24th International Conference
ENGINEERING MECHANICS 2018
Svratka, Czech Republic, May 14 –17, 2018
Paper **#208,** pp. 589–592, doi: 10.21495/91-8-589

IMPLEMENTING DISTRIBUTION MIXTURES IN DEVELOPING RELIABILITY MODEL FOR PAPER MACHINE

A. Neubauer[*], L. Knopik[**], M. Szubartowski[***], K. Migawa[†]

Abstract: *This paper analyzes reliability of paper machine. Probability distributions for repair time and worthiness time for paper machine and its main subsystems are examined. The usefulness of the application of a mixture of probability distributions of a random variable for the description of the paper machine repair time distribution has been demonstrated.*

Keywords: empirical distribution, gamma distribution, good of fit test, availability

1. Introduction

Paper machines belong to a group of machines that require high reliability and availability to operate. This paper analyzes the repair time and worthiness time for the paper machine and its main subsystems: the automation subsystem, the electrical subsystem and the mechanical subsystem. Basic statistical analysis was performed for both the entire paper machine and for its individual subsystems and the compatibility of the empirical distribution of the analyzed random variables with the gamma distribution was examined. For the analyzed paper machine data (for all subsystems total) and for each of the subsystems listed, compliance with the gamma distribution was found. The EM (Expectation-Maximization) algorithm was used to assess the values of the parameters of the mixture of probability distributions of the entire paper machine. The basic facts about the EM algorithm are presented in papers (Bishop, 2006, and Bohning, 2000, and Everitt et al., 1981, and McLachan et al., 2000, and Tettrington et al., 1985). In order to develop the results presented in this work, the problems and methods presented in particular in the paper (Dempster et al., 1977) were used.

2. Statistical analysis

Table 1 shows the results of the calculation of basic statistics for repair time of the paper machine and its main subsystems. The second and third columns contain mean values and standard deviations (STD) respectively. Analysis of these values leads to the conclusion that the paper machine's subsystems generate very diverse mean values and standard deviations of repair time. This proves that the data set analyzed jointly for all subsystems of the paper machine (marked "All" in the tables) is a statistically heterogeneous set. The fourth column contains the number of statistical samples n. In the fifth and sixth columns, the estimated values of α and β parameters of the gamma distribution of the examined random variables are given. The seventh column contains the sum of square (Σ^2) deviations of the empirical cumulative distribution function from the distribution function of the gamma distribution with the parameters given earlier. The eighth column contains the statistics of λ-Kolmogorov (λ-Kol) for the test

[*] RNDr. Andrzej Neubauer, DSc.: Nicolaus Copernicus University, Gagarina Street 11, 87-100 Toruń, Poland, aneub@umk.pl

[**] Assoc. Prof. Leszek Knopik, PhD.: Faculty of Management, UTP University of Science and Technology, Fordońska Street 430, 85-790 Bydgoszcz, Poland, knopikl@utp.edu.pl

[***] RNDr. Mirosław Szubartowski, DSc.: Management Business and Service, Fordońska Street 40, 85-719 Bydgoszcz, Poland, analityk@karor.com.pl

[†] Assoc. Prof. Klaudiusz Migawa, PhD.: Faculty of Mechanical Engineering, UTP University of Science and Technology, Al. prof. S. Kaliskiego 7, 85-789 Bydgoszcz, Poland, klaudiusz.migawa@utp.edu.pl

of compatibility with the gamma distribution. The last column contains p-values for the significance level for the compatibility test. All analyzed statistical samples have distributions consistent with the gamma distribution, because the calculated values of the λ-Kolmogorov statistics satisfy the inequality of λ-Kol < 1.36. The value 1.36 is read from the distribution tables of the Kolmogorov distribution at the significance level $p = 0.05$.

Table 1. Values of basic repair time statistics for paper machine and its subsystems.

System	Mean	STD	n	α	β	Σ^2	λ-Kol	p-value
All	101.02	265.10	917	0.9900	72.19	0.0070	1.1872	0.121
Automation	59.43	60.62	387	1.1020	48.21	0.0009	0.3627	0.999
Electrical	97.70	157.50	305	0.2484	314.79	0.0009	0.4375	0.993
Mechanical	177.06	487.65	225	0.4345	283.43	0.0010	0.2592	0.999

Table 2 gives the results of calculations to Table 1 including time of paper machine's worthiness. Worthiness time values are significantly greater than the time of repair. This is visible for the mean value, standard deviation and β parameter of the gamma distribution. All distributions are consistent with the gamma distribution. For the entire paper machine (All) results, compatibility with the gamma distribution is much better than in the case of repair time.

Table 2. Values of basic statistics for time of paper machine worthiness and that of its subsystems.

System	Mean	STD	n	α	β	Σ^2	λ-Kol	p-value
All	4474.12	6158.2	916	0.4600	10094	0.0020	0.6657	0.670
Automation	10748.93	13417.9	386	0.6532	16749	0.0008	0.3257	0.999
Electrical	13603.92	20044.9	304	0.5152	25794	0.0039	0.4906	0.996
Mechanical	18422.29	22927.3	224	0.6493	28790	0.0024	0.3554	0.999

Table 3 lists availability rate values for paper machine and for all subsystems tested. It is noteworthy that all subsystems have a very high availability rate more than 0.99. The availability rate for paper machine is 0.977921 and is clearly lower than the availability rate for each of its subsystems.

Table 3. Values of paper machine availability rate as well as its subsystems.

System	Availability
All	0.977921
Automation	0.994501
Electrical	0.992869
Mechanical	0.990480

3. Paper machine repair time model

The relatively poor compatibility of the repair time distribution with the gamma distribution results from the fact that the population of the repair time value is a heterogeneous set. In this case, a mixture of distributions is used. In order to simplify the considerations in this paper, a mixture of repair time distributions is examined for subsystems: automation, electrical and mechanical. Then the probability density f (x) for the mixture of three distributions of random variable has the following form:

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + p_3 f_3(x), \quad (1)$$

where: p_1, p_2, p_3 are fractions (shares) of the mixture components, the following equality is true $p_1 + p_2 + p_3 = 1$. Fractions p_1, p_2, p_3 may be assessed on the basis of the number of statistical samples. On the basis of data from Tables 1 and 2, the following was obtained: $p_1 = 387 / 917 = 0.422028$, $p_2 = 304 / 917 = 0.332606$ and $p_3 = 225 / 917 = 0.245365$. Functions $f_1(x)$, $f_2(x)$, $f_3(x)$ are probability densities for repair time of subsystems: automation, electrical and mechanical. Formula (1) results in mean values being true for equation:

$$ET = p_1\, ET_1 + p_2\, ET_2 + p_3\, ET_3, \tag{2}$$

where: ET is the mean value of paper machine repair time, ET_1 is the average value of the repair time for the automation subsystem, ET_2 is the mean of the repair time for the electrical subsystem, ET_3 is the average value of the repair time for the mechanical subsystem. It can be seen that equality (2) is true for fractions p_1, p_2, p_3 calculated above and for the mean values from Table 1.

The main purpose of this work is to show that the use of a mixture of gamma distributions of a random variable results in a much better estimation of the repair time distribution of a complex technical object, e.g. a paper machine, than when compared to a gamma distribution.

The mixture of repair time distributions for the paper machine subsystems (1) depends on the 9 values of the p_i, α_i, β_i parameters, where i = 1, 2, 3. The fractions p_1, p_2, p_3 are known and calculated above 6 values of the parameters α_i, β_i remain to be determined. The values of these parameters are determined by the EM algorithm used for mixtures of random variable distributions (Dempster et al., 1977). The calculated values of these parameters are shown in Table 4.

Table 4. Values of the parameters of gamma distribution mixture of repair time for paper machine subsystems repair time.

System	**α_i**	**β_i**	**p_i**
Automation	4.808943	13.9681	0.422028
Electrical	0.422578	374.0054	0.332606
Mechanical	0.341098	90.6831	0.245365

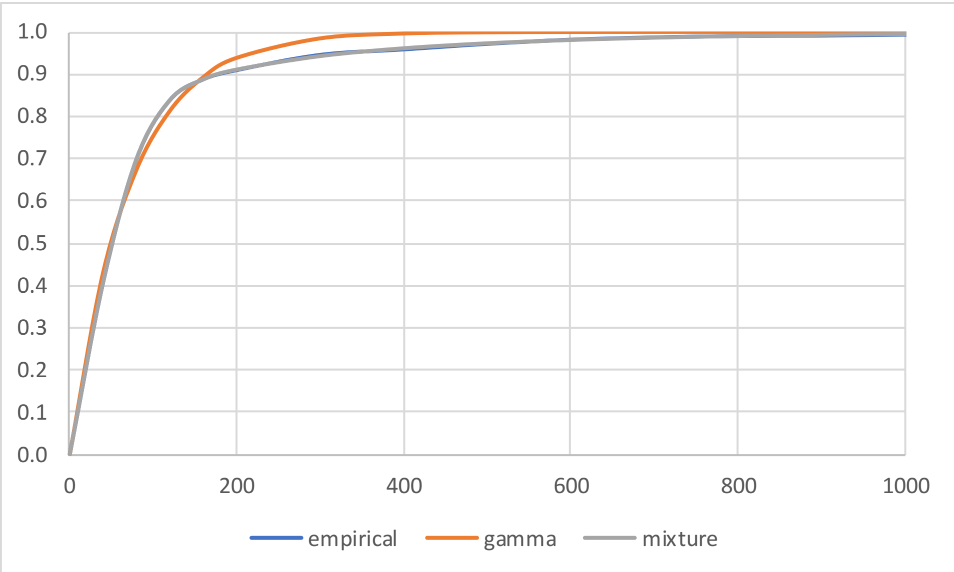

Fig. 1: Charts of distribution functions of paper machine repair time: empirical distribution, gamma distribution, gamma distributions mixture.

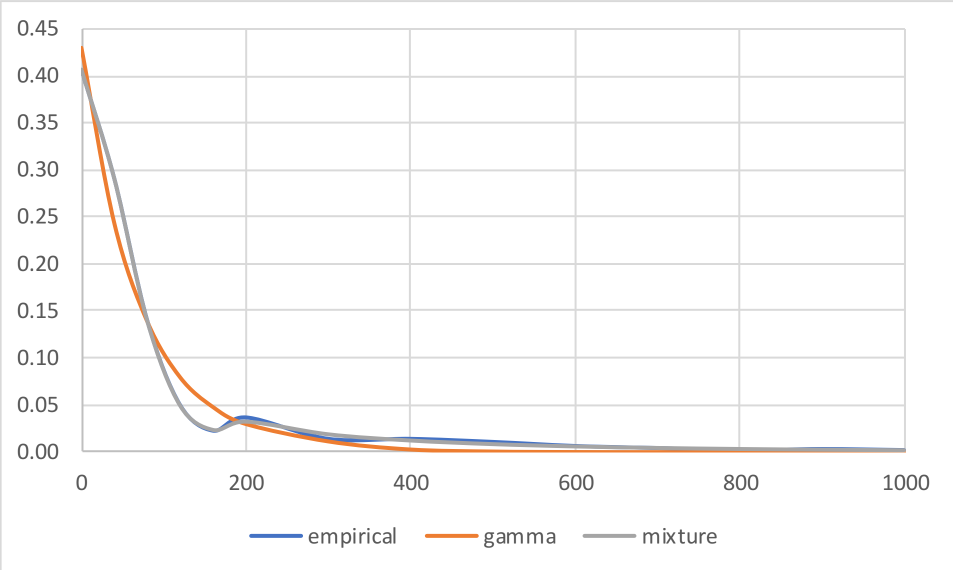

Fig. 2: Charts of paper machine repair time probability density functions: empirical distribution, gamma distribution, gamma distributions mixture.

For the mixture of repair time schedules (1) with the parameters from table 4, the sum of squared deviations Σ^2 = 0.000018 and the value of statistics λ-Kolmogorowa λ-Kol = 0.088877 were calculated. The comparison of the values obtained with the respective values from table 1 shows a very high efficiency of the mixture application for the construction of a paper machine repair time model. In addition, the compatibility of the empirical distribution of repair time with mixture distribution (1) was tested using the χ^2-Pearson conformity test. If the mixture was used, χ^2 = 2.51, p-value = 0.77 was obtained, while for parameters from the first row of table 1 χ^2 = 22.98, p-value = 0.004 was calculated. The analysis of the results of the χ^2-Pearson conformity test confirms the high compatibility of the empirical distribution of the repair time with the distribution of the given mixture (1), in contrast to the lack of compatibility with the gamma distribution. In the figures below, there are diagrams for the distribution function (Fig. 1) and the probability density function (Fig. 2) of the paper machine repair time for: empirical distribution, gamma distribution and gamma distributions mixture. The charts of distribution function of empirical distribution as well as of gamma distributions mixture are practically the same.

4. Conclusions

In this paper, on the basis of a large set of operational data, the usefulness of mixtures of probability distributions of random variable in development of probabilistic models applicable in the analysis of the reliability of complex technical objects, e.g. paper machines, has been shown. The mixture of distributions of repair time of paper machine subsystems (automation, electrical and mechanical subsystems) was analyzed. On the basis of the obtained results, it can be concluded that the mixture of analyzed random variables determined in the paper is a better model of paper machine repair time than a model developed with the use of only one distribution (gamma distribution).

References

Bishop, C.M. (2006) Pattern recognition and mixture Learning. Springer.

Bohning, D. (2000) Computer assisted analysis of mixture and application, meta-analysis, disease mapping and other. Chapman & Hall.

Dempster, A.P., Larid, N.M. and Rubin, D.B. (1977) Maximum likelihood from incomplete data via EM algorithm. Journal of the Royal Statistical Society, 39(1), pp. 1-38.

Everitt, B. and Hand, D.J. (1981) Finite mixture distribution. Chapman & Hall.

McLachan, G.J. and Peel, D. (2000) Finite mixture models. Wiley. New York.

Tettrington, D.M., Smith, A.F.M. and Makov, U.E. (1985) Statistical analysis of finite distributions. Wiley.