Article citation info:
Kołowrocki K, Soszyńska-Budny J. Uniformity testing of complex technical system operation process statistical data sets. Diagnostyka. 2016;17(1):13-20

DIAGNOSTYKA, 2016, Vol. 17, No. 1

ISSN 1641-6414
e-ISSN 2449-5220

UNIFORMITY TESTING OF COMPLEX TECHNICAL SYSTEM OPERATION PROCESS STATISTICAL DATA SETS

Krzysztof KOŁOWROCKI, Joanna SOSZYŃSKA-BUDNY

Akademia Morska w Gdyni, 81-225 Gdynia, ul. Morska 81-87,
fax: 48 58 6206701, email: katmatkk@am.gdynia.pl, joannas@am.gdynia.pl

Summary

The paper is concerned with the methods for statistical data uniformity testing and identifying unknown parameters of a general probability model of a complex technical system operation process and their practical application. The general model of a complex technical system operation process is constructed. The procedure of statistical data sets uniformity testing is proposed and applied to the empirical realizations of the system operation processes sojourn times at the operation states coming from the realizations of a maritime ferry operation process collected during spring and winter into separate two sets of data. After that, the identification of the maritime ferry technical system operation process is performed and moreover the identified process is applied to its operation characteristics prediction.

Keywords: operation process; data uniformity testing; identification; prediction; maritime transport.

TESTOWANIE JEDNORODNOŚCI DANYCH STATYSTYCZNYCH PROCESU EKSPLOATACJI ZŁOŻONEGO SYSTEMU TECHNICZNEGO

Streszczenie

Artykuł dotyczy metod testowania jednorodności danych statystycznych oraz identyfikacji nieznanych parametrów ogólnego modelu probabilistycznego procesu eksploatacji złożonego systemu technicznego oraz ich praktycznego zastosowania. Skonstruowany jest model ogólny procesu eksploatacji złożonego systemu technicznego. Zaproponowana jest procedura testowania jednorodności zbiorów danych statystycznych i zastosowana do empirycznych realizacji czasów przebywania procesu eksploatacji systemu w stanach eksploatacyjnych pochodzących z realizacji procesu eksploatacji promu morskiego zebranych w czasie wiosny i zimy w dwóch oddzielnych zbiorach danych. Następnie, przeprowadzona jest identyfikacja procesu eksploatacji systemu technicznego promu morskiego, a ponadto zidentyfikowany proces eksploatacji jest zastosowany do promocji jego charakterystyk eksploatacyjnych.

Słowa kluczowe: proces eksploatacji; testowanie jednorodności; identyfikacja; predykcja; transport morski.

1. INTRODUCTION

The general joint model linking the system reliability model with the model of its operation process is constructed in [1] and [2]. To apply this general model practically to the evaluation and prediction of real complex technical systems reliability it is necessary to elaborate the statistical methods concerned with determining the unknown parameters of the proposed model. Particularly, concerning the system operation process, the methods of estimating the probabilities of the initials system operation states, the probabilities of transitions between the system operation states and the distributions of the sojourn times of the system operation process at the particular operation states should be proposed. The methods of testing the hypotheses concerned with the conditional sojourn times of the system operation process at particular operation states should be also elaborated. In the case when the statistical data are coming from different experiments, before the system operation identification, the investigation of these data uniformity is necessary.

2. MATHEMATICAL MODEL OF COMPLEX TECHNICAL SYSTEM OPERATION PROCESS

We assume that the system during its operation process is taking $v, v \in N$, different operation states

$z_1, z_2, \ldots, z_v$. Further, we define the system operation process $Z(t)$, $t \in <0,+\infty)$, with discrete operation states from the set $\{z_1, z_2, \ldots, z_v\}$. Moreover, we assume that the system operation process $Z(t)$ is a semi-Markov process [1]-[9] with the conditional sojourn times θ_{bl} at the operation states z_b when its next operation state is z_l, $b, l = 1,2,...,v$, $b \neq l$.

Under these assumptions, the system operation process may be described by [10]:

– the vector $[p_b(0)]_{1xv}$ of the initial probabilities

$$p_b(0) = P(Z(0) = z_b), \quad b = 1,2,...,v,$$

of the system operation process $Z(t)$ staying at the operation states at the moment $t = 0$;

– the matrix $[p_{bl}]_{vxv}$ of probabilities

$$p_{bl}, \quad b, l = 1,2,...,v, \quad b \neq l,$$

of the system operation process $Z(t)$ transitions between the operation states z_b and z_l;

– the matrix $[H_{bl}(t)]_{vxv}$ of conditional distribution functions

$$H_{bl}(t) = P(\theta_{bl} < t), \quad b, l = 1,2,...,v, \quad b \neq l,$$

of the system operation process $Z(t)$ conditional sojourn times θ_{bl} at the operation states.

The mean values of the conditional sojourn times θ_{bl} of the system operation process $Z(t)$ are given by

$$M_{bl} = E[\theta_{bl}] = \int_0^{\infty} t dH_{bl}(t), \quad b,l = 1,2,...,v, \quad b \neq l. \quad (1)$$

From the formula for total probability, it follows that the unconditional distribution functions of the sojourn times θ_b, $b = 1,2,...,v$, of the system operation process $Z(t)$ at the operation states z_b, $b = 1,2,...,v$, are given by [1]-[2]

$$H_b(t) = \sum_{l=1}^{v} p_{bl} H_{bl}(t), \quad b = 1,2,...,v. \quad (2)$$

Hence, the mean values $E[\theta_b]$ of the system operation process $Z(t)$ unconditional sojourn times θ_b, $b = 1,2,...,v$, at the operation states are given by

$$M_b = E[\theta_b] = \sum_{l=1}^{v} p_{bl} M_{bl}, \quad b = 1,2,...,v, \quad (3)$$

where M_{bl} are defined by the formula (1).

The limit values of the system operation process $Z(t)$ transient probabilities at the particular operation states

$$p_b(t) = P(Z(t) = z_b), \quad t \in < 0,+\infty), \quad b = 1,2,...,v, \quad (4)$$

are given by [1]-[9]

$$p_b = \lim_{t \to \infty} p_b(t) = \frac{\pi_b M_b}{\sum_{l=1}^{v} \pi_l M_l}, \quad b = 1,2,...,v, \quad (5)$$

where M_b, $b = 1,2,...,v$, are given by (3), while the steady probabilities π_b of the vector $[\pi_b]_{1xv}$ satisfy the system of equations

$$\begin{cases} [\pi_b] = [\pi_b][p_{bl}] \\ \sum_{l=1}^{v} \pi_l = 1. \end{cases} \quad (6)$$

Other interesting characteristics of the system operation process $Z(t)$ possible to obtain are its total sojourn times $\hat{\theta}_b$ at the particular operation states z_b, $b = 1,2,...,v$, during the fixed system operation time. It is well known [2], [5] that the system operation process total sojourn times $\hat{\theta}_b$ at the particular operation states z_b, for sufficiently large operation time θ, have approximately normal distributions with the expected value given by

$$\hat{M}_b = E[\hat{\theta}_b] = p_b \theta, \quad b = 1,2,...,v, \quad (7)$$

where p_b are given by (5).

3. PROCEDURE OF EXPERIMENTAL STATISTICAL DATA UNIFORMITY ANALYSIS

We consider test λ [2] that can be used for testing whether two independent samples of realizations of the conditional sojourn times at the operation states of the system operation process are drawn from the population with the same distribution. We assume that we have two independent samples of non-decreasing ordered realizations

$$\theta_{bl}^{1k}, \quad k = 1,2,...,n_{bl}^{1}, \text{ and } \theta_{bl}^{2k}, \quad k = 1,2,...,n_{bl}^{2}, \quad (8)$$

of the sojourn times

$$\theta_{bl}^{1} \text{ and } \theta_{bl}^{2}, \quad b,l \in \{1,2,...,v\}, \quad b \neq l,$$

respectively composed of n_{bl}^1 and n_{bl}^2 realizations and we mark their empirical distribution functions by

$$H_{bl}^1(t) = \frac{1}{n_{bl}^1} \# \{k : \theta_{bl}^{1k} < t, k \in \{1,2,...,n_{bl}^1\}\}, \quad t \ge 0, \quad (9)$$

and

$$H_{bl}^2(t) = \frac{1}{n_{bl}^2} \# \{k : \theta_{bl}^{2k} < t, k \in \{1,2,...,n_{bl}^2\}\}, \quad t \ge 0, \quad (10)$$

Then, according to Kolmogorov-Smirnov theorem [7], the sequence of distribution functions given by the equation

$$Q_{n_1 n_2}(\lambda) = P(D_{n_1 n_2} < \frac{\lambda}{\sqrt{n}}) \quad (11)$$

defined for $\lambda > 0$, where

$$n_1 = n_{bl}^1, \ n_2 = n_{bl}^2, \ n = \frac{n_1 n_2}{n_1 + n_2}, \quad (12)$$

and

$$D_{n_1 n_2} = \max_{-\infty < t < +\infty} \left| H_{bl}^1(t) - H_{bl}^2(t) \right|, \quad (13)$$

is convergent, as $n \to \infty$, to the limit distribution function

$$Q(\lambda) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2}, \ \lambda > 0. \quad (14)$$

The distribution function $Q(\lambda)$ given by (14) is called λ distribution and its Tables of values are available.

The convergence of the sequence $Q_{n_1 n_2}(\lambda)$ to the λ distribution $Q(\lambda)$ means that for sufficiently large n_1 and n_2 we may use the following approximate formula

$$Q_{n_1 n_2}(\lambda) \cong Q(\lambda). \quad (15)$$

Hence, it follows that if we define the statistic

$$U_n = D_{n_1 n_2} \sqrt{n}, \quad (16)$$

where $D_{n_1 n_2}$ is defined by (13), then by (14) and (15), we have

$$P(U_n < u) = P(D_{n_1 n_2} \sqrt{n} < u) = P(D_{n_1 n_2} < \frac{u}{\sqrt{n}})$$

$$= Q_{n_1 n_2}(u) \cong Q(u) \text{ for } u > 0. \quad (17)$$

This result means that in order to formulate and next to verify the hypothesis that the two independent samples of the realizations of the system operation process conditional sojourn times

$$\theta_{bl}^1 \text{ and } \theta_{bl}^2, \ b,l \in \{1,2,...,\nu\}, \ b \ne l,$$

at the operation state z_b when the next transition is to the operation state z_l are coming from the population with the same distribution, it is necessary to proceed according to the following scheme:

- to fix the numbers of realizations n_{bl}^1 and n_{bl}^2 in the samples;
- to collect the realizations (8) of the conditional sojourn times θ_{bl}^1 and θ_{bl}^2 of the system operation process in the samples;
- to find the realization of the empirical distribution functions $H_{bl}^1(t)$ and $H_{bl}^2(t)$ defined by (9) and (10) respectively, in the following forms:

$$H_{bl}^1(t) = \begin{cases} \frac{n_{bl}^{11}}{n_{bl}^1} = 0, & t \le \theta_{bl}^{11} \\ \frac{n_{bl}^{1k}}{n_{bl}^1}, & \theta_{bl}^{1k-1} < t \le \theta_{bl}^{1k}, \ k = 2,3,...,n_{bl}^1, \\ \frac{n_{bl}^{1n_{bl}^1+1}}{n_{bl}^1} = 1, & t \ge \theta_{bl}^{1n_{bl}^1} \end{cases} \quad (18)$$

where

$$n_{bl}^{11} = 0, \ n_{bl}^{1n_{bl}^1+1} = n_{bl}^1, \quad (19)$$

and

$$n_{bl}^{1k} = \# \{j : \theta_{bl}^{1j} < \theta_{bl}^{1k}, \ j \in \{1,2,...,n_{bl}^1\}\},$$
$$k = 2,3,...,n_{bl}^1, \quad (20)$$

is the number of the sojourn time θ_{bl}^1 realizations less than its realization θ_{bl}^{1k}, $k = 2,3,...,n_{bl}^1$,

$$H_{bl}^2(t) = \begin{cases} \frac{n_{bl}^{21}}{n_{bl}^2} = 0, & t \le \theta_{bl}^{21} \\ \frac{n_{bl}^{2k}}{n_{bl}^2}, & \theta_{bl}^{2k-1} < t \le \theta_{bl}^{2k}, \ k = 2,3,...,n_{bl}^2, \\ \frac{n_{bl}^{2n_{bl}^2+1}}{n_{bl}^2} = 1, & t \ge \theta_{bl}^{2n_{bl}^2} \end{cases} \quad (21)$$

where

$$n_{bl}^{21} = 0, \ n_{bl}^{2n_{bl}^2+1} = n_{bl}^2, \quad (22)$$

and

$$n_{bl}^{2k} = \# \{j : \theta_{bl}^{2j} < \theta_{bl}^{2k}, \ j \in \{1,2,...,n_{bl}^2\}\},$$
$$k = 2,3,...,n_{bl}^2, \quad (23)$$

is the number of the sojourn time θ_{bl}^2 realizations less than its realization θ_{bl}^{2k}, $k = 2,3,...,n_{bl}^2$;

- to calculate the realization of the statistic u_n defined by (16) according to the formula

$$u_n = d_{n^1_{bl} n^2_{bl}} \sqrt{n}, \quad (24)$$

where

$$d_{n^1_{bl} n^2_{bl}} = \max \{ d^1_{n^1_{bl} n^2_{bl}}, d^2_{n^1_{bl} n^2_{bl}} \}, \quad (25)$$

$$d^1_{n^1_{bl} n^2_{bl}}$$

$$= \max \{ \left| H^1_{bl}(\theta^{1k}_{bl}) - H^2_{bl}(\theta^{1k}_{bl}) \right|, k \in \{1,2,...,n^1_{bl}\} \}, \quad (26)$$

$$d^2_{n^1_{bl} n^2_{bl}}$$

$$= \max \{ \left| H^1_{bl}(\theta^{2k}_{bl}) - H^2_{bl}(\theta^{2k}_{bl}) \right|, k \in \{1,2,...,n^2_{bl}\} \}, \quad (27)$$

$$n = \frac{n^1_{bl} n^2_{bl}}{n^1_{bl} + n^2_{bl}}; \quad (28)$$

- to formulate the null hypothesis H_0 in the following form:
  H_0 : The samples of realizations (8) are coming from the populations with the same distributions;
- to fix the significance level α of the λ test;
- to read from the Tables of λ distribution the value $u = \lambda_0$ such that the following equality holds

$$P(U_n < u) = Q(u) = Q(\lambda_0) = 1 - \alpha; \quad (29)$$

- to determine the critical domain in the firm $(u, +\infty)$;
- to compare the obtained value u_n of the realization of the statistics U_n with the read from Tables value $u = \lambda_0$;
- to decide on the formulated hypothesis H_0 in the following way: if the value u_n does not belong to the critical domain, i.e.

$$u_n \le u,$$

then we do not reject the hypothesis H_0, otherwise if the value u_n belongs to the critical domain, i.e.

$$u_n > u,$$

then we reject the hypothesis H_0.

In the case when the null hypothesis H_0 is not rejected we may join the statistical data from the considered two separate sets into one new set of data and if there are no other sets of statistical data we proceed with the data of this new set in the way described in [8]. Otherwise, if there are other sets of statistical data we select the next one of them and perform the procedure of this section for data from this set and data from the previously formed new set. We continue this procedure up to the moment when the store of the statistical data sets is exhausted.

4. MARITIME FERRY OPERATION PROCESS UNIFORMITY TESTING

We use the two-sample λ test described in Section III to verify the hypotheses that spring and winter realizations of the maritime ferry [7] conditional sojourn times at the operation states are from the populations with the same distribution. For instance, the procedure of testing the uniformity of data collected at the operation state z_1 when the next operation state was z_2 is as follows: For spring and winter data, the conditional sojourn times θ^1_{12} and θ^2_{12} have the empirical distribution functions

$$H^1{}_{12}(t) = \begin{cases} 0, & t \le 15, \\ 1/42, & 15 < t \le 20, \\ 2/42, & 20 < t \le 25, \\ 3/42, & 25 < t \le 33, \\ 4/42, & 33 < t \le 35, \\ 5/42, & 35 < t \le 37, \\ 6/42, & 37 < t \le 40, \\ 8/42, & 40 < t \le 43, \\ 9/42, & 43 < t \le 44, \\ 13/42, & 44 < t \le 45, \\ 14/42, & 45 < t \le 46, \\ 15/42, & 46 < t \le 47, \\ 17/42, & 47 < t \le 50, \\ 18/42, & 50 < t \le 52, \\ 19/42, & 52 < t \le 53, \\ 21/42, & 53 < t \le 55, \\ 22/42, & 55 < t \le 57, \\ 23/42, & 57 < t \le 58, \\ 24/42, & 58 < t \le 59, \\ 26/42, & 59 < t \le 60, \\ 27/42, & 60 < t \le 61, \\ 29/42, & 61 < t \le 62, \\ 30/42, & 62 < t \le 63, \\ 32/42, & 63 < t \le 65, \\ 33/42, & 65 < t \le 67, \\ 34/42, & 67 < t \le 68, \\ 35/42, & 68 < t \le 71, \\ 36/42, & 71 < t \le 72, \\ 38/42, & 72 < t \le 75, \\ 39/42, & 75 < t \le 78, \\ 40/42, & 78 < t \le 84, \\ 41/42, & 84 < t \le 97, \\ 1, & t > 97; \end{cases} \qquad H^2{}_{12}(t) = \begin{cases} 0, & t \le 12, \\ 1/40, & 12 < t \le 15, \\ 3/40, & 15 < t \le 18, \\ 4/40, & 18 < t \le 19, \\ 5/40, & 19 < t \le 20, \\ 6/40, & 20 < t \le 25, \\ 7/40, & 25 < t \le 33, \\ 9/40, & 33 < t \le 34, \\ 10/40, & 34 < t \le 36, \\ 11/40, & 36 < t \le 37, \\ 14/40, & 37 < t \le 40, \\ 15/40, & 40 < t \le 41, \\ 16/40, & 41 < t \le 44, \\ 17/40, & 44 < t \le 46, \\ 18/40, & 46 < t \le 48, \\ 20/40, & 48 < t \le 50, \\ 21/40, & 50 < t \le 53, \\ 22/40, & 53 < t \le 55, \\ 23/40, & 55 < t \le 57, \\ 24/40, & 57 < t \le 59, \\ 25/40, & 59 < t \le 60, \\ 27/40, & 60 < t \le 61, \\ 28/40, & 61 < t \le 62, \\ 29/40, & 62 < t \le 63, \\ 30/40, & 63 < t \le 65, \\ 34/40, & 65 < t \le 67, \\ 35/40, & 67 < t \le 69, \\ 36/40, & 69 < t \le 75, \\ 38/40, & 75 < t \le 80, \\ 39/40, & 80 < t \le 90, \\ 1, & t > 90. \end{cases}$$

The null hypothesis is H_0 : The winter and spring data at the operation state z_1 when the next operation state was z_2 are from the population with the same distribution.

To verify this hypothesis we will apply the two-sample λ test at the significance level $\alpha = 0.05$. Using the above empirical distributions we form a common Table 1 composed of all their values.

Table 1. Joint empirical distribution function

$t_k = \theta_{12}^{1k} \vee \theta_{12}^{2k}$	$H^1{}_{12}(t_k)$	$H^2{}_{12}(t_k)$	$\lvert H^1{}_{12}(t_k) - H^2{}_{12}(t_k)\rvert$
12	0	0	0
15	0	1/40	0.025
18	1/42	3/40	0.051
19	1/42	4/40	0.076
20	1/42	5/40	0.101
25	2/42	6/40	0.102
33	3/42	7/40	0.104
34	4/42	9/40	0.129
35	4/42	10/40	0.156
36	5/42	10/40	0.131
37	5/42	11/40	0.156
40	6/42	14/40	0.207
41	8/42	15/40	0.185
43	8/42	16/40	0.209
44	9/42	16/40	0.186
45	13/42	17/40	0.115
46	14/42	17/40	0.092
47	15/42	18/40	0.093
48	17/42	18/40	0.045
50	17/42	20/40	0.095
52	18/42	21/40	0.096
53	19/42	21/40	0.073
55	21/42	22/40	0.05
57	22/42	23/40	0.051
58	23/42	24/40	0.052
59	24/42	24/40	0.029
60	26/42	25/40	0.006
61	27/42	24/40	0.032
62	29/42	28/40	0.009
63	30/42	29/40	0.011
65	32/42	30/40	0.012
67	33/42	34/40	0.064
68	34/42	35/40	0.065
69	35/42	35/40	0.042
71	35/42	36/40	0.067
72	36/42	36/40	0.043
75	38/42	36/40	0.005
78	39/42	38/40	0.021
80	40/42	38/40	0.002
84	40/42	39/40	0.023
90	41/42	39/40	0.001
97	41/42	1	0.024
>97	1	1	0

In Table 1, the values t_k are joint together all realizations

$$\theta_{12}^{1k},\ k = 1,2,\ldots,n_{12}^{1},\ \text{and}\ \theta_{12}^{2k},\ \ k = 1,2,\ldots,n_{12}^{2},$$

of the conditional sojourn times θ_{12}^{1} and θ_{12}^{2}, i.e. they are all discontinuity points of the empirical distribution function $H^1{}_{12}(t)$ and $H^2{}_{12}(t)$ were they have jumps in their values $H^1{}_{12}(t_k)$ and $H^2{}_{12}(t_k)$. Next, according to (25)-(27), from Table 1, we get

$$d_{42\,40} = \max_{t_k} \left| H^1{}_{12}(t_k) - H^2{}_{12}(t_k) \right| \cong 0.209,$$

and according to (28)

$$n_{12} = \frac{42 \cdot 40}{42 + 40} = 20.48.$$

Thus, the realization u_n of the statistics (24) is

$$u_n = d_{42\,40}\sqrt{n_{12}} = 0.209\sqrt{20.48} \cong 0.946.$$

From the table of the λ distribution for the significance level $\alpha = 0.05$, according to (29), we get the critical value $\lambda_0 = u \cong 1.36$.
Since

$$u_n \cong 0.946 < u = 1.36,$$

then we do not reject the null hypothesis H_0.
After proceeding in an analogous way with data in the remaining operation states we can obtain the same conclusions that the sprig data sets and the winter data sets are from the populations with the identical distributions.

5. STATISTICAL IDENTIFICATION OF MARITIME FERRY OPERATION PROCESS

To identify all parameters of the considered maritime ferry operation process [7] the statistical data coming from this process is needed. The joint statistical data that has been collected during spring and winter are:
- the number of the ship operation process states $\nu = 18$;
- the ferry operation process observation time $\Theta =$ 82 days;
- the number of the ferry operation process realizations $n(0) =$ 82;
- the vector of realizations of the numbers of the ferry operation process staying at the operation states z_b at the initial moment $t = 0$

$$[n_b(0)]_{1x18} = [82,0,\ldots,0];$$

- the matrix of realizations n_{bl} of the numbers of the ferry operation process $Z(t)$ transitions from the state z_b into the state z_l during the observation time $\Theta = 82$ days

$$[n_{bl}]_{18x18} = \begin{bmatrix} 0\ 82\ \ 0 \ldots 0\ \ 0 \\ 0\ \ 0\ 82 \ldots 0\ \ 0 \\ \ldots \\ 0\ \ 0\ \ 0 \ldots 0\ 82 \\ 82\ 0\ \ 0 \ldots 0\ \ 0 \end{bmatrix};$$

- the vector of realizations of the total numbers of the ferry operation process transitions from the operation state z_b during the observation time $\Theta = 82$ days

$$[n_b]_{18x1} = [82, 82, ..., 82]^T.$$

On the basis of the above statistical data it is possible to evaluate
- the vector of realizations

$$[p(0)] = [1, 0, 0 \ldots, 0, 0],$$

of the initial probabilities $p_b(0)$, $b = 1,2,...,18$, of the ferry operation process transients at the operation states z_b at the moment $t = 0$
- the matrix of realizations

$$[p_{bl}] = \begin{bmatrix} 0\,1\,0 \ldots 0\,0 \\ 0\,0\,1 \ldots 0\,0 \\ \ldots \\ 0\,0\,0 \ldots 0\,1 \\ 1\,0\,0 \ldots 0\,0 \end{bmatrix}, \qquad (30)$$

of the transition probabilities p_{bl}, $b, l = 1,2,...,18$, of the system operation process $Z(t)$ from the operation state z_b into the operation state z_l.
The statistical data allow that applying the same methods as in [8], we may verify the hypotheses about the conditional distribution functions $H_{bl}(t)$ of the maritime ferry operation process sojourn times

$$\theta_{bl},\ b = 1,2,...,17,\ l = b+1 \text{ and } b = 18,\ l = 1$$

at the state z_b while the next transition is to the state z_l on the base of their joint realizations θ_{bl}^j, $j = 1,2,...,82$. For instance, the conditional sojourn time θ_{12} has a normal distribution with the density function

$$h_{12}(t) = \frac{1}{18.256\sqrt{2\pi}} \exp[-\frac{(t - 51.415)^2}{666.563}]$$

for $t \in (-\infty, \infty)$.
Next for the verified distributions, the mean values

$$M_{bl} = E[\theta_{bl}],\ b, l = 1,2,...,18,\ b \neq l,$$

of the system operation process $Z(t)$ conditional sojourn times at the operation states defined by (1) can be determined:

$$M_{12} = 51.415,\ M_{34} = 36.176,\ M_{67} = 37.268,$$

$$M_{78} = 6.807,\ M_{89} = 19,\ M_{910} = 46.614,$$

$$M_{1011} = 2.829,\ M_{1112} = 4.459,\ M_{1213} = 25.091,$$

$$M_{1314} = 513.689,\ M_{1415} = 51.182,$$

$$M_{1516} = 33.807. \qquad (31)$$

In the remaining cases, because of lack of sufficiently extensive empirical data, the mean values $M_{bl} = E[\theta_{bl}]$ can be estimated by application the formula for the empirical mean [8] giving the following their approximate values:

$$M_{23} = 2.533,\ M_{45} = 52.393,\ M_{56} = 530.188,$$

$$M_{1617} = 4.448,\ M_{1718} = 5.473. \qquad (32)$$

6. MARITIME FERRY OPERATION PROCESS PREDICTION

After applying (3) and the results (31)-(32), the unconditional mean sojourn times of the maritime ferry operation process at the particular operation states are:

$$M_1 = 51.415,\ M_2 = 2.533,\ M_3 = 36.176,$$

$$M_4 = 52.393,\ M_5 = 530.188,\ M_6 = 37.268,$$

$$M_7 = 6.807,\ M_8 = 19,\ M_9 = 46.614,$$

$$M_{10} = 2.829,\ M_{11} = 4.459,\ M_{12} = 25.091,$$

$$M_{13} = 513.689,\ M_{14} = 51.182,\ M_{15} = 31.807,$$

$$M_{16} = 4.448,\ M_{17} = 5.473,\ M_{18} = 18.039. \qquad (33)$$

Considering (30) in the system of equations (6), we get its following solution

$$\pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5 = \pi_6 = \pi_7 = \pi_8 =$$

$$\pi_9 = \pi_{10} = \pi_{11} = \pi_{12} = \pi_{13} = \pi_{14} = \pi_{15} = \pi_{16} = \pi_{17} =$$

$$= \pi_{18} \cong 0.056$$

Hence and from (33), after applying (5), it follows that the limit values of the maritime ferry operation process transient probabilities at the operation states z_b, $b = 1,2,...,18,$ are:

$$p_1 = 0.036, \quad p_2 = 0.002, \quad p_3 = 0.025, \quad p_4 = 0.036,$$

$$p_5 = 0.368, \quad p_6 = 0.026, \quad p_7 = 0.005, \quad p_8 = 0.013,$$

$$p_9 = 0.032, \; p_{10} = 0.002, \; p_{11} = 0.003, \; p_{12} = 0.017,$$

$$p_{13} = 0.356, \; p_{14} = 0.036, \; p_{15} = 0.023, \; p_{16} = 0.003,$$

$$p_{17} = 0.004, \; p_{18} = 0.013,$$

Substituting the above transient probabilities at operation states into (7), we can get the mean values of the maritime ferry operation process total sojourn times at the particular operation states during for instance $\theta = 1$ year:

$$\hat{M}_1 = 13.14, \; \hat{M}_2 = 0.73, \quad \hat{M}_3 = 9.13, \quad \hat{M}_4 = 13.14,$$

$$\hat{M}_5 = 134.32, \; \hat{M}_6 = 9.49, \; \hat{M}_7 = 1.83, \; \hat{M}_8 = 4.75,$$

$$\hat{M}_9 = 11.68, \; \hat{M}_{10} = 0.73, \; \hat{M}_{11} = 0.10, \; \hat{M}_{12} = 6.21,$$

$$\hat{M}_{13} = 129.94, \; \hat{M}_{14} = 13.14, \; \hat{M}_{15} = 8.40,$$

$$\hat{M}_{16} = 1.10, \; \hat{M}_{17} = 1.46, \; \hat{M}_{18} = 4.75 \text{ days.}$$

7. CONCLUSIONS

The way of the uniformity testing of statistical data coming from different sets of realizations of the same complex technical system operation process before joining them into one common set of data and identifying its unknown operation parameters and prognosis its operation characteristics was presented and practically applied. The results of its application to the empirical data uniformity testing and the parameters identifying of the maritime ferry operation process and the operation characteristic prognosis justifies the proposed methods and procedures practical importance in everyday practice concerned with the complex transportation systems operation processes identification and prediction.

REFERENCES

[1] Kołowrocki K. *Reliability of Large and Complex Systems*. Amsterdam, Boston, Heidelberd, London, New York, Oxford, Paris, San Diego, San Francisko, Singapore, Sidney, Tokyo, Elsevier, 2014.

[2] Kolowrocki K., Soszynska J. *Safety and risk evaluation of Stena Baltica ferry in variable operation conditions*. Electron J Reliab Risk Anal: Theory Appl, vol. 2, No 4, 168-180, 2009.

[3] Ferreira F., Pacheco, A. *Comparison of level-crossing times for Markov and semi-Markov processes*. Stat & Probab Lett vol. 77(2), 151-157, 2007.

[4] Glynn P.W., P.J. Haas P. J. *Laws of large numbers and functional central limit theorems for generalized semi-Markov processes*. Stoch Model vol. 22(2), 201-231, 2006.

[5] Grabski F. *Semi-Markov Models of Systems Reliability and Operations Analysis*. System Research Institute, Warsaw. Polish Academy of Science, (in Polish), 2002.

[6] Kołowrocki K., Soszyńska J. *A general model of industrial systems operation processes related to their environment and infrastructure*. J Pol Saf Reliab Assoc, Summer Safety & Reliability Seminars, vol. 2(2), pp. 223-226, 2008.

[7] Kołowrocki K., Soszyńska-Budny J. *Reliability and Safety of Complex Technical Systems and Processes: Modeling-Identification-Prediction-Optimization*. London, Dordrecht, Heildeberg, New York, Springer, 2011.

[8] Kołowrocki K., Soszyńska J., *Statistical identification of complex technical system operation process*. Diagnostyka, Vol. 16, No.4 (2015).

[9] Limnios N., Oprisan G. *Semi-Markov Processes and Reliability*. Birkhauser, Boston, 2005.

[10] Kołowrocki K., Soszyńska J. *Methods and algorithms for evaluating unknown parameters of operation processes of complex technical systems*. Proc. 3rd Summer Safety and Reliability Seminars – SSARS 2009, Gdańsk-Sopot, vol. 2, pp. 211-222, 2009.

[11] Mercier S. *Numerical bounds for semi-Markovian quantities and application to reliability*. Methodol and Comput in Appl Probab. 10(2), pp. 179-198, 2008.

Received 2015-09-15
Accepted 2016-02-15
Available online 2016-02-22

Krzysztof KOLOWROCKI is Titular Professor and the Head of Mathematics Department at the Faculty of Navigation of Gdynia Maritime University. His field of interest is mathematical modeling of safety and reliability of complex systems and processes. He has published several books and over 300 scientific articles and papers. He is the Chairman of Polish Safety and Reliability Association and the Co-Editor of Journal of Polish Safety and Reliability Association, Summer Safety and Reliability Seminars. His home site can be found at: http://www.am.gdynia.pl/~katmatkk/

Joanna SOSZYŃSKA-BUDNY is University Professor at Department of Mathematics of the Faculty of Navigation of Gdynia Maritime University. Her field of interest is mathematical modelling of safety and reliability of complex systems at variable operation conditions. She has published 2 books and over 100 papers in scientific journals and conference proceedings. She is the Member of Management Boart of Polish Safety and Reliability Association and the Co-Editor of Journal of Polish Safety and Reliability Association, Summer Safety and Reliability Seminars.