cost of wind power plant (WPP), RUR/kW;
N_{WPU} – installed capacity of WPP, kW;
e – unit cost of storage facility (SF), RUR/kW;
N_{SU} – installed capacity of SF, kW;
f – unit cost of additional power plant (DPP) construction, RUR/kW;
N_{DPP} – installed capacity of DPP, kW;
g – fuel cost for DPP, RUR/kg;
B_f – specific fuel consumption by DPP, kg/kWh;
W – required power generation for the calculation period T, kWh;
T – calculation period corresponding to cyclic recurrence of wind activity at the considered site, hours;
$T_{i/w}$– duration of calm period at the calculation period, hours;
n – recurrence number of periods T during the WHPC service life.

Effectiveness is assessed on the basis of the following simplifications and assumptions.

The unit cost includes the cost of land allotted for facilities to be constructed.

Only the costs that differ in the options compared are calculated. Therefore, the proceeds from electricity trade that are assumed to be equal and the fixed costs for operation of compared power facilities are not taken into account.

Discounting costs during the service life of power facilities, which can hardly influence the basic compared option are not considered. Their service life is taken equal, i.e. 30 years.

If the cost of plants compared proves to be almost equal, preference should be given to WHPC owing to the indicated additional effects of its use.

The hydropower plant that is considered in this statement differs from the traditional run-of-river plant in the essential decrease of the probabilistic nature of water inflow to the reservoir. Here the necessary water volume is provided to a sufficiently high degree by installation of additional wind-driven pumps that deliver the needed volume of water at the periods of sufficient wind speed.

Expression (13) is of universal character to compare WHPC to any types of alternative options. Therefore, in the right-hand side there are zero values for the plants that are not used in the corresponding option.

The financial efficiency can be assessed by calculating the payback period:

$$T_{payback} = S/P = S/((C - Z)W_{year}),$$

where S – expenditures for the corresponding project (the right- or left-hand side of expression (13));
P – annual profit from produced electricity sales;
C – electricity price in the energy market;
Z – electricity production cost;
W_{year} – volume of annual electricity sales.

Preliminary analysis of the field of WHPC use has shown that these systems prove to be attractive in the range of power consumption from 30–50 kW to 10–15 MW. In this case at low loads it is possible to have small ponds filled by two or three wind-driven pumps that deliver water at a height of 20-30 m rather than conventional reservoirs. WHPCs of higher capacity will require water heads of 100–150 m and higher (considering employment of diversion schemes). WHPCs of larger capacity become irrational because of vast areas required for allocation of wind-driven pumps and a reservoir.

An example of the comparative technical and economic evaluation of WHPC and an alternative power supply option (an additional diesel power plant as the cheapest alternative option) is presented below.

Example of comparative assessment of technical and economic effectiveness of WHPC and an alternative option

The authors consider power supply to a coastal area of Lake Baikal [8, 9].The annual regular maximum load is 650 kW. In accordance with (1) the irregular maximum load is taken equal to 1.09*650 = 710 kW with σ_L=0.03. The required power consumption for the calculation period T considering losses and auxiliary power supply will make up 1254144 kWh. The required HPP capacity considering uninterrupted power supply (Ᵽ = 0.996) will amount to 800 kW. Lake Baikal is the water source. The necessary reservoir capacity is 0.00593 km^3. Wind speed in the area allows the calculation period T to be taken equal to 3 months and the total time of energy-effective wind strength during the calculation period T_w to be taken equal to 1.6 months.

Two options are studied:

1. Wind hydropower system.
2. Wind power plant with an additional diesel power plant.

According to calculations in the first option the wind-driven pump **capacity** should be 2000 kW.

In the second option the wind power plant capacity equals 1100 kW, the diesel power plant capacity –1100 kW (considering power supply reliability).

The technical and economic analysis was carried out based on the following averaged economic indices (see (13)):

a = 70000 RUR/kW;
b= 100 RUR/ m^3;
c = 5000 RUR/ kW;
d = 51200 RUR/ kW;
f = 10000 RUR/ kW;
g = 30 RUR/ kg;
B_f = 0.4 kg / kWh;
n = 120 – number of occurrences of periods T during the power supply system life (30 years).

The costs of construction and operation of the considered power supply options are calculated based on (13):

1. $aN_{WP} + bV_{total} + cN_{HPP} = 70000 \cdot 2000 + 100 \cdot 0.00593 \cdot 10^9 + 5000 \cdot 800 = 140000000 +$

   $+ 593000000 + 4000000 = 0.737$ billion RUR.

2. $dN_{WPP} + eN_{SF} + fN_{DPP} + gB_f W \dfrac{T_{calm}}{T} n = 51200 \cdot 1100 + 0 + 10000 \cdot 1100 +$

   $+ 30 \cdot 0.4 \cdot 1254144 \cdot \dfrac{31025}{2190} \cdot 120 = 56320000 + 0 + 11000000 + 845258680 = 0.913$ billion RUR.

The calculations show that the reservoir is the most expensive structure of WHPS (86.7 %). Hence, special attention should be paid to decrease of its construction costs. To make the calculations more accurate it is necessary first of all to estimate the real value of b because of its impact on the cost of the first option. The WHPC competitiveness will increase with the fuel price rise (the costs of fuel delivery to remote areas are not taken into account and they are comparable to fuel cost and even exceed it).

The main conclusion from the indicated calculations is: effectiveness of the WHPC option depends on reservoir parameters. The smaller is the reservoir capacity, and correspondingly the unit costs, the more profitable will be the WHPC option. The indicated reservoir parameters can be decreased by increasing the head H, shortening the calculation period T and maximum possible utilization of natural relief elements for reservoir construction at the specific site.

Conclusion

1. In the context of increasing shortage of fossil fuel resources and topicality of environmental problems the necessity of using renewable energy resources rises.
2. Investigations in the field of the most economical and technologically expedient renewable energy sources for specific areas result in designs of different systems, WHPC as an example.
3. The design works and commissioning of such a system can be realized on the basis of the technique for determination of its technical and economic effectiveness. The technique is to solve a great number of problems: from choice of the primary WHPC link – wind-driven pump to the final result – generation of power of the required quality for its reliable supply to consumers.
4. The technique should be applied as a tool for assessment of the efficiency of using the suggested system. All sorts of difficulties cannot be overcome successfully without flexible consideration of wind energy utilization forms.
5. Parameters of the required reservoir depend on the electric load, on the one hand, and the wind speed in the area of WHPC construction, on the other hand. The calculation period of reservoir drawdown is chosen based on the wind speed in the considered area. Therewith T_w is always shorter than T, and the shorter is T and the longer is T_w, the smaller will be the reservoir capacity and hence, the cheaper will be the WHPC construction.
6. The paper suggests a sequence of the WHPC calculation, choice of its basic parameters including power supply reliability and technical and economic effectiveness. The sequence of calculation is universal, i.e. it is applicable to any conditions of WHPC operation.
7. In general, WHPC plays a part of “distributed” generation that is defined as power generation at the point of its consumption. In this case the power losses and the costs of its transmission by regional power grids are excluded. Power supply reliability improves.
8. Availability of even a weak tie line with the power grid enhances flexibility, reliability and effectiveness of the local power supply system. Power quality in the considered area improves considerably. Besides, in this case excess power can be supplied to the common system network.
9. WHPC as distributed generation is of diversification character, allowing the variety of plants on renewable energy resources that utilize wind energy to be increased.

References

1. Golitsyn M.V., Golitsyn A.M., Pronina N.V. Alternative energy carriers. – M.: Nauka, 2004. – 159 p. (in Russian).
2. Nikolaev V.G. Trend towards world wind energy development. // Energiya: ekonomika, tekhnika, ekologiya. 2011. P .10–18 (in Russian).
3. Sinyugin V.Yu., Magruk V.I., Rodionov V.G. Pumped storage plants in the modern electric power industry. – M.: ENAS, 2008. – 358 p. (in Russian).

4. Power system design handbook./V.V. Ershevich, A.N. Zeiliger, G.A. Illarionov et al. Ed. by S.S. Rokotyan and I.M. Shapiro. 3d edition, revised and enlarged. – M.: Energoatomizdat, 1985. – 352 p. (in Russian).
5. Ventsel E.S. Probability theory. – M.: Vysshaya shkola, 2002. – 575 p. (in Russian)
6. Asarin A.E., Bestuzheva K.N. Hydraulic and energy calculations. – M.: Energoatomizdat, 1986. – 224 p. (in Russian).
7. Karol L.A. Hydraulic storage. – M.: Energiya, 1975. – 208 p. (in Russian).
8. Rychkov M.A. Choice of rated capacity and number of hydropower units at mini-HPP considering reliability. // System energy studies. Issue 40. – Irkutsk: ISEM SO RAN, 2010. P. 130–137 (in Russian).
9. Rychkov M.A. Determination of reservoir capacity for wind-hydro power complex. // System energy studies. Issue 41. – Irkutsk: ISEM SO RAN, 2011. P. 90–98 (in Russian).

STEP STRESS ACCELERATED LIFE TESTING PLAN FOR TWO PARAMETER PARETO DISTRIBUTION

Mustafa Kamal

•

Department of Statistics & Operations Research
(Aligarh Muslim University, Aligarh-202002, India)
E-mail: kamal19252003@gmail.com

Shazia Zarrin

•

Department of Statistics & Operations Research
(Aligarh Muslim University,Aligarh-202002, India)
E-mail: shaziazarrin@gmail.com

Arif-Ul-Islam

•

Department of Statistics & Operations Research
(Aligarh Muslim University, Aligarh-202002, India)
E-mail: arifislam2@yahoo.com

ABSTRACT

In Accelerated life testing if the accelerated test stress level is not high enough then many of the test items will not fail during the available time and one has to be prepared to handle a lot of censured data. To avoid such type of problems, a better way is step-stress ALT. In Step-stress ALT all test items are first tested at a specified constant stress for a specified period of time and then Items which are not failed will be tested at next higher level of stress for another specified time and so on until all items have failed or the test stops for other reasons. In this paper simple step stress pattern of ALT assuming that the lifetime of a product at any constant level of stress follow a two parameter Pareto distribution is considered. The maximum likelihood and asymptotic confidence interval estimate of the parameters are obtained. Optimal step stress ALT plan is proposed by minimizing the asymptotic variance of the MLE of the 100 P^{th} percentile of the lifetime distribution at normal stress condition. A simulation study is also performed to analyse the performance of parameter estimates.

KEYWORDS: Cumulative Exposure Model; Maximum Likelihood Estimation Method; Fisher Information Matrix; Asymptotic Confidence Intervals; Simulation Study.

1 INTRODUCTION

Accelerated life testing (ALT) is a quick way to obtain information about the life distribution of a material, component or product. In Accelerated life testing (ALT) items are subjected to conditions that are more severe than the normal ones, which yields shorter life but, hopefully, do not change the failure mechanisms. Some assumptions are needed in order to relate the life at high stress levels to life at normal stress levels in use. Based on these assumptions, the life distribution under normal stress levels can be estimated. Such way of testing reduces both time and cost.

Three types of stress loadings are usually applied in accelerated life tests: constant stress, step stress and Progressive-stress. Constant stress is the most common type of stress loading. Every item is tested under a constant level of the stress, which is higher than normal level. In this kind of testing, we may have several stress levels, which are applied for different groups of the tested items. This means that every item is subjected to only one stress level until the item fails or the test is stopped for other reasons. In Step-stress loading, the test items are subjected to successively higher levels of stress at pre-assigned test times. All items are first subjected to a specified constant stress for a specified period of time. Items that do not fail will be subjected to a higher level of stress for another specified time. The level of stress is increased step by step until all items have failed or the test stops for other reasons. Progressive-stress loading is quite like the step stress testing with the difference that the stress level increases continuously.

Failure data obtained from ALT can be divided into two categories: complete (all failure data are available) or censored (some of failure data are missing). Complete data consist of the exact failure time of test units, which means that the failure time of each sample unit is observed or known. In many cases when life data are analysed, all units in the sample may not fail. This type of data is called censored or incomplete data. Due to different types of censoring, censored data can be divided into time-censored (or type I censored) data and failure-censored (or type II censored) data. Time censored (or type I censored) data is usually obtained when censoring time is fixed, and then the number of failures in that fixed time is a random variable. Failure censored (or type II censored) data is obtained when the test is terminated after a specified number of failures, and then time to obtain that fixed number of failures is a random variable.

Simple step-stress ALT, where only one change of stress occurs, proposed by Nelson (1980) has been widely studied and referred as the Cumulative Exposure (CE) model. Many studies regarding SSALT planning based on the CE Model, have been performed. Miller and Nelson (1983) presented the optimum simple SSALT model. Bai *et al.* (1989) and Bai and Chun (1991) extended this model to the case where a prescribed censoring time is involved. Many authors also have provided the studies for statistical inference model for SSALT based on CEM; e.g., see Xiong (1998), Watkins (2001), Zhao and Elsayed (2005), Balakrishnan *el al.* (2009), Yeo and Tang (1999), Xiong and Ji (2004) and Xiong and Milliken (1999). Khamis and Higgins (1998) proposed a new model for SSALT as an alternative to the CEM, which is based on a time transformation of the exponential CEM. Most of works using the K-H model are concentrated on the optimal design plan for SSALT. Alhadeed and Yang (2002) provided the optimal plan for a simple SSALT using K-H model when the shape parameter is unknown.

More recently Lu and Rudy (2002) have dealt with the Weibull CE model under the inverse power law in the simple SSALT. McSorley, Lu and Li (2002) have shown the properties of the maximum likelihood (ML) estimators of parameters in the Weibull CE model with a log-linear function of stress on three-step SSALT data. Gounu, Sen and Balakrishnan (2004) tackled the optimal stress change points for multiple-step SSALT based on minimizing the asymptotic confidence interval of MLE of the mean life at design stress. Wu, Lin and Chen (2006) discussed the ALT with progressively Type-I group-censored exponential data. Balakrishnan and Han (2008) considered modification for censoring scheme in small sample sizes. Fan, Wang and Balakrishnan (2008) discussed the maximum likelihood (ML) estimation and Bayesian inference in group data ALT models under the relationship between the failure rate and the stress variables is linear under Box-Cox transformation. Al-Masri and Al-Haj Ebrahem (2009) derived the optimum times of changing stress level for simple step-stress plans under a cumulative exposure model assuming that the life time of a test unit follows a log-logistic distribution with known scale parameter by minimizing the asymptotic variance of the maximum likelihood estimator of the model parameters at the design stress with respect to the change time. Hassan and Al-Ghamdi (2009) obtained the optimal times of changing stress level for simple stress plans under a cumulative exposure model using the Lomax distribution for a wide range of values of the model parameters. Xu and Fei (2012) introduced and compared the four basic models for step-stress accelerated life testing: cumulative exposure model (CEM), linear cumulative exposure model (LCEM), tampered random variable model (TRVM), and tampered failure rate model (TFRM). Limitations of the four models are also introduced for better use of the models.

In this paper the two-parameter Pareto distribution as a lifetime model under simple-step-stress ALT is considered. Maximum likelihood estimates of parameters and their asymptotic confidence intervals are obtained. The performance of the estimates is evaluated by a simulation study with different pre-fixed values of parameters.

2 THE MODEL

2.1 The Pareto Distribution

The concept of this distribution was first introduced by Vilfredo Pareto (1897) in his well-known economics text "*Cours d'Economie Politique*".

The two parameter forms of Pareto probability density function (pdf), cumulative distribution function (CDF), the reliability function (RF) and the hazard rate (HR) with shape parameter α and scale parameter θ given respectively by

$$f(t;\theta,\alpha)=\frac{\alpha\theta^{\alpha}}{(\theta+t)^{\alpha+1}};\qquad t>0,\theta>0,\alpha>0 \tag{2.1}$$

$$F(t)=1-\frac{\theta^{\alpha}}{(\theta+t)^{\alpha}};\qquad t>0,\theta>0,\alpha>0 \tag{2.2}$$

$$R(t)=\frac{\theta^{\alpha}}{(\theta+t)^{\alpha}} \tag{2.3}$$

$$h(t)=\frac{\alpha}{\theta+t} \tag{2.4}$$

The hazard rate (HR) is a decreasing function as $t>0$ and an increasing function as $t<0$.

2.2 Assumptions and Test Procedure

1. There two stress levels x_1 and x_2 $(x_1 < x_2)$.
2. The failure time of a test unit follows a two-parameter Pareto distribution at every stress level.
3. A random sample of n identical products is placed on test under initial stress level x_1 and run until time τ, and then the stress is changed to x_2 and the test is continued until all products fail.
4. The lifetimes of the products at each stress level are i.i.d.
5. The scale parameter is a log-linear function of stress. That is, $\log\theta(x_i)=a+bx_i$, $i=1,2$ where a and b are unknown parameters depending on the nature of the product and the test method. Therefore, the lifetime of a test product at lower stress x_1 is longer than at higher stress x_2.
6. The Pareto shape parameter α is constant, i.e. independent of stress.
7. A cumulative exposure model holds, that is, the remaining life of test items depends only on the current cumulative fraction failed and current stress regardless of how the fraction accumulated. Moreover, if held at the current stress, items will fail according to the CDF of stress, but starting at the previously accumulated fraction failed, for more detail on CE Model see Nelson (1990). According to cumulative exposure model the CDF in step-stress ALT are given by

$$F(t)=\begin{cases}F_1(t) & 0\le t<\tau\\ F_2(t-\tau+\tau') & \tau\le t<\infty\end{cases}$$

where the equivalent starting time, τ', is a solution of $F_1(\tau) = F_2(\tau')$ solving for τ', then $\tau' = \frac{\theta_2}{\theta_1}\tau$ and now the CDF is of the form

$$F(t) = \begin{cases} F_1(t), & 0 < t < \tau \\ F_2\left(\frac{\theta_2}{\theta_1}\tau + t - \tau\right), & \tau \le t < \infty \end{cases} \tag{2.5}$$

and corresponding pdf is obtained as as

$$f(t) = \begin{cases} f_1(t), & 0 < t < \tau \\ f_2\left(\frac{\theta_2}{\theta_1}\tau + t - \tau\right), & \tau \le t < \infty \end{cases} \tag{2.6}$$

From the assumptions of cumulative exposure model and the equation (2.2), the CDF of a test product failing according to Pareto distribution under simple step-stress test is given by

$$F(t) = \begin{cases} 1 - \frac{\theta_1^\alpha}{(\theta_1 + t)^\alpha}, & 0 < t < \tau \\ 1 - \frac{\theta_2^\alpha}{\left[\theta_2 + \tau\left(\frac{\theta_2}{\theta_1} - 1\right) + t\right]^\alpha} & \tau \le t < \infty \end{cases} \tag{2.7}$$

The PDF corresponding to (2.6) becomes

$$f(t) = \begin{cases} \frac{\alpha\theta_1^\alpha}{(\theta_1 + t)^{\alpha+1}}, & 0 < t < \tau \\ \frac{\alpha\theta_2^\alpha}{\left[\theta_2 + \tau\left(\frac{\theta_2}{\theta_1} - 1\right) + t\right]^{\alpha+1}} & \tau \le t < \infty \end{cases} \tag{2.8}$$

2.3 Objective of Study

For pre-fixed sample size n and the testing stress levels x_1 and x_2, the first objective is estimating the parameters a, b and α in a simple step-stress accelerated life test. The second objective is to obtain the optimal stress changing time τ which minimizes the asymptotic variance of the MLE of the P^{th} percentile of the lifetime distribution at normal stress condition $t_p(x_0)$.

3 ESTIMATION PROCEDURE

3.1 Point Estimates

Here the maximum likelihood method of estimation is used because ML method is very robust and gives the estimates of parameter with good statistical properties. In this method, the estimates of parameters are those values which maximize the sampling distribution of data. However, ML estimation method is very simple for one parameter distributions but its implementation in ALT is mathematically more intense and, generally, estimates of parameters do not exist in closed form, therefore, numerical techniques such as Newton Method, Some computer programs are used to compute them.

For obtaining the MLE of the model parameters, let t_{ij}, $j = 1,2,\cdots n_i$, $i = 1,2$ be the observed failure times of a test unit j under stress level i, where n_1 denotes the number of units failed at the low stress x_1 and n_2 denotes the number of units failed at higher stress level x_2. Therefore, the likelihood function for two-parameter Pareto distribution for simple step stress pattern can be written in the following form

$$L(\theta_1, \theta_2, \alpha) = \prod_{j=1}^{n_1} \frac{\alpha\theta_1^{\alpha}}{(\theta_1 + t_{1j})^{\alpha+1}} \prod_{j=1}^{n_2} \frac{\alpha\theta_2^{\alpha}}{\left[\theta_2 + \tau\left(\frac{\theta_2}{\theta_1} - 1\right) + t_{2j}\right]^{\alpha+1}} \tag{3.1}$$

The log-likelihood function $l\ (= \log L)$ corresponding to equation (3.1) can be rewritten as

$$\begin{aligned} l = \log L = n\log\alpha + n_1\alpha\log\theta_1 + n_2\alpha\log\theta_2 - (\alpha+1)\sum_{j=1}^{n_1}\log(\theta_1 + t_{1j}) \\ - (\alpha+1)\sum_{j=1}^{n_2}\log\left[\theta_2 + \tau\left(\frac{\theta_2}{\theta_1} - 1\right) + t_{2j}\right] \end{aligned} \tag{3.2}$$

where $n = n_1 + n_2$.

Now by using the relation $\log\theta(x) = a + bx_i$, $i = 1,2$ for the scale parameter θ, in (3.2), the likelihood function becomes,

$$l = n\log\alpha + n_1\alpha(a + bx_1) + n_2\alpha(a + bx_2) - (\alpha+1)\sum_{j=1}^{n_1}\log(e^{a+bx_1} + t_{1j})$$

$$-(\alpha+1)\sum_{j=1}^{n_2}\log\left[e^{a+bx_2} + \tau\left(\frac{e^{a+bx_2}}{e^{a+bx_1}} - 1\right) + t_{2j}\right] \tag{3.3}$$

Differentiating (3.3) partially with respect to a, b and α, we get

$$\frac{\partial l}{\partial a} = n\alpha - (\alpha+1)\sum_{j=1}^{n_1}\frac{e^{a+bx_1}}{[e^{a+bx_1} + t_{1j}]} - (\alpha+1)\sum_{j=1}^{n_2}\frac{e^{a+bx_2}}{[e^{a+bx_2} + \tau(e^{b(x_2-x_1)} - 1) + t_{2j}]} \tag{3.4}$$

$$\frac{\partial l}{\partial b} = n_1x_1\alpha + n_2x_2\alpha - (\alpha+1)\sum_{j=1}^{n_1}\frac{x_1e^{a+bx_1}}{[e^{a+bx_1} + t_{1j}]} - (\alpha+1)\sum_{j=1}^{n_2}\frac{x_2e^{a+bx_2} + \tau(x_2 - x_1)e^{b(x_2-x_1)}}{[e^{a+bx_2} + \tau(e^{b(x_2-x_1)} - 1) + t_{2j}]} \tag{3.5}$$

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + n_1(a + bx_1) + n_2(a + bx_2) - \sum_{j=1}^{n_1}\log[e^{a+bx_1} + t_{1j}] - \sum_{j=1}^{n_2}\log[e^{a+bx_2} + \tau(e^{b(x_2-x_1)} - 1) + t_{2j}] \tag{3.6}$$

From (3.6) the maximum likelihood estimates of α is given by the following equation:

$$\hat{\alpha} = \frac{n}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)} \tag{3.7}$$

where,

$$\psi_1 = \sum_{j=1}^{n_1} \log[e^{a+bx_1} + t_{1j}],$$

$$\psi_2 = \sum_{j=1}^{n_2} \log[e^{a+bx_2} + \tau(e^{b(x_2-x_1)} - 1) + t_{2j}].$$

By substituting for α into (3.4) and (3.5), the system equations are reduced into the following two non-linear equations:

$$\begin{aligned}\frac{\partial l}{\partial a} = &\frac{n^2}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)} \\ &-\left(\frac{n}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)} + 1\right)\sum_{j=1}^{n_1} \frac{e^{a+bx_1}}{[e^{a+bx_1} + t_{1j}]} \\ &-\left(\frac{n}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)} + 1\right)\sum_{j=1}^{n_2} \frac{e^{a+bx_1}}{[e^{a+bx_2} + \tau(e^{b(x_2-x_1)} - 1) + t_{2j}]}\end{aligned} \tag{3.8}$$

$$\begin{aligned}\frac{\partial l}{\partial b} = &\left(\frac{n}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)}\right)(n_1x_1 + n_2x_2) \\ &-\left(\frac{n}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)} + 1\right)\sum_{j=1}^{n_1} \frac{x_1e^{a+bx_1}}{[e^{a+bx_1} + t_{1j}]} \\ &-\left(\frac{n}{\psi_1 + \psi_2 - n_1(a + bx_1) - n_2(a + bx_2)} + 1\right)\sum_{j=1}^{n_2} \frac{x_2e^{a+bx_2} + \tau(x_2 - x_1)e^{b(x_2-x_1)}}{[e^{a+bx_2} + \tau(e^{b(x_2-x_1)} - 1) + t_{2j}]}\end{aligned} \tag{3.9}$$

Since (3.8) and (3.9) are non linear equations, their solutions are numerically obtained by using Newton Raphson method. They are solved simultaneously to obtain a and b. Then by substitution in (3.7) an estimate of α is easily obtained.

3.2 Interval Estimates

According to large sample theory, the maximum likelihood estimators, under some appropriate regularity conditions, are consistent and normally distributed. Since ML estimates of parameters are not in closed form, therefore, it is impossible to obtain the exact confidence intervals, so asymptotic confidence intervals based on the asymptotic normal distribution of ML estimators instead of exact confidence intervals are obtained here.

The Fisher-information matrix composed of the negative second partial derivatives of log likelihood function can be written as

$$F = \begin{bmatrix} -\frac{\partial^2 l}{\partial a^2} & -\frac{\partial^2 l}{\partial a \partial b} & -\frac{\partial^2 l}{\partial a \partial \alpha} \\ -\frac{\partial^2 l}{\partial b \partial a} & -\frac{\partial^2 l}{\partial b^2} & -\frac{\partial^2 l}{\partial b \partial \alpha} \\ -\frac{\partial^2 l}{\partial \alpha \partial a} & -\frac{\partial^2 l}{\partial \alpha \partial b} & -\frac{\partial^2 l}{\partial \alpha^2} \end{bmatrix}$$

The elements of information matrix F are:

$$\frac{\partial^2 l}{\partial a^2} = -(\alpha+1)\sum_{j=1}^{n_1}\frac{[e^{a+bx_1}+t_{1j}]e^{a+bx_1}-e^{2(a+bx_1)}}{[e^{a+bx_1}+t_{1j}]^2}$$

$$-(\alpha+1)\sum_{j=1}^{n_2}\frac{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]e^{a+bx_2}-e^{2(a+bx_2)}}{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]^2}$$

$$\frac{\partial^2 l}{\partial b^2} = -(\alpha+1)\sum_{j=1}^{n_1}\frac{[(x_1^2\, e^{a+bx_1})t_{1j}]}{[e^{a+bx_1}+t_{1j}]^2}$$

$$-(\alpha+1)\sum_{j=1}^{n_2}\frac{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}][x_2^2 e^{a+bx_2}+\tau(x_2-x_1)^2 e^{b(x_2-x_1)}]}{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]^2}$$

$$-\frac{[x_2 e^{a+bx_2}+\tau(x_2-x_1)e^{b(x_2-x_1)}]^2}{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]^2}$$

$$\frac{\partial^2 l}{\partial \alpha^2} = -\frac{n}{\alpha^2}$$

$$\frac{\partial^2 l}{\partial a \partial b} = \frac{\partial^2 l}{\partial b \partial a} = -(\alpha+1)\sum_{j=1}^{n_1}\frac{[(x_1\, e^{a+bx_1})t_{1j}]}{[e^{a+bx_1}+t_{1j}]^2}$$

$$-(\alpha+1)\sum_{j=1}^{n_2}\frac{[e^{a+bx_2}+[\tau(x_1 e^{b(x_2-x_1)}-x_2)+x_2 t_{2j}]}{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]^2}$$

$$\frac{\partial^2 l}{\partial a \partial \alpha} = \frac{\partial^2 l}{\partial \alpha \partial a} = n - \sum_{j=1}^{n_1}\frac{e^{a+bx_1}}{[e^{a+bx_1}+t_{1j}]} - \sum_{j=1}^{n_2}\frac{e^{a+bx_1}}{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]}$$

$$\frac{\partial^2 l}{\partial \alpha \partial b} = \frac{\partial^2 l}{\partial b \partial \alpha} = n_1 x_1 + n_2 x_2 - \sum_{j=1}^{n_1}\frac{x_1 e^{a+bx_1}}{[e^{a+bx_1}+t_{1j}]} - \sum_{j=1}^{n_2}\frac{x_2 e^{a+bx_2}+\tau(x_2-x_1)e^{b(x_2-x_1)}}{[e^{a+bx_2}+\tau(e^{b(x_2-x_1)}-1)+t_{2j}]}$$

The asymptotic variance-covariance matrix of $\hat{a},\hat{b}$ and $\hat{\alpha}$ is obtained by inverting the Fisher-information matrix that is

$$\Sigma = \begin{bmatrix} -\dfrac{\partial^2 l}{\partial a^2} & -\dfrac{\partial^2 l}{\partial a \partial b} & -\dfrac{\partial^2 l}{\partial a \partial \alpha} \\ -\dfrac{\partial^2 l}{\partial b \partial a} & -\dfrac{\partial^2 l}{\partial b^2} & -\dfrac{\partial^2 l}{\partial b \partial \alpha} \\ -\dfrac{\partial^2 l}{\partial \alpha \partial a} & -\dfrac{\partial^2 l}{\partial \alpha \partial b} & -\dfrac{\partial^2 l}{\partial \alpha^2} \end{bmatrix}^{-1} = F^{-1}$$

$$= \begin{bmatrix} AVar(\hat{a}) & ACov(\hat{a}\hat{b}) & ACov(\hat{a}\hat{\alpha}) \\ ACov(\hat{b}\hat{a}) & AVar(\hat{b}) & ACov(\hat{b}\hat{\alpha}) \\ ACov(\hat{\alpha}\hat{a}) & ACov(\hat{\alpha}\hat{b}) & AVar(\hat{\alpha}) \end{bmatrix}$$

Now, the two-sided approximate $100\lambda\%$ confidence limits for population parameters $\hat{a}, \hat{b}$ and $\hat{\alpha}$ can be constructed as

$$\left[\hat{a} \pm Z_{\lambda} \sqrt{AVar(\hat{a})}\right]$$
$$\left[\hat{b} \pm Z_{\lambda} \sqrt{AVar(\hat{b})}\right]$$
$$\left[\hat{\alpha} \pm Z_{\lambda} \sqrt{AVar(\hat{\alpha})}\right]$$

4 OPTIMAL TEST PLAN

The optimum criterion here is to find the optimum stress change time τ. Since the accuracy of ML method is measured by the asymptotic variance of the MLE of the $100P^{th}$ percentile of the lifetime distribution at normal stress condition $t_p(x_0)$, therefore the optimum value of the stress change time will the value which minimizes the asymptotic variance of the MLE of $t_p(x_0)$.

The $100P^{th}$ percentile of a distribution $F()$ is the age t_p by which a proportion of population fails Nelson (1990). It is a solution of the equation $P = F(t_p)$, therefore the $100P^{th}$ percentile for Pareto distribution is

$$t_p = \frac{\theta\left\{1-(1-P)^{1/\alpha}\right\}}{(1-P)^{1/\alpha}}$$

The $100P^{th}$ percentile for Pareto distribution at use condition is

$$t_p(x_0) = \frac{\exp(a+bx_0)\left\{1-(1-P)^{1/\alpha}\right\}}{(1-P)^{1/\alpha}}$$

Now the asymptotic variance of MLE of the $100P^{th}$ percentile at normal operating conditions is given by

$$AVar(t_p(\tilde{x}_0)) = \left[\frac{\partial t_p(\hat{x}_0)}{\partial \hat{a}}, \frac{\partial t_p(\hat{x}_0)}{\partial \hat{b}}, \frac{\partial t_p(\hat{x}_0)}{\partial \hat{\alpha}}\right] \Sigma \left[\frac{\partial t_p(\hat{x}_0)}{\partial \hat{a}}, \frac{\partial t_p(\hat{x}_0)}{\partial \hat{b}}, \frac{\partial t_p(\hat{x}_0)}{\partial \hat{\alpha}}\right]^{-1}$$

The optimum stress change time τ will be the value which minimizes $AVar(t_p(\hat{x}_0))$.

5 SIMULATION STUDY

To evaluate the performance of the method of inference described in present study, several data sets with sample sizes n=100, 200,...,500 are generated for from two-parameter Pareto distribution. The values for true parameters and stress combinations are chosen to be $a = 0.5, b = 0.2$ $\alpha = 1.5$ and $(x_1, x_2) = (2,4),(3,5)$. The estimates and the corresponding summary statistics are obtained by the present Step Stress ALT model and the Newton iteration method. For different given samples and stresses combinations with $a = 0.5, b = 0.2$ and $\alpha = 1.5$, the ML estimates , asymptotic variance, the asymptotic standard error (SE), the mean squared error (MSE) and the coverage rate of the 95% confidence interval for a, b and α are obtained. Table-1 and 2 summarize the results of the estimates for a, b and α. The numerical results presented in Table-1 and 2 are based on 1000 simulation replications.

Table1: Simulations results based on Step stress with $a = 0.5, b = 0.2$ $\alpha = 1.5$ and $(x_1, x_2) = (2,4)$

Sample Size *n*	**Parameter**	**MLE**	**Variance**	**SE**	**MSE**	**95% Asymptotic CI Coverage**
	a	0.51069	0.01349	0.01368	0.01360	0.94282
100	b	0.20353	0.00084	0.00083	0.00085	0.95030
	α	1.52507	0.04789	0.04737	0.04851	0.94984
	a	0.50734	0.00614	0.00653	0.00619	0.95866
200	b	0.20098	0.00038	0.00040	0.00038	0.95766
	α	1.50620	0.02194	0.02290	0.02198	0.95595
	a	0.50339	0.00425	0.00428	0.00426	0.94789
300	b	0.20132	0.00027	0.00027	0.00027	0.95075
	α	1.50997	0.01561	0.01527	0.01571	0.95090
	a	0.50258	0.00302	0.00318	0.00303	0.95290
400	b	0.20096	0.00019	0.00020	0.00019	0.96192
	α	1.50735	0.01095	0.01141	0.01100	0.96200
	a	0.50323	0.00239	0.00255	0.00241	0.95795
500	b	0.20059	0.00015	0.00016	0.00015	0.95595
	α	1.50444	0.00895	0.00908	0.00897	0.95595

Table2: Simulations results based on Step stress with $a = 0.5, b = 0.2$ $\alpha = 1.5$ and $(x_1, x_2) = (3,5)$

Sample Size *n*	**Parameter**	**MLE**	**Variance**	**SE**	**MSE**	**95% Asymptotic CI Coverage**
	a	0.51740	0.01333	0.01404	0.01363	0.94964
100	b	0.20057	0.00040	0.00040	0.00040	0.95066
	α	1.50444	0.02299	0.02258	0.02301	0.95159
	a	0.50652	0.00793	0.00888	0.00797	0.94478
200	b	0.20171	0.00027	0.00027	0.00028	0.94979
	α	1.51273	0.01582	0.01521	0.01598	0.94887
	a	0.50692	0.00578	0.00633	0.00583	0.95030
300	b	0.20072	0.00020	0.00020	0.00020	0.94726
	α	1.50522	0.01152	0.01133	0.01155	0.94736
	a	0.50487	0.00443	0.00501	0.00446	0.96146
400	b	0.20067	0.00015	0.00015	0.00015	0.95740
	α	1.50506	0.00872	0.00900	0.00874	0.95740
	a	0.50257	0.00279	0.00298	0.00279	0.95595
500	b	0.20084	0.00011	0.00011	0.00011	0.95682
	α	1.50453	0.00555	0.00568	0.00557	0.95095

6 SUMMARY AND CONCLUDING REMARKS

This paper deals with parameter estimation of Pareto distribution under simple step stress ALT plan. The MLEs of the model parameters were obtained. The MLEs, the asymptotic variance and covariance of model parameters were obtained. Based on the asymptotic normality, the coverage rate of 95% confidence intervals of the model parameters are obtained. Optimal plan for step stress ALT is also determined by minimizing the asymptotic variance of the MLE of the $100\,P^{th}$ percentile of the lifetime distribution at normal stress condition.

From results in Table 1 and 2, it is observed that $\hat{a},\hat{b}$ and $\hat{\alpha}$ estimates the true parameters a, b and α quite well respectively with relatively small mean squared errors. The estimated standard error also approximates well the sample standard deviation. For a fixed a, b and α we find that as n increases, variance, standard error and the mean squared errors of $\hat{a},\hat{b}$ and $\hat{\alpha}$ get smaller. This is because that a larger sample size results in a better large sample approximation. It is also noticed that the coverage probabilities of the asymptotic confidence interval are close to the nominal level and do not change much across the five different sample sizes. In short, it is reasonable to say that the present step stress ALT plan works well and has a promising potential in the analysis of accelerated life testing.

REFERENCES

Alhadeed, A. A. and Yang, S. S. (2002): Optimal Simple Step-Stress Plan for Khamis-Higgins Model, IEEE Transactions on Reliability, vol. 51, 212-215.

Al-Masri, A.Q. and Al-Haj Ebrahem, M. (2009): Optimum Times for Step-Stress Cumulative Exposure Model Using Log-Logistic Distribution with Known Scale Parameter, Austrian Journal of Statistics, vol. 38, 59–66.

Bai, D.S., Chun, Y.R. (1991): Optimum simple step-stress accelerated life tests with competing causes of failure, IEEE Transactions on Reliability, vol.40, 622-627.

Bai, D.S., Kim, M.S. and Lee, S.H. (1989): Optimum simple step-stress accelerated life tests with censoring, IEEE Transactions on Reliability, vol. 38, 528-532.

Balakrishnan, N. and Han, D. (2008): Optimal step-stress testing for progressively Type-I censored data from exponential distribution, Journal of Statistical Planning and Inference, vol. 138, 4172-4186.

Balakrishnan, N., Beutner, E., and Kateri, M. (2009): Order Restricted Inference for Exponential Step-Stress Models, IEEE Transactions on Reliability, vol. **58**, 132-142.

Fan T.H., Wang W.L. and Balakrishnan, N. (2008): Exponential progressive step-stress life testing with link function based on Box-Cox transformation, Journal of Statistical Planning and Inference, vol. 138, 2340-2354.

Gouno, E., Sen, A. and Balakrishnan, N. (2004): Optimal step-stress test under progressive Type-I censoring, IEEE Transactions on Reliability, vol. 53, 383-393.

Hassan, A. S. and Al-Ghamdi, A. S. (2009): Optimum Step Stress Accelerated Life Testing for Lomax Distribution, Journal of Applied Sciences Research, vol. 5, 2153-2164.

Khamis, I.H. and Higgins, J.J. (1998): A new model for step-stress testing, IEEE Transactions on Reliability, vol.47, 131-134.

Lu, M.W. & Rudy, R.J. (2002): Step-stress accelerated test, International Journal of Materials & Product Technology, vol. 17, 425-434.

McSorley, E.O., Lu, J.C. and Li, C.S. (2002): Performance of parameter-estimates in step-stress accelerated life-tests with various sample-sizes, *IEEE Transactions on Reliability*, vol. 51, 271-277.

Miller, R. and Nelson, W. (1983): Optimum simple step stress plans for accelerated life testing, IEEE Transactions on Reliability, vol. 32, 59-65.

Nelson, W. (1980): Accelerated life testing step-stress models and data analysis, IEEE Transactions on Reliability, vol. 29, 103-108.

Nelson, W. (1990): Accelerated Testing: Statistical Models, Test Plans and Data Analyses, John Wiley & Sons, New York.

Pareto, V. (1897): *Cours d'Economie Politique*, 2. F. Rouge, Lausanne, Switzerland.

Watkins, A. J. (2001): Commentary: inference in simple step-stress models, IEEE Transactions on Reliability, vol. 50, 36-37.

Wu, S.J., Lin, Y.P. and Chen, Y.J. (2006): Planning step-stress life test with progressively type I group-censored exponential data, Statist. Neerlandica, vol. 60, 46-56.

Xiong, C. (1998): Step stress model with threshold parameter, *Journal of Statistical Computation and Simulation*, vol. 63, 349–360.

Xiong, C. and Ji, M. (2004): Analysis of grouped and censored data from step-stress life test, IEEE Transactions on Reliability, vol. 53, 22–28.

Xiong, C. and Milliken, G. A. (1999): Step-stress life-testing with random stress-change times for exponential data, IEEE Transactions on Reliability, vol. 48, 141-148.

Xu, H.Y. and Fei, H. (2012): Models Comparison for Step-Stress Accelerated Life Testing, Communications in Statistics - Theory and Methods, vol. 41, 3878-3887.

Yeo, K.P. and Tang, L.C. (1999): Planning step-stress life-test with a target accelerated factor, IEEE Transactions on Reliability, vol. 48, 61-67.

Zhao, W. and Elsayed, E. A. (2005): A General accelerated life model for step-stress testing. IIE Transactions, vol. **37**, 1059-1069.

OPTIMAL REDUNDANCY IN SYSTEMS WITH MULTI-LEVEL UNITS

Igor Ushakov

ABSTRACT

Method of Universal Generating Function (UGF) was introduced in [1]-[5] and got further fundamental developing in [7]-[8]. Here we give an example how method of UGF can be implemented to solution problems of optimal redundancy for systems consisting of multi-level units.

The Method of Universal Generating Functions (U-functions) was introduced in [1]-[5]. Before detailed consideration of this method, let us remark that applying to the optimal redundancy problem this method represents a modification of the Kettelle's Algorithm [1] conveniently arranged for calculations with the use of computer.

Detailed description of the UGF method can be found in [1]-[5] and [5].

For the Reader's convenience, we begin with a numerical example that can explain the idea of the problem solution more transparently than general arguing. The final description of the algorithm is given at the end.

<u>Example</u>. Consider a simplest series system of two units (see figure below).

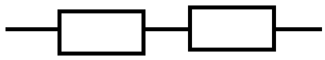


Figure 1. Series system consisting of two units.

However, each unit itself is not a simple binary element but multistate element that is characterized by several levels of performance. Performance may be measured various physical values. Effectiveness of such system operation depends on levels of performance of Unit-1 and Unit-2.

Let units are characterized by the following parameters:

Unit-1

Level of performance (W_1)	Probability p_1	Cost of a single unit
100%	p_{11}=Pr{ W_1=100%}=0.9	c_1=1
70%	p_{12}=Pr{ W_1=100%}=0.05	
40%	p_{13}=Pr{ W_1=100%}=0.04	
0%	p_{14}=Pr{ W_1=100%}=0.01	

Unit-2

Level of performance (W_1)	Probability p_2	Cost of a single unit
100%	P_{21}=Pr{ W_2=100%}=0.8	c_2=2
80%	P_{22}=Pr{ W_2=80%}=0.18	
20%	P_{23}=Pr{ W_2=20%}=0.01	
0%	P_{24}=Pr{ W_2=0%}=0.01	

Assume that performance effectiveness of each unit can be improved by using simple redundancy and that each moment of time unit performance is equal to the performance of the best component of the redundant group. Thus, behavior of Unit-1, consisting of the main component and single redundant element, can be depicted as in Figure 2.

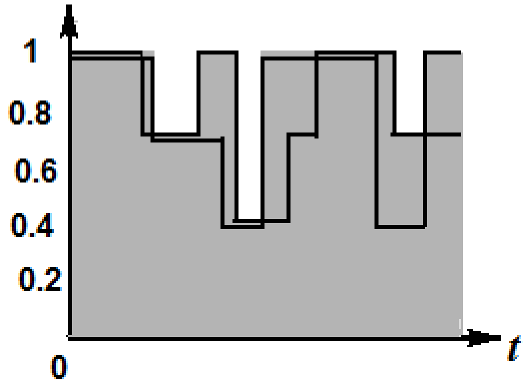


Figure 2. A realization of stochastic behavior of Unit-1, consisting of two elements, main and redundant. The shadowed area denotes the behavior of the Unit-1.

For Unit-2 analogous process is presented in Figure 3.

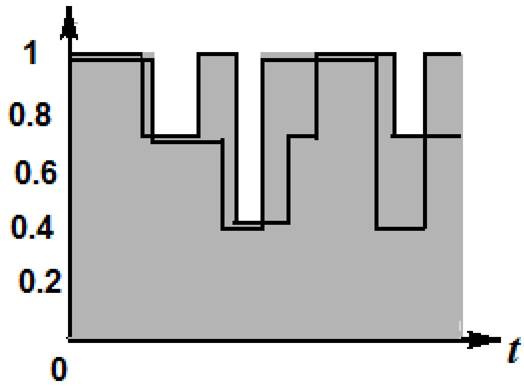


Figure 3. A realization of stochastic behavior of Unit-2, consisting of two elements, main and redundant. The shadowed area denotes the behavior of the Unit-2.

Further, assume that the entire system (series connection of Unit-1 and Unit-2) is characterized by the worst level of effectiveness of its units at each moment of time. In Figure 4, one can see the system behavior for the case when both units consist of a single main element.

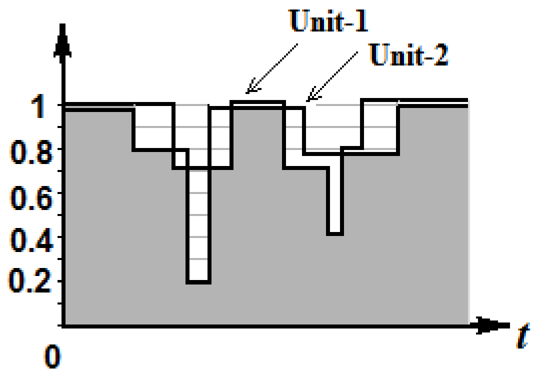


Figure 4. A realization of stochastic behavior of the entire system when both its units consist of a single main element. The shadowed area denotes the behavior of the system.

Let the problem is to find optimal redundant elements allocation for solving two optimal redundancy problems:

(1) Direct problem: Find such an allocation of redundant elements than delivers average level of the system performance not less than W^0 with minimum possible cost of redundant elements;
(2) Inverse problem: Find such an allocation of redundant elements than delivers maximum possible level of system performance under condition that the total expenses on redundant elements do not exceed C^0 units of cost.

Now consider construction of dominating sequence during the optimization process. (For details about dominating sequence, see [1] or [2].) In principle, one has to construct a table of type that presented below and choose members of dominating sequence.

Table 1. Construction of dominating sequence.

		Number of redundant elements for Unit-1			
		0	**1**	**2**	**…**
	0	X=(0, 0) P(0, 0) W(0,0) C(0, 0)	X=(1, 0) P(1, 0) W(1, 0) C(1, 0)	X=(2, 0) P(2, 0) W(2, 0) C(2, 0)	**…**
	1	X=(0, 1) P(0, 1) W(0,1) C(0, 1)	X=(1, 1) P(1, 1) W(1, 1) C(1, 1)	X=(2, 1) P(2, 1) W(2, 1) C(2, 1)	**…**
Number of redundant elements for Unit-2	**2**	X=(0, 2) P(0, 2) W(0, 2) C(0, 2)	X=(1, 2) P(1, 2) W(1, 2) C(1, 2)	X=(2, 2) P(2, 1) W(2, 2) C(2, 2)	**…**
	…	**…**	**…**	**…**	**…**

Further discussion will be provided in terms of Universal Generating Functions. As one sees, in this case we deal with quadruplets of type:

{**Vector of units' variants; Discrete distribution of performance levels; System cost**}.

The problem complicates due to necessity of calculations because Probabilities of performance levels and Performance levels are not numbers but vectors that needed special calculations. This aspect will be demonstrated below. Here we would like to note that there is no necessity to calculate quadruplets for all cells of Table 1. Fortunately, we can use the property of Kettelle Algorithm: members of dominating sequences are located around table's diagonal and corresponding cells form simply connected area. It allows to use "dichotomy tree" procedure, i.e. avoid unnecessary calculations by cutting non-perspective branches (see Figure 5).

Indeed, consider bordering cells around simple connected area (they marked with sign "x".). There is no dominating cells in area located upper the right border, and there is no dominating cells in area located lower the left border.

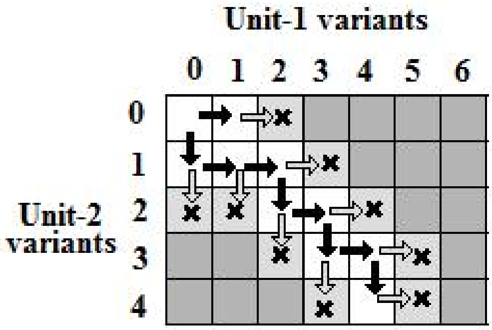


Figure 5. Example of excluding non-perspective branches. Black arrows are members of dominating sequence, Grey arrows are trial test that led to non-perspective variants marked by "x". All cells marked with dark grey cannot contain dominating quadruplets.

Thus, in this case calculations occur to be sufficiently compact. However, as we mentioned above some special calculations for each redundant group have to be done.

Let us consider a numerical example.

In accordance with described above calculating procedure, one has to consider first variant (0, 0), i.e just Unit-1 and Unit-2 with no redundancy at all, and find quadruple, In this case resulting solution will be:

$\{0;\ [(p_{11}, W_{11}), (p_{12}, W_{12}), (p_{13}, W_{13}), (p_{14}, W_{14})]; c_1\} \otimes \{0;\ [(p_{21}, W_{21}), (p_{22}, W_{22}), (p_{23}, W_{23}), (p_{24}, W_{24})]; c_2\} = \{0 \underset{\rightarrow}{\otimes} 0;\ [(p_{11}, W_{11}), (p_{12}, W_{12}), (p_{13}, W_{13}), (p_{14}, W_{14})] \underset{UGF}{\otimes} \{0;\ [(p_{21},W_{21}), (p_{22}, W_{22}), (p_{23}, W_{23}), (p_{24}, W_{24})];\ c_1 \underset{+}{\otimes} c_2\}.$

Here we use the following operators:

$\underset{\rightarrow}{\otimes}$ is an operator of forming a vector, i.e. $j \underset{\rightarrow}{\otimes} k = (j,k)$;

$\underset{UGF}{\otimes}$ is an operator equivalent to the U-function, i.e. $\left(\sum_{j\in A} p_j z^{W_j}\right) \underset{UGF}{\otimes} \left(\sum_{k\in B} p_k z^{W_k}\right) = \sum_{\forall j,\forall k} p_j \cdot p_k z^{W_j \underset{\min}{\otimes} W_k}$,

where, in turn, $W_j \underset{\min}{\otimes} W_k = \min(W_j, W_k)$; $c_1 \underset{+}{\otimes} c_2$ is operator of summation, i.e. $c_1 \underset{+}{\otimes} c_2 = c_1 + c_2$.

Numerical results are presented in Table 2.

This leads to the following final result:

$P^{(0.0)}(W_{syst}=100\%) = 0.72$;
$P^{(0.0)}(W_{syst}=80\%) = 0.171$;
$P^{(0.0)}(W_{syst}=70\%) = 0.04+0.0095=0.0495$;
$P^{(0.0)}(W_{syst}=40\% = 0.032+0.0076=0.0396$;
$P^{(0.0)}(W_{syst}=20\%) = 0.009+0.0005+0.004=0.0099$;
$P^{(0.0)}(W_{syst}=0\%) = 0.008+0.0019+0.0001+0.0001+0.0009+0.0005+0.0004= 0.0201$.

Cost of additional units in this case equals 0. As one can easily calculate, the average level of the system performance is equal to

$W_{syst}^{(0,0)} = 0.72 + 0.171 \cdot 0.8 + 0.0497 \cdot 0.7 + 0.0396 \cdot 0.5 + 0.0095 \cdot 0.2 \approx 0.9092.$

Table 2. Step 1 of the process of optimization

(0, 0)		Unit-2					
		p_{21}=0.8 $W_{21}^{(0)}$**=100%**	**p_{22}=0.18** $W_{22}^{(0)}$**=80%**			**p_{23}=0.01** $W_{23}^{(0)}$**=20%**	**p_{24}=0.01** $W_{24}^{(0)}$**=0%**
Unit-1	**p_{11}=0.9** $W_{11}^{(0)}$ **=100%**	$p_{21}\cdot p_{11}$= 0.72 $\min(W_{21}^{(0)},W_{11}^{(0)})$ =100%	$p_{22}\cdot p_{11}$= 0.171 $\min(W_{22}^{(0)},W_{11}^{(0)})$ =80%			$p_{23}\cdot p_{14}$= 0.009 $\min(W_{23}^{(0)},W_{11}^{(0)})$ =20%	$p_{24}\cdot p_{14}$= 0.009 $\min(W_{21}^{(0)},W_{11}^{(0)})$ =0%
	p_{12}=0.05 $W_{12}^{(0)}$ **=70%**	$p_{21}\cdot p_{12}$= 0.04 $\min(W_{21}^{(0)},W_{12}^{(0)})$ =70%	$p_{22}\cdot p_{12}$=0.0095 $\min(W_{22}^{(0)},W_{12}^{(0)})$ =70%			$p_{23}\cdot p_{14}$= 0.0005 $\min(W_{23}^{(0)},W_{12}^{(0)})$ =20%	$p_{24}\cdot p_{14}$= 0.0005 $\min(W_{21}^{(0)},W_{11}^{(0)})$ =0%
	p_{13}=0.04 $W_{13}^{(0)}$ **=40%**	$p_{21}\cdot p_{13}$= 0.032 $\min(W_{21}^{(0)},W_{13}^{(0)})$=40%	$p_{22}\cdot p_{13}$=0.0076 $\min(W_{22}^{(0)},W_{13}^{(0)})$ =40%			$p_{23}\cdot p_{14}$= 0.0004 $\min(W_{23}^{(0)},W_{13}^{(0)})$ =20%	$p_{24}\cdot p_{14}$= 0.0004 $\min(W_{21}^{(0)},W_{11}^{(0)})$ =0%
	p_{14}=0.01 $W_{14}^{(0)}$ **=0%**	$p_{21}\cdot p_{14}$= 0.008 $\min(W_{21}^{(0)},W_{14}^{(0)})$ =0%	$p_{22}\cdot p_{14}$= 0.0019 $\min(W_{22}^{(0)},W_{14}^{(0)})$ =0%			$p_{23}\cdot p_{14}$= 0.0001 $\min(W_{23}^{(0)},W_{14}^{(0)})$ = 0%	$p_{24}\cdot p_{14}$= 0.0001 $\min(W_{23}^{(0)},W_{14}^{(0)})$ = 0%

Now let's make trial steps to the neighbor cells: check cells (1, 0) and (0, 1). Let us start with cell (1, 0) in accordance with Figure 4. First find performance levels distribution for Unit-1 consisting of two elements, main and redundant.

Table 3. Forehand calculation of performance levels distribution for Unit-1, consisting of two elements, main and redundant.

		Element-1			
		p_{11}=0.9 $W_{11}^{(0)}$**=100%**	**p_{12}=0.05** $W_{12}^{(0)}$**=70%**	**p_{13}=0.04** $W_{13}^{(0)}$)**=40%**	**p_{14}=0.01** $W_{14}^{(0)}$**=0%**
Element-1	**p_{11}=0.9** $W_{11}^{(0)}$ **=100%**	$(p_{11})^2$=0.81 $W_{11}^{(0)}$=100%	$p_{12}\cdot p_{11}$=0.045 $\max(W_{12}^{(0)},W_{11}^{(0)})$ =100%	$p_{13}\cdot p_{11}$=0.036 max $(W_{13}^{(0)},W_{11}^{(0)})$ =100%	$p_{14}\cdot p_{11}$=0.009 max $(W_{14}^{(0)},W_{11}^{(0)})$ =100%
	p_{12}=0.05 $W_{12}^{(0)}$ **=70%**	$p_{11}\cdot p_{12}$=0.045 max $(W_{11}^{(0)},W_{12}^{(0)})$ =100%	$(p_{12})^2$=0.025 max$(W_{12}^{(0)}, W_{12}^{(0)})$ =70%	$p_{13}\cdot p_{12}$=0.002 max $(W_{13}^{(0)},W_{12}^{(0)})$ =70%	$p_{14}\cdot p_{12}$=0.0005 max $(W_{14}^{(0)},W_{12}^{(0)})$) =70%
	p_{13}=0.04 $W_{32}^{(0)}$ **=40%**	$p_{11}\cdot p_{13}$=0.036 max $(W_{11}^{(0)},W_{32}^{(0)})$ =100%	$p_{12}\cdot p_{13}$=0.002 max $(W_{12}^{(0)},W_{32}^{(0)})$ =70%	$(p_{13})^2$=0.0016 $W_{32}^{(0)}$=40%	$p_{14}\cdot p_{13}$=0.0004 max $(W_{14}^{(0)},W_{32}^{(0)})$ =40%
	p_{14}=0.01 $W_{14}^{(0)}$**=0%**	$p_{11}\cdot p_{14}$=0.009 max $(W_{11}^{(0)},W_{14}^{(0)})$ =100%	$p_{12}\cdot p_{14}$=0.0005 max $(W_{12}^{(0)},W_{14}^{(0)})$ =70%	$p_{13}\cdot p_{14}$=0.0004 max $(W_{13}^{(0)},W_{14}^{(0)})$ =40%	$(p_{14})^2$=0.0001 $W_{14}^{(0)}$=0%

On the basis of this table, one gets for Unit-1 the following distribution

$\Pr\{W_1^{(1)} = 100\%\} = P_{11}^{(1)}= (p_{11})^2+2p_{11}\cdot(p_{12}+ p_{13}+ p_{14}) = 0.81+2\cdot (0.045+0.036+0.009)=0.99$;

$\Pr\{W_1^{(1)} = 70\%\} = P_{12}^{(1)}=(p_{12})^2+2 \cdot p_{12} \cdot (p_{13}+ p_{14})=0.025+2\cdot 0.025 (0.002+0.0005)=0.0075$;

$\Pr\{W_1^{(1)} = 40\%\} = P_{13}^{(1)}=(p_{13})^2+2p_{13}\cdot p_{14}=0.0016+2\cdot 0.0016\cdot 0.0004\approx 0.0016$:

$\Pr\{W_1^{(1)} = 0\%\} = P_{14}^{(1)} = (p_{14})^2=0.0001$.

Using these results, one can compile Table 4 that gives performance levels distribution for the system characterized by vector of redundant elements $X = (1, 0)$.

Table 4. Step 3 of the optimization process.

(1, 0) $C_{system}=c_1 = 1$		**Unit-2**			
		p_{21}=0.8 $W_{21}^{(0)}$=100%	p_{22}=0.19 $W_{22}^{(0)}$=80%	p_{23}=0.01 $W_{23}^{(0)}$=20%	p_{24}=0.01 $W_{24}^{(0)}$=0%
Unit -1	$P_{11}^{(1)}$=0.99 $W_{11}^{(1)}$ =100%	$p_{21}\cdot P_{11}^{(1)}$= 0.792 min ($W_{21}^{(0)}, W_{11}^{(1)}$) =100%	$p_{22}\cdot P_{11}^{(1)}$= 0.188 min($W_{22}^{(0)}, W_{11}^{(1)}$) =80%	$p_{23}\cdot P_{11}^{(1)}\approx$0.01 min($W_{23}^{(0)}, W_{11}^{(1)}$) =20%	$p_{24}\cdot P_{11}^{(1)}\approx$0.01 min($W_{24}^{(0)}, W_{11}^{(1)}$)) =0%
	$P_{12}^{(1)}$=0.0075 $W_{12}^{(1)}$ =70%	$p_{21}\cdot P_{12}^{(1)}$= 0.006 min($W_{21}^{(0)}, W_{12}^{(1)}$) =70%	$p_{22}\cdot P_{12}^{(1)}\approx$0.0014 min($W_{22}^{(0)}, W_{12}^{(1)}$) =70%	$p_{23}\cdot P_{12}^{(1)}\approx$0.0001 min($W_{23}^{(0)}, W_{12}^{(1)}$) =20%	$p_{24}\cdot P_{12}^{(1)}$=0.0001 min($W_{24}^{(0)}, W_{12}^{(1)}$) =0%
	$P_{13}^{(1)}$=0.0016 $W_{13}^{(1)}$ =40%	$p_{21}\cdot P_{13}^{(1)}$0.0013 min($W_{21}^{(0)}, W_{13}^{(1)}$) =40%	$p_{22}\cdot P_{13}^{(1)}\approx$0.0003 min($W_{22}^{(0)}, W_{13}^{(1)}$) =40%	$p_{23}\cdot P_{13}^{(1)}\approx$0 min($W_{23}^{(0)}, W_{13}^{(1)}$) =20%	$p_{24}\cdot P_{13}^{(1)} \approx 0$ min($W_{24}^{(0)}, W_{13}^{(1)}$) =0%
	$P_{14}^{(1)}$=0.0001 $W_{14}^{(1)}$ =0%	$p_{21}\cdot P_{14}^{(1)}\approx$0.0001 min($W_{21}^{(0)}, W_{14}^{(1)}$) =0%	$p_{22}\cdot P_{14}^{(1)}\approx$0 min($W_{22}^{(0)}, W_{14}^{(1)}$) =0%	$p_{23}\cdot P_{14}^{(1)}\approx$0 min($W_{23}^{(0)}, W_{14}^{(1)}$) =0%	$p_{24}\cdot P_{14}^{(1)}\approx$0 min($W_{24}^{(0)}, W_{14}^{(1)}$) =0%

This leads to the following final result:

$P^{(1.0)}(W_{syst}=100\%) =0.792$;

$P^{(0.0)}(W_{syst}=80\%) =0.188$;

$P^{(0.0)}(W_{syst}=70\%) =0.006+0.0014=0.0074$;

$P^{(0.0)}(W_{syst}=40\% =0.0013+0.0003=0.0016$;

$P^{(0.0)}(W_{syst}=20\%) =0.01+0.0001=0.0101$;

$P^{(0.0)}(W_{syst}=0\%) =0.008+0.0019+0.0001+0.0001+0.0009+0.0005+0.0004= 0.0201$.

Cost of additional units in this case equals 1. Average system's performance level equals

$W_{syst}^{(1,0)} = 0.792 + 0.188\cdot 0.8 + 0.0497\cdot 0.7 + 0.0396\cdot 0.4 + 0.0095\cdot 0.2 \approx 0.9502$.

Then try another neighbor cell, namely (0, 1). Beforehand, one has to perform an additional calculation of performance levels distribution for Unit-2 consisting of two elements, main and redundant.

Table 5. Forehand calculation of performance levels distribution for Unit-2, consisting of two elements, main and redundant.

		Element-2			
Element-2		p_{21}=0.8 $W_{21}^{(0)}$=100%	p_{22}=0.19 $W_{22}^{(0)}$=80%	p_{23}=0.01 $W_{23}^{(0)}$=20%	p_{24}=0.01 $W_{24}^{(0)}$=0%
	p_{21}=0.8 $W_{21}^{(0)}$=100%	$(p_{21})^2$=0.64 $W_{21}^{(0)}$=100%	$p_{22} \cdot p_{21}$=0.045 max $(W_{22}^{(0)}, W_{21}^{(0)})$ =100%	$p_{23} \cdot p_{21}$=0.036 max $(W_{23}^{(0)}, W_{21}^{(0)})$ =100%	$p_{24} \cdot p_{21}$=0.008 max $(W_{24}^{(0)}, W_{21}^{(0)})$ =100%
	p_{22}=0.19 $W_{21}^{(0)}$=80%	$p_{21} \cdot p_{22}$=0.152 max $(W_{21}^{(0)}, W_{21}^{(0)})$ =100%	$(p_{22})^2$=0.0361 $W_{22}^{(0)}$=80%	$p_{23} \cdot p_{22}$=0.0002 max $(W_{23}^{(0)}, W_{21}^{(0)})$ =80%	$p_{24} \cdot p_{22}$=0.0002 max $(W_{24}^{(0)}, W_{21}^{(0)})$ =80%
	p_{23}=0.01 $W_{23}^{(0)}$=20%	$p_{21} \cdot p_{23}$=0.008 max $(W_{21}^{(0)}, W_{23}^{(0)})$ =100%	$p_{22} \cdot p_{23}$=0.0002 max $(W_{22}^{(0)}, W_{23}^{(0)})$ =80%	$(p_{23})^2$=0.0001 $W_{23}^{(0)}$=20%	$p_{24} \cdot p_{23}$=0.0001 max $(W_{24}^{(0)}, W_{23}^{(0)})$ =20%
	p_{24}=0.01 $W_{24}^{(0)}$=0%	$p_{21} \cdot p_{24}$=0.008 max $(W_{21}^{(0)}, W_{24}^{(0)})$ =100%	$p_{222} \cdot p_{24}$=0.0002 max $(W_{22}^{(0)}, W_{24}^{(0)})$ =70%	p $p_{23} \cdot p_{24}$=0.0001 max $(W_{23}^{(0)}, W_{24}^{(0)})$ =20%	$(p_{24})^2$=0.0001 $W_{24}^{(0)}$=0%

On the basis of this table, one gets for Unit-2, consisting of two elements, the following distribution

$\Pr\{W_2^{(1)} = 100\%\} = P_{21}^{(1)} = (p_{21})^2 + 2p_{21} \cdot (p_{22} + p_{23} + p_{34}) = 0.64 + 2 \cdot 0.8 \cdot (0.045+0.036+0.008) \approx 0.7709;$

$\Pr\{W_2^{(1)} = 80\%\} = P_{22}^{(1)} = (p_{22})^2 + 2 \cdot p_{22} \cdot (p_{23} + p_{24}) = 0.0361 + 2 \cdot 0.0361\ (0.0002+0.0002) \approx 0.0361;$

$\Pr\{W_2^{(1)} = 20\%\} = P_{23}^{(1)} = (p_{13})^2 + 2p_{13} \cdot p_{14} = 0.0001+0.0001+0.0001 = 0,0003:$

$\Pr\{W_2^{(1)} = 0\%\} = P_{24}^{(1)} = (p_{14})^2 = 0.0001.$

After such preparations, one can construct a table with system's performance levels distribution for the system configuration characterized by vector of redundant elements $X = (0, 1)$. This leads to the following final result:

$P^{(1.0)}(W_{syst}=100\%) = 0.6038;$
$P^{(0.0)}(W_{syst}=80\%) = 0.0325;$
$P^{(0.0)}(W_{syst}=70\%) = 0.0386+0.0018=0.0404;$
$P^{(0.0)}(W_{syst}=40\% = 0.0308+0.0014=0.0322;$
$P^{(0.0)}(W_{syst}=20\%) \approx 0.0003;$
$P^{(0.0)}(W_{syst}=0\%) = 0.0077+0.0004+0.0001 \approx 0.0082.$

Cost of additional units in this case equals 2 units of cost. Average system's performance level equals

$$W_{syst}^{(0,1)} = 0.6038 + 0.0.0325 \cdot 0.8 + 0.0404 \cdot 0.7 + 0.0322 \cdot 0.4 + 0.0003 \cdot 0.2 \approx 0.671.$$

Table 6. Step 4 of the optimization process.

(0, 1) **C**$_{system}$=c_2 = 2		**Unit-2**					
		$P_{21}^{(1)}$=**0.7709** $W_{21}^{(1)}$=**100%**	$P_{22}^{(1)}$=**0.0361** $W_{22}^{(1)}$=**80%**			$P_{23}^{(1)}$=**0.0003** $W_{23}^{(1)}$=**20%**	$P_{24}^{(1)}$=**0.0001** $W_{24}^{(1)}$=**0%**
Unit-1	p_{11}=**0.9** $W_{11}^{(0)}$=**100%**	$P_{21}^{(1)} \cdot p_{11} \approx 0.6038$ min $(W_{21}^{(1)}, W_{11}^{(0)})$ =100%	$P_{22}^{(1)} \cdot p_{11} \approx 0.0325$ min$(W_{22}^{(0)}, W_{11}^{(1)})$ =80%			$P_{23}^{(1)} \cdot p_{11} \approx 0.0003$ min$(W_{23}^{(0)}, W_{11}^{(1)})$ =20%	$P_{24}^{(1)} \cdot p_{11} \approx 0.0001$ min$(W_{24}^{(0)}, W_{11}^{(1)})$ =0%
	p_{12}=**0.05** $W_{12}^{(0)}$=**70%**	$P_{21}^{(1)} \cdot p_{12}$ $p_{12}p_{12} \approx 0.0386$ min$(W_{21}^{(0)}, W_{12}^{(1)})$ =70%	$P_{22}^{(1)} \cdot p_{12} \approx 0.0018$ min$(W_{22}^{(0)}, W_{12}^{(1)})$ =70%			$P_{23}^{(1)} \cdot p_{12} \approx 0$ min$(W_{23}^{(0)}, W_{12}^{(1)})$ =20%	$P_{24}^{(1)} \cdot p_{12} \approx 0$ min$(W_{24}^{(0)}, W_{12}^{(1)})$ =0%
	p_{13}=**0.04** $W_{32}^{(0)}$=**40%**	$P_{21}^{(1)} \cdot p_{13}p_{12} \approx 0.0308$ min$(W_{21}^{(0)}, W_{13}^{(1)})$ =40%	$P_{22}^{(1)} \cdot p_{13} \approx 0.0014$ min$(W_{22}^{(0)}, W_{13}^{(1)})$ =40%			$P_{23}^{(1)} \cdot p_{13} \approx 0$ min$(W_{23}^{(0)}, W_{13}^{(1)})$ =20%	$P_{24}^{(1)} \cdot p_{13} \approx 0$ min$(W_{24}^{(0)}, W_{13}^{(1)})$ =0%
	p_{14}=**0.01** $W_{14}^{(0)}$=**0%**	$P_{21}^{(1)} \cdot p_{14} \approx 0.00771$ min$(W_{21}^{(0)}, W_{14}^{(1)})$ =0%	$P_{22}^{(1)} \cdot p_{14} \approx 0.0004$ min$(W_{22}^{(0)}, W_{14}^{(1)})$ =0%			$P_{23}^{(1)} \cdot p_{14} \approx 0$ min$(W_{23}^{(0)}, W_{14}^{(1)})$ =0%	$P_{24}^{(1)} \cdot p_{14} \approx 0$ min$(W_{24}^{(0)}, W_{14}^{(1)})$ =0%

Thus, for vector (1, 0) one has additional cost equal 1 and $W_{syst}^{(1,0)} \approx 0.9502$ and for vector (0, 1) corresponding values equal to 2 and 0.671, so system configuration (1, 0) is dominating over configuration (0, 1), since higher average performance level delivers with less expenses. It means that all vectors of type (0, *k*) are excluded from further analysis.

The next cells, for which current trials have to be done, are cells (1, 1) and (2, 0) in accordance with self-explanatory Figure 6.

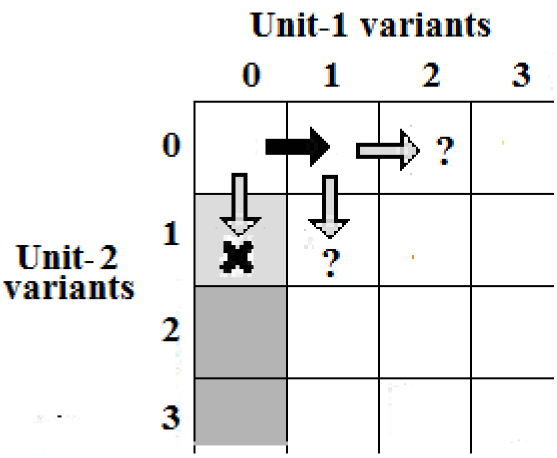


Figure 6. Directions of further analysis of cells

The next cells under investigation are (2, 0) and (1, 1), one can see from Figure 7.

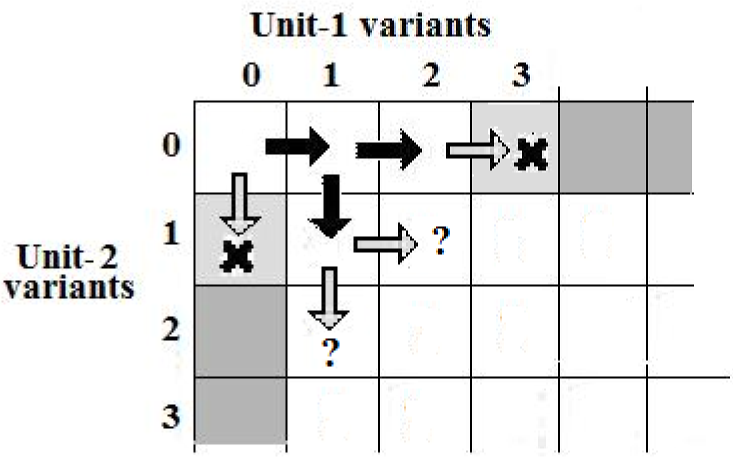


Figure 7. Further development of checking cells

Avoiding simple, however cumbersome calculations, let us present only final results (see Table 5).

Table 7. Costs and levels of performance for different vectors of redundant elements.

		Unit-1: Number of redundant elements						
		0	1	2	3	4	5	6
Unit-2: Number of redundant elements	0	C=0 W= 90.16	C=1 W= 94.26	C=2 W= 94.57				…
	1	C=2 W= 94.68	C=3 W= 99.16	C=4 W= 99.50	C=5 W= 99.53			…
	2	C=4 W= 95.03	C=5 W= 99.54	C=6 W= 99.89	C=7 W= 99.92	?		…
	3		C=7 W= 99.61	C=8 W= 99.95	?			….
	4							….
	…	…	….	…	…	…	…	…
	Legend: light grey color – dominated cells, dark grey color – non-prospective variants.							

Probably, the last table needs some explanations. System without redundant elements initially has average level of performance (W) equals 90.16%. Next phase of calculation is checking neighbor cells to cell (0, 0), i.e. (1, 0) and (0, 1). After adding a redundant element of the 1st type, one gets W= 94.26% and after adding a redundant element of the 2nd type, one gets W= 94.68%. Both cells contain dominating vectors of redundant elements. Next phase of trials are vectors (2, 0), (1, 1) and (0, 2). Vectors (1, 1) gives W=94.57 with total cost of redundant elements C=3. Vector (2, 0) is dominated by vector (0, 1) since possesses lower value with the same expenses for redundant elements. Therefore all vectors of type (3, 0), (4, 0) and so on, are excluded from further trials. Vector (1, 1) is dominating.

Next phase is trial of neighbor cells to the currently existing cells with dominating vectors, These cells are (2, 1), (1, 2) and (0, 3) (Remind that vector (3, 0) is excluded as dominated one.) As one can see from Table 5, vector (2, 1) dominates over vector (0, 2), so all vectors of type (0, 3), (0, 4) and soon are excluded from further trials. Vectors (1, 2) and (2, 1) belong to the dominating sequence of vectors.

Such trials and selection of dominating vectors continued until appearance of first vector with the average level of performance higher than required value of W^o for the direct problem of optimal redundancy, or until total expense of all redundant elements are nor exceed given value C^o for the inverse problem. These comments become absolutely transparent if one take a look on Figure 8.

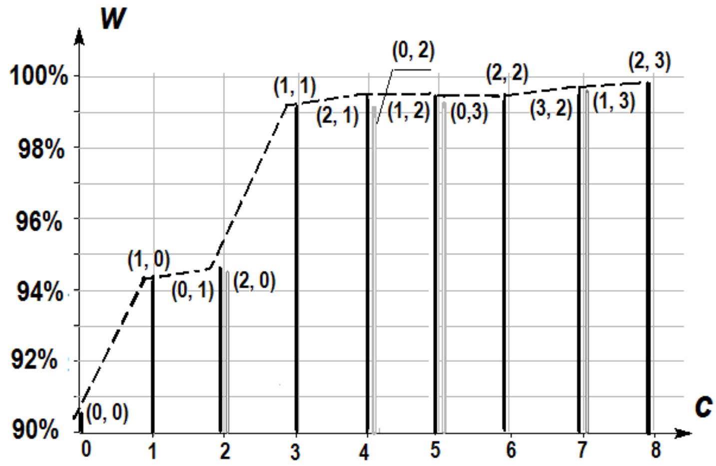


Figure 8. Depiction of the process of compiling the dominating sequence

From Table 5, one can see that optimal solution for requirement that the average level of system performance is not less than W^o =0.999 is delivered by vector (3,2), and the total expenses of redundant elements is 7 cost units. For the total expenses on redundant elements limited by $C^o \leq 4$ cost units, one gets maximum possible solution as vector (1, 2) that characterizes by W=99.54%.

It is interesting what happens with the optimal solution if one changes costs of elements> Let us assume that for the same system cost of a single redundant element of the 1st type is c_1=2 and the cost am element of the 2nd type c_2=1.

Table 8. Costs and levels of performance for different vectors of redundant elements for new elements' costs.

		Unit-1: Number of redundant elements						
		0	1	2	3	4	5	6
Unit-2: Number of redundant elements	0	C=0 W=90.156	C=2 W= 94.26456					…
	1	C=1 W= 94.68072	C=3 W= 99.16318	C=5 W= 99.50462				…
	2	C=2 W= 95.02683	C=4 W= 99.54058	C=6 W= 99.88507	C=8 W= 99.9156			…
	3	C=3 W= 95.08558	C=5 W= 99.60514	C=7 W= 99.9502	?			….
	4		C=6 W= 99.61784	?				….
	…	…	….	…	…	…	…	…
	Legend: light grey color – dominated cells, dark grey color – non-prospective variants.							

In this case optimal solutions found from Table 6 are: For the direct problem vector (2, 3), for which W=99.95% and total expenses on redundant elements are equal to 7 cost units, and for inverse problem the solution is (1, 2), for which W=99.54% and total expenses C=4. For inverse problem, the solutions coincide with each other in both cases.

Solution of optimal redundancy problems for system consisting of several multilevel units seems a bit cumbersome. However, let us note that all enumerative methods like dynamic programming practically unsolvable without computerizing calculations. Numerical example above was solved with the help of a simple programs using Microsoft Excel.

For complex systems consisting of *n* multiple multistate units, one can compile a simple program for a mainframe computer. The algorithm should include the following steps.

i. FIRST STEP

1. Take an *n*-dimensional vector of redundant elements $X^{(0)} = (x_1^{(0)} = 0, x_2^{(0)} = 0, ..., x_n^{(0)} = 0)$.
2. Perform calculations to get initial pair of values $(W_{syst}^{(0)}, C_{syst}^{(0)})$ (see Table.2).
3. Put calculated pair $(W_{syst}^{(0)}, C_{syst}^{(0)})$ into list of dominating solutions,

ii. SECOND STEP

4. Generate vectors $X_i^{(1)}$ such that each of them distinguishes from $X^{(0)}$ by changing number of elements of Unit-*i* on one, i.e. $X_i^{(1)} = (x_1^{(0)} = 0, x_2^{(0)} = 0, ..., x_i^{(0)} = 1, ..., x_n^{(0)} = 0)$.

iii. THIRD STEP

5. For each $X_i^{(1)}, \overline{i = 1, n}$, calculate new values of $P_{ik_i}^{(1)}$, for all k_i where k_i is the number of performance levels of Uniy-*I* (see Tables 3 and 5).
6. Perform corresponding calculations for getting *n* pairs $(W_1^{(1)}, C_1^{(1)}), (W_2^{(1)}, C_2^{(1)}), ..., (W_n^{(1)}, C_n^{(1)}),$ for all vectors (see Tables 4 and 6).
7. Analyze all pairs obtained in previous point to form a set $G^{(1)}$ that includes only dominating vectors $X_i^{(1)}$.
8. Return to the 3rd step, using vectors belonging to set $G^{(1)}$

Stopping rules:

(a) For direct optimization problem, choose such a vector $X_i^{(k)}$ among $G^{(k)}$ that was obtained at the *k*-th step of the optimization process that delivers $\min_{\forall i \in G^{(k)}} C_i^{(1)}$ for all $W_i^{(k)} \geq W^0$.

(b) For inverse optimization problem, choose such a vector $X_i^{(k)}$ among $G^{(k)}$ that was obtained at the *k*-th step of the optimization process that delivers $\max_{\forall i \in G^{(k)}} W_i^{(1)}$ for all $C_i^{(k)} \leq C^0$..

References

1. Kettele, J. D. , Jr. (1962) Least-coast allocation of reliability investment. *Operations Research*, No. 2.
2. Ushakov I. (1986) A universal generating function.Sov J Comput Syst Sci;Vol.24.

3. Ushakov I. (1987) Optimal standby problem and a universal generating function. Sov J Comput Syst Sci; vol.25.
4. Ushakov, I.A. (1988) Solving of Optimal Redundancy Problem by Means of a Generalized Generating Function. Journal of Information Processes and Cybernetics (Germany), Vol.24, No.4.
5. Ushakov, I.A. (2000) The Method of Generating Sequences. *European Journal of Operational Research*, 2000, Vol. 125/2.
6. Gnedenko, B.V., and Ushakov I.A. (1995) Probabilistic Reliability Engineering. John Wiley & Sons.
7. Levitin, G. (2005)T he Universal Generating Function in Reliability Analysis and Optimisation. Springer-Verlag.
8. Lisnianski A., and Levitin G. (2003) *Multi-state system reliability. Assessment, optimization and applications,* World Scientific.

MEAN RESIDUAL LIFE CRITERIA OF FIRST PASSAGE TIME OF SEMI-MARKOV PROCESS BASED ON TOTAL TIME ON TEST TRANSFORMS

V. M. Chacko

•

Department of Statistics, St. Thomas College
Thrissur, Kerala-680001
Email: chackovm@gmail.com

M. Manoharan

•

Department of Statistics
University of Calicut, Kerala-673635, Kerala
Email: mano30@rediffmail.com

ABSTRACT

Mean residual life criteria of first passage time of semi-Markov process is considered. Properties of transition probability functions when using scaled Total Time on Test (TTT) transform for some criteria of mean residual life are discussed. Application to Multistate reliability system is also addressed.

1 INTRODUCTION

First passage times of appropriate stochastic process have often been used to represent times to failure of devices or systems which are subject to shocks and wear, random repair time and random interruptions during their operations. The life distribution properties of these processes have therefore been widely investigated in Multistate system reliability and maintenance literature. The life distributions involved in devices or systems have several interesting properties such as increasing mean residual life (IMRL), decreasing mean residual life (DMRL), etc. The total time on test (TTT) transform is used as a tool for identification of failure distribution model in binary system. Marshell and Shaked (1983), (1986) and Shantikumar (1984) considered processes with new better than used (NBU) first passage times. Belzunce et al. (2002) derived, for the uniformizible, continuous time Markov process, conditions in terms of discrete uniformized Markov chain for the second order NBU and NBU based on laplace transformation classes.

Karasu and Ozekici (1989) studied NBUE and new worse than used in expectation (NWUE) properties of increasing Markov processes and Markov Chains. Lam (1992) considered the NBUE and NWUE properties of Markov renewal processes.

Use of TTT transform for the identification of failure rate models is discussed by Barlow and Campo (1975). Later, Klefsjo (1982) presented some relationship between the TTT transform and other ageing properties (with their duals) of random variable, eg. decreasing mean residual life (DMRL), NBU, NBUE, harmonically new better than used in expectation (HNBUE) and heavy tailedness. Abouammoh and Khalique (1987) further discussed properties of scaled TTT transform for some criteria of the mean residual life such as decreasing mean residual life average (DMRLA), decreasing harmonic mean residual life average (DHMRLA), new better than used harmonic mean residual life average (NBUHMRLA), and new better than used mean residual life average (NBUMRLA).

But when we consider a complex system whose performance process is Markov or semi-Markov, we need the knowledge of DMRLA properties or other relevant ageing properties for applying suitable maintenance and repair/replacement policies. The identification of failure rate model of a system whose performance process is Markov/semi-Markov will be helpful to the engineers and designers for applying suitable maintenance and repair or replacement policies, since identification of failure rate model using TTT describes new methods for analyzing nonnegative observations. Chacko et.al (2010) discussed use of TTT transform in identifying failure rate model of semi-Markov reliability system.

In this paper, we consider a semi-Markov process whose first passage time distribution is DMRL or DMRLA or NBUHMRL or NBUMRL (with their duals). We consider the reliability function based on the transition probability function in the upstates.

This paper is arranged as follows. Section 2 describes various ageing properties of a lifetime random variable. Section 3 recall the existing results for identification of failure rate model of random variables based on TTT. In Section 4, we introduce some sufficient conditions for the MRL criteria of the semi-Markov process based on TTT built from transition probability function. An illustrative example for multistate system is given in section 5. Conclusions are given at the last section.

2. AGEING OF A LIFETIME RANDOM VARIABLE

The concept of ageing is very important in reliability theory. 'No ageing' means the age of a component has no effect on the distribution of residual life time. 'Positive ageing' describes the situation where residual lifetime tends to decrease, in some probabilistic sense, with increasing age of the component. On the other hand, 'Negative ageing' has an opposite effect on the residual lifetime.

Let $R(t) = 1 - F(t)$ be the survival or reliability function of a lifetime random variable. Let X_t be the random variable representing the residual life time of a unit which has attained the age t. Then the respective distribution function and survival function are $F_x(t)$ and $R_x(t)$. Next it is seen that $R_x(t) = R(t+x) /R(t)$. This is the conditional probability that the unit survived up to time t, will not fail before additional x units of time. Further $R_0(t) = R(t)$.

By positive ageing we mean the phenomenon where by an older system has shorter remaining life time in some statistical sense than a newer or younger one.

That is $R_x(t) = R(t+x)/ R(t) < R_0(t)$ or
$R_x(t) = R(t+x)/R(t)$ is decreasing in t.
Similarly, for negative ageing, $R_x(t) = R(t+x)/R(t)$ is increasing in t.
Obliviously, any study of the phenomenon of ageing is to be based on $R_x(t)$ and functions related.

2.1 Failure Rate Function

The conditional failure rate or failure rate at time t is defined as

$$\lambda(t) = \lim_{x\to 0} \frac{F(t+x)-F(t)}{xR(t)},$$

so that, $\lambda(t) = \frac{f(t)}{R(t)}$ when F is absolutely continuous and f(t) is the probability density function of F(t). The failure rate function has been extensively studied in literature since it is very important parameter in reliability theory.

Another important order is mean residual life order. The definition of mean residual life is given below.

2.2 Mean Residual Life

Let T_D be the lite time. The mean residual life is defined as

$$\mu(t) = E(T_D - t|T_D > t) = \frac{\int_t^\infty R(x)dx}{R(t)}$$

if $t < \infty$ and zero otherwise.

The failure rate function $\lambda(t)$ will be continuous and twice differentiable for all t > 0 with the exception of the exponential distribution.

3. TOTAL TIME ON TEST TRANSFORM

Total time on test (TTT) transform is a fundamental tool in reliability investigation.

Let X has distribution F. Given a sample of size n from the non-negative random variables X, let $X_{(1)} \le X_{(2)} \le ... \le X_{(k)} \le ... \le X_{(n)}$ be the sample. TTT to the rth failure from distributions F and is, $T(X_{(r)}) = nX_{(1)} + (n-1)(X_{(2)} - X_{(1)}) + ... + (n-r+1)(X_{(r)} - X_{(r-1)}) = \sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)}$

Define

$H_n^{-1}(r/n) = \frac{1}{n} T(X_{(r)})$ and $H_n^{-1}(r/n) = \int_0^{F_n^{-1}(r/n)} (1 - F_n(u))du$,

where $F_n(u) = \begin{cases} 0 & u < X_{(i)} \\ i/n & X_{(i)} \le u < X_{(i+1)} \\ 1 & X_{(n)} > u \end{cases}$, $F_n^{-1}(x) = \inf\{x : F_n(x) \ge u\}$

The fact that $F_n(u) \to F(x)$ a.s. implies, by Glivenko Cantelli Theorem, $\lim_{n\to\infty, r/n \to t} \int_0^{F_0^{-1}(r/n)} (1 - F_n(u))du = \int_0^{F^{-1}(t)} (1 - F(u))du \quad t \in [0,1]$.

We define TTT transform of F as $H_F^{-1}(t) = \int_0^{F^{-1}(t)} (1 - F(u))du \ t \in [0,1]$.

3.1. Model Identification

TTT is a very important index in reliability for the model identification of lifetime data. TTT test plots are useful for analyzing non-negative data. Using these plots incomplete data can be analyzed and there is a theoretical basis for such an analysis. We can define ageing properties in terms of TTT transforms. TTT transforms permits us to classify distributions according to their failure rate. Total time on test plots also permits the comparison of distribution functions with respect to their failure rate and they can be used to find a model for the data under study. Thus in many aspects TTT is very important part in the study of reliability. But the existing results are limited to the case of random variables. We extend the study to the case of semi-Markov process in the next section, for the first passage time random variable.

Let $G(x) = 1 - \exp(-x/\theta), \quad x, \theta \ge 0$ be the exponential distribution with mean μ. Then

$H_G^{-1}(t) = \int_0^{G^{-1}(t)} e^{-x/\theta} dx = \int_0^{G^{-1}(t)} \theta dG(x) = \theta t$ and scaled TTT,

$$\phi(t) = \frac{H_G^{-1}(t)}{H_G^{-1}(1)} = t, \quad t \in [0,1]. \tag{3.1}$$

The scaled TTT of the Exponential distribution is a 45^0 line on $[0,1]$. The normalized total time on test is the boundary between the corresponding transforms of IFR and DFR distributions. TTT that permits to classify distributions according to their failure rate is that its slope evaluated at *t* = *F*(*x*) is the reciprocal of the failure rate at *X*.

$$\frac{d}{dt}H_F^{-1}(t)|_{t=F(x)} = \frac{(1-t)}{f[F^{-1}(t)]}|_{t=F(x)} = \frac{1-F(x)}{f(x)} = \frac{1}{\lambda(x)}, \tag{3.2}$$

where λ is the failure rate of *F*.

Now we consider the following results in Barlow and Campo (1975).

Proposition 3.1 *F IFR(DFR))* $\Rightarrow \frac{H_F^{-1}(t)}{H_F^{-1}(1)}$ *is concave or (convex) in* $t \in [0,1]$.

Proposition 3.2 *F IFRA(DFRA) implies* $\frac{H_F^{-1}(t)}{tH_F^{-1}(1)} \uparrow (\downarrow)$ *in* $t \in [0,1]$.

Proposition 3.3 *F NBU(NWU))* $\Rightarrow$ *slope of* $\frac{H_F^{-1}(t)}{H_F^{-1}(1)}$ *is larger(smaller) at the origin than at any other t,* $0 < t < 1$.

For *F* NWU reverse the direction of inequalities. Now we consider a characterization for NBUE (NWUE) and DMRL (IMRL), see Klefsjo (1982).

Theorem 3.1 *A life distribution F is NBUE(NWUE) if and only if* $\phi(t) \le (\ge) t$ *for* $0 \le t \le 1$.
Theorem 3.2 *A life distribution F is DMRL(IMRL) if and only if* $Q(t) = (1-\phi(t))/(1-t)$ *is decreasing (increasing) for* $0 \le t \le 1$.

The following are some important results in model identification of a univariate lifetime random variable.

Theorem 3.3 (Abouammoh and Khalique (1987)) Let F and $\phi_F(t)$ be as in above theorem, then we have the following

1. F is DMRL (IMRL) if and only if $1-\phi_F(t)-(1-t)\phi(t) \le (\ge) 0, 0 < t \le 1$
2. F is NBUMRL(NWUMRL) if and only if $\phi_F(t) \ge (\le) t, 0 \le t \le 1$.
3. F is DMRLA (IMRLA) if and only if $1/t\int_0^t (1-\phi_F(x))/(1-x)dx$
is decreasing (increasing) for 0 < t < 1.
4. F is NBAMRL (NWAMRL) if and only if
$\int_0^t (1-\phi_F(x))/(1-x)dx \le (\ge) t$ for $0 \le t < 1$..
5. F is DHMRLA (IHMRLA) if and only if $1/t\int_0^t (1-x)/(1-\phi_F(x))dx$
is increasing (decreasing) in t for 0 < t < 1.
6. F is NBUHMRL(NWUHMRL) if and only if
$\int_0^t (1-x)/(1-\phi_F(x))dx \ge (\le) t$ for $0 \le t < 1$.

But the existing results are limited to the case of random variables. We extend the study to the case of semi-Markov process in the next section, for the first passage time random variable.

4. MEAN RESIDUAL LIFE CRITERIA OF A SEMI-MARKOV SYSTEM

We are concerned with a multistate system (MSS) having $M + 1$ states 0,1,…,M where '0' is the best state and '*M*' is the worst state, see Barlow and Wu (1978) for details of MSSs. At time zero the system begins at its best state and as time passes the system begins to deteriorate. It is assumed that the time spent by the system in each state is random with arbitrary sojourn time distribution. The system stays in some acceptable states for some time and then it moves to unacceptable (down) state. The first time at which the MSS enters the down state after spending a random amount of time in acceptable states is termed as the first passage time (failure time) to the down state of the MSS. We study the aging properties of the first passage time distribution of the MSS modeled by the semi-Markov process $\{Y_t, t \geq 0\}$. In the MSS with states *{0,1,…,k,k+1,…,M}* where *{0,1,…,k}* is the acceptable states, the sojourn time between state '*i*' to state '*j*' is assumed to be distributed with arbitrary distribution F_{ij}.

4.1 First Passage time and Reliability Function

Let *E = {0,1,…,M}* be a set representing the state of the MSS and probability space with probability function *P*, on which we define a bivariate time homogeneous Markov chain $(X,T) = \{X_n, T_n, n \in \{0,1,2,...\}\}$, X_n takes values of *E* and T_n on the half real line $R^+ = [0,\infty)$, with $0 \leq T_1 \leq T_2 \leq ... \leq T_n \leq ...$ Put $U_n = T_n - T_{n-1}$ for all $n \geq 1$. This Markov process is called a Markov renewal process (MRP) with transition function, the semi-Markov kernel, Q = [Q_{ij}], where $Q_{ij}(t) = P[X_{n+1} = j, U_n \leq t \mid X_n = i], i, j \in E, t \geq 0$ and $Q_{ii}(t) = 0, i \in E, t \geq 0$.

Now we consider the semi-Markov process (SMP), as defined in Pyke (1961). It is the generalization of Markov process with countable state space. SMP is a stochastic process which moves from one state to another of a countable number of states with successive states visiting form a Markov chain, and that the process stays in a given state a random length of time, the distribution of which may depend on this state as well as on the one to be visited in the next. Define $Z_t = X_{N_t}, N_t = \sup\{n, T_n = U_1 + U_2 + ... + U_n \leq t\}$, it is the semi-Markov process associated with the MRP defined above. In terms of *Z*, the times $T_1, T_2, ...$ are successive times of transitions for *Z*, and $X_0, X_1, ...$ are successive states visited.

If Q has the form $Q_{ij}(t) = P[X_{n+1} = j \mid X_n = i][1 - e^{-\lambda(i)t}], i, j \in E, t \geq 0,$ for some function λ(*i*), $j \in E$ then the process Z_t is a Markov process. That is, in a Markov process, the distributions of the sojourn times are all exponential independent of the next state. The word *semi*-Markov comes from the somewhat limited Markov property which *Z* enjoys, namely, that the future of *Z* is independent of its past given the present state provided the "present" is the time of jump. Let I_{ij} =indicator function of $\{i = j\}$. Define the transition probability that system occupied state $j \in E$ at time *t* > 0, given that it is started at state *i* at time zero, as, $i, j \in E, t \geq 0$

$$p_{ij}(t) = P[Z_t = j \mid Z_0 = i] = P[X_{N_t} = j \mid X_0 = i] = h_i(t) I_{ij} + Q * P(t)(i,j),$$

where $h_i(t) = 1 - \sum_k Q_{ik}(t)$, $P(t) = [p_{ij}(t)]$ and $Q * P(t)(i,j) = \sum_{k \in E} \int_0^t Q_{ik}(dx) p_{kj}(t-x)$

To obtain the reliability function of the semi-Markov system described above, we must define a new process, *Y* with state space $U \cup \nabla$, where *U* denotes set of all up states *{*0*;* 1*; :::; k}* and ∇ is the absorbing state in which all the states *{k +* 1*, …;M}* of the system is united. Let T_D denote the time of first entry to the down states of *Z* process.

That is, $Y_t = Z_t(\omega)$ if $t < T_D(\omega)$ and $Y_t = \nabla$ if $t \geq T_D(\omega)$.

Let $1=(1,1,...,1)^1$, a unit row vector with appropriate dimension. The process Y_t is a semi-Markov process with semi-Markov kernel

$$\begin{bmatrix} \overbrace{Q_{11}(t)}^{Up} & \overbrace{Q_{12}(t)}^{Down} \\ 0 & 0 \end{bmatrix}$$

We denote $\alpha = (\overbrace{\alpha(0),...,\alpha(k)}^{Up,\alpha_1}, \overbrace{\alpha(k+1),...,\alpha(M)}^{Down,\alpha_2})$ where $\alpha(i) = P(Y_0 = i)$.

The reliability function is

$$\begin{aligned} R(t) &= P[\forall u \in [0,t], Z_u \in U] = P[Y_t \in U] = \sum_{j\in U} P[Y_t = j] \\ &= \sum_{i\in U}\sum_{j\in U} P[Y_t = j, Y_0 = i] \\ &= \sum_{i\in U}\sum_{j\in U} p_{ij}(t)\alpha(i). \end{aligned}$$

4.2. Model Identification

In order to identify the failure rate behavior of a semi-Markov system based on the transition probability function, we define the TTT based on transition probability function in up states as follows. Let F be the first passage time distribution of a semi-Markov system, define

$$H_{p_{ij}}^{-1}(t) = \int_0^{F^{-1}(t)} p_{ij}(u)du, \forall i,j \in U, t \in [0,1]$$

where

$$\begin{aligned} F^{-1}(t) &= \inf\{(1-R(t)) \geq t\} \\ &= \inf\left\{x : \left(1 - \sum_{i\in U}\sum_{j\in U} p_{ij}(x)\alpha(i)\right) \geq t\right\} \\ &= \inf\left\{x : \left(\sum_{i\in U}\sum_{j\in U} p_{ij}(x)\alpha(i)\right) \leq 1-t\right\} \end{aligned}$$

But

$$\begin{aligned} H_F^{-1}(t) &= \int_0^{F^{-1}(t)} \sum_{i\in U}\sum_{j\in U} p_{ij}(u)du, \forall i,j \in U, t \in [0,1] \\ &= \sum_{i\in U}\sum_{j\in U} \alpha(i) \int_0^{F^{-1}(t)} p_{ij}(x)dx \\ &= \sum_{i\in U}\sum_{j\in U} \alpha(i) H_{p_{ij}}^{-1}(t) \end{aligned}$$

Then

$$H_F^{-1}(1) = \sum_{i\in U}\sum_{j\in U} \alpha(i) \int_0^{F^{-1}(1)} p_{ij}(x)dx = \sum_{i\in U}\sum_{j\in U} \alpha(i) H_{p_{ij}}^{-1}(1)$$

$$\frac{H_F^{-1}(t)}{H_F^{-1}(1)} = \frac{\sum_{i\in U}\sum_{j\in U} \alpha(i) \int_0^{F^{-1}(t)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U} \alpha(i) \int_0^{F^{-1}(1)} p_{ij}(x)dx}, \quad t \in [0,1].$$

Chacko and Manoharan (2009) proved the following if $p_{ij}(x)$ is monotonic increasing or decreasing.

Proposition 4.1 *The first passage time distribution of a semi-Markov system F is IFR if* $\frac{H_F^{-1}(t)}{H_F^{-1}(1)} \le t$ *and concave in* $t \in [0,1], \forall i, j \in U$.

Proposition 4.2 *The first passage time distribution of a semi-Markov system F is DFR if* $\frac{H_F^{-1}(t)}{H_F^{-1}(1)} \ge t$ *and* $\frac{H_{p_{ij}}^{-1}(t)}{H_{p_{ij}}^{-1}(1)}$ *convex in* $t \in [0,1], \forall i, j \in U$.

Remark 4.1 *The constant failure rate model arises when it is both IFR and DFR. Therefore we must have* $\frac{H_{p_{ij}}^{-1}(t)}{H_{p_{ij}}^{-1}(1)} = t, \quad \forall i, j \in U$.

Proposition 4.3 *The first passage time distribution of a semi-Markov system F IFRA implies* $\sum_{i\in U}\sum_{j\in U}\frac{d}{dt}H_{p_{ij}}^{-1}(t) \ge 1, \quad t \in [0,1]$.

Proposition 4.4 *The first passage time distribution of a semi-Markov system F DFRA implies* $\sum_{i\in U}\sum_{j\in U}\frac{d}{dt}H_{p_{ij}}^{-1}(t) \le 1, \quad t \in [0,1]$.

Proposition 4.5 *F NBU(NWU)) implies*

$$\left(\sum_{i\in U}\sum_{j\in U}\frac{d}{dt}H_{p_{ij}}^{-1}(t)\big|_{t=0} - \sum_{i\in U}\sum_{j\in U}\frac{d}{dt}H_{p_{ij}}^{-1}(t)\right) \ge (\le) 0, \quad t \in [0,1], \quad i, j \in U.$$

Proposition 4.6 *The first passage time distribution of a semi-Markov system F is NBUE* $\frac{H_{p_{ij}}^{-1}(t)}{H_{p_{ij}}^{-1}(1)} \le t, \quad \forall i, j \in U, \quad t \in [0,1]$.

Proposition 4.7 *The first passage time distribution of a semi-Markov system F is NWUE if* $\frac{H_{p_{ij}}^{-1}(t)}{H_{p_{ij}}^{-1}(1)} \ge t, \quad \forall i, j \in U, \quad t \in [0,1]$.

Proposition 4.8 *The first passage time distribution of a semi-Markov system F is DMRL(IMRL) if*

$$\frac{dH_{p_{ij}}^{-1}(t)}{dt} \ge (\le) 1, \quad \forall i, j \in U, \quad t \in [0,1].$$

We prove the following.

Theorem 4.1. Let F and $\phi_F(t)$ be as in above theorem, then we have the following

1. F is DMRL (IMRL) if

$$\left(\int_0^{F^{-1}(1)} p_{ij}(x)dx - \int_0^{F^{-1}(t)} p_{ij}(x)dx\right)/(1-t) \le (\ge) \frac{d\int_0^{F^{-1}(t)} p_{ij}(x)dx}{dt}, \quad \forall i, j \in U, \quad t \in [0,1].$$

Proof:

$$\left(\int_0^{F^{-1}(1)} p_{ij}(x)dx - \int_0^{F^{-1}(t)} p_{ij}(x)dx\right)/(1-t) \le (\ge) \frac{d\int_0^{F^{-1}(t)} p_{ij}(x)dx}{dt}, \quad \forall i, j \in U, \quad t \in [0,1].$$

It implies, for

$$\left(\int_0^{F^{-1}(1)} p_{ij}(x)dx - \int_0^{F^{-1}(t)} p_{ij}(x)dx\right) - (1-t)\frac{d\int_0^{F^{-1}(t)} p_{ij}(x)dx}{dt} \le (\ge) 0, \quad \forall i,j \in U, \quad t \in [0,1].$$

Taking over U and dividing both sides by $\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx$, we get

$$\left(1 - \frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}\right) - (1-t)\frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)d\int_0^{F^{-1}(t)} p_{ij}(x)dx}{dt\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx} \le (\ge) 0, \quad \forall i,j \in U, \quad t \in [0,1].$$

By theorem (3.3), we proved the result.

2. F is NBUMRL (NWUMRL) if $\left(\frac{\int_0^{F^{-1}(t)} p_{ij}(x)dx}{\int_0^{F^{-1}(1)} p_{ij}(x)dx}\right) \ge (\le) t, i, j \in U, \quad t \in [0,1]$.

Proof: Suppose $\left(\frac{\int_0^{F^{-1}(t)} p_{ij}(x)dx}{\int_0^{F^{-1}(1)} p_{ij}(x)dx}\right) \ge (\le) t, i, j \in U, \quad t \in [0,1]$

$$\Rightarrow \left(\int_0^{F^{-1}(1)} p_{ij}(x)dx - t\int_0^{F^{-1}(t)} p_{ij}(x)dx\right) \ge (\le) 0 \quad \forall i,j \in U, \quad t \in [0,1]$$

$$\Rightarrow \left(\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx - t\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx\right) \ge (\le) 0, \quad t \in [0,1]$$

$$\Rightarrow \left(\frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx}\right) \ge (\le) t, \quad t \in [0,1]$$

By theorem (3.3) we proved the result.

3. F is DMRL (IMRL) if $\left(\frac{\int_0^{F^{-1}(1)} p_{ij}(x)dx}{t}\right)$, $t \in [0,1]$ is decreasing (increasing) in t, $i, j \in U, \quad t \in [0,1]$.

Proof: Suppose if $\left(\frac{\int_0^{F^{-1}(1)} p_{ij}(x)dx}{t}\right)$, $t \in [0,1]$ is decreasing (increasing) in t, $i, j \in U, \quad t \in [0,1]$.

Then $\left(\frac{\int_0^{F^{-1}(t)} p_{ij}(x)dx}{\int_0^{F^{-1}(1)} p_{ij}(x)dx}\right) \ge (\le) t, i, j \in U, \quad t \in [0,1]$

$$\Rightarrow \left(\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx - t\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx\right) \ge (\le) 0, \quad t \in [0,1]$$

$$\Rightarrow \left(1 - \frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx} \right) \le (\ge) 1-t, \quad t\in[0,1]$$

$$\Rightarrow \tfrac{1}{1-t}\left(1 - \frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx} \right) \le (\ge) 1, \quad t\in[0,1].$$

Rate of increase of $\Rightarrow \tfrac{1}{1-t}\left(1 - \frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx} \right), \quad t\in[0,1]$ is smaller (larger) than that of t.

By theorem (3.3) we proved the result.

4. F is NBAMRL (NWAMRL) if

$$\left(1/t\int_0^{F^{-1}(1)} p_{ij}(x)dx \ge (\le) \int_0^{F^{-1}(t)} p_{ij}(x)dx \right) \quad \forall i,j\in U, \quad t\in[0,1]$$

Proof: Suppose, $\left(1/t\int_0^{F^{-1}(1)} p_{ij}(x)dx \ge (\le) \int_0^{F^{-1}(t)} p_{ij}(x)dx \right) \quad \forall i,j\in U, \quad t\in[0,1]$

$$\Rightarrow \left(\frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx} \right) \ge (\le) t, \quad t\in[0,1]$$

$$\Rightarrow \tfrac{1}{1-t}\left(1 - \frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx} \right) \le (\ge) 1, \quad t\in[0,1]$$

By theorem (3.3) we proved result.

5. F is DHMRLA (IHMRLA) if $\left(\frac{\int_0^{F^{-1}(1)} p_{ij}(x)dx}{t} \right), \quad t\in[0,1]$, is increasing (decreasing) in t, $i,j\in U, \quad t\in[0,1]$.

Proof: Given, $\left(1/t\int_0^{F^{-1}(1)} p_{ij}(x)dx \ge (\le) \int_0^{F^{-1}(t)} p_{ij}(x)dx \right) \quad \forall i,j\in U, \quad t\in[0,1]$

$$\Rightarrow \tfrac{1}{1-t}\left(1 - \frac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx} \right) \le (\ge) 1, \quad t\in[0,1]$$

$$\Rightarrow 1 \le (\ge) \frac{1-t}{\left(1-\dfrac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(1)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx}\right)}, \quad t\in[0,1]$$

The rate of increase of $\int_0^t \frac{1-u}{\left(1-\dfrac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(u)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx}\right)}, \quad t\in[0,1]$ is larger than that of t.

$$\Rightarrow 1/t\int_0^t \frac{1-u}{\left(1-\dfrac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(u)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx}\right)}, \quad t\in[0,1] \text{ is increasing in t.}$$

6. F is NBUHMRL (NWUHMRL) if

$$\left(1/t\int_0^{F^{-1}(1)} p_{ij}(x)dx \ge (\le)\int_0^{F^{-1}(t)} p_{ij}(x)dx\right) \quad \forall i,j\in U, \quad t\in[0,1].$$

Proof: $\left(1/t\int_0^{F^{-1}(1)} p_{ij}(x)dx \ge (\le)\int_0^{F^{-1}(t)} p_{ij}(x)dx\right) \quad \forall i,j\in U, \quad t\in[0,1]$

$$\Rightarrow \int_0^t \frac{1-u}{\left(1-\dfrac{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(u)} p_{ij}(x)dx}{\sum_{i\in U}\sum_{j\in U}\alpha(i)\int_0^{F^{-1}(t)} p_{ij}(x)dx}\right)} \ge (\le)\int_0^t dt, \quad t\in[0,1].$$

Hence by theorem (3.3), we proved the result.

5 APPLICATION AND ILLUSTRATIVE EXAMPLE

We are concerned with a multistate system (MSS) having M+1 states 0, 1, ...,M where ’0’ is the best state and ’M’ is the worst state, see Chacko and Manoharan (2009) for details of MSSs. At time zero the system begins at its best state and as time passes the system begins to deteriorate. It is assumed that the time spent by the system in each state is random with arbitrary sojourn time distribution. The system stays in some acceptable states for some time and then it moves to unacceptable (down) state. The first time at which the MSS enters the down state after spending a random amount of time in acceptable states is termed as the first passage time (failure time) to the down state of the MSS. Major application of the above results is in maintenance and repair of complex systems such as age and block replacement policies. A variety of applications in maintenance and replacement policies of a binary system can be seen in Barlow and Proschan (1996).

Example 5.1 *Consider a Markov process in continuous time and discrete state space {1,2,...,M} given in Doob (1953), p.241.* The system starts in state '1' at time zero and as it enters 'M', it remains there. Consider the intensity matrix, $Q=[q_{ij}]$ with entries

$q_{ij} = 0, \quad i \in \{1,2,...,M-1\}, j \neq i+1, q_{ii+1} = q, \quad q_{iM} = 0.$

The Kolmogorov's system of differential equation becomes,
for $p_{ij}(t-u) = P[Y_t = j \mid Y_u = i], 0 \le u < t$ and we take $u = 0$,

$p^1_{ik}(t) = -qp_{ik}(t) + qp_{i+1k}(t), i < M, \quad p_{Mk}(t) = 0$

with initial conditions, $p_{ik}(0) = \delta_{ik}$, the indicator of *{i=k}*. Then,

$p_{Mk}(t) = 0, \quad k \neq M, p_{MM}(t) = 1$

and it is easily verified that the solution is

$$p_{ik}(t) = \begin{cases} 0, \quad k < i \\ \dfrac{(qt)^{k-i} e^{-qt}}{\Gamma(k)}, \quad i \le k < M \\ e^{-qt}[e^{qt} - 1 - qt - ... - \dfrac{(qt)^{M-i-1}}{\Gamma(M-i)}], \quad k = M \end{cases}$$

Here the process is of monotone paths. Now consider $\forall i, j \in \{0,1,...,M-1\}$

$$H^{-1}_{p_{ij}}(1) = \int_0^\infty \frac{(qt)^{k-i} e^{-qt}}{\Gamma(k)} dt = \frac{q^{k-i}}{\Gamma(k)} \int_0^\infty t^{k-i} e^{-qt} dt = \frac{\Gamma(k-i+1)}{\Gamma(k)q}$$

Therefore

$$\frac{H^{-1}_{p_{ij}}(t)}{H^{-1}_{p_{ij}}(1)} = \frac{\Gamma(k)q}{\Gamma(k-i+1)} \int_0^{F^{-1}(t)} \frac{(qu)^{k-i} e^{-qu}}{\Gamma(k)} du$$

$$= \frac{q^{k-i+1}}{\Gamma(k-i+1)} \int_0^{F^{-1}(t)} u^{k-i} e^{-qu} du$$

This is an increasing function of *t* and bounded by 1. Therefore $\dfrac{dH^{-1}_{p_{ij}}(t)/dt}{H^{-1}_{p_{ij}}(1)}$ increases *in t* $\forall i, j \in U$.

6. CONCLUSIONS

The identification of the failure rate model of first passage time distribution of a semi-Markov process is discussed. The results are applicable to systems like power generation system whose performance is measured in terms of productivity or capacity and having more than two levels of performance. Preventive or corrective maintenance can be applied to the MSS if we have the knowledge regarding its failure behavior, since type of the failure rate is an important parameter for the maintenance and replacement policies.

REFERENCES

1. Barlow, R. E. and Proschan, F. (1975) Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, New York.
2. R. E. Barlow and Campo, R. A. (1975) Total Time On Test Processes And Application To Failure Data Analysis , reserch report No. University of California.
3. Belzunce, F., Ortega, E. M. and Ruizon, J. M. (2002) Ageing Properties of First-Passage Times of Increasing Markov Processes, Adv. Appl. Prob., 34, 241-259.

4. Klefsjo, B. (1981) On Ageing Properties And Total Time on Test Transforms, *Scand. Jr. Statistics*., 9, 37-41
5. Chacko, V. M. and Manoharan, M. (2009) Ageing properties of first passage time distribution of multistate reliability system, ProbStat Forum: e-journal, 2, 22-30.
6. Chacko, V. M., Praveena, P. C. and M. Manoharan (2010) Ageing properties of a semi-Markov system and Total Time on test transforms, In. Trans. Mathematical Science and Computers, 14 Vol 3, 257-275. (Presented at ICM2010, Hyderbad, India).
7. Doob, J. L. (1953) Stochastic Process, John Wiley and sons, New York.
8. Karasu, I. and Ozekici, S. (1989) NBUE and NWUE properties of increasing Markov processes,J. Appl. Prob., 27, 827-834.
9. Lam, C. Y. T. (1992) New Better than used in expectation processes, J. Appl. Prob., 29, 1116-1128.
10. Marshell and Shaked (1983) New better than used processes, Adv. Appl. Prob., 15, 601-615.
11. Marshell and Shaked (1986) NBU processes with general state space, Math. Oper. Res., 11, 95-105.
12. Pyke, R. (1961) Markov renewal processes: definitions and preliminary properties, Ann. Math.Stat., 32, 1231-1242.
13. Shantikumar (1984) Processes with NBU first passage times, Adv. Appl. Prob., 16, 667-686.

COMPLEX DELTA - FUNCTION

Smagin V.A.
●
e-mail: va_smagin@mail.ru

ABSTRACT

Smagin V.A. the brief review of a history of introduction of delta - function on a complex plane. The proof of the mathematical form of complex delta - function is given. The example of application of complex delta - function for a presence of stationary value alternation casual process of accumulation with the information income and the charge is given.

Introduction

From a history of mathematics, V.M.Kalinin, professor SPBSU in his book " my formulas " [1] 1996, referring to N.A.Lebedev, the professor of academy him A.F.Mozhajskiy, the outstanding expert in the theory of a complex variable, writes the following (so-called citations and formulas of the author of the book further are resulted):

« Delta - function $\delta(x)$ is invented not by mathematics. Her has opened for the first time, probably, Oliver Heaviside, but wide and almost at physics and engineering she has received a general recognition after her has entered into practice Pol Dirak, without the reference to compatriot Heaviside. Mathematics have recognized its right on existence only as functional, making functions its value in zero. The big and deep theory of the generalized functions has appeared. At a statement of the classical mathematical analysis delta - function (equal to zero everywhere except for a point the area under the diagram of this function is considered zero where she is equal $+\infty$, and to equal unit), - is usually ignored. The antipathy which « functions » had to similar classical mathematics is quite clear. For example, my generation mathematics and physics has grown on V.I.Smirnov's textbooks and G.M.Fihtengol'ts where delta - function is not mentioned at all ».

Now delta - function in the modern analysis is used widely. She is defined on a material axis and connected to concepts of function of distribution of probabilities and characteristic function. By the way, function of distribution to present [2]:

$$F(x) = F_1(x) + F_2(x) + F_3(x), \tag{1}$$

where $F_1(x) -$ the function of jumps growing only in points of accounting set of points of breaks of function $F(x), F_2(x) -$ continuous function, possible which points of growth form set of zero measure Lebeg, and its increment on this set is equal to an increment $F(x)$ on it. $F_2(x)$ is not absolutely continuous, the derivative of her is equal all points or to zero, or infinity. $F_3(x) -$ absolutely continuous function, with usual derivative - density of probability $f(x)$. For absolutely continuous functions characteristic function is used:

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx. \tag{2}$$

Let step Heaviside, a constant a with probability 1 is given. Its characteristic function

$$\varphi(t) = e^{iat}. \tag{3}$$

To it corresponds derivative - delta - function $\delta(t-a)$.

Whether there is an analogue of function Heaviside and delta - function on a complex plane? The given question can seem inappropriate, but the answer to him appears positive [1].

« As all of us know, frequently, while remain in material area, the true mathematical nature of any phenomenon remains latent. As a classical example similarity of properties of trigonometrical and hyperbolic functions serves. On a complex plane all at once becomes simple and clear: actually it is the same functions which have been written down in a little bit changed system of coordinates».

« The similar phenomenon is found out and with delta - function. If her to define on a complex plane for analytical function $f(x)$ by the formula

$$f(z)=\int_S f(\varsigma)\delta(\varsigma-z)d\varsigma,\ \forall z\in D. \tag{4}$$

For a contour S laying in the field of D analyticity of function $f(z)$ and bypassing a point z, that, obviously $\delta(z)=\dfrac{1}{2\pi iz}$, and any mystery or a paradoxically in it is not present. The unnatural kind she gets at attempt to drive her from a complex plane on a material axis, thus there are all well-known approximations of delta – function as narrow language or the extended bell. In material area it is possible to enter delta – function axiomatic. Advantage of such approach that logically is not required to define this concept, having attributed it to initial, initial, indefinable. Its properties, for example, $\delta(x)=\dfrac{d}{dx}\varepsilon(x)$ for individual step Heaviside are set only.

Its decomposition in a number and in integral Fourier will give

$$\delta(x)=\frac{1}{\pi}\left(\frac{1}{2}+\sum_{k=1}^{\infty}\cos kx\right),\ \delta(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{tx}dt. \tag{5}$$

Divergence of a lines and integral does not interfere with their use as divergence асимптотических numbers does not prevent their applications. Introduction thus deltas - functions in the mathematical analysis transforms a class piece smooth functions into a class of differentiable functions, and for them the formula of integration in parts takes place».

«The first consequence from such approach for me was rather unexpected. All of us have got used to that numbers Fourier explosive piece smooth functions badly converge, and they cannot be differentiated term by term. It appeared incorrect. Numbers Fourier at term by term differentiation automatically allocate deltas - functions as their numbers in points of break, and after their reduction true formulas turn out. It is the easiest to illustrate it the simple example frequently included in textbooks for an interval $(-\pi,\pi)$:

$$\frac{\pi\cos ax}{2\sin a\pi}=\frac{1}{2a}+\sum_{n=1}^{\infty}(-1)^n\frac{n\cos nx}{a^2-n^2},$$

$$\frac{\pi\sin ax}{2\sin a\pi}=\frac{1}{2a}+\sum_{n=1}^{\infty}(-1)^n\frac{n\sin nx}{a^2-n^2}. \tag{6}$$

The second formula turns out from the first differentiation on x. But also the first turns out differentiation from the second if to take into account, that from jumps in points $(2k+1)$ appears at the left composed

$$-\pi\sum_{k=-\infty}^{\infty}\delta(x-(2k+1)\pi)=-\left(\frac{1}{2}+\sum_{n=1}^{\infty}(-1)^n\cos nx\right),\ -\infty<x<\infty. \tag{7}$$

It is necessary to reduce superfluous composed.

Accurate application of delta - function also allows to remove (take off) vain, erected on Heaviside: it is considered, that a source of many mistakes at formal application of operational method Heaviside for the decision of the linear differential equations with constant factors is that fact, that operators of differentiation p and p-1 do not switch integration. The authorship of this statement takes up G.Dzheffris. It repeats and in R.Kurant's well-known textbook and D.Gil'bert " Methods of mathematical physics " from which I shall bring the citation: " the Basis of a method is made with introduction of operators of differentiation and integration, p and p^{-1}, as mutual - return operations. We shall enter into consideration for functions of time t at $t > 0$ operators of integration p^{-1} and differentiation p by equality

$$p^{-1} f(t) = g(t) = \int_0^t f(\tau) d\tau, \quad pg(t) = f(t) = \frac{dg}{dt} \quad . \tag{8}$$

For construction of calculation with the rules appropriate to rules of algebra, importance represents that fact, that operators p and p^{-1} are mutual - are return or, symbolically, that

$$pp^{-1} = p^{-1} p = 1 \, . \tag{9}$$

To provide this parity, we should enter the following restriction: the operator p can be applied only to such functions $g(t)$, for which $g(0) = 0$. Otherwise we would have:

$$p^{-1} pg = \int_0^t \frac{dg(\tau)}{d\tau} = g(t) - g(0), \quad pp^{-1} g = \frac{d}{dt} \int_0^t g(\tau) d\tau = g(t) \, , \tag{10}$$

hence

$$p^{-1} pg \neq pp^{-1} g'' \, . \tag{11}$$

Actually Heaviside all has made that operators p and p^{-1} switched: it has invented delta - function which I write down in the designations accepted now, with properties

$$\frac{d}{dt} \varepsilon(t) = \delta(t) \quad , \quad \sum_{-\infty}^{\infty} g(\tau) \delta(\tau) d\tau = g(0) \, ,$$

Also has defined a class of so-called originals, considering, that for all elements $g(t)$ of this class $g(0) = 0$ at $t < 0$. Thus for switching operators p also p^{-1} are not present necessity to demand, that $g(+0) = 0$: switching takes place and without this condition. Only it is necessary to take into account, that because of gallop in zero of function $g(t)$ at differentiation appears delta - function:

$$pg(t) = g'(t) + g(+0)\delta(t) \, ,$$

где

$$g'(t) = \begin{cases} \frac{dg}{dt}, t > 0, \\ 0, t < 0, \end{cases}$$

therefore

$$p^{-1} pg = \int_0^t [g`(\tau) + g(+0)\delta(\tau)] d\tau = g(t) = pp^{-1} g \text{ ».} \tag{12}$$

However, use of the complex delta - function offered by V.M.Kalinin in the applied analysis inconveniently enough.

Function Heavisaide and delta - function on a complex plane. It is proved, that delta - function is equal to a derivative from individual function. Thus, function Heavisaide is integrated for delta - function Dirak:

$$U(t) = \int_{-\infty}^{t} \delta(z) dz . \quad (13)$$

In the book [2] constant with probability unit equal a, is submitted as characteristic function $\varphi(t) = e^{iat}$. Hence, the appropriate delta - function will be equal $\delta(t-a)$.

Characteristic function of a normal amount with average m and a deviation σ is equal:

$$\varphi_1(t) = e^{-\frac{\sigma_1^2 t^2}{2} + im_1 t} . \quad (14)$$

Function (14) is submitted depending on a material variable t that is marked at it by an index 1. We shall present similar function depending on an imaginary variable. Sizes of a population mean will become im_2, a deviation – $i\sigma_2$, and characteristic function-

$$\varphi_2(t) = e^{\frac{\sigma_2^2 t^2}{2} - mt} . \quad (15)$$

The sum of two independent normal amounts has normal distribution with average, equal to the sum of average, and a dispersion equal to the sum of dispersions. Really,

$$e^{-\frac{\sigma_1^2 t^2}{2} + im_1 t} e^{\frac{\sigma_2^2 t^2}{2} - mt} = e^{-\frac{\sigma_1^2 - \sigma_2^2}{2} t^2 + i(m_1 + im_2)t} . \quad (16)$$

Now let absolute values $|\sigma_1| = |\sigma_2|$ then the right part (16) will be equal to size

$$e^{i(m_1 + im_2)t} , \quad (17)$$

which represents characteristic function of individual angular step Heaviside on a complex plane with indexes of axes $(1, i)$. Hence, the appropriate delta - function on a complex plane in a point (x_0, y_0) can be submitted as

$$\delta_C(z) = \delta_C(z - x_0 - iy_0) . \quad (18)$$

The sum of casual deviations from this point (x_0, y_0) represents a casual vector with the specified distribution (16) and characteristic function (17) under condition of $\sigma_1, \sigma_2 \to 0$. Private acknowledgement of it is that fact, that the density of one-dimensional normal distribution at it $\sigma \to 0$ represents delta - function in one-dimensional material distribution to axes.

Example. We shall consider an example 2 of [3] with elimination of discrepancies and in more detail. On a plane $(1, i)$, $i = \sqrt{-1}$ it is observed alternating process of restoration with incomes and charges. Duration of serviceability and restoration of full serviceability we shall define in density of probabilities $f_X(t), f_Y(t)$, and density of probabilities of sizes of the income and the charge- $g_R(t), g_K(t)$, all of them are concentrated on $[0 \le t < \infty)$. All random variables are in pairs independent.

Realization of process has the following interpretation. From the beginning of coordinates the object starts to function, after the expiration of time it(he) refuses and instantly acts on restoration. During casual time of serviceability it accumulates a random variable of the income. This income can represent quantity of the advanced, saved up information, cost etc. During casual time of restoration the object can accumulate a random variable of the charge, alternating measuring the same dimension, as dimension of the income. Thus the income can be both positive, and negative. After end of the first cycle of functioning, process renews in the second cycle and so on can proceed indefinitely. So, we deal with process of restoration together with process of accumulation [6]. In figure 1 provisional realization of considered processes submitted.

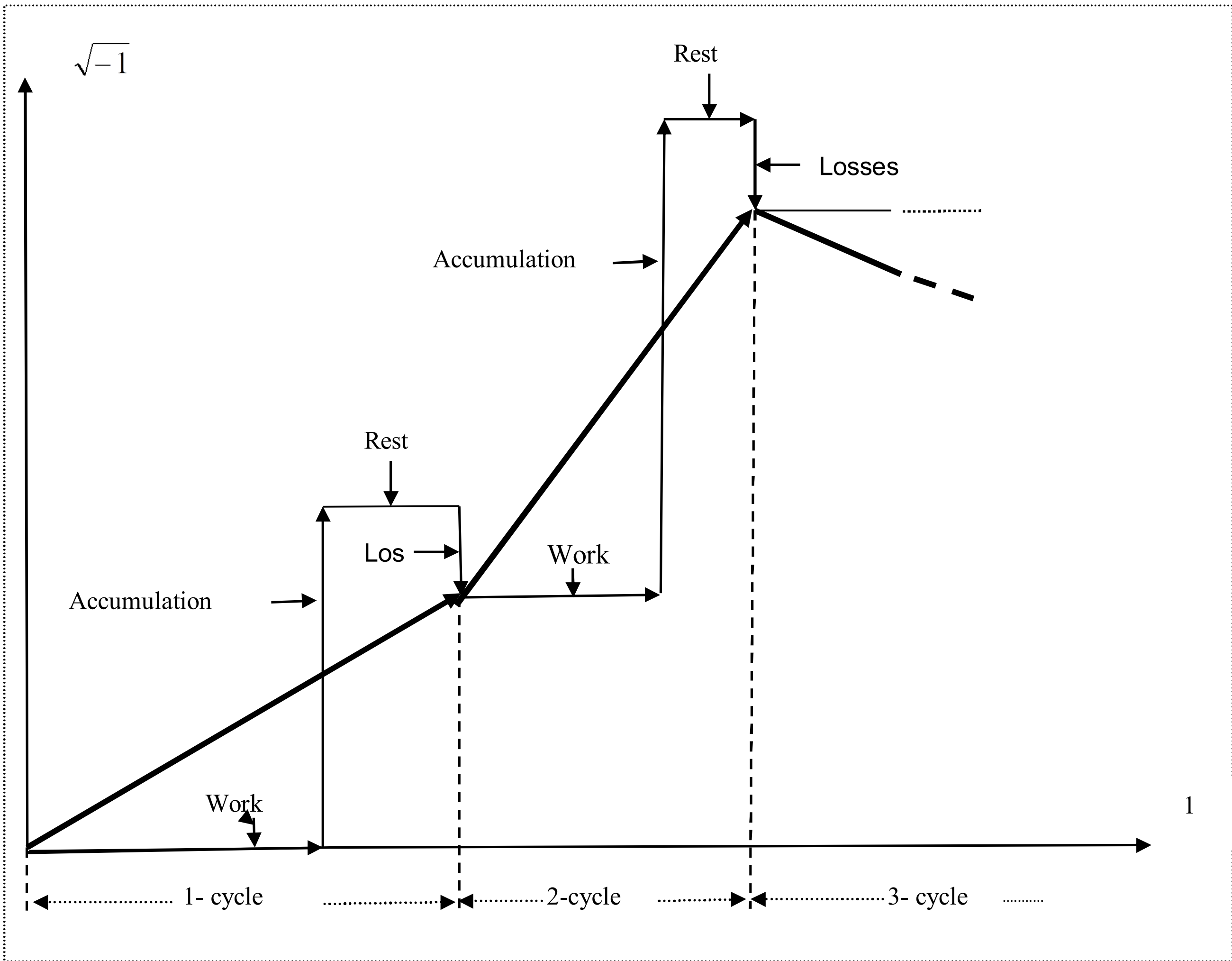


Figure 1. Realization alternating process with accumulation and the charge

Our problem is to find the stationary decision for complex alternating casual process with accumulation.

Let two density of probabilities of material variables $a_X(x), b_Y(y)$, with x, y, concentrated are given on $[0,t)$. Using them and (18) we shall construct the following expression (3):

$$\varphi_Z(z) = \int_0^\infty \int_0^\infty \delta(z - x - iy) a_X(x) b_Y(y) dx dy = \int_0^{z/i} a_X(x - iy) b_Y(y) dy = -i \int_0^z a_X(x) b_Y(i(x - z)) dx . \quad (21)$$

Under the semantic contents (21) it is density of probability of a vector $x + iy$ on planes $(1, \sqrt{-1})$, analogue « complex density ». Having applied to her transformation Laplace, we shall receive:

$$\varphi_Z^*(z) = \int_0^\infty e^{-us} \varphi_Z(u) du = \varphi_X^*(s) \varphi_Y^*(is), \quad (22)$$

where $*, s -$ a symbol and a variable of transformation Laplace. Expression (22) can be received, using as well characteristic functions.

In conditions of our example we shall have:

$$\varphi_X^*(s) = f_X^*(s) f_R^*(is), \quad \varphi_Y^*(s) = f_Y^*(s) f_K^*(is) . \quad (23)$$

For a presence of the stationary decision it is applicable known expression from the theory of restoration for alternating process [6]:

$$K_\Gamma^*(s) = \frac{1-\varphi_X^*(s)}{s[1-\varphi_X^*(s)\varphi_Y^*(s)]}. \tag{24}$$

This decision of known integrated equation Voltera for function of readiness of the theory of reliability in transformation Laplace.

The stationary decision (24), at $t \to \infty$ we shall find, having applied one of theorems Taubere type:

$$K_\Gamma = \lim_{t\to\infty} K_\Gamma(t) = \lim_{s\to 0} sK_\Gamma^*(s). \tag{25}$$

As a result of the decision we shall receive:

$$K_\Gamma = \frac{\nu_X + i\nu_R}{(\nu_X+\nu_Y)+i(\nu_R+\nu_K)}. \tag{26}$$

Where ν_X – average time of non-failure operation of object, ν_Y – average time of restoration of object after his refusal, ν_R – average size of target accumulation by object during his efficient condition in a cycle, ν_K – average size of the charge of the received accumulation in a cycle.

For an illustration of the numerical decision we shall accept the following initial sizes: $\nu_X = 100\,h., \nu_Y = 10\,h., \nu_R = 20\,units\,\inf ormation, \nu_K = 10\,units\,\inf ormation$. From (26) it is received $\mathrm{Re}(K_\Gamma) = 0{,}892; \mathrm{Im}(K_\Gamma) = -0{,}062;$ size of the module $M = \sqrt{(\mathrm{Re}(K_\Gamma))^2 + (\mathrm{Im}(K_\Gamma))^2} = 0{,}894$. We shall find value of a phase

$$tg(\psi) = \frac{\nu_Y\nu_R - \nu_X\nu_K}{\nu_X(\nu_X+\nu_R)+\nu_R(\nu_R+\nu_K)}, \psi = a\tan(-\frac{2}{29}) = -0{,}069\,rad. = -3{,}945\deg.$$

Under condition of when the income and the charge are absent $\nu_R = \nu_K = 0$, value $K_\Gamma = 0{,}909$. The relative mistake of calculation will make < 2 %. If to put $\nu_X = 0, \nu_Y = 0,$ then $K_\Gamma = \nu_R/(\nu_R+\nu_K) = 0{,}667$. It means, that the share of the income under the attitude to the sum of the income and the charge will make $\approx 67\%$. Thus, alongside with an estimation of readiness of object on an axis 1, it is possible to receive an estimation of profitability of use of object. It justifies application of the complex approach in research of more complex model in comparison with models [6].

Let's make one remark concerning the given example. At the decision of the practical problems connected to information processes, it is natural to believe, that the size of information work should be directly connected to time of serviceability and restoration of object. In our example it is possible to confirm this statement the following. If average time before refusal of object equally ν_X, the average size of the income can be expressed as $\nu_R = I_R\nu_X$, and average size of the charge- $\nu_K = I_K\nu_Y$, where I_R, I_K sizes of the income and the charge of object in unit of time. We shall put, for example, $I_R = 5$ units of the information in one hour, and $I_K = 2$ unit of the information in one hour. Then sizes of average values of the income and the charge will be equal: $\nu_R = 5\cdot 100 = 500\,un., \nu_K = 2\cdot 10 = 20\,un.$, and relative private receptions of the net profit – $\nu_R/(\nu_R+\nu_K) = 0{,}962$. Thus, dependence of a share of the general profit of object can be taken into account due to communication of an operating time and restoration with information productivity of both processes. This remark does not exclude an opportunity of application and other, more complex models "income - charge".

Conclusion

In given article acknowledgement of expression for delta - function on the complex density, entered by the author [3] is given. Examples of her use in the analysis in the same place are given. Also it is necessary to note, that for the first time the term « the complex probability » was entered by D.R. Cox for the decision not Markov problems of the theory of reliability [4]. However, problem to remain a question on introduction complex probabilistic measures in the analysis and a graphic representation of the complex functions determined on a complex variable. Expansion of complex numbers is the field quaternion's Hamilton, forming not to switch algebra with division above a field of real numbers. Thus everyone quaternion can be submitted as

$$a = a_0 + ia_1 + ja_2 + ka_3 = (a_0 + ia_1) + (a_2 + ia_3)\,j, \qquad (19)$$

Where a_0, a_1, a_2, a_3 real numbers, and $i, j, k -$ special quaternions, forming together with the valid unit basis of four-dimensional space and satisfying the following system of equality:

$$\begin{cases} i^2 = j^2 = k^2 = -1, \\ jk = -kj = i,\ ki = -ik = j,\ ij = -ji = k \end{cases}. \qquad (20)$$

Use of complex numbers and quaternions in probability theory for the description of casual multivariate processes and decisions of the appropriate scientific and technical problems, in our opinion, has the big prospects and demands deeper studying. By the way, now opportunities of performance of mathematical operations even with complex numbers on a computer, unfortunately, are extremely limited [5]. The example of the decision of a problem given in article by definition of stationary value of function of readiness of object with accumulation of quantity of the information evidently enough shows advantages of application of the complex analysis in probability theory.

Reference

1. Kalinin V.M. « My formulas ». – Publ. SPBSU him. M.V.Lomonosov, 1966.
2. Dugue D. Theoretical and applied statistics. – Transl. from French V.M.Kalinin. – M: the Science. – 1972. – 383 p.
3. Smagin V.A. Probabilistic the analysis of a complex variable. – AVT. – 1999, 2. – pp. 3-13.
4. Cox D.R. A use of complex probabilities in the theory of stochastic processes. – Proc. Soc., V. 51. – 1955.
5. Gatsenko O.J., Smagin V.A. Elements probabilistic the analysis of a complex variable. – Saint Petersburg, VCA him. A.F.Mozhajskogo. – 1998. – 18 p.
6. Cox D.R. Renewal theory. - London: Methuen and Co Ltd, New York: John Wiley and Sons Inc. - 1961. - 300 p. (the Russian edition in B.V.Gnedenko's translation).

PROBABILISTIC MODEL OF ECOLOGICAL CONSEQUENCES OF RAILROAD ACCIDENTS

Mykhailo D. Katsman, Candidate of Technical Sciences (Ph.D.), e-mail: vnbsp@sw.uz.gov.ua,
Viktor K. Myronenko, Doctor of Technical Sciences, e-mail: viktor.myronenko@yandex.ua
State University for Transport Economy and Technologies, Kyiv (Ukraine)
•
Nikolaj I. Adamenko, Doctor of Technical Sciences, e-mail: nikolajadamenko@mail.ru
V.N. Karazin Kharkiv National University, Kharkiv (Ukraine)

ABSTRACT

The paper discusses the processes of inertial reacting and self-regulation of the environment impacted by hazards of railway accidents involving dangerous goods and the queuing system Markovian model is proposed to determine the probable consequences of such accidents development.
Key words: railway transport, dangerous goods, accidents, ecosystem self-regulation, combustion, explosion, system inertia, queuing system, Markovian process, mathematical model.

1 INTRODUCTION

Rail transport carries a large number of dangerous goods with different properties, which in case of accidents may affect the environment. Therefore, it is clear that at all levels of dangerous goods transportation due attention should be given to environmental protection, thus ensuring the human life protection. Implementing appropriate measures should provide balance, stability and flexibility of natural systems, the violation of which can lead to serious negative consequences and environmental disasters.

The problems of stability, equilibrium, homeostasis of ecosystems and the biosphere are central to modern environmental science [1].

According to the laws of the biosphere, the basic principles and laws of human and biosphere evolution, their interdependence, the main causes of the ecological crisis are ill-conceived and erroneous human actions, which result not only dying species, but also destroys the ability of natural systems and components for the restoration and self-regulation.

The main feature of the biosphere and ecosystems is the ability of the environment to adapt to intense anthropogenic influence that reflects the concept of “environmental capacity” of natural systems.

In the recent literature on ecology there is a very large number of interpretations and definitions of the term, each of which reveals only part of attributes and properties that reflect the ability of ecosystems, and hence the biosphere as a whole to self-preservation, self-regulation and self-healing.

The environmental capacity of the ecosystem is understood as the maximum amount of energy or matter that may be involved in the circulation per unit of time without significant violations of its structure and stable operation [1].

The definition of ecosystem capacity is based on “substance-energy” approach, in which the functioning of ecosystems is considered as a process of transformation of energy and substance that are coming from the environment and returning to the same environment. In the process of transformation of energy and substance they turn into forms that provide a continuous circulation flow of substances in the ecosystem, its poise and balance in the biosphere, which is necessary to maintain stable operation of the structures and relationships that were formed in the ecosystem.

Thus, the ecological capacity characterizes the capability of the ecosystem to transform energy and matter that come into it, into the forms that carry on the ecosystem biological cycle, passing in historically determined path [1].

The ecosystem ability to cleanse itself is associated with the ability of ecosystems to involve in circulation slightly more matter than that which they pulled in before anthropogenic influences.

If using ecological capacity to characterize the potential ability of ecosystems to adapt and sustain stability during human impacts, the self cleaning ability is likely to consist of several stages, such as the inertia of the system towards the external negative influences, opposition to the transformation of its physical and chemical properties to the extent that can lead to catastrophic state, or attenuation of negative processes in the environment and preservation of the original ecosystem status, are all consequences of this ability, and describe the results of operation of the system in certain specific circumstances.

The sad history of railway accidents and disasters draws one's attention to one of their features – sometimes with inscrutable reasons they occur, although seemingly nothing led to them, while in other cases – on the contrary. Why so? Let's try to look for analogies with the processes occurring in living ecosystems.

2 FEATURES OF TYPICAL RAILWAY ACCIDENTS WITH DANGEROUS GOODS

Consideration of typical rail rolling stock accidents involving dangerous goods showed that the initial conditions of such accidents, in particular, are a leak, spill or release of gaseous and liquid hazardous substances caused by depressurization of rail tank cars or containers, destruction of pipelines, valve failures, emergency damage holes, etc. [2].

The formation of explosive concentrations zones in accidents involving dangerous substances is affected by two types of parameters: the parameters of the leakage source and meteorological and topographical parameters [1,3].

The intensity of the leakage of gas, vapor and liquid properties are due to such sources characteristics as leakage geometric size, velocity of a combustible substance, its concentration, temperature and pressure in the middle of the container or tank, the density and quantity of the liquid phase, evaporation and others.

Dimensions of clouds that are formed at leakage of combustible gas or vapor depend on the velocity of runoff and dispersion.

The intensity of leakage increases with the speed of leakage of combustible material and with increasing concentration of flammable gas or vapor in the combustible material, which is released.

At high velocity of outflow gas and steam can form a conical jet, which is pulling the air inside causing the ability to "self-dilution". The level of explosiveness of gas mixture that is formed in this way does not depend on wind speed.

At low outflow speeds, or when jet speed decreases or any interference occurs, that causes the "self-dilution" of gas mixture explosiveness level and it is dependent on the speed of air [3].

Evaporation of combustible liquid depends mainly on vapor pressure and specific vaporization heat of combustible material.

If the vapor pressure is unknown, to determine the mixture explosiveness, its boiling point and flash temperatures are used. Explosive mixture can not exist if the flash point exceeds the maximum temperature of combustible substances. The lower is the flash point, the larger is explosive zone.

It should be noted that the temperature of the flash is not an exact physical quantity. Some fluids are not characterized by parameters such as flash point, although they may form an explosive gas mixture. In such cases, the established value of liquid temperature corresponding to the concentration of vapor at the lower concentration limit the outbreak (ie with a minimum content of combustible material in a homogeneous mixture with an oxidant at which flame propagation is possible at any distance from the source of ignition) is compared with maximum temperature of the liquid.

For a certain amount of leakage of combustible material, the lower the minimum concentration limit of flame propagation, the larger explosive area [3].

Research on air pollution and formation of zones of explosive concentrations have shown that such processes are greatly influenced by meteorological and topographical characteristics of the area, namely, air speed, real air density, volumetric concentration of gas (vapor), wind direction, humidity, precipitation, pressure, terrain, etc., which can either increase the size of explosive concentration zones or slow down the process of their formation and reduction of such zones size to the minimum [4].

3 FUNDAMENTALS OF PHYSICAL AND CHEMICAL PROCESSES OF COMBUSTION AND EXPLOSION OF DANGEROUS GOODS

Experience of eradication of railway failures and accidents shows that the greatest threat to people, rolling stock, railway infrastructure and the environment are those that are accompanied by fire of dangerous goods [2].

Let us consider the basic physical and chemical processes of combustion and explosion of hazardous materials in different aggregate states. First of all, it should be noted that combustion is a complex, rapidly leaking chemical transformations, which is accompanied by a significant amount of heat and bright glow. In most cases, burning is a result of exothermic oxidation of substances capable of burning (fuel) by oxidant (oxygen, chlorine, etc.). Some other processes are also considered as burning and are associated with the rapid transformation and thermal or chain acceleration of the processes: the decomposition of explosives, ozone, interaction of barium oxide with carbon dioxide; decomposition of acetylene, etc. [5].

Combustion is a complex of interrelated chemical and physical processes, the most important of which are heat and mass transfer [6]. The most common feature is the ability of fire burning flame that arose to move throughout the mixture by heat transfer or diffusion of active particles from the combustion zone into a new mix.

In the first case the heat transfer is realized, and in the second case the diffusion mechanism of flame propagation takes place. Typically, combustion occurs in combined thermal diffusion mechanism. It is important that combustion is characterized by critical conditions (mixture composition, pressure, temperature and geometric size of the system) for the emergence and spread of flame. In all cases, the combustion is characteristic of three typical stages: emergence, spread and flame extinction [7,8].

Depending on the physical state of fuel and oxidizer there are three types of combustion:

- Homogeneous combustion of gases and vapor flammable substances in the medium of gaseous oxidizer;
- Heterogeneous combustion of liquid and solid combustibles in medium of gaseous oxidizer (kind of heterogeneous combustion is the combustion of liquid fuels in liquid oxidizer);
- Combustion of explosives and powder.

Depending on the speed of flame propagation combustion is divided into deflagration that flows at subsonic velocities and detonation, which is distributed with supersonic velocities.

In its turn, subsonic combustion is divided into laminar and turbulent. Laminar burning speed depends on the mixture composition, initial values of pressure and temperature, as well as the kinetics of chemical reactions in the flame. Speed of propagation of turbulent flames, in addition to these factors, also depends on the flow velocity, the degree and scale of turbulence [6].

The explosion is a process of rapid discharge of large amounts of energy. The blast explosive (or explosive) mixture fills the volume where energy discharge occurred, the mixture is converted into highly heated gas at the high pressure. This gas affects the environment with great force, causing the formation of a blast wave. Destruction caused by an explosion is due to the action of such a wave. As the distance from the explosion mechanical action of the blast wave weakens.

Consider basic processes of combustion gases.

The first stage of combustion – ignition – is the initiation of the initial fire burning in the fuel mixture. Found that ignition of flammable gas mixture may be in their contact with hot surfaces (eg., gas exit from tank hole that is in the fire zone), or the appearance of sparks or flames in the middle of the mixture, as it may happen in a situation of liquefied hydrocarbon gases leakage, followed by leakage source fire.

Ignition of a combustible gas mixture resulting from the collision with the red-hot surface of the container is provided if the surface temperature exceeds the value of the temperature of ignition ($T_{\inf}$). The nature of the process is such that when the surface temperature (T_{sur}) is not sufficient for the process of progressive fuel mixture heating and self accelerating reaction, then the exothermic heat of transformation is given back to the cold mixture. If $T_{\inf} \succ T_{sur}$, then progressive self-heating occurs in the fuel mixture and at some distance from the heated surface the combustible mixture temperature becomes greater than T_{sur} that leads to the formation of a primary combustion chamber.

In addition, the ignition temperature depends on the nature of the container surface material faced by gas mixture for at red-hot surface ignition of gases, the catalytic properties of the surface are activated. Thus, if the catalytic effect found in branched chain reaction, the critical ignition temperature decreases, and vice versa, when the interaction of the gas with the surface leads to breaking the chain reaction, the greater ignition temperature T_{sur} is needed.

Ignition temperature changes also depending on the initial values of the mixture pressure – pressure reduction leads to an increase of $T_{\inf}$ [6, 7].

A great threat to the rolling stock, railway transport infrastructure and the environment is a situation where an electrical discharge occurs in a leakage zone of dangerous goods in gaseous aggregate state.

The emergence of electric discharge in combustible gas leads to ionization of the gas and transforms it into a plasma. This process is accompanied by a strong heating of ionized zone. In the discharge channel, the temperature exceeds 10000 K [7].

However, not any electrical discharge results in the emergence of fire flame in combustible environment. Flames arises only when the energy released during the discharge exceeds the value of the minimum ignition energy. In other cases, the fireplace flame does not occur.

Heating by electric discharge of an initial volume of combustible gas mixture causes additional heat by chemical conversion. Redistribution of heat pulse energy in a combustible mixture makes the energy of chemical reactions added together with the energy of the initial pulse. Increasing the size of the heating sector is accompanied by increasing the total amount of heat produced, and share of chemical reaction energy in it.

If the effect of an electric spark to combustible mixture led to involvement in chemical transformation enough of combustible material and temperature of the process of volume increasing of the heated mixture is committed to the combustion temperature, the system is set stationary.

The heat that is given from the reaction zone into fresh mix is offset by the heat produced during the reaction and there is a steady flame front.

If the distance from the flame front to the place of spark increases, the influence of initial momentum to the process that develops becomes less significant.

Thus, stable flame front is formed in the case when the energy level is sufficient to heat up to the temperature of combustion a spherical volume of combustible mixture, the critical radius r_{CR} should be several times larger than the characteristic width of the laminar flame zone δ_{FL} [9]:

$$r_{CR} \geq 3{,}7\delta_{FL} . \quad (1)$$

In this condition the mix layers surrounding the area that is burning, have enough time to catch fire before the volume around hot spark gets cooled.

If equation (1) is not satisfied, then the stationary regime is not established for the heat output from the reaction zone exceeds the heat produced inside the zone, the combustible mixture is cooled, and the reaction that occurred in the area of discharge stops.

Another common phenomenon that accompanies accidents involving dangerous goods during their transportation by rail is spontaneous ignition.

The essence of ignition is a sharp increase in the rate of exothermic reactions, resulting in the burning of substances in the absence of a source of ignition.

It should be noted that in many theoretical studies, investigations of combustion processes often do not distinguish between the terms "ignition" and "spontaneous ignition". In papers devoted to fire and explosion hazard the term "ignition" is used for the process of forced ignition, i.e. initiating combustion by highly heated source of ignition, and the concept of "spontaneous combustion" for the processes of flame burning in the absence of such sources [7].

The condition of thermal ignition is to ensure that the initial self heating of a fuel mixture resulting from the oxidation reaction must exceed a certain critical value [10]:

$$\Delta T \geq RT_0 / E , \quad (2)$$

where R is the universal gas constant; T_0 is the temperature of the cooled mixture, K; E is the activation energy.

The time during which the reacting system is getting a heating which is defined by (2), is called the spontaneous combustion induction time. Induction period depends on the composition of the mixture, its initial temperature and pressure. Induction period is of practical importance when combustible gas-air mixture is exposed to low-power source of ignition (spark). Spark, getting into this mixture heats a mixture of volume and at the same time the spark is cooled. Thus, if the induction time is longer than cooling time, the ignition will not occur.

It is found that thermal ignition occurs more easily, the higher are the reaction rate and temperature of combustion, and the less are the heat transfer speed and pre-explosion heating [6, 7].

Particular attention is given to the temperature dependence of the spontaneous combustion on the fuel mixture composition. If the mixture has a small amount of combustible material and there is an excess of air, the ignition of the mixture is not possible. Also, the presence in the mixture of excess fuel and shortness of air, too, making impossible the ignition of the mixture.

The spread of flame is worth special attention. Combustion initiation of gas mixture at one point leads to heating the neighboring layers of mixture, which starts the chemical conversion. Combustion of these layers entails initiating combustion of further layers and so on, until the complete burnout of combustible mixture. Thus, after the ignition, the flammable mixture burns by layers. Combustion zone moves across the mixture, providing flame propagation.

The area in which the chemical transformation occurs and there is intense warming of gas that burns, is called the flame front.

Before the flame front that is moving, there is a mixture of fresh mixture (not yet burned), and behind the front there are the products of combustion.

If fresh mixture moves toward the flame front at a speed equal to the speed of flame propagation, the flame will be fixed [11, 12].

Since the chemical transformation is highly dependent on temperature, bulk gas combustion is carried out in the area where the temperature is close to fresh mixture combustion temperature (T_B), so the length of time (τ) of the mixture stay in the combustion zone [12] is:

$$\tau = \tau_0 e^{E/RT_B} \quad (3)$$

And flame propagation velocity [12] is given by:

$$U = B_0 e^{E/RT_B}, \quad (4)$$

where B_0 is the value which depends on the properties of the mixture.

Redistribution of flame heat released by the reaction takes place, for heating of fresh mixture and partially coming into the surrounding environment. If the heat loss will be higher than a certain critical value, the progressive decrease in temperature and its attenuation will take place.

Taking into account the mutual influence of heat losses from the combustion zone, the combustion temperature and flame propagation velocity, the basic theory tenets can be formulated that limit the spread of flame. From this theory it follows that the condition for the possibility of flame propagation through combustible mixture is predicted by the relation [12]:

$$T_{CR} = T_{TH} - (RT_{TH}^2 / E), \quad (5)$$

where T_{CR} is the threshold value of T_B; T_{TH} is the theoretical combustion temperature.

The limit flame propagation speed U is given by [12]:

$$U_{CR} = U_{MAX} / \sqrt{e}. \quad (6)$$

Equation (5) indicates that the flame can not spread through the fuel mixture when the temperature is lower than the theoretical value that exaggerates (RT_{TH}^2 / E).

During the combustion of gases in open space the reaction products freely expand and pressure remains almost constant. Combustion in a closed volume is accompanied by increased pressure. The maximum pressure of the explosion in a closed volume is defined by thermodynamic properties of the combustible mixture and heat losses from the combustion zone.

Based on the above, it can be assumed that the nature of the phenomena that accompany the combustion of gases, tends to lag the ability to respond to external factors, and in some circumstances can completely prevent these factors to keep the development of the combustion process, until its termination.

The largest quantity of dangerous goods carried by rail is goods that are flammable liquids.

Consider basic processes that accompany the burning of liquids.

Burning of liquids is a complex physico-chemical process that takes place when mutual influence of kinetic, thermal and hydrodynamic phenomena occurs. Burning of liquids occurs in the gas phase. As a result of evaporation of the liquid surface a steam jet is formed and mixed with the air oxygen and chemical interaction ensures the formation of the combustion zone.

Burning zone is a thin layer of glowing gases, which come from the surface of the liquid flammable vapors and oxygen diffused from the air. Stoichiometric mixture is formed (i.e. such that has no excess of either fuel or oxidizer) which is burned in a split of second.

Shape and size of flame of burning liquids depend on the diameter of the tank (hole in the unit), which is burning. Flame height increases with the diameter of the reservoir holes [13].

The flame above the surface of the combustible liquid is stable, if there is a defined speed of coming fuel and oxygen.

Rate of fuel input depends on its vapor pressure above the liquid, and hence on its temperature. The lowest temperature of the liquid (T_B) in which the flame arose, and will not go out, is called flashpoint temperature.

Established that the ignition temperature is determined by formula [13]:

$$P_B = A/(D_0 \beta T_B), \quad (7)$$

Where P_B is saturated vapor pressure at the temperature of liquid ignition; A is a fixed device value; D_0 is diffusion coefficient of vapor in air; β is oxygen stoichiometric coefficient.

The process of liquid burning is also characterized by burnout speed. Burnout speed is not a physical or a chemical constant; it depends on the properties of flammable liquids, tank (holes) diameter and the conditions of heat and mass transfer in the fire zone.

Like the processes of gases combustion, during combustion of flammable liquids a tendency is observed when the combustion processes slowdown response time to external factors that cause their burning despite burning process. Under certain conditions, untill a significant slowdown or even, to a complete termination.

A significant proportion of goods transported by rail are solid combustible materials.

Combustion of solids differs from combustion of gases by the presence of stage of decomposition and gasification.

Combustion among gaseous oxidizer often comes as a result of ignition of volatile pyrolysis products. Converting solid combustible material into products of combustion is not concentrated only in the area of the flame.

Combustion of solids has a multistage nature. Under the influence of external heat the solid phase is heating, which is accompanied by decomposition and release of gaseous products. Then these products are ignited and burned. Heat from the torch that is formed affects the solid surface, causing revenues to the combustion zone of new portions of combustible gases.

Model of solid substance burning presupposes such zones [14]:

- Heating of the condensed phase. Thermoplastic materials are melting in this zone. The thickness of this zone is defined by the coefficients of thermal conductivity and burning rate and is about 3 mm;
- Pyrolysis or reaction zone in the condensed phase, where gaseous combustible substances are formed;
- Pre-flame zone in the gas phase, where a combustible mixture is formed;
- Flame zone or reaction zone in the gas phase, where the pyrolysis products are conversed into the gaseous products of combustion;
- Combustion products zone.

The intensity of the reactions that occur in the surface layer of the solid and heat exchange conditions of gaseous decomposition products with the environment define the processes of combustion – spontaneous ignition or ignition.

In case of spontaneous ignition, the warmth that comes to the surface of the solid from the heat source is uniformly distributed throughout the thickness at the surface layer, which corresponds to the characteristic size of the material. With ignition from an external source which is the warmed external layer, where the heterogeneous reaction is occuring the layer thickness is substantially less than characteristic size of the material.

Thus, as in the process of combustion gases and liquids, the combustion of solids phenomenon of inertia of combustion processes is also observed. Under certain conditions, the system “solid material that burns – environment”, tends to inhibition of combustion processes, until their termination.

Some classes of dangerous goods transported by rail can form powders in accidents.

The powder mixture burning process is determined by the heat transfer mechanism in the flame front. There are several theories that explain the pattern of flame spread by means of conductive, radiative and conductive-radiative heat transfer from the combustion zone into the fresh mixture.

For organic systems, heat transfer is mostly carried out by conductivity and convection. Because of low fuel gasification temperature, and narrow zones of combustion the predominant mechanism of heat transfer is thermal conductivity of the gas. Effect of gravity on powder-gas mixture combustion is found in particles settling down under gravity, which leads to a relative velocity of phases in the fresh mixture, as heated products of combustion are affected by Archimedean force.

Model of the flame front in this case we can apply in some way like this. Under the influence of heat flow from high temperature zone where powder cloud burns, particles have time to evaporate before ignition. Flame front spreads in homogeneous gaseous mixture of fuel vapor and air. Reaction of fuel with oxidizer flows in the kinetic region, obeying certain thermal theory laws.

The movement of the flame front leads to partial scattering of fresh mixture near the leading points of flame. This gas phase (oxidizer) is dissipated to a greater extent than condensed phase (fuel), resulting in phases having relative velocity and thus changing the fuel and oxidizer ratio in the flame front.

Increasing the fuel concentration is accompanied by the growth of the flame velocity in these areas, causing further growth of convex parts of the flame front and lagging concave regions.

This effect leads to the fact that the flames can spread throughout suspension mixture with average fuel concentration below the concentration limits of flat flame front propagation in the gas mixture. Approximate estimates show that the minimum content of fuel in homogeneous mixtures with an oxidizing agent, which may spread the flames to mixture at any distance from the source of ignition is about two times less than the same minimum content of fuel gas mixtures of the same substances. This property of organic substances suspension mixture is detected since the particles diameter of 10 microns [15].

Particularly noteworthy are combustion processes of natural fuels suspension mixture, which account for a substantial proportion of goods transported by rail (coal, peat, some fertilizers).

Solid fossil fuels differ from most chemicals the presence of three components: the flying particles, coke and ash. The processes of ignition and flame propagation of each of these components have certain features.

Flying share of solid fuel is a gaseous component released from the fuel during heating without oxidant. Coke in its composition is similar to carbon. Speed of coke burning is much lower than burning rate of volatile particles.

In this regard, participation of coke in powder explosions of natural fuels is negligible. In the ash, which is part of the mineral fuels, a number of components are contained that can participate in combustion (alkali metals, Pirita and pyrite). But ash, in general, plays the role of an inert material.

Explosions of solid fuels suspension mixtures are typical thermal explosions. Flame propagation in mixtures is a result of heat transfer from combustion products into the fresh mixture. Heat can be transmitted by different mechanisms depending on the particle size, concentration, composition and parameters of gas medium and other factors [15,16].

Unlike combustion processes in gas mixtures suspension mixtures of natural fuels are constrained by the duration of particles heating and the possibility of fuel oxidation reaction to occur in kinetic as well as in the diffuse field. In general, the temperature of the particles differs from the temperature of the ambient gas both in the area of chemical interaction and in the area of heating [16].

Mostly recognized is the model of flame propagation in suspension mixtures of fossil fuels particles, which was proposed in [16].

According to this model, the maximum speed of flame propagation U_{FL} at sufficiently large thickness of the front is:

$$U_{FL} = \frac{\sigma T_E^4}{c\rho\mu(T_S - T_0)}, \tag{8}$$

where T_E is effective radiation temperature of the flame front; σ is constant Stefan - Boltzmann; c, ρ, μ are respectively, volumetric heat capacity, density and concentration of the solid phase; T_S is the temperature of spontaneous combustion; T_0 is initial temperature of the mixture.

The initial period of flame propagation in suspension mixtures is characterized by hopping rate, due to the size of initially inflamed area, duration of heating to a temperature of spontaneous combustion, which depends on the thickness of the flame front radiating, and dust particles burning.

Based on the above, it is natural to assume that the phenomenon of dust burning is accompanied by inertia, that is able, under certain conditions, to slow down combustion process, or stop it completely.

4 MATHEMATICAL MODEL OF THE ECOLOGICAL SYSTEM

In order to forecast the state of the environment a large number of mathematical and simulation models is used. To build such models, differential equations are used that describe the various physical and chemical processes of pollution spread in soil, rivers and reservoirs under various boundary conditions, taking into account the spread of contamination, given weather conditions, power of pollution sources and physical properties of the underlying surface (its relief, development, forest areas, etc.). Methods of linear regression analysis, pattern recognition, image sequential regression and others [17, 18] are also widely used.

In our view, next to the above models, to predict the consequences of accidents involving rail transport of dangerous goods, models of queuing theory are also quite useful. Having examined in this paper the processes of accidents development with dangerous goods of different physical state, we can conclude that the environment has some lag of response to an external hazardous accident factors. That prevents the environment to change its condition and behavior due to the properties of self-support and self-regulation that, under certain conditions, may lead to inhibition of catastrophic processes until their termination.

Let us consider the ecological system of “emergency rolling stock – the environment” as a Markovian queuing system (QS). In Markovian queuing system all flows of events (arrival of customers) that lead the QS from one state to another are stationary Poisson flows. This means that the time intervals between adjacent events in the flow have exponential distribution with parameter λ equal to the intensity of the corresponding flow (or reciprocal value of time interval between events).

In this QS an arrival flow of “customers” enters the system – subsequent portions of hazardous accident factors (HAF) that impact the QS in with intensity λ. As such “portions” can be seen, for example, some smallest quantities of hazardous liquid or gas escaping under great pressure from the holes in the tank, creating (or not creating) an explosive gas-air mixture, etc. Then, the arrival flow rate λ can be defined as the reciprocal expected time to reach an explosive concentration of the mixture.

Inertial action time of “emergency rolling stock – environment” system has exponential distribution with intensity ν, and the time of self-healing (recovery ability of the system to return to the original safe state) has a parameter of intensity μ. In this sense μ is the reciprocal average time of system self-healing, while ν is the reciprocal average time lag (delay of system responses to HAF). Service of a portion customer in that QS consists of two phases.

The essence of the service is that a portion arriving to the first phase reach a critical concentration value and after that servicing the next portion customer is refused and the portion is moved into the second phase of service (QS-2).

The first phase of the HAF portion service is a single-channel queuing system with queue and a service channel “heating-up”, which is considered in [19]. In this case, the “heating-up” type of service channel is the realization of the inertial properties of the system. The graph of this QS-1 is shown in Figure 1.

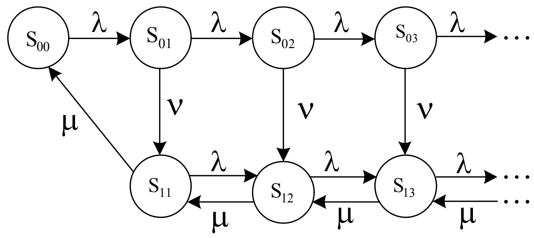


Figure 1. Graph of QS-1 states with queue and "heating-up" service channel

The states of QS-1 (Fig. 1) are as follows:

S_{00} - Channel free, not heated;

S_{01} - One portion of HAF arrived and is waiting till the channel is heated; system inertia is in action;

S_{11} - The channel is heated, one portion of HAF is being serviced, no queues;

S_{02} - The channel is being heated up; there are two portions of HAF in a queue;

....

S_{0k} - The channel is being heated up; there are k portions of HAF in a queue;

S_{1k} - One portion of HAF is being served in a channel; there are $(k-1)$ portions of HAF in a queue, etc.

The system of equations for the final probabilities P_{lk} is as follows:

$$\begin{cases} \lambda P_{00} = \mu P_{11} \\ (\lambda + \nu) P_{01} = \lambda P_{00} \\ (\lambda + \mu) P_{11} = \nu P_{01} + \mu P_{12} \\ (\lambda + \mu) P_{02} = \lambda P_{01} \\ (\lambda + \mu) P_{12} = \nu P_{02} + \lambda P_{11} + \mu P_{13} \\ \dots\dots\dots\dots\dots\dots \\ (\lambda + \nu) P_{1,k} = \lambda P_{0,(k-1)} \\ (\lambda + \mu) P_{1,k} = \nu P_{0,k} + \lambda P_{1,(k-1)} + \mu P_{1,(k+1)};\dots \end{cases} \tag{9}$$

In the second phase of the service when the next $(l+1)$-st portion of HAF is rejected, because the number of customers in the QS-1 exceeds the limit (one portion), it comes for the service in the QS-2.

QS-2 is a single-channel queuing system with failure (Fig. 2).

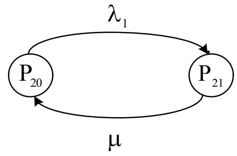


Figure 2. Graph of QS-2 states

The equations for the final probabilities will be the following:

$$\begin{cases} \mu P_{21} = \lambda_1 P_{20} \\ P_{20} + P_{21} = 1 \\ \lambda_1 = \lambda P_{12} \end{cases} . \tag{10}$$

Hence the probability of catastrophic consequences of the accident is:

$$P_{21} = \frac{\lambda_1}{\lambda_1 + \mu} = \frac{1}{1 + \frac{\mu}{\lambda_1}} = \frac{1}{1 + \frac{\mu}{\lambda P_{12}}} . \tag{11}$$

Figures 3a, 3b and 3c show dependences of probability of catastrophic consequences of accidents against the intensity of the recovery processes of the system μ at different values of inertia ν and arrival flow λ_1. These figures with upper and lower indices in the intensity or inertia values, for example ν^{IV}_3, mean the relevant series of computational experiments.

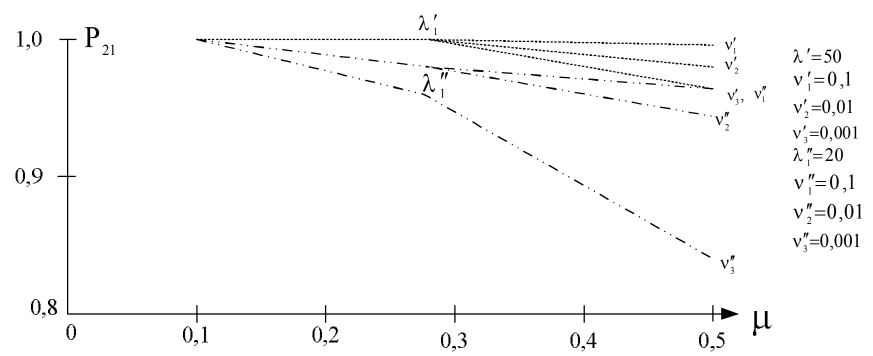


Figure 3a. Probability of catastrophic consequences of accident P_{21} against the system self-recovery intensity μ for large rate of HAF arrival flow λ_1 and small values of inertia of the system ν.

Figure 3a shows that for large values of HAF portions flow intensity at QS-2 entry, that were rejected in the QS-1, significantly higher than the intensity of response on the violation of its equilibrium ν (for large values of the average inertia time of the system) and with increasing intensity of recovery processes μ (reducing the value of the mean recovery time), the probability of the catastrophic consequences of the accident for example, considered is somewhat reduced, but still remains quite high.

Figure 3b presents dependence of probability P_{21} on intensity of recovery processes μ at moderate λ_1 and small ν values.

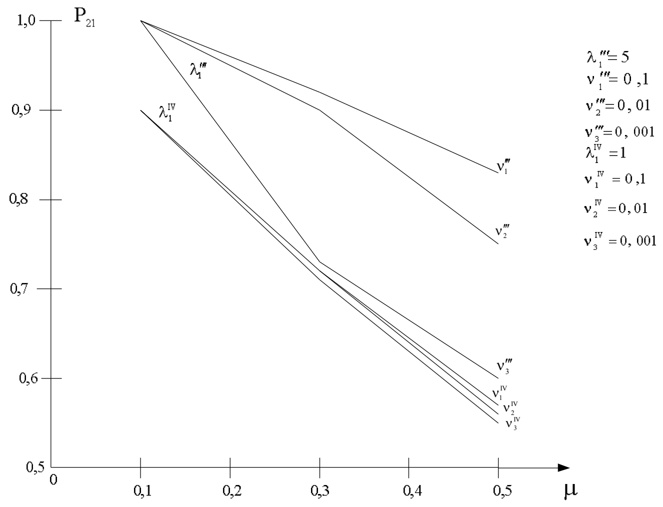


Figure 3b. Probability of catastrophic consequences of accident P_{21} against the system self-recovery intensity μ for moderate intensity of HAF arrival λ_1 and small values of system inertia ν .

Examination of the graphs in Figure 3b shows that at moderate values of arrival flow rate λ_1, the low intensity of the system inertia ν and increase the intensity of recovery μ , the probability P_{21} tends to decrease. The largest decrease of the probability of catastrophic consequences is having place with decreasing HAF arrival intensity to QS-2. For example, given λ^{IV}_1= 1, ν^{IV}_3= 0.001 and at μ= 0.5 the value of P_{21}= 0.545. Meanwhile, the probability of catastrophic consequences of accidents for example in question is still considerable.

With increasing ν values (reducing the average time of the system inertia) and decreasing the average time recovery (increase of μ), the probability P_{21} is significantly reduced and becomes insignificant as λ_1 decreases (Fig. 3c).

The calculations show that at λ_1^{VII}= 20, λ_1^{VIII}= 50 and ν= 1, 2, 3 with an increase, the probability P_{21} is from 0.005 at μ= 0.1 to 0.004 at μ= 0.5, ie catastrophic consequence is practically impossible.

Similar conclusions can be drawn and at λ_1^{IX} = 20, λ_1^{X} = 50 and ν= 10, 20, 30, when at certain μ changes the probability P_{21} in the highest value does not exceed 0.4, and the lowest is 0.01.

Note that in the examples of this mathematical model application (graphs in Fig. 3), somewhat “abstract”, dimensionless (relative) values of HAF arrival flow and service rate are used. This enabled us to focus on the demonstration of a new theoretical approach proposed by the authors for this class of problems. In certain practical applications of this theoretical approach and mathematical model one should apply appropriate values and their parameters. Features of the practical application of the model for different conditions that occur during transportation of dangerous goods by rail and other transport modes obviously require some specific research and scientific analysis.

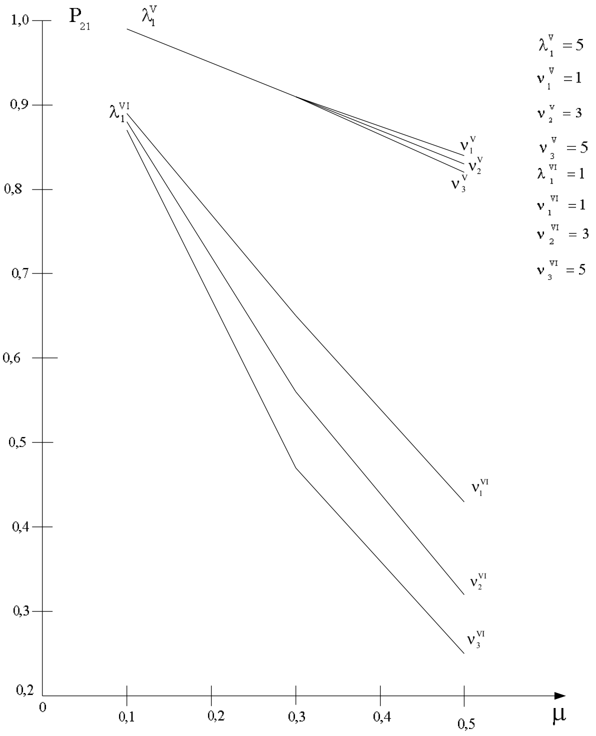


Figure 3c. Probability of catastrophic consequences of accident P_{21} against the system self-recovery intensity μ for moderate intensity of HAF arrival flow λ_1 and system inertia ν .

5 CONCLUSIONS

1. Natural processes and anthropogenic factors of hazards in rail rolling stock accidents as having properties similar to natural ecosystems against hazards of accidents, such as inertia and self-recovery are considered in this study on the basis of general theoretical provisions of ecosystems, combustion and explosion theories, probability and queuing theory.

2. Formal description of hazards development process in rail accidents is done on the basis of mathematical tools of queuing theory, which is used in other applications (communications, transportation, etc.), to simulate conditions and quantify the factors that characterize accidents involving dangerous goods carried by rail.

3. Specific numerical examples are proposed and examined, based on the mathematical model of a two-phase queuing system and the conditions of disastrous effects are analyzed depending on the intensity of hazardous accident factors, inertia and self-recovery properties of systems “emergency rolling stock – the environment”.

6 REFERENCES

1. Lavrik V.I., Bogolyubov V.M., Poletaeva L.M. (2010). *Simulation and prediction of environmental conditions*: a textbook. Kyiv: “Academy” Publishing Center.

2. Katsman M.D. et al. *Elimination of fires on the railway*: Manual (2006). Kyiv: Osnova Publishers.

3. NAPB B.03.002-2007. *Rules determining the categories of buildings, houses and outdoor installations by explosion and fire hazard.* (2007). Ministry of Emergency Situations of Ukraine. Order issued 03.12.2007, № 833. Kyiv.
4. Berland M.E. *Prediction and control of air pollution* (1985). Leningrad: Gidrometizdat.
5. Baratov A.M., Pchelintsev V.A. *Fire safety*. (1997). Moscow: ACB Publishing House.
6. Williams F.A. *Combustion Theory.* (1974). Moscow: Nauka.
7. Baratov A.N. *Fire safety. Danger of explosion. Reference book* (1987) Moscow: Chemistry.
8. Spalding D.B. *Fundamentals of the theory of combustion.* (1959). Moscow: State Energy Publishing House.
9. Schetinkov E.S. *Physics of gas combustion*. (1965). Moscow: Nauka.
10. Semenov N.N. *Chain reactions*. (1986). Moscow: Nauka.
11. Frank-Kamenetsky D.A. *Diffusion and heat transfer in chemical kinetics*. (1967). Moscow: Nauka.
12. Zeldovich J.B. et al. *The mathematical theory of combustion and explosion* (1980) Moscow: Nauka.
13. Blinov V.I., Khudyakov G.N. *Diffuse burning of liquids*. (1961) Moscow: Publishing House of the USSR Academy of Sciences.
14. Monakhov V.T. *Research methods of fire danger substances*. (1972) Moscow: Chemistry.
15. Baker W. et al. *Explosive phenomena. Assessment and Implications: In 2 books*. (1986). Moscow: Mir.
16. Todes O.M. (1973) et al. *The radiation mechanism of the formation and development of the flame front in aerodisperse systems.* Reports USSR Academy of Sciences, Vol. 213, № 2. – p. 321-324.
17. Krapivin V.F. et al. *Mathematical modeling of global biospheric processes* (1982). Moscow: Nauka.

18. Akimov V.A. et al. *Fundamentals of risk analysis and management in the fields of natural and technological* (2004) Moscow: Business Express.
19. Wentsel E.S., Ovcharov L.A. *Applied problems in probability theory*. (1983). Moscow: Radio and communication.

INLIER PRONESS IN NORMAL DISTRIBUTION

K. Muralidharan[1] and Arti Khabia [2]

•

Department of Statistics, Faculty of Science
The Maharajah Sayajirao University of Baroda, Vadodara 390 002, India.
[1]Email: lmv_murali@yahoo.com
[2]Email: artimkhabia@yahoo.com

ABSTRACT

Inliers in a data set are subset of observations not necessarily all zeroes, which appears to be inconsistent with the remaining data set. They are either the resultant of instantaneous or early failures usually encountered in life testing, financial, clinical trial and many other studies. We study the estimation of inliers in Normal distribution. The masking effect problem for correctly identifying the inliers is also discussed. An illustration and a real life example is presented with detailed discussions.

Key Words: inliers; optimal estimating equations; mixture distribution; MLEs; asymptotic distribution; early failures; Schwarz's Information criterion; modified likelihood test.

1. Introduction

The normal distribution is a very important statistical model occurring in many natural phenomena, such as measurement of height, blood pressure, lengths of objects produced by machines, etc. Usually normal distributions are symmetrical with a single central peak at the mean (average) of the data. But many times we may get normal distribution as mixture of two groups. For example the life time of an electronic item will have two sets of observations, where one set of data may have zero or small life times due to instantaneous or early failures (together called inliers) and the other set contains positive life times called target life times. This may create two symmetrical curved graphs, where the mean of inliers group is much less than the mean of target group. Such failures usually discard the assumption of a unimodal distribution and hence the usual method of modeling and inference procedures may not be accurate in practice. Usually, these situations are handled by modifying commonly used parametric models suitably incorporating inconsistent observations. The modified model is then a non-standard distribution and we call such models as inliers prone models.

Normal mixture distributions are arguably the most important mixture models, and also the most technically challenging. The likelihood function of the normal mixture model is unbounded based on a set of random samples, unless an artificial bound is placed on its component variance parameter. There has been extensive research on finite normal mixture models, but much of it addresses merely consistency of the point estimation or useful practical procedures, and many results require undesirable restrictions on the parameter space.

The first formal treatment for inliers is discussed in Muralidharan and Lathika (2004). Some recent studies on inlier model related problems in exponential distribution are by Kale and Muralidharan (2000), Muralidharan and Kale (2007, 2008) and Muralidharan and Arti (2008) and the references contained therein.

The object of this paper is to consider the problems associated with the inliers detection in normal distribution as given in (1.1) as the distribution has many potential applications in life testing experiments with instantaneous and early failures. A two parameter normal family has the probability density

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2, \quad -\infty < x < +\infty, -\infty < \mu < +\infty, \quad \sigma > 0 \qquad (1.1)$$

In section 2, we present various inlier prone models and their estimation belonging to (1.1). Section 3 deals with an illustrative example, where the estimates of the parameters under various models are discussed. The inliers detection using information criterion is presented in Section 4. In section 5 we list down two statistical tests useful to detect whether all observation belong to single normal population or they belong to mixture of two normal populations. The masking effect of the inliers is presented in the last section.

2. Inlier(s) prone models and estimation

2.1 Normal with instantaneous failures

In a parametric model for Failure time distribution (FTD) we start with a family of FTD $\Im$ = $\{F(x, \theta), x \geq 0, \theta \in \Omega \subset R_m\}$ where the form of the distribution function (df) is known except for labeling parameter, *m* dimensional θ and *F* is absolutely continuous function with probability density function (pdf), $f(x,\theta)$ with respect to Lebesgue measure. The basic problem is to infer about unknown θ or a suitable functions thereof say $\psi(\theta)$, on the basis of a random sample of size *n* on the observable random variable say $X_1, X_2, ..., X_n$. The occurrence of instantaneous failures when some items put on test giving $X_i = 0$ is quite common in electronic component and life testing situations. Note that because of the limited accuracy of measuring failure time it is possible that we record $X_i = 0$ for some units although $P(X_i = 0 \mid \theta) = 0$. To accommodate such instantaneous failures, the model $\Im$ is modified to model G $=\{G(x,\theta,p), x \geq 0, \theta \in \Omega, 0 < p < 1\}$, where

$$G(x,\theta,p)=\begin{cases}1-p, & x=0\\ 1-p+pF(x,\theta), & x>0\end{cases} \qquad (2.1)$$

where $F(x,\theta)$ is according to normal distribution and p is the mixing proportion. The estimation of parameters in the above model is straight forward and depends on only the positive observations in the model.

2.2 Normal with early failures

If early failures are nominally reported as $X = \delta$ then the distribution function of the modified model G_1 is given by

$$G_1(x,p,\theta)=\begin{cases}0, & x<\delta\\ 1-p+pF(\delta,\theta), & x=\delta\\ 1-p+pF(x,\theta), & x>\delta\end{cases} \qquad (2.2)$$

The corresponding probability density function is given by

$$g_1(x,p,\theta)=\begin{cases}0, & x<\delta\\ 1-p+pF(\delta,\theta), & x=\delta\\ pf(x,\theta), & x>\delta\end{cases} \tag{2.3}$$

The likelihood of this model can be written as

$$L(x,p,\theta)=\left[1-p+pF(\delta,\theta)\right]^r\left(p\left[1-F(\delta,\theta)\right]\right)^{n-r}\prod_{x_i>\delta}\frac{f(x_i,\theta)}{1-F(\delta,\theta)} \tag{2.4}$$

That is, the likelihood of the sample under $g_1\in G_1$ is the product of the likelihoods of r and the conditional likelihood of the sample given r which is same as the likelihood of $(n-r)$ observations coming from the truncated version of $f\in\Im$ (or $g_1\in G_1$) restricted to (δ, ∞). Since r is binomial with probability of success given by $1-p+pF(\delta,\theta)$, the distribution is complete for fixed θ and $p\in[0,1]$. Therefore, the optimal estimating equation for θ ignoring p is the conditional score function given r or $\frac{\partial \ln L_r}{\partial\theta}=0$, where $L_r=\prod_{x_i>\delta}\frac{f(x_i,\theta)}{1-F(\delta,\theta)}$. Maximum likelihood (ML) equations corresponds to two parameter normal models are given as

$$\ln L=r\ln\left[1-p\bar{F}(\delta,\theta)\right]+(n-r)[\ln p-\ln\sigma_1]-\frac{1}{2}\sum_{x_i>\delta}\frac{(x_i-\theta)^2}{\sigma_1^2} \tag{2.5}$$

$$\frac{\partial \ln L}{\partial p}=0\Rightarrow\frac{-r\bar{F}(\delta,\theta,\sigma_1)}{1-p\bar{F}(\delta,\theta,\sigma_1)}+\frac{(n-r)}{p}=0 \tag{2.6}$$

$$\frac{\partial \ln L}{\partial\theta}=0\Rightarrow\frac{-rp\frac{\partial}{\partial\theta}\bar{F}(\delta,\theta,\sigma_1)}{1-p\bar{F}(\delta,\theta,\sigma_1)}+\sum_{r+1}^{n}\left(\frac{x_i-\theta}{\sigma_1^2}\right)=0 \tag{2.7}$$

and

$$\frac{\partial \ln L}{\partial\sigma_1}=0\Rightarrow\frac{-rp\frac{\partial}{\partial\sigma_1}\bar{F}(\delta,\theta,\sigma_1)}{1-p\bar{F}(\delta,\theta,\sigma_1)}-\left(\frac{n-r}{\sigma_1}\right)+\sum_{r+1}^{n}\frac{(x_i-\theta)^2}{\sigma_1^3}=0 \tag{2.8}$$

Here equations (2.7) and (2.8) may be solved simultaneously. The above equations give reasonably good estimates of the parameters for *δ* fixed.

2.3 Normal with nearly instantaneous failures

Let $F(x)$ and $R(x)=1-F(x)$ denote the cumulative distribution function and the survival function of the mixture, respectively. The component distribution functions and their Survival functions are $F_i(x)$ and $R_i(x)=1-F_i(x)$ respectively, $i=1,2$. The failure rate of a lifetime distribution is defined as $h(x)=\frac{f(x)}{R(x)}$ provided the density exists. Instead of assuming an instant or an early failures to occur at a particular point, as in the original model of Lai et.al. (2007), we now represent this model as a mixture of the generalized Dirac delta function and the 2-parameter normal as opposed to a mixture of a singular distribution with normal. Thus the resulting modification gives rise to a density function:

$$f(x) = p\delta_d(x - x_0) + q\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{1}{2}\left(\frac{x-\theta}{\sigma_1}\right)^2\right) \quad , p+q=1,\ 0<p<1 \qquad (2.9)$$

$$\sigma_1 > 0,\ -\infty < \theta < +\infty$$

Where

$$\delta_d(x - x_0) = \begin{cases} \frac{1}{d}, & x_0 \le x \le x_0 + d \\ 0, & o.w. \end{cases} \qquad (2.10)$$

for sufficiently small *d*. Here *p* is the mixing proportion and $p > 0$. Also note that

$$\delta(x - x_0) = \lim_{d \to 0} \delta_d(x - x_0) \qquad (2.11)$$

where $\delta(.)$ is the Dirac delta function. We may view the Dirac delta function as approximately normal distribution having a zero mean and standard deviation that tends to 1 (see Strichartz (1994) and Li and Wong (2008) for details). For fixed value of *d*, (2.10) denotes a uniform distribution over an interval $[x_0, x_0 + d]$ so the modified model is now effectively a mixture of a normal with a uniform distribution. Instead of including a possible instantaneous failure in the model (2.10) allows for a possible "near instantaneous" failure to occur uniformly over a very small time interval. Note that the case $x_0 = 0$ corresponds to instantaneous failures, whereas $x_0 \neq 0$ (but small) corresponds to the case with early failures. The survival function and failure rate functions can be obtained as follows: Since $f(x) = p\, f_1(x) + q\, f_2(x)$ and $F(x) = p\, F_1(x) + q\, F_2(x)$. We have,

$$R(x) = 1 - F(x) = p + q - pF_1(x) + q\, F_2(x) = pR_1(x) + qR_2(x) \qquad (2.12)$$

and the corresponding failure rate function as

$$h(x) = \frac{pf_1(x) + qf_2(x)}{pR_1(x) + qR_2(x)} \qquad (2.13)$$

where

$$R_1(x) = \begin{cases} 1, & 0 \le x < x_0 \\ \frac{d + x_0 - x}{d}, & x_0 \le x \le x_0 + d \\ 0, & x > x_0 + d \end{cases} \qquad (2.14)$$

and $$R_2(x) = 1 - F_2(x) \qquad x > x_0 + d \qquad (2.15)$$

Similarly, the failure rates for each component is given by

$$h_1(x)=\begin{cases}0, & 0\le x<x_0\\ \dfrac{1}{d+x_0-x}, & x_0\le x\le x_0+d\\ \infty, & x>x_0+d\end{cases} \tag{2.16}$$

and

$$h_2(x)=\frac{\dfrac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\dfrac{1}{2}\left(\dfrac{x-\theta}{\sigma_1}\right)^2\right)}{1-F_2(x)} \tag{2.17}$$

Consider the special case of model (2.9) whereby $x_0=0$. The model may be called the normal with “nearly instantaneous failure” model. In this case, (2.16) can be simplified as

$$h_1(x)=\begin{cases}\dfrac{1}{d-x}, & 0\le x\le d\\ \infty, & x>d\end{cases} \tag{2.18}$$

and (2.14) simplifies to

$$R_1(x)=\begin{cases}\dfrac{d-x}{d}, & 0\le x\le d\\ 0, & x>d\end{cases} \tag{2.19}$$

Thus the normal model with “nearly instantaneous failure” occurring uniformly over [0, d] has

$$R(x)=\begin{cases}\dfrac{p(d-x)}{d}+q\left[1-F_2(x)\right], & 0\le x\le d\\ q\left[1-F_2(x)\right], & x>d\end{cases} \tag{2.20}$$

and

$$h(x)=\begin{cases}\dfrac{p}{p(d-x)+dq\left(1-F_2(x)\right)}+\left[1-\dfrac{dp}{p(d-x)+dq\left(1-F_2(x)\right)}\right]\dfrac{f_2(x)}{R_2(x)}, & o\le x\le d\\ \dfrac{qf_2(x)}{R_2(x)}, & x>d\end{cases} \tag{2.21}$$

respectively. The plots for reliability and failure functions are presented in figures 1 to 3 below.

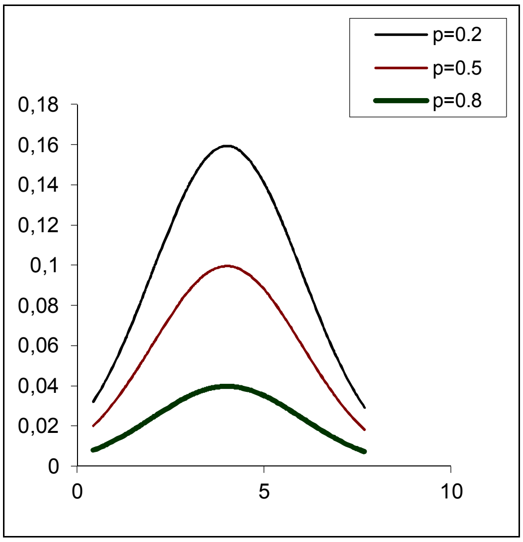


Figure 1. Density function

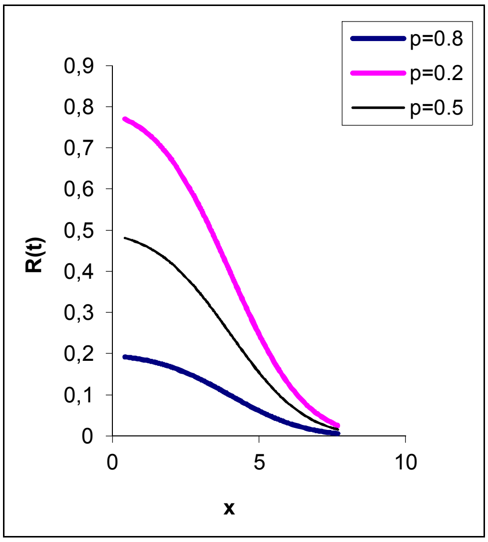


Figure 2. Reliability function

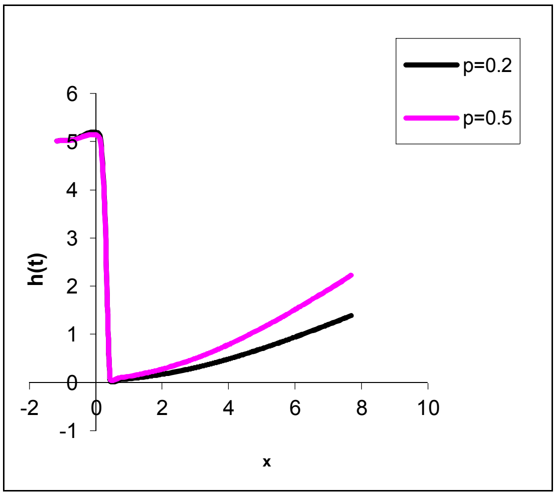


Figure 3. Failure distribution for $\mu = 4$ and $\sigma = 2$

3. An illustrative example

This example is due to Vannman (1991). A batch of wooden boards is dried by a particular chemical process and the object of the experiment is to compare two processes as regards the extent of deformation of boards due to checking. The measure of damage to the board is the checking area x defined as $x = \frac{l\bar{d}}{hl_0}100$, where l is the length of the check, $\bar{d}$ is the mean depth of the check, h is the thickness of the board area and l_0 is the length of the board. Thus x is the check area measured as percentage of the board area. The boards are dried at the same time under different schedule and under some climatic conditions. When drying boards not all of them will get the checks and a typical sample of wood contain several observations with $x_i = 0$ or $x_i > 0$ but relatively small compared to the rest of the checks. These observations will correspond to instantaneous failures or early failures. Note that the larger the number of instantaneous failures better is the process. We below reproduce the data of Schedule 1 and 2 of Experiment 3. The estimates are presented in Table 1.

E-3, S-1: $x_i = 0$, i=1,2,…,13 and the other positive observations arranged in increasing order are 0.08, 0.32, 0.38, 0.46, 0.71, 0.82, 1.15,1.23, 1.40, 3.00, 3.23, 4.03, 4.20, 5.04, 5.36, 6.12, 6.79, 7.90, 8.27, 8.62, 9.50, 10.15, 10.58 and 17.49.

E-3, S-2: $x_i = 0$, i=1,2,…,17 and the other 20 positive observations arranged in increasing are 0.02, 0.02, 0.02 0.04, 0.09, 0.23, 0.26, 0.37, 0.93, 0.94, 1.02, 2.23, 2.79, 3.93, 4.47, 5.12, 5.19, 5.39, 6.83 and 8.22.

Table 1: Estimation for instantaneous failure, early failures and nearly instantaneous

Schedule		Instantaneous	Early failures	Nearly instantaneous
1 (δ=1.5)	$\hat{\theta}$	4.867917	7.352	5.076087
	$\hat{\sigma}_1$	4.398309	3.745867	4.374601
2 (δ=0.9)	$\hat{\theta}$	2.43900	3.919167	3.042500
	$\hat{\sigma}_1$	2.606334	2.390099	2.581076

4. Inliers detection using Information criterion

Denoting the parameter of X by $\alpha_i = \theta,\ i = 1,2,......n$. We consider the following model of no inliers in the Model as

$$\text{Model(0): } \alpha_i = \theta,\ i = 1,2,......n \tag{4.1}$$

and the model with r inliers as

$$\text{Model}(\mathrm{r}) : \alpha_i = \begin{cases} \phi, & 1 \le i < r \\ \theta, & r+1 \le i < n \end{cases} \tag{4.2}$$

where r, $1 \le r \le$ n-1, is the unknown index of the inliers. Model(0) may also be interpreted as having all observations from the target distribution F with common parameter θ.

Suppose that the life times of $X_{(1)}, X_{(2)}, ..., X_{(n)}$ is sequence of independent random variables with normal distribution having unknown mean θ. According to the procedure, the model(0) is selected with no inliers if $SIC(0) < \min_{1 \le r \le n-1} SIC(r)$. And the model(r) is selected if $SIC(0) > \min_{1 \le r \le n-1} SIC(r)$. Here SIC is the Schwartz Information criterion. Thus we have

$$SIC(0) = 2n \log \sigma_1 + \sum_{i=1}^{n} \left(\frac{x_{(i)} - \theta}{\sigma_1} \right)^2 + p \log n \tag{4.3}$$

and

$$SIC(r) = 2r\log\sigma_0 + 2(n-r)\log\sigma_1 + \sum_{i=1}^{r}\left(\frac{x_{(i)}-\phi}{\sigma_0}\right)^2 + \sum_{i=r+1}^{n}\left(\frac{x_{(i)}-\theta}{\sigma_1}\right)^2 + p\log n \quad (4.4)$$

The estimate of inliers say r is such that $SIC(r) = \min_{1\le k\le n} SIC(k)$. The above procedure is implemented through other information criteria's like the Bayesian Information criterion: $BIC = -\ln L(\Theta) + \frac{0.5\mathrm{p}\ln(n)}{n}$ and the Hannan-Quinn criterion given by: HQ= $-\ln L(\Theta)+$ p $\ln[\ln(n)]$, where $L(\Theta)$ the maximum likelihood function and p is the number of free parameters that need to be estimated under the model. The method is illustrated through numerical examples in the later sections.

5. Testing of hypothesis

Here we are interested to test the hypothesis that, whether sample observations belong to inliers population from $N(\phi,\sigma_0^2)$ against the hypothesis that it belongs to target population from $N(\theta,\sigma_1^2)$, assuming $\sigma = \sigma_0 = \sigma_1$. Equivalently, the hypothesis can be written as H_0: $\mu = \phi$ versus H_1: $\mu = \theta$. Below we discuss two computationally simple test procedures to detect inliers in a model.

5.1 Modified likelihood ratio test

The study of the modified likelihood approach to finite normal mixture models with a common and unknown variance in the mixing components and a test of the hypothesis of a homogeneous model versus a mixture on two or more components were done by Chen and Kalbfleisch (2005).

We define $M_1 : \{F(x)/x \sim N(\theta,\sigma^2)\}$, That is, all observations come from target population and $M_2 : \{F(x) = (1-p)F_1(x) + pF_2(x)\}$, That is, the observations comes from a mixture of two normal distributions, with $F_1(x)$ and $F_2(x)$ are distribution functions of inliers and target populations respectively, as defined in previous sections.

We want to test null hypothesis $H_0 : p = 1$ against $H_0 : p < 1$ or in other words a test of the hypothesis $X \in M_1$ versus $X \in M_2$ then ordinary LRT statistics is given by

$$\ln\lambda = 2\left[\sup_{\theta,X\in M_2} \ln(\phi,\theta,X) - \sup_{\phi,X\in M_1} \ln(\theta,X)\right] \quad (5.1)$$

Due to non-regularity of the finite mixture models ln λ does not have usual chi-squared distribution. Therefore, we modify the likelihood as

$$m\ln(\phi,\theta,X) = \ln(\phi,\theta,X) + C\ln\{4p(1-p)\} \quad (5.2)$$

where C is a positive constant. The purpose of the penalty term $C\ln\{4p(1-p)\}$ is to restore regularity to the problem by avoiding estimates of p on or near the boundary. Let $\ln\left(\hat{\theta}, X\right)$ maximizes $m\ln(\theta, X)$ for $X \in M_1$ and $\ln\left(\hat{\phi}, \hat{\theta}, X\right)$ maximizes $m\ln(\phi, \theta, X)$ for $X \in M_2$. Then the modified likelihood ratio statistic is

$$\ln \hat{\lambda} = 2\left[\ln\left(\hat{\phi}, \hat{\theta}, X\right) - \ln\left(\hat{\theta}, X\right)\right] \tag{5.3}$$

The null hypothesis is rejected for large values of ln €, where ln € follows $\chi^2_{(2)}$ distribution.

5.3. Most powerful test

The most powerful test for testing $H_0 : \mu = \phi$ against $H_0 : \mu = \theta$ where μ is the mean of normal population and p known is given by

$$\psi(x) = \begin{cases} 1, & \dfrac{P_1(x)}{P_0(x)} > C\alpha \\ 0, & \dfrac{P_1(x)}{P_0(x)} < C\alpha \end{cases} \tag{5.4}$$

where $P_1(x)$ and $P_0(x)$ are likelihood functions under distribution of target population $\Im$ and inlier population G respectively, and $C\alpha$ is such that $P_{H_0}[\psi(x)] = \alpha$, where α is the level of significance. We reject H_0 for large values of the ratio $\dfrac{P_1(x)}{P_0(x)}$. Also, the value of $C\alpha$ is obtained as $C_\alpha = \phi + \sigma z_\alpha$, after some numerical computation.

6. Simulation Study

To illustrate the method of identifying inliers model we have generated 15 independent random samples, where 5 of them are from normal distribution with mean $\phi = 4$ and $\sigma_0{}^2 = 2$, and remaining ten observations from normal distribution with parameter mean $\theta = 20$ and $\sigma_1{}^2 = 3$. The observations are 1.44852, 3.667636, 3.949972, 5.548854, 6.017887, 17.61194, 19.26654, 20.09814, 20.23482, 20.36071, 20.64048, 21.08915, 21.26954, 22.53701 and 24.23439.

The identification is done as follows: Evaluate for each fixed r the maximum likelihood equation €$_r$, and then consider € being that value of r for which likelihood is maximum. The estimates are presented in table 2. It is interesting to note that the likelihood is maximum corresponds to *r=5*, which is expected. The corresponding estimates of the parameters are $\hat{\phi}$ = 4.126574, σ_0=1.80372 and $\hat{\theta}$ =20.73427, σ_1= 1.783219.

Table 2: The Likelihood and Information criterions

r	*L*	*SIC*	*BIC*	*HQ*
2	-38.1951	69.8294	-3.4621	-2.6464
3	-34.5019	62.4430	-3.3604	-2.5447
4	-31.2064	55.8519	-3.2600	-2.4443
5	**-20.7104**	**34.8599**	**-2.8501**	**-2.0344**
6	-26.0540	45.5470	-3.0796	-2.2639
7	-28.5460	50.5312	-3.1709	-2.3552
8	-30.9970	55.4332	-3.2533	-2.4376
9	-33.0941	59.6274	-3.3188	-2.5031
10	-34.9391	63.3174	-3.3730	-2.5573
11	-36.6837	66.8065	-3.4218	-2.6061
12	-38.4748	70.3887	-3.4694	-2.6537
13	-39.6796	72.7984	-3.5003	-2.6846

Clearly *SIC*(0) = 58.4562 > $SIC(5) = \min_{1 \le r \le n} SIC(r) = 34.85999$. A similar conclusion can be drawn in the case of *BIC* and *HQ*. Next, we carried out an experiment with 1000 samples each of size 15 and number of inliers as 3, 4, 5 and 6 each with $\phi = 3$ and $\theta = 6, 9, 12, 15$. The table 3 entitled power of *SIC* procedure presents the number of times the SIC procedure correctly identified the number of inliers as proportion to total number of samples. The values clearly indicate the effectiveness of the method in detecting the inliers.

Table 3. Power of SIC procedure

θ / ϕ \ r	2	3	4	5
3	0.570	0.720	0.700	0.550
4	0.460	0.480	0.490	0.440
5	0.460	0.460	0.460	0.462
6	0.410	0.420	0.430	0.410

6.1. Numerical Example

We recall the Vannman (1991) data example discussed in section 3 to illustrate the identification of inliers using information criterions. The computed value *SIC*(0) = 99.45467 and below in Table 4, the value of likelihood, *SIC*(*r*) and modified likelihood ratio for different values of *r* are given for different information criterions.

Clearly, *SIC*(0) = 99.45467 > *SIC*(9) = min *SIC*(*r*) = 53.87482. Also the likelihood is maximum for *r = 9*. The corresponding estimates of the parameter are $\hat{\phi}$ = 0.727778, σ_0 =

0.456858 and $\hat{\theta}$ =7.352, σ_1= 3.745867. For modified likelihood ratio test also the maximum $\ln\hat{\lambda}$ is attained at *r = 9*.

One of the important problems while detecting the inliers is the masking effect, where masking effect is defined as the loss of power due to wrong detection of more than one inliers. This is discussed in the next section.

Table 4. Estimates of parameters for various values of *r*.

r	Likelihood	*SIC*	*BIC*	*HQ*	$\ln\hat{\lambda}$
2	-39.7964	85.94886	-3.55136	-2.52751	13.50582
3	-36.1743	78.70478	-3.45593	-2.43208	20.74989
4	-32.7564	71.86888	-3.35668	-2.33283	27.5858
5	-30.8971	68.15026	-3.29824	-2.27439	31.30442
6	-28.6342	63.62454	-3.22218	-2.19833	35.83014
7	-27.5317	61.41942	-3.18292	-2.15907	38.03526
8	-25.6430	57.64209	-3.11185	-2.08800	41.81259
9	**-23.7594**	**53.87482**	**-3.03556**	**-2.01171**	**45.57985**
10	-27.4743	61.30473	-3.18083	-2.15698	38.14995
11	-28.1648	62.68569	-3.20565	-2.18180	36.76899
12	-29.3104	64.97688	-3.24552	-2.22167	34.47779
13	-29.6057	65.56758	-3.25555	-2.23170	33.88709
14	-30.5163	67.38864	-3.28584	-2.26199	32.06603
15	-31.1017	68.55955	-3.30484	-2.28099	30.89513
16	-32.0722	70.50050	-3.33557	-2.31172	28.95417
17	-33.2247	72.80552	-3.37087	-2.34702	26.64915
18	-35.0261	76.40824	-3.42367	-2.39982	23.04643
19	-36.5309	79.41796	-3.46574	-2.44189	20.03672
20	-37.8073	81.97070	-3.50008	-2.47623	17.48397
21	-39.3469	85.04991	-3.54000	-2.51615	14.40476
22	-40.8648	88.08568	-3.57785	-2.55400	11.369

7. Masking effect on tests for inlier(s)

Suppose $X_1, X_2 ....X_n$ be sequence of *n* independent random variables with some known FTD. Under the null hypothesis H_0 these random variables are identically distributed with df *F* whereas under alternative hypothesis H_1, discordant observations (inliers) arise from population df *G*. The df of *G* is assumed to be of same form as that of *F* with a change in location or scale parameter by an unknown quantity λ. This parameter is called discordancy parameter, measuring the degree of discordancy. Under H_1 it is assumed one of the observations follows df *G*. Let T(x) be a test statistics to detect a single discordant observation with critical region *A(n,α)*. Due to lack of information about the number of discordant observations present in the sample, however, the true situation may not be specified by H_1 and more than one discordant observation may be present in the sample. In such cases a test statistics T(x) suggested for detection of a single discordant, may fail to detect a single inlier as discordant even when additional discordant observations are present in the sample. Such a phenomenon is called masking effect.

All tests for detecting a single inlier, H_o against H_1 are based on symmetric functions of observations or on functions of order statistics. In the k-inlier model, the joint distribution of order statistics $X_{(1)}, X_{(2)}, \ldots X_{(n)}$ is same as that under the exchangeable model introduced by Kale (1975) where it is assumed that any set $X_{i_1}, X_{i_2}, \ldots X_{i_k}$ has priori equal probability of being independent and identically distributed as G_λ and the remaining (*n-k*) observation are distributed as *F*, the distribution function of target population.

In exchangeable model $X_{(1)}, X_{(2)}, \ldots X_{(n)}$ has minimum posterior probability of coming from G_λ such that $\frac{\partial G_\lambda}{\partial F}$ is the decreasing function in X. The limiting masking effect (Bendre and Kale 1987) can be studied by assuming $X_{(1)}, X_{(2)}, \ldots X_{(k)}$ correspond to observation coming from $N\left(\mu-\lambda\sigma,\sigma^2\right)$ and then taking limit as $\lambda \to \infty$. In the above condition, the joint probability is defined as

$$h\left(x_{(1)}, x_{(2)} \ldots x_{(n)}\right) = \frac{k!(n-k)!}{\varphi_\lambda\left(1,2,3\ldots\ldots k\right)} \prod_{i=1}^{k} g_\lambda\left(x_i\right) \prod_{i=k+1}^{n} f\left(x_i\right) , \quad -\infty < x_{(1)} < x_{(2)} \ldots < x_{(n)} < \infty \tag{7.1}$$

Also f and g_λ are probability density functions of $N\left(\mu,\sigma^2\right)$ and $N\left(\mu-\lambda\sigma,\sigma^2\right)$ respectively. Thus masking effect on any test statistics *T(x)* with critical region *A(n,α)*, we have

$$\lim_{\lambda\to\infty} P\left[T\left(x\right) \in A\left(n,\alpha\right) / L_{sk}\right] = \lim_{\lambda\to\infty} \int_{A(n,\alpha)} h\left(x_{(1)}, x_{(2)}, \ldots\ldots x_{(n)}\right) dx_{(1)} \ldots\ldots dx_{(n)} \tag{7.2}$$

Thus under the labeled slippage model, L_{sk} as $\lambda \to \infty$, $x_{(n-k+1)}, x_{(n-k+2)}, \ldots\ldots x_{(n)}$ behave as order statistics of a sample of size (*n-k*) from $N\left(\mu,\sigma^2\right)$ and $x_{(1)}, x_{(2)}, \ldots\ldots x_{(k)}$ diverge to zero. However if $T\left(x_{(1)}, x_{(2)}, \ldots\ldots x_{(k)}\right)$ is a function whose distribution does not depend on λ then T converges in distribution to a proper random variable as $\lambda \to \infty$.

7.1 Limiting masking effect

In line with Grubb's test, for a single inlier, we propose the test

$$G = \frac{\sum_{i=2}^{n}\left(x_{(i)} - \bar{x}_n\right)^2}{\sum_{i=1}^{n}\left(x_{(i)} - \bar{x}\right)^2} \quad where \ \bar{x}_n = \frac{\sum_{i=2}^{n} x_{(i)}}{n-1} \quad and \ \bar{x} = \frac{\sum_{i=1}^{n} x_{(i)}}{n} \tag{7.3}$$

and the maximum studentized residual *T* as

$$T = \frac{\frac{(n-1)}{n}}{\left[\sum \frac{\left(x_{(i)} - \bar{x}_{(n)}\right)^2}{\left(x_{(n)} - \bar{x}_{(1)}\right)^2} + \frac{(n-1)}{n}\right]^{\frac{1}{2}}} \tag{7.4}$$

Since under L_{s1} corresponds to the inlier observation coming from $N\left(\mu - \lambda\sigma, \sigma^2\right)$ and $\frac{\left(x_{(i)} - \bar{x}_{(n)}\right)^2}{\left(x_{(1)} - \bar{x}_{(n)}\right)^2} \to 0$ in probability as $\lambda \to \infty$ for $i = 2,3,4 \ldots\ldots n$ and therefore $T \to \left[\frac{n-1}{n}\right]^{\frac{1}{2}}$ in probability as $\lambda \to \infty$. Hence as $\lambda \to \infty$, $\lim P_1^G(\lambda) = 1$, where $P_1^G(\lambda)$ is the power function of Grubb's test. To study $\lim P_2^G(\lambda) = \lim P\left[T < t_{n,\alpha} \mid L_{sk}\right]$ *as* $\lambda \to \infty$ we write

$$T = \frac{Y_{(1)} - \frac{k}{n}}{\left[\sum Y_{(1)}^2 - 2k \frac{\sum Y_{(i)}}{n} + \frac{k^2}{n}\right]^{\frac{1}{2}}} \tag{7.5}$$

where

$$Y_{(i)} = \frac{\left(x_{(i)} - \bar{x}_{(k)}\right)}{\left(\bar{x}'_{(n-k+1)} - \bar{x}_{(k)}\right)} \qquad i = 1, 2, \ldots\ldots n \tag{7.6}$$

With $\bar{x}'_{(n-k+1)}$ is the mean of $x_{(k+1)}, x_{(k+2)}, \ldots\ldots x_{(n)}$ and $\bar{x}_k$ is the mean of $x_{(1)}, x_{(2)}, \ldots\ldots x_{(k)}$. Therefore $Y_{(i)} \to 0$ in probability for $i = 1,2 \ldots\ldots k$ because the numerator of $Y_{(i)}$ is a proper random variable, while denominator diverges to infinity. For $i = 1,2,\ldots\ldots k$, we observe that $Y_{(i)} - 1 = \frac{\left(x_{(i)} - \bar{x}'_{(n-k+1)}\right)}{\left(\bar{x}_{(n-k+1)} - \bar{x}_{(k)}\right)}$ is such that the numerator has a distribution independent of λ and therefore converges to a proper random variable, but denominator diverges to infinity and hence $Y_{(i)} \to 1$ in probability as $\lambda \to \infty$.

Therefore under L_{sk} as $\lambda \to \infty$, $T \to \left[\frac{(n-k)}{nk}\right]^{\frac{1}{2}}$ and

$$\lim P_2^G(\lambda) = \begin{cases} 1, & \left[\frac{(n-k)}{nk}\right]^{\frac{1}{2}} < t_{n,\alpha} \\ 0 & o.w. \end{cases} \tag{7.7}$$

Thus Grubb's test is free from the limiting masking effect for $\left[\frac{(n-k)}{nk}\right]^{\frac{1}{2}} \geq t_{n,\alpha}$ and the performance of the test depends on the sample size n and the number of inliers. In general $t_{n,\alpha}$ is a decreasing function of the sample size and hence for large *n* with moderate *k* the test is free from the limiting masking effect. Table 5, presents the maximum number of inliers in a sample of size *n* up to which Grubb's test is free from the limiting masking effect.

Table. 5 Maximum inliers accommodated by Grubb's test

α	n =10	n = 15	n = 20	n = 25
0.01	1	1	1	2
0.05	1	2	2	2
0.10	1	2	2	3

From the table, it is observed that for large sample size more number of inliers may be accommodated.

References

Akaike, H. (1974). A new look at the Statistical identification model. *IEEE Trans. Auto. Control,* 19, 716-723.

Bendre S.M. and Kale B.K. (1987). Masking effect on test for outliers in normal samples, *Biometrica,*74(4), 891-896.

Chen, J., Kalbfleisch J.D.(2005). Modified likelihood ratio test in finite mixture models with a structural parameter, *Journal of Statistical Planning and Inference* 129, 93-107.

Lai, C. D., Khoo, B. C., Muralidharan, K. and Xie, M. (2007). Weibull model allowing nearly instantaneous failures. *J. Applied Mathematics and Decision Sciences* Article ID 90842, 11 pages.

Kale, B. K. (1975). Trimmed means and the method of maximum likelihood. *Applied Statistical Publishing House, Amsterdam,* 177-185.

Kale, B. K. and Muralidharan, K. (2007). Masking effect of inliers. *J. Indian Statistical Association,* 45(1), 33-49.

Kale, B. K. and Muralidharan, K. (2008). Maximum Likelihood estimation in presence of inliers. *Journal of Indian Society for Probability and Statistics*, 10, 65-80.

Li, Y. T., and Wong, R. (2008). Integral and series representations of the Dirac delta function, *Commu. Pure Appl. Analysis*. 7 (2): 229–247.

Muralidharan, K and Lathika, P. (2004). The concept of inliers. Proceedings of the First Sino-International Symposium on Probability, Statistics and Quantitative Management., Taiwan, October, 77-92.

Muralidharan K. and Arti M. (2008). Analysis of instantaneous and early failures in Pareto distribution, *Journal of statistical theory and Applications,* Vol 7, 187-204.

Strichartz, R. (1994). *A Guide to Distribution Theory and Fourier Transforms*, CRC Press, ISBN 0-8493-8273-4.

Titterington, D.M., Smith, A., Makov, U.E.(1985). *Statistical Analysis of finite mixture distribution.* John Wiley and Sons, New York.

Vannman. K. (1991). Comparing samples from nonstandard mixtures of distributions with Applications to quality comparison of wood. Research report 1991:2 submitted to Division of Quality Technology, Lulea University, Lulea, Sweden.

APPLICATION OF CHEBYSHEV- AND MARKOV-TYPE INEQUALITIES IN STRUCTURAL ENGINEERING

K. Balaji Rao[1]. M.B. Anoop[2]. Nagesh R. Iyer[3]

•

[1]Chief Scientist, Head, Risk and Reliability of Structures Group, [2]Senior Principal Scientist, [3]Director
(CSIR-Structural Engineering Research Centre, Taramani, Chennai, Tamil Nadu, India)

e-mail: balajiserc@gmail.com

ABSTRACT

This paper aims at bringing out the usefulness of Chebyshev- and Markov- type inequalities in structural engineering design decision making. By examining whether the bounds arising from Chebyshev - type inequality (associated with these are weak upper bound probabilities) enclose the respective experimental values for deflections of six ferrocement I-beams and web shear fatigue life of a steel plate girder it is inferred that the bounds and the associated probabilities estimated are realistic and hence can be used in structural engineering design decision making. The paper also presents some recent developments in application of Markov type inequalities (which are due to Steliga and Szynal (2010)) for estimation of bounds on probability of an event sought. The importance of such bounds in structural engineering applications is brought out. It is shown from the results of Monte Carlo simulation that the bounds on probability of an event, estimated using the method presented by Steliga and Szynal, are sharp. One of the important advantages of the bounds presented by Steliga and Szynal (2010) is that the original (hidden/internal) random variable need not have well defined moments. Possible engineering applications are also pointed out.

Keywords: Chebyshev inequality, Markov inequality, deflection, fatigue life

1 CHEBYSHEV INEQUALITY- SOME PRELIMINARIES

Let X be a random variable representing an action or response quantity. Example of action quantity can be load (or loading intensity), external bending moment or external traction force. The response quantity can be deflection, rotation, warping, strain, crack width. In most engineering applications we may not be knowing the actual probability density function (pdf) of X; yet, we will be asked to answer questions like $P[g(x) \geq r]$=?. It may be noted that $g(X)$ is a function of random variable and r is a specified value. Such decision making probabilities are required in limit state design of structural components (viz. Bolotin, 1969).

In the face of non-availability of pdf of X can we make probabilistic inferences about $P[g(X) \geq r]$. It can be shown that (viz. Gnedenko, 1976)

$$P[g(X) \geq r] \leq \frac{E[g(X)]}{r} \tag{1}$$

Let $g(x) = [x - \mu]^2$; $r = z^2\sigma^2$ where μ and σ are mean and standard deviation of X. According to Chebyschev's inequality this probability computed from (Gnedenko, 1976)

$$P[(X-\mu)^2 \geq z^2\sigma^2] \leq \frac{E[(X-\mu)^2]}{z^2\sigma^2}$$
$$\text{(or) } P[(X-\mu)^2 \geq z^2\sigma^2] \leq \frac{1}{z^2} \quad (2)$$

Therefore, Chebyshev's inequality gives weak upper bound on the desired probability. We note from Eq.(2) that z should be greater than or equal to 1 since if $z < 1$ (though positive) the interpretation as the value of probability will not be proper. Also, since we are not having any information on type of pdf of X the bounds will be weak. Efforts have been made in the literature to sharpen the bounds and to determine two-sided bounds and also, to determine bounds for multivariate case (with- and without- correlation effect).

While Eq. (2) gives the one-sided bound, let us consider two-sided case. When the distribution is symmetrical about the mean, the symmetrical bounds around the mean are given by (Steliga and Szynal, 2010),

$$P[k_1 < \times < k_2] \geq 1 - \frac{4\sigma^2}{(k_2-k_1)^2} \quad \forall(k_2-k_1) > 2\sigma \quad (3)$$

When the pdf of the random variable is not symmetrical about the mean, the bounds are given by,

$$P[k_1 < \times < k_2] \geq \frac{4[(\mu-k_1)(k_2-\mu)-\sigma^2]}{(k_2-k_1)^2} \quad (4)$$

Where $k_1 < \mu < k_2; \quad (\mu-k_1)(k_2-\mu) > \sigma^2$

2 APPLICATIONS

In this section two example problems demonstrating the use of Chebyshev inequalities in determining the weak upper bound probabilities, those required for engineering decision making, are presented. One of the highlights of these examples is, to infuse confidence in engineering applications, to compare the results with the respective experimental values.

Example 1: In this example an attempt has been made to estimate the weak upper bound probabilities on random central deflection of ferrocement I-beams used for roofing in low-cost housing. This example is considered since the test data on central deflection, at different stages of loading, was available for six specimens. These specimens were tested at the structural engineering laboratory of Indian Institute of Science, Bangalore, in 1980s. The details of tests and the test results are available in (Prakash Desayi and Balaji Rao (1988), Prakash Desayi and Balaji Rao (1993), Balaji Rao (1990)). Also, an effort was made to determine statistical properties of deflections using Monte Carlo simulation technique. More details about basic random variables considered and details of simulation are presented in Balaji Rao (1990). The final results of simulation (viz. mean and standard deviations of deflection) for six specimens considered here, at different stages of loading, are presented in Table 1. Also presented in this table are experimental values of central deflections. The weak upper bound probabilities associated with bounds of lengths 2.25σ, 2.5σ, 2.75σ and 3σ are computed using Chebyshev inequalities. These probabilities are computed for two conditions : (a) assuming that the pdf of deflection, at different stages of loading, are symmetrical about the mean, and, (b) assuming that the pdf of deflection is unknown or unsymmetrical about the mean. The values of the bounds and their corresponding probabilities are presented in Table 1 typically for first two interval lengths. Since a bound of length 3σ is very often used in engineering decision making, the same are compared for the cases (a) and (b) in Figs.

1 – 6 for the specimens considered. Also shown in these figures are the experimentally observed deflections. From these figures it is observed that at almost all stages of loading, the estimated bounds contain the observed deflection suggesting that the estimated weak upper bound probabilities are acceptable and can be used in engineering decision making. If it is felt that the length of interval of 3σ is high, from Table 1 it is noted that, at higher stages of loading, even though the lengths of interval are small, the bounds enclose the experimentally observed deflections. It may be noted that the weak upper bound probabilities for the two intervals presented in Table 1 vary between 20 to 21% (which is small though). *These observations suggest that the Chebyshev's inequalities can be used for engineering decision making.*

Table 1. Bounds of different interval lengths and their comparison with experimental results

Specimen designation	Applied load (kN)	Exp. deflection (mm)	Results of Monte Carlo simulation, Balaji Rao (1990)		Bounds – symmetrical (length of interval =2.25 σ)[1]		Bounds – unsymmetrical (length of interval = 2.25σ)[2]	
			Mean, μ (mm)	Standard deviation, σ (mm)	Lower bound (μ – 1.125σ) (mm)	Upper bound (μ + 1.125σ) (mm)	Lower bound (μ – σ) (mm)	Upper bound (μ + 1.25σ) (mm)
MI1	2	0.038	0.015	0.004	0.011	0.019	0.012	0.019
	4	0.069	0.033	0.020	0.010	0.056	0.012	0.058
	8	0.216	0.317	0.157	0.140	0.493	0.160	0.513
	10	0.407	0.555	0.184	0.348	0.762	0.371	0.785
	15	0.9	1.154	0.246	0.878	1.431	0.908	1.462
MI2	1	0.035	0.024	0.005	0.018	0.030	0.018	0.035
	1.5	0.075	0.036	0.009	0.026	0.046	0.027	0.054
	8	1.273	1.542	0.378	1.117	1.967	1.164	2.297
	10.41	1.965	2.336	0.460	1.818	2.854	1.876	3.257
	13	2.809	3.190	0.549	2.370	4.013	2.641	4.287
MI3	2	0.021	0.015	0.003	0.011	0.019	0.011	0.020
	4.5	0.065	0.035	0.018	0.015	0.055	0.018	0.058
	10	0.261	0.422	0.204	0.192	0.651	0.217	0.677
	15	0.728	1.047	0.270	0.743	1.350	0.777	1.384
	17.62	1.02	1.375	0.304	1.033	1.717	1.071	1.755
MI4	1	0.036	0.020	0.005	0.015	0.025	0.015	0.026
	1.5	0.05	0.030	0.007	0.023	0.038	0.023	0.039
	8	1.4	1.083	0.363	0.675	1.491	0.720	1.536
	9.9	1.45	1.645	0.419	1.173	2.116	1.225	2.169
	10	1.63	1.673	0.423	1.196	2.148	1.249	2.201
MI5	3	0.04	0.020	0.005	0.015	0.025	0.016	0.026
	4	0.062	0.027	0.007	0.012	0.034	0.020	0.035
	15	0.645	0.699	0.286	0.378	1.021	0.413	1.057
	16	0.701	0.820	0.299	0.483	1.156	0.520	1.193
	17.79	0.825	1.035	0.321	0.674	1.396	0.714	1.436
MI6	2	0.077	0.035	0.008	0.026	0.045	0.027	0.045
	3	0.12	0.053	0.013	0.038	0.068	0.040	0.070
	9	0.73	0.731	0.478	0.193	1.268	0.253	1.328
	9.6	0.932	0.915	0.516	0.334	1.496	0.399	1.560
	14	2.8	2.356	0.678	1.593	3.119	1.677	3.204

Note: 1,2 – associated weak upper bound probabilities are 0.21 and 0.20, respectively

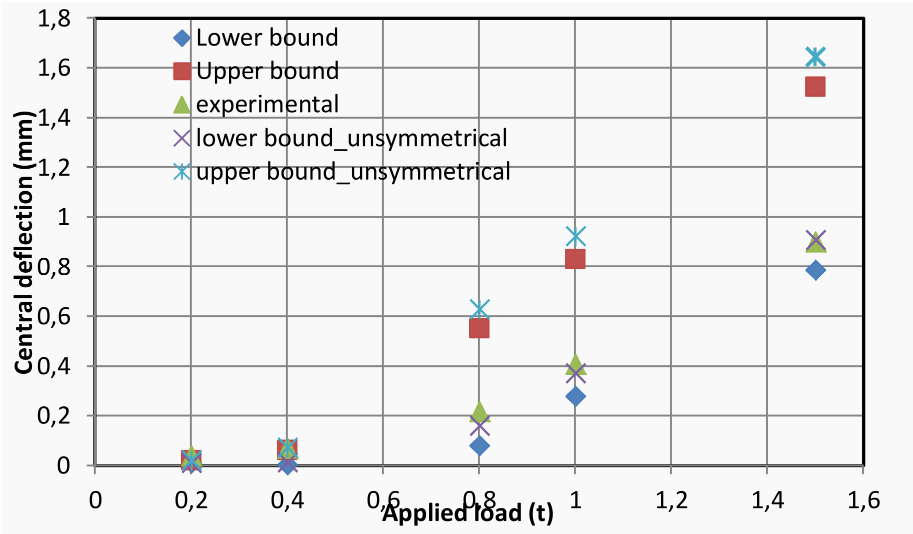


Figure 1. Load versus deflection plot for specimen MI1 with (μ -1.5σ, μ +1.5 σ) symmetrical bounds (associated minimum probability 0.56) and (μ - σ, μ + 2.0 σ) for unsymmetrical bounds (associated minimum probability 0.44) with experimental values

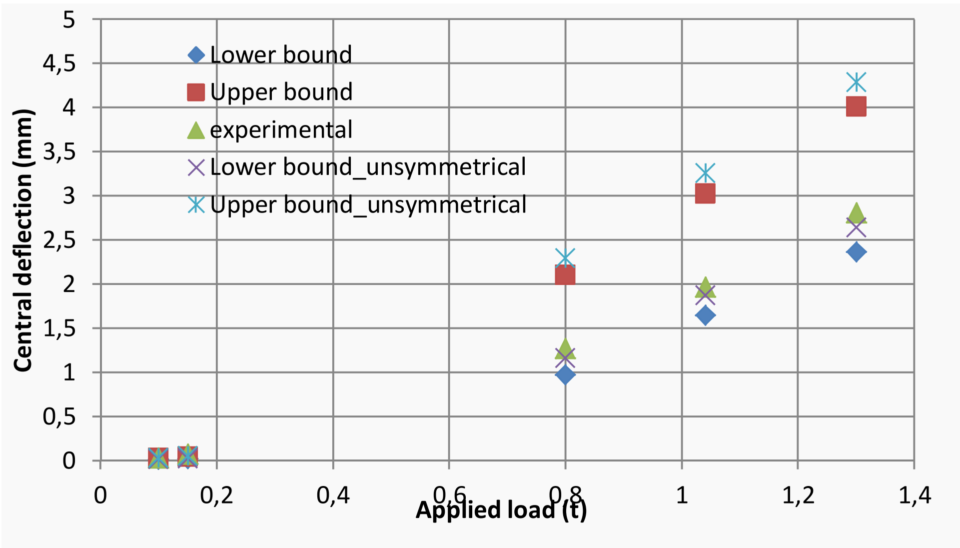


Figure 2. Load versus deflection plot for specimen MI2 with (μ -1.5σ, μ +1.5 σ) symmetrical bounds (associated minimum probability 0.56) and (μ - σ, μ + 2.0 σ) for unsymmetrical bounds (associated minimum probability 0.44) with experimental values

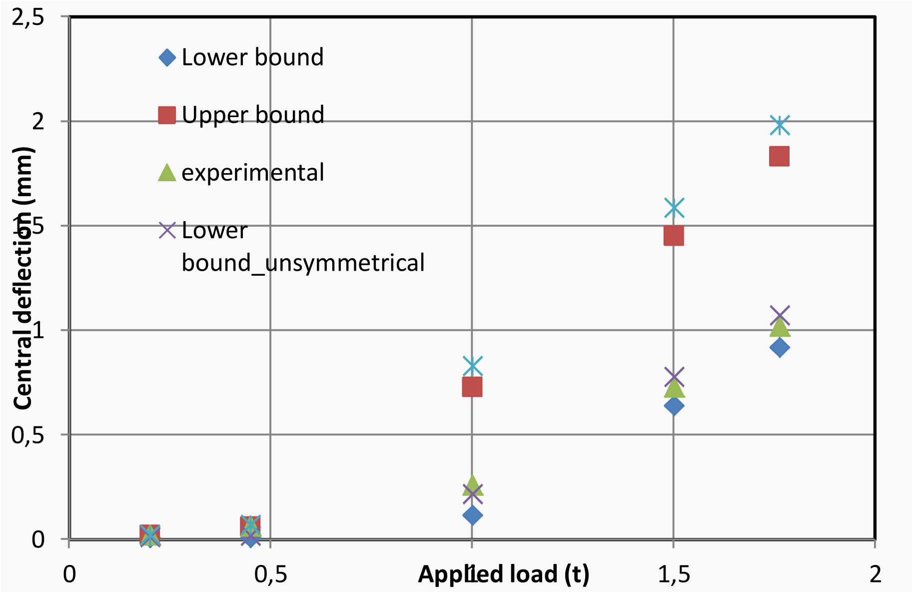


Figure 3. Load versus deflection plot for specimen MI3 with (μ -1.5σ, μ +1.5 σ) symmetrical bounds (associated minimum probability 0.56) and (μ - σ, μ + 2.0 σ) for unsymmetrical bounds (associated minimum probability 0.44) with experimental values

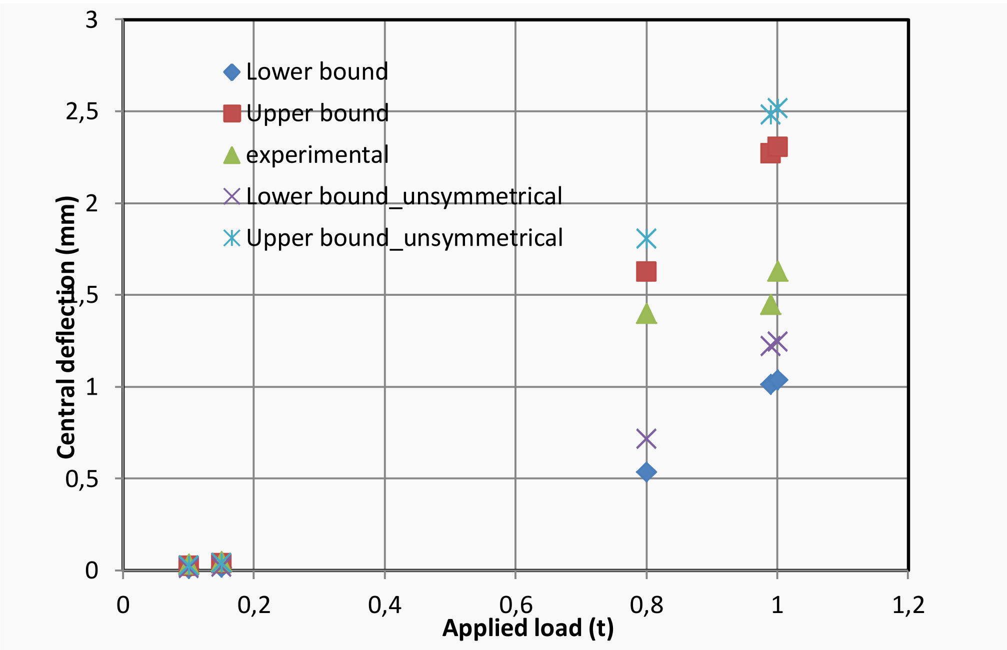


Figure 4. Load versus deflection plot for specimen MI4 with (μ -1.5σ, μ +1.5 σ) symmetrical bounds (associated minimum probability 0.56) and (μ - σ, μ + 2.0 σ) for unsymmetrical bounds (associated minimum probability 0.44) with experimental values

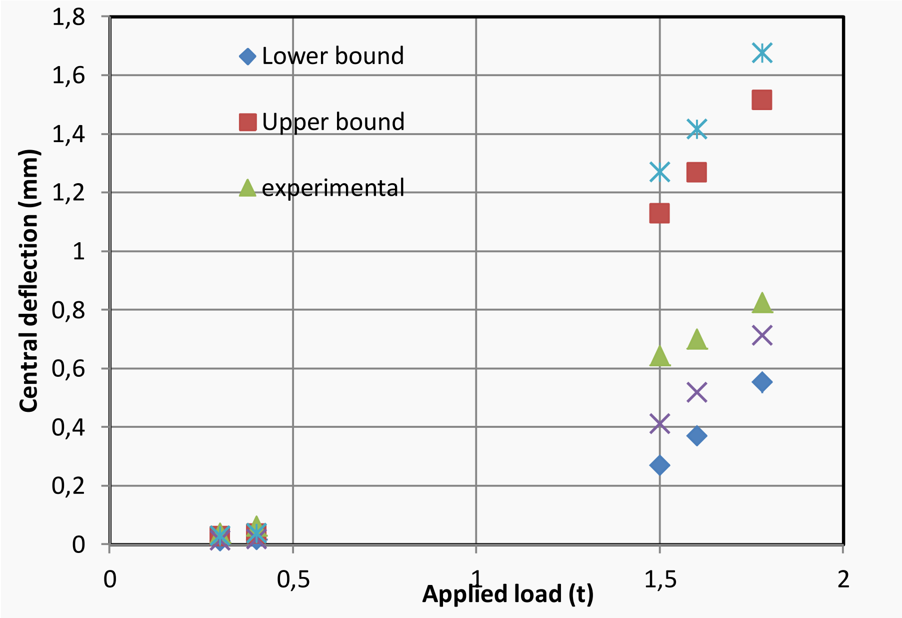


Figure 5. Load versus deflection plot for specimen MI5 with (μ -1.5σ, μ +1.5 σ) symmetrical bounds (associated minimum probability 0.56) and (μ - σ, μ + 2.0 σ) for unsymmetrical bounds (associated minimum probability 0.44) with experimental values

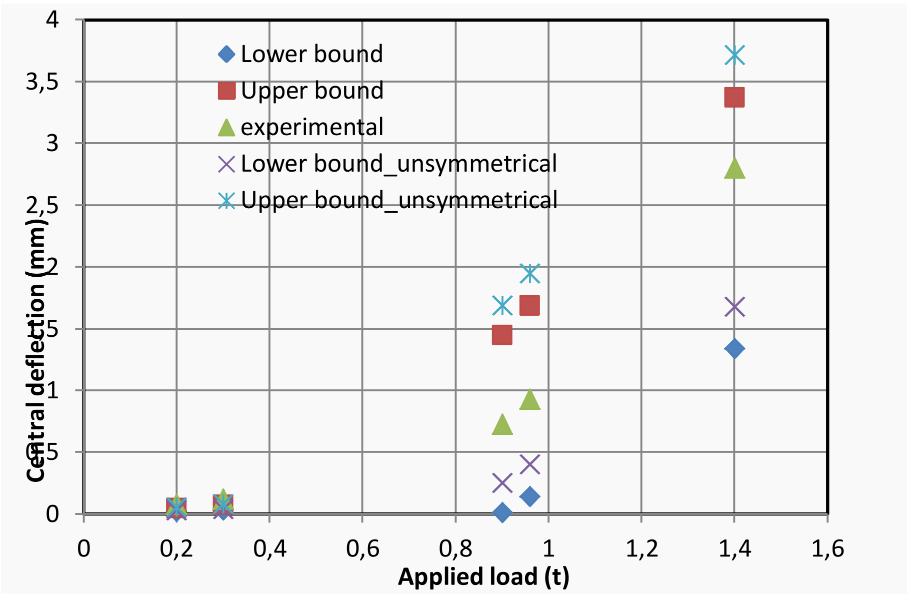


Figure 6. Load versus deflection plot for specimen MI6 with (μ -1.5σ, μ +1.5 σ) symmetrical bounds (associated minimum probability 0.56) and (μ - σ, μ + 2.0 σ) for unsymmetrical bounds (associated minimum probability 0.44) with experimental values

Example 2: This example shows how Chebyshev's inequality will help in fatigue resistant design of steel plate girders of a plate girder bridge. More details of this problem can be found in Balaji Rao and Anoop (2013).

The basic equation used in predicting the fatigue life using S – N approach is given by,

$$N_f = C\left(\Delta\sigma\right)^{-b} \tag{5}$$

where N_f is the number of load cycles to the fatigue limit and $\Delta\sigma$ is the applied stress range, C and b are the material parameters, known as the fatigue strength coefficient and the fatigue strength exponent, respectively.

It is known that the number of cycles to failure (i.e. fatigue life, N_f), at a given applied stress range is a random variable. A typical plot showing the same is presented in Fig. 7. It may be noted that the nature of pdf and the statistical properties of N_f may depend on stress range. While it is desirable to establish the nature of these probability distributions using fatigue tests, it is expensive and time consuming. The median and the 5% and 95% fractiles of fatigue life computed using the transformation of variable technique, at different applied stress ranges are shown in the figure. More details of the probabilistic analysis of the fatigue life of the plate girder are presented in Balaji Rao et.al. (2013). Also shown in this figure are experimental fatigue lives reported in literature. Except in few cases, experimental scatter is enclosed by the estimated bounds. Let us apply the Chebyshev's inequality to determine the bounds on fatigue life. At the applied stress range of 270 MPa the mean (μ) and standard deviation (σ) of N_f are respectively 7.471E+05 and 3.447E+05. Assuming the bounds to be symmetrical and $(k_2\text{-}k_1) = 3\sigma$, the probability that the fatigue life will be between (2.3005E+05, 12.6415E+05) is equal to or greater than 0.556. On the other hand if the distribution is unsymmetrical about the mean, even though we may keep $(k_2\text{-}k_1) = 3\sigma$, assuming the bounds to be (4.024E+05, 14.365E+05), that satisfies the conditions associated with Eq. (4), then the probability that the bound will contain actual life will be equal to greater than 0.444. This value of probability is less than the case when the bounds are symmetrical for the reason that in order to assume that the bounds are symmetrical we should have had more justification/confidence and this gets embedded in our predictions.

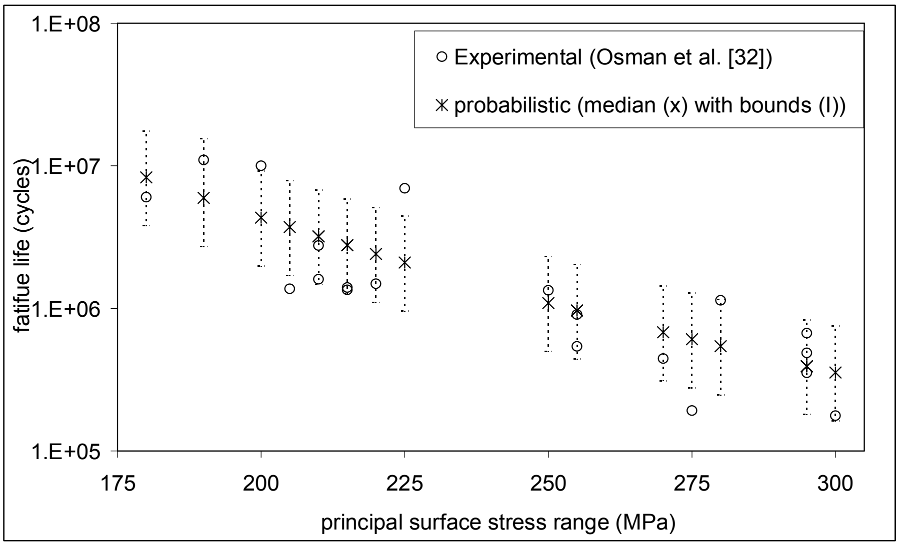


Figure 7. Comparison of results of probabilistic analysis with experimentally observed values of fatigue life for plate girders reported in literature

2 SOME RECENT DEVELOPMENTS

Extension of Markov-type inequalities to a class of random variables without moment condition requirement was proposed by Steliga and Szynal (2010). This is a flexible formulation that enables the computation of bounds on probabilities for the random variable which does not have well defined moments. Some of the examples of random variables that belong to this category are alpha stable random variables. These random variables have fractional moments and the existence of the order of moments depends on the value of the exponent. Recent R&D at CSIR-SERC has revealed that for describing the variations in some of the engineering quantities, alpha-stable distributions are more appropriate (Balaji Rao and Anoop, 2012). Hence, the results of Steliga and Szynal (2010) are important and the same are considered further.

Steliga and Szynal (2010) assumed the following conditions: (1) X is a positive random variables ($X \geq 0$ a.s.), (2) $\boldsymbol{G}$ is a class of all positive, strictly increasing functions g with $g(0) = 0$. Let $\boldsymbol{N}$ be set of integers. Making use of Markov inequalities, the following inequality has been proved by them.

For a given $\varepsilon > 0$ and k in $\boldsymbol{N}$

$$E\left[\frac{g(kX)}{g(kX)+g((k-1)X+\varepsilon)}I[X \geq \varepsilon]\right] \leq P[X \geq \varepsilon] \leq 2E\left[\frac{g(kX)}{g(kX)+g((k-1)X+\varepsilon)}I[X \geq \varepsilon]\right] \leq 2E\left[\frac{g(kX)}{g(kX)+g((k-1)X+\varepsilon)}\right] \quad (6)$$

Where I[.] denotes the indicator function. The above bounds are valid only when X is a positive random variable. The computation of bounds does not require moments of X. However, we should know the functional form (strictly increasing) of $g(.)$ and an idea about the realizations of X. We should also be in a position to determine the expected values of $g(.)$. The strictly increasing function g(.) can be formulated based on phenomenological modeling involving X.

Possible structural engineering application: For example, the average (smoothed) roof load-displacement curve of a moment resisting frame subjected to lateral loads, obtained using pushover analysis, over the range of engineering design interest, can be considered as strictly increasing function of load. In a gravity controlled experiment, the variations observed in roof displacements of nominally similar frames can be attributed to the variations in dimensions and strengths of materials. These variations can be aggregated into overall rigidity/compliance of the frame (X, which is always positive, and thus satisfying the condition required to estimate the bounds using the above equation). We can always assign an acceptable probability distribution to the deflection, generate random values of deflection following the assigned pdf and indirectly estimate the rigidity and examine the indicator function. This exercise would circumvent us from directly generating random realizations of X (as already indicated this random variable may not have well defined moments). Still we will be able to compute the bounds on the required probability with regard to X. This also suggests that the formulations presented by Steliga and Szynal (2010) can be used in an inverse problem to characterize, probabilistically, the internal variable. However, the condition $k \in N$ may perhaps need to relaxed through proper formulations.

Computation of bounds on probability of required event using Eq. (6) requires X to be positive. Let X be the random variable which is real valued. Then, the bounds presented above for positive random variables can be used by substituting $|X|$ in the place of X. Let us consider some special cases wherein does X take on negative values whose moments may or may not exist and a strictly increasing function of X, $g(X)$, is observable and whose distribution is known (in such a case X can be considered as a hidden/internal variable whose value is inferred based on a physical relationship $g(.)$ and X).

The following functional form is assumed for ε >0. We will now consider special cases and provide necessary bounds for *P(X ≥ ε)*.

$(a) \quad g(|x|) = |x|^m, \quad m \in N$

$(b) \quad g(|x|) = |x|^r, \quad 0 < r < \infty$

In order to compute denominator in Eq. (6), we need the following bounds (Gut, 2005), whenever $x > 0$ and $y > 0$,

$$(x+y)^r \le \begin{cases} x^r + y^r & \text{for } 0 \le r \le 1 \\ 2^{r-1}(x^r + y^r) & \text{for } r \ge 1 \end{cases} \qquad (7)$$

The following are the bounds derived by Steliga and Szynal (2010). These would be useful in engineering applications some of which are pointed out in the next section.

$$LB_k^I(r;\varepsilon) = 2E\left[\frac{(k|X|)^r}{(k|X|)^r + [(k-1)|X| + \varepsilon]^r} I(|X| \ge \varepsilon)\right], \quad k \in N, \quad r > 0$$

$$LB_k^I(m;\varepsilon) = 2E\left[\frac{(k|X|)^m}{(k|X|)^m + [(k-1)|X| + \varepsilon]^m} I(|X| \ge \varepsilon)\right], \quad k, m \in N$$

$$MB_k^I(r;\varepsilon) = 2E\left[\frac{(k|X|)}{(k|X|) + [(k-1)|X| + \varepsilon]} I(|X| \ge \varepsilon)\right]^r, \quad k \in N, \quad 0 < r \le 1$$

$$MB_k^I(r;\varepsilon) = 2^r E\left[\frac{(k|X|)}{(k|X|) + [(k-1)|X| + \varepsilon]} I(|X| \ge \varepsilon)\right]^r, \quad k \in N, \quad r \ge 1$$

The predictive power of the above equations, characterised by sharpness of bounds, for $g(x) = |x|^r$, when the underlying random variable is following normal and lognormal distributions are presented in Tables 1 and 2, respectively, for different values of *k* and *r*. To estimate the lower and upper bounds, Monte Carlo simulation technique involving 10^5 simulation cycles are used. The actual probabilities whose bounds are being estimated are presented in foot note of these tables. *From the results presented in these tables it is inferred that the bounds developed by Steliga and Szynal are tight and can be used in the engineering applications for making probability statements about the internal variable which is responsible for generating the random observable g(X).* However, it should be noted that the assumptions made in deriving the bounds should be satisfied. As stated earlier, one of the limitations, perhaps, in engineering application would be the need imposed by $k \in N$.

Possible structural engineering application: In many engineering problems, *X* may take negative values and also moments of *X* may not exist. For example, recent studies by Balaji Rao and Anoop (2012), at CSIR-SERC, have shown that the description of evolution of surface strain field of a reinforced concrete flexural member follows a Levy process. Accordingly, at any stage of loading, the fluctuations in surface strains may be described using an alpha-stable distribution. It is known that handling such random variables can be difficult and may be desirable to make probabilistic inferences of these variables based on the probabilistic variations in observables such as deflections and/or crackwidths which are functions of internal variables such as strains. This study is being furthered at CSIR-SERC.

Table 2. Results of simulation (N = 10^5 cycles) for the bounds on $P[|X| \ge 2]^*$ for *X* being normally distributed with mean = 0.0 and standard deviation = 1.0; $g(x) = |x|^r$

k	*r*	¾	½	¼	1/9	1/16
1	LBI(k;r;ε) MBI(k;r; ε)	0.04852617 0.05764368	0.04759212 0.067207573	0.04665248 0.078375767	0.04612905 0.085371748	0.04594573 0.087966528
2	LBI(k;r; ε) MBI(k;r; ε)	0.04702721 0.05591246	0.04658857 0.065866515	0.04614941 0.077596523	0.04590531 0.08499524	0.04581986 0.087748455
4	LBI(k;r;∈) MBI(k;r; ε)	0.04634894 0.05511564	0.04613601 0.0652417	0.04592302 0.077229053	0.04580468 0.084816504	0.04576326 0.087644691
9	LBI(k;r; ε) MBI(k;r; ε)	0.04598943 0.05469043	0.04589629 0.064906321	0.04580315 0.07703066	0.0457514 0.084719696	0.04573329 0.087588427
16	LBI(k;r; ε) UBI(k;r; ε)	0.04586632 0.05454438	0.04581421 0.064790817	0.04576211 0.076962149	0.04573316 0.084686215	0.04572303 0.087568958
25	LBI(k;r; ε) MBI(k;r; ε)	0.04580979 0.05447725	0.04577653 0.064737677	0.04574326 0.076930597	0.04572478 0.084670787	0.04571832 0.087559985
36	LBI(k;r; ε) MBI(k;r; ε)	0.0457792 0.05444092	0.04575613 0.064708895	0.04573307 0.076913499	0.04572025 0.084662424	0.04571577 0.08755512

Note: * Actual probability value = 0.0455

Table 3. Results of simulation (N = 10^5 cycles) for the bounds on $P[X \geq 3]^*$ for X being lognormally distributed with mean = 2.0 and standard deviation = 0.50^{**}; $g(x) = |x|^r$

k	r	¾	½	¼	1/9	1/16
1	LBI(k;r; ε) MBI(k;r; ε)	0.03954956 0.04701278	0.0390874 0.055247321	0.03862401 0.064931017	0.03836626 0.071029536	0.03827603 0.07329724
2	LBI(k;r; ε) MBI(k;r; ε)	0.03882517 0.04616696	0.03860355 0.054587084	0.03838181 0.064544453	0.03825858 0.070841976	0.03821545 0.073188447
4	LBI(k;r; ε) MBI(k;r; ε)	0.03848589 0.04576672	0.03837727 0.054272124	0.03826864 0.064358563	0.03820829 0.070751383	0.03818716 0.073135818
9	LBI(k;r; ε) MBI(k;r; ε)	0.03830327 0.04555033	0.03825551 0.054101171	0.03820776 0.06425727	0.03818123 0.070701911	0.03817194 0.073107056
16	LBI(k;r; ε) MBI(k;r; ε)	0.03824028 0.04547556	0.03821352 0.054041991	0.03818676 0.06422214	0.03817189 0.070684736	0.03816669 0.073097067
25	LBI(k;r; ε) MBI(k;r; ε)	0.03821129 0.04544111	0.0381942 0.054014712	0.0381771 0.064205936	0.0381676 0.07067681	0.03816427 0.073092457
36	LBI(k;r; ε) MBI(k;r; ε)	0.03819559 0.04542245	0.03818372 0.053999923	0.03817186 0.064197148	0.03816527 0.070672511	0.03816297 0.073089956

Note : *Actual probability value = 0.0382; **parameters of lognormal: - lamda = 0.66261654; exi = 0.246068276;

3 SUMMARY

This paper aims at bringing out the usefulness of Chebyshev- and Markov- type inequalities in structural engineering design decision making. By examining whether the bounds arising from Chebyshev - type inequality (associated with these are weak upper bound probabilities) encloses the respective experimental values for: (a) prediction of central deflection of six ferrocement I-beams, and, (b) fatigue life of a steel plate girder of a plate-girder bridge, against the limit state of web shear buckling, it is inferred that the bounds and the associated probabilities estimated are realistic and hence can be used in structural engineering design decision making. The paper also presents recent developments in determination of inequalities of the type of Markov, which are due to Steliga and Szynal (2010). The importance of such bounds in structural engineering applications is brought out. It is shown from the results of Monte Carlo simulation that the bounds on probability of an event sought, estimated using the method presented by Steliga and Szynal, are sharp. One of the important advantages of the bounds presented by Steliga and Szynal (2010) is that the original (hidden/internal) random variable need not have well defined moments. Possible engineering applications are also pointed out.

ACKNOWLEDGEMENT

One of the authors, KBR, is grateful to Prof. Prakash Desayi, formerly professor, Department of Civil Engineering, Indian Institute of Science, Bangalore, under whose guidance he did his Ph.D. The authors thank the Director, CSIR-SERC, for giving permission for publishing this paper.

REFERENCES

1. Balaji Rao, K and Anoop, M.B 2012. Why do we need probability distributions with fat tails to describe the surface strain evolution in reinforced concrete flexural members?. *Meccanica*, Springer, DOI 10.1007/S11012-012-9681-8.
2. Balaji Rao, K., Anoop, M. B., Raghava, G., Prakash, M. and Rajadurai, A 2013. Probabilistic fatigue life analysis of welded steel plate railway bridge girders using S-N curve approach. *Journal of* Risk and Reliability, *Part O, IMecH*, UK – Accepted for Publication
3. Balaji Rao. K 1990. *Studies on Reliability of Reinforced Concrete Beams in Cracking and Ferrocement Elements in Tension and Flexure*. Ph. D thesis submitted to Department of Civil Engineering, Indian Institute of Science, Bangalore, 872 pp.
4. Bolotin, V. V 1969. *Statistical Methods in Structural Mechanics.* Holden-Day, Inc., USA.
5. Gnedenko, B.V 1976. *The theory of probability*. English Translation, Mir Publishers, Moscow.
6. Gut, A 2005. *Probability: A graduate course*. Springer texts in Statistics, Springer, New York.
7. Prakash Desayi and Balaji Rao, K 1993. Estimation of Spacing and Width of Cracks in Ferrocement Monolithic I-Joists. *Journal of Ferrocement*, Vol.23, No.1, pp:1-13.
8. Prakash Desayi and Balaji Rao, K 1988. Prediction of Cracking and Ultimate Moments, and, Load-Deflection Behaviour of Ferrocement Elements. *Third International Symposium on Ferrocement*, New Delhi, December 8-10, *Proceedings*, pp:90-98.
9. Steliga, K and Szynal, D 2010. On Markov-type inequalities. *International Journal of Pure and Applied Mathematics*, Vol. 58, No. 2, pp:137 – 152.

ISSN 1932-2321