Journal of POLISH CIMAC
Faculty of Ocean Engineering & Ship Technology
GDAŃSK UNIVERSITY OF TECHNOLOGY

ASSESSMENT OF FAILURE DISTRIBUTIONS OF MARINE POWER PLANTS FUEL OIL SYSTEMS GROUP

Zbigniew Matuszak, Grzegorz Nicewicz

Maritime Academy of Szczecin
ul. Wały Chrobrego 1-2, 70-500 Szczecin, Poland
tel.: +48 91 4809414, +48 91 4809442
e-mail: zbimat@am.szczecin.pl, nicze1@wp.pl

Abstract

For the failure moments and the time between failures of the marine power plants fuel oil systems, an attempt of defining the failure distributions has been made. The analysis was based on the observations of the failure of marine power plants fuel oil systems elements. Due to the character of the statistical data, the analysis deals only with relatively simple distributions, most frequently used in the reliability theory. Failures to the marine power plant systems of 10 ships owned by the Polish Steamship Company of Szczecin was the subject of a statistical data analysis. All the ships differed in respect to their place and time of construction as well as their technical parameters. Data on their failures refer to the fuel oil system. The data on marine power plants failures were collected in similar conditions, that is, they were supplied by an engine crew member working in the marine power plant. The data on the failures of particular marine power plant systems were obtained accordingly to the test schedule [N, W, T], which means that N renewable objects were the subject of the test within the time T. Since the recovery time of the damaged system appeared negligibly short, when compared to the time of the test, it was assumed that consecutive recoveries overlap the failure moments. The statistical analysis dealt with moments $t_1 \le t_2 \le ... \le t_n$ *of the particular systems' consecutive failures and the length of time intervals* τ_n *between the objects' consecutive failures. In order to carry out the distribution estimation, the statistical package STATISTICA was used.*

***Keywords**: marine power plant, fuel oil system, failure distributions*

1. Introduction

There has been carried out a statistical analysis of the data concerning marine power plants systems failures on ten ships owned by the Polish Steamship Company of Szczecin. The ships were given symbols from S1 to S10. All the ships differed in respect to their place and time of construction as well as their technical parameters. The data on failures cover the following six marine power plant systems: lubricating oil system - LOS, sea water cooling system - SWCS, fresh water cooling system - FWCS, fuel oil system - FOS, compressed air system - CAS and steam system - SS [2, 3].

The data on marine power plants failures were collected in similar conditions, that is, they were supplied by an engine crew member working in the marine power plant.

The data on the failures of particular marine power plant systems were obtained accordingly to the test schedule [*N*, *W*, *T*], which means that *N* renewable objects were the subject of the test within the time *T*. Since the recovery time of the damaged system appeared negligibly short, when compared to the time of the test, it was assumed that consecutive recoveries overlap the failure moments.

The statistical analysis dealt with moments $t_1 \le t_2 \le ... \le t_n$ of the particular systems' consecutive failures and the length of time intervals τ_n between the objects' consecutive failures.

The following has been assumed [2, 3]:
1) on ships S1 - S5 time is measured from the moment of the first failure repair;
2) for ships S6 - S10 time intervals τ_n between the consecutive failures do not include the time between the beginning of sea voyage and the first failure;
3) time between the last failure and the end of the observation, that is after 180 days, was taken into consideration.

Total numbers of failures in particular systems of ten tested ships' marine power plants have been presented in table 1.

Table 1. Total number of failures to particular marine power plant systems of all the tested ships

Name of the system	Total number of failures										
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Sum
Lubricating oil system - LOS	3	1	5	6	2	3	21	8	8	9	66
Sea water cooling system - SWCS	7	3	4	4	2	11	5	8	8	5	57
Fresh water cooling system - FWCS	6	2	6	3	2	7	8	9	5	7	55
Fuel oil system - FOS	8	12	5	12	6	19	23	17	7	12	121
Compressed air system - CAS	2	1	2	2	2	5	5	8	3	6	36
Steam system - SS	4	6	3	2	3	10	4	9	3	6	50
Sum	30	25	25	29	17	55	66	59	34	45	385

On the basis of the table, most failures, that is 30% of all of them, can be pointed out in the fuel oil system. It is to be noticed, that among all the tested ships, the prevailing number of failures occurred in case of three of them, that is, S6, S7 and S8.

On the basis of the obtained field data on the system failures, for particular systems, there was built a mathematical model of distribution of time to failure and time between failures.

Due to the type of the statistical data, the most favored distributions appeared relatively simple ones, most frequently applied to the reliability theory. Thus, the possibility of building a model based on the following distributions was sequentially examined:
- the exponential distribution, with probability density function

$$f(t) = \lambda e^{-\lambda t} \quad \text{dla} \quad t \ge 0 \ , \tag{1}$$

- the Weibull distribution , with probability density function

$$f(t) = \alpha \beta^{-\alpha} t^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^{\alpha}} \quad \text{dla} \quad t > 0 \ , \tag{2}$$

- the log-normal distribution, with probability density function

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} \quad \text{dla} \quad t > 0 \ , \tag{3}$$

- the gamma distribution, with probability density function

$$f(t) = \frac{\beta^{\alpha} t^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta t} \quad \text{dla} \quad t > 0 \, . \tag{4}$$

The results of the Kruskal-Wallis test, aided by the program *STATISTICA 8.0*, for the statistical data of fuel oil system failure moments t_n concerning all the 10 tested ships, at the assumed significance level α=0,05, turn out to come from one general population (for space reasons the analysis has not been enclosed). However, when considering the statistical data concerning the time between the fuel oil system failures (τ_n) of the 10 analyzed transport ships at the significance level α=0,05, the hypothesis that they come from one general population needs to be rejected.

Thus, multiple comparisons of rank means (*post-hoc* tests) for the data of the time between failures (τ_n) in reference to each pair of the 10 ships, using the program *STATISTICA 8.0* have been performed. For each comparison, there have been computed '*z'* values and '*p*' values. On the basis of the obtained results it has been concluded that data from ships S3, S5 and S7 are significantly different from each other. Therefore, it was assumed that the data could belong to various populations. However, to the data from S1, S2, S4, S6, S8, S9 and S10 there was applied the Kruskal-Wallis test, whose results point out that there are no bases for rejecting the hypothesis that they come from one general population at the significance level α=0,05. A broader presentation of the above analyses goes far beyond the frames of the paper and would require a separate publication, what has initially been done in [1].

On the basis of the obtained results, data on fuel oil system failures of S1, S2, S4, S6, S8, S9, S10 concerning both consecutive failure moments (t_n) and time between failures time ($\tau_{n,}$) need to be treated as coming from one general population. Therefore, an attempt of estimation of the obtained empirical distributions by means of the relatively simple theoretical exponential, gamma, log-normal and Weibull distributions, most frequently applied to the reliability theory. In case the exponential, gamma, or log-normal distributions have been accepted as the models, the program *STATISTICA 8.0* enables adjustment and application of tests of goodness of fit. But in reference to the Weibull distribution, it only allows for the estimation of the distribution parameters by the Maximum Likelihood test. Such an analysis has been carried out by means of the statistical package *STATGRAPHICS,* enclosed in [2].

2. Tests of goodness of fit

Tests of goodness of fit allow for the verification of the null hypothesis that an analyzed random variable has a distribution which belongs to the appropriate distribution group. The oldest test of goodness of fit appears Pearson's chi-square test, first presented in 1900 [4]. The test can not be applied to small size samples (n<50) [4, 5]. In case of small size samples the Kolmogorov-Smirnov test is used.

Program *STATISTICA* [4, 5] performs the chi-square test by default on the basis of the number of the observed and expected frequencies. Categories containing expected frequencies lower than 5 are grouped to make classes of higher frequency. The number of degrees of freedom for chi-square test statistics is computed in the following way:

df = *number of categories* - *number of parameters* - 1,

where:

- *number of categories* refers to the number of classes in the frequencies table, where the expected frequencies are higher than 5;
- *number of parameters* refers to the parameters of the appropriate theoretical distribution.

If the chi-square test outcome is marked with "df adjusted", it means that to compute chi-square test statistic, the program joined the categories with expected frequencies lower than 5. In particular, the neighboring categories are joined until their frequency becomes higher than 5.

If the test demonstrates statistical significance (that is, *p*-value is lower than the assumed significance level, usually 0,05), the hypothesis, that the observed data are liable to a specific distribution, is rejected [4, 5].

The Kolmogorov-Smirnov test can be computed, or not, depending upon the settings on the chart *Options*. In case of two samples, the Kolmogorov-Smirnov test is the test of statistical significance differences between them. Just like the test for one sample, the test statistic compares cumulative distribution functions, in this case it deals with cumulative distribution functions of two samples (e.g. the observed and the simulated values). Big difference between the cumulative distribution functions shows that the data come from two different populations. The test may turn out useful when building a model for assessment whether the expected results (based on the simulation input data) differ from the observed ones. A significant difference between the expected and observed outcomes usually reveals the model's insufficiency for demonstrating dependencies between the input and output data [4, 5].

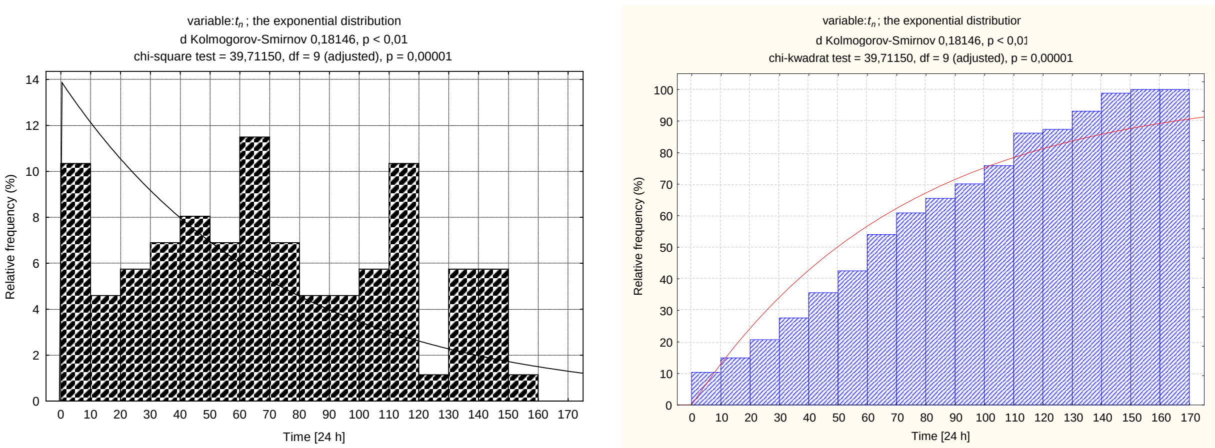
In the *Kolmogorov-Smirnov test* group there may be set options for Kolmogorov-Smirnov test. Calculations can be done on the basis of categorized data by clicking the button *Yes (categorized)* - (computing is faster) or using raw data (the procedure is slower; button *Yes (continuous)* must be pressed to make use of this method). In the first case the program calculates the value of D-max statistic on the basis of grouped data, whereas in the second case, at every point it performs calculations of cumulative expected frequencies based on the ordered data. Kolmogorov D-max statistic is the biggest of the absolute differences between the observed and expected cumulative values of the distribution. If the Kolmogorov-Smirnov test shows statistical significance (that is, the p –value appears lower than the assumed significance level), we may reject the hypothesis assuming that the observed data are subject to the hypothetical distribution.

Using the *STATISTICA 8.0* program (chart *Distribution fitting*), the verifications of null hypothesis of fitting the distribution of the failure moments (t_n) and the time between failures (τ_n) in marine power plants fuel oil systems with the following theoretical distributions: exponential, gamma and log-normal were carried out; for this purpose Kolmogorov-Smirnov (continuous) test was applied. For reasons of space the *STATISTICA 8.0* spreadsheets with the observed and expected frequencies and the outcome of the tests of goodness of fit data with theoretical distributions fit have not been included. The results of the tests of goodness of fit have been shown on histograms and cumulative histograms of marine power plants fuel oil systems failure moments (t_n) and the time between failures (τ_n) with probability density functions and cumulative distribution functions of the discussed theoretical distributions.

3. Results of the tests of goodness of fit of the consecutive failure moments

The performance of the Kolmogorov-Smirnov test for the exponential distribution, has resulted in its following parameters: number of categories: 17; lower limit: 0; upper limit: 170; lambda: 0,0139; observed mean: 71,7130; observed variance: 1907,3664.

Histogram and cumulative histogram of the consecutive marine power plants fuel oil systems failure moments with the probability density function of the exponential distribution and the cumulative distribution function of the exponential distribution and with the Kolmogorov-Smirnov test results inserted have been shown in fig. 1.



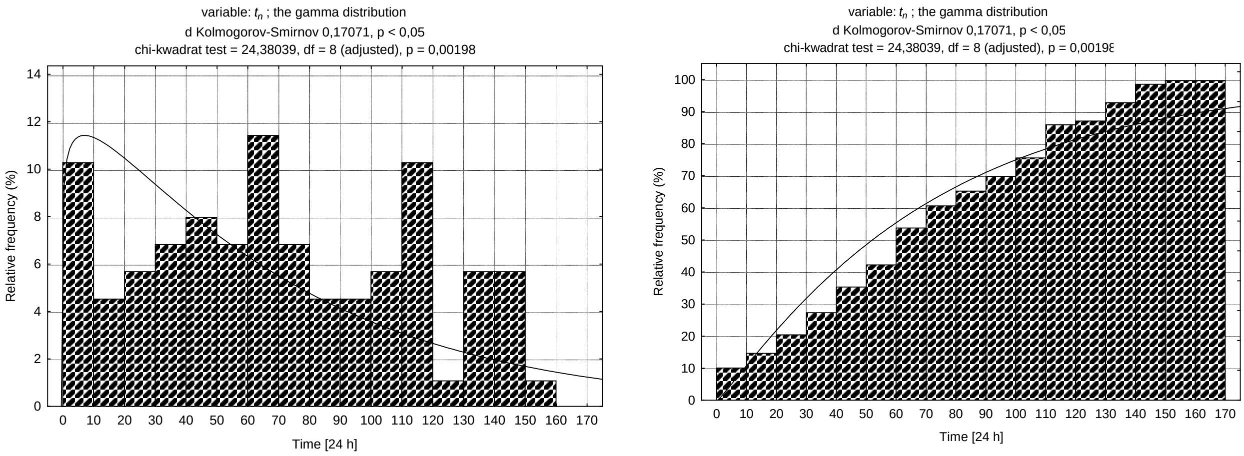
a) *b)*

Fig. 1. Histogram (a) and cumulative histogram (b) of the consecutive failure moments of marine power plants fuel oil systems with probability density function of the exponential distribution (full line) (a) and cumulative distribution function of the exponential distribution (full line) (b) and with the results of the Kolmogorov-Smirnov test inserted

There are no bases for accepting the null hypothesis that the empirical distribution of the marine power plants fuel systems consecutive failure moments t_n complies with the exponential distribution.

The Kolmogorov-Smirnov test applied to the gamma distribution has resulted in its following parameters: number of categories: 17; lower limit: 0; upper limit: 170; scale parameter: 64,6862, shape parameter: 3,2067; observed mean: 71,7130; observed variance – 1907,3664.

Histogram and cumulative histogram of the consecutive marine power plants fuel oil systems failure moments with the probability density function of the exponential distribution and the cumulative distribution function of the exponential distribution and with the Kolmogorov-Smirnov test results, inserted, have been shown in fig. 2.



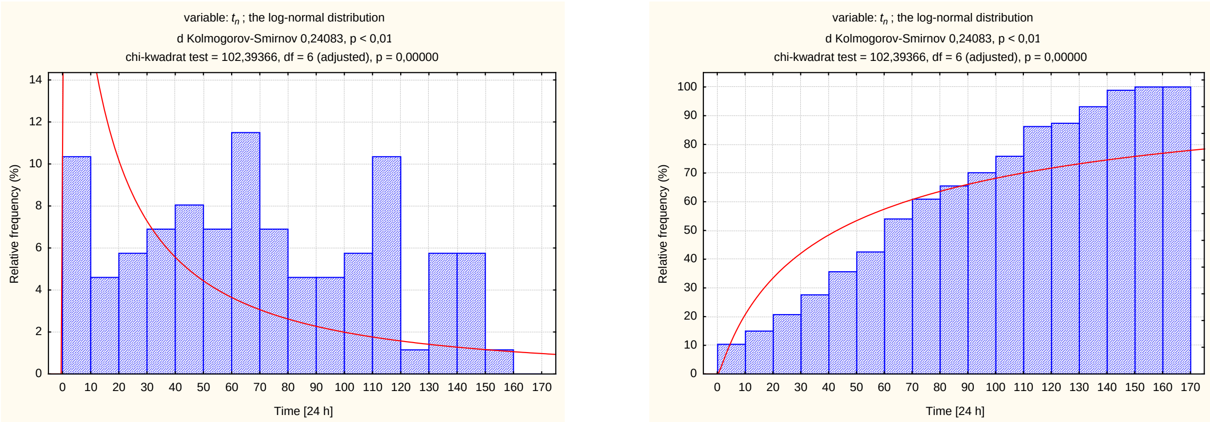
a) *b)*

Fig. 2. Histogram (a) and cumulative histogram (b) of the consecutive failure moments of marine power plants fuel oil systems with probability density function of the gamma distribution (full line) (a) and cumulative distribution function of the gamma distribution (full line) (b) and with the results of the Kolmogorov-Smirnov test inserted

There are no bases for accepting the null hypothesis assuming that the empirical distribution of the marine power plants fuel oil systems consecutive failure moments t_n complies with the gamma distribution.

The Kolmogorov-Smirnov test applied to the log-normal distribution has resulted in its following parameters: number of categories:17; lower limit: 0; upper limit: 170; mean (M): 3,7573; variance: 3,20669; observed mean: 71,7130, observed variance: 1907,3664.

Histogram and cumulative histogram of the consecutive marine power plants fuel oil systems failure moments with the probability density function of the log-normal distribution and the cumulative distribution function of the log-normal distribution and with the Kolmogorov-Smirnov test results, inserted, have been shown in fig. 3.



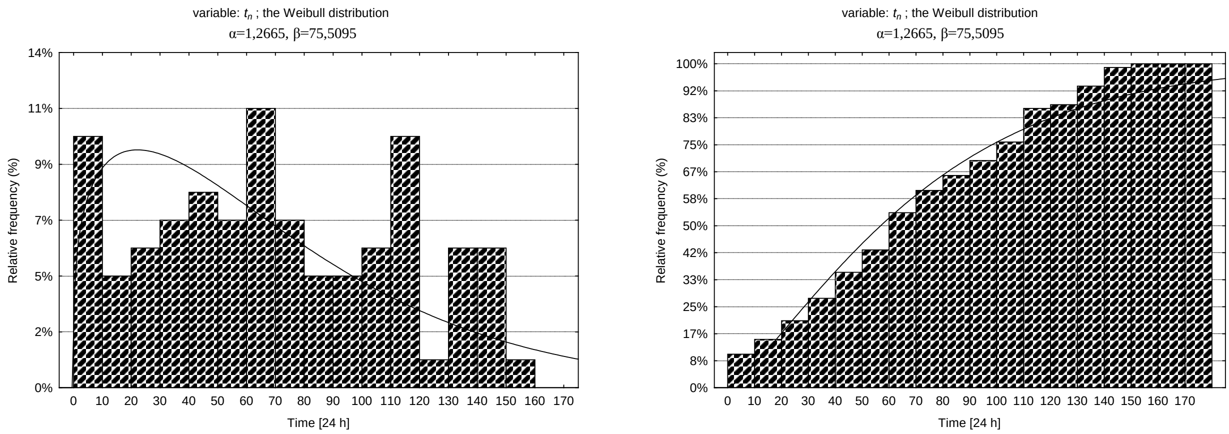
a) *b)*

Fig. 3. Histogram (a) and cumulative histogram (b) of the consecutive failure moments of marine power plants fuel oil systems with probability density function of the log-normal distribution (full line) (a) and cumulative distribution function of the log-normal distribution (full line) (b) and with the results of the Kolmogorov-Smirnov test inserted

There are no bases for accepting the null hypothesis that the empirical distribution of the marine power plants fuel systems consecutive failure moments t_n complies with the log-normal distribution.

To define the goodness of fit of the obtained empirical distribution of the marine power plants fuel oil systems' consecutive failure moments t_n with the Weibull distribution, a graphical method (the window *Graphs 2D>Histograms 2D* of the program *STATISTICA 8.0*) have been used. Parameters of the Weibull distribution α, β have been estimated by the Maximum Likelihood test. The obtained parameters for the distribution are as follows: number of categories: 17; lower limit: 0; upper limit: 170; α=1,2665; β=75,5095.

Histogram and cumulative histogram of the consecutive marine power plants fuel oil systems failure moments with the probability density function of the Weibull distribution and the cumulative distribution function of the Weibull distribution have been shown in fig. 4.



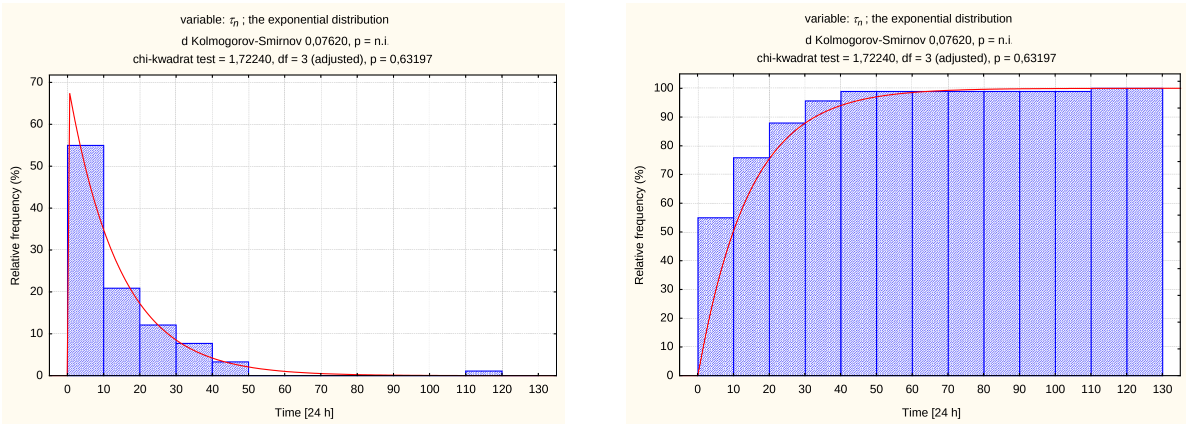
a) *b)*

Fig. 4. Histogram (a) and cumulative histogram (b) of the consecutive failure moments of marine power plants fuel oil systems with probability density function of the Weibull distribution (full line) (a) and cumulative distribution function of the Weibull distribution (full line) (b) with the distribution parameters α=1,2665, β=75,5095

4. Results of the tests of goodness of fit of the time between failures

The Kolmogorov-Smirnov test applied to the exponential distribution has resulted in its following parameters: number of categories: 13; lower limit: 0; upper limit: 130; lambda: 0,0704; observed mean: 14,1979, observed variance: 247,5795.

Histogram and cumulative histogram of the time between the failures of marine power plants fuel oil systems with the probability density function of the exponential distribution and the cumulative distribution function of the exponential distribution and with the Kolmogorov-Smirnov test results, inserted, have been shown in fig. 5.



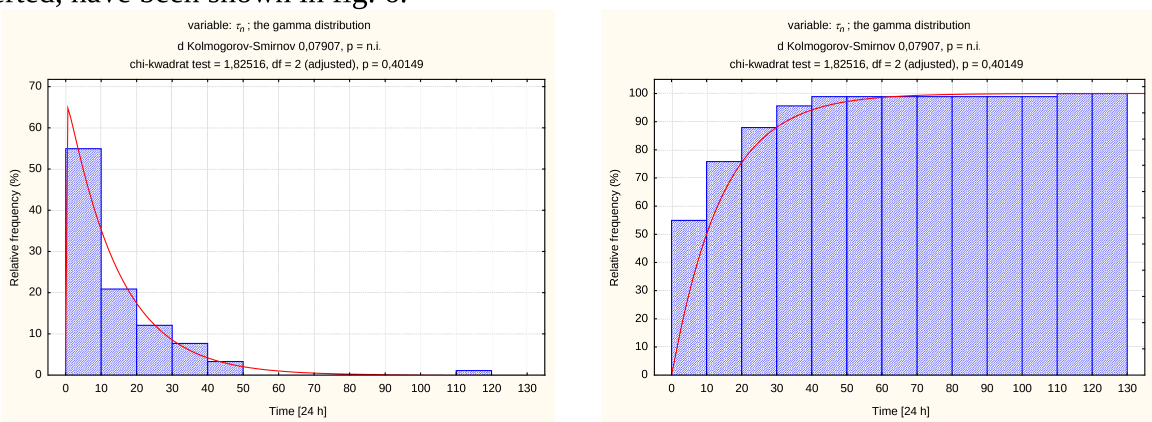
a) *b)*

Fig. 5. Histogram (a) and cumulative histogram (b) of the time between failures of marine power plants fuel oil systems with probability density function of the exponential distribution (full line) (a) and cumulative distribution function of the exponential distribution (full line) (b) and with the results of the Kolmogorov-Smirnov test inserted

There are no bases for rejecting the null hypothesis that the empirical distribution of the time between failures τ_n of the marine power plants fuel oil systems complies with the exponential distribution.

The Kolmogorov-Smirnov test applied to the gamma distribution has resulted in its following parameters: number of categories: 13; lower limit: 0; upper limit: 130; scale parameter: 13,8525; shape parameter: 1,6149; observed mean: 14,1979; observed variance: 247,5795.

Histogram and cumulative histogram of the time between the failures of marine power plants fuel oil systems with the probability density function of the gamma distribution and the cumulative distribution function of the gamma distribution and with the Kolmogorov-Smirnov test results, inserted, have been shown in fig. 6.



a) *b)*

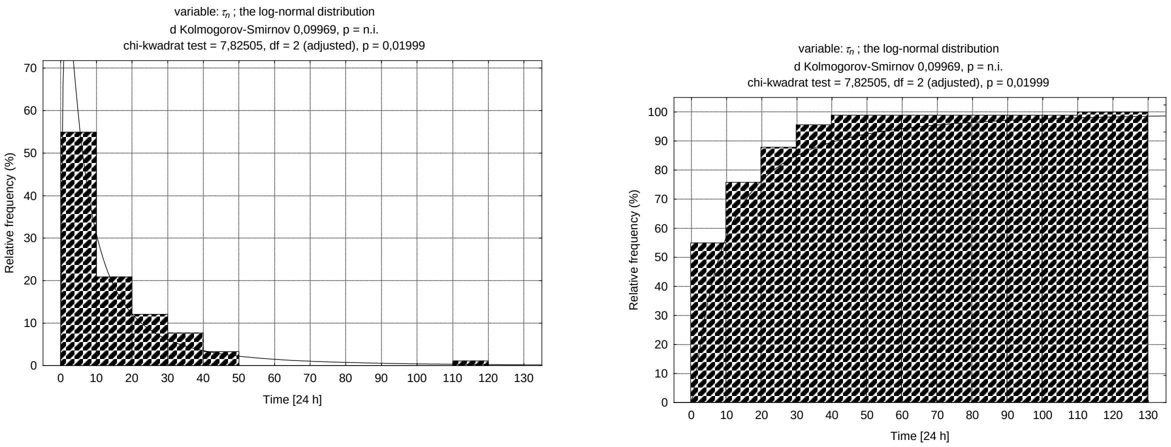
Fig. 6. Histogram (a) and cumulative histogram (b) of the time between failures of marine power plants fuel oil systems with probability density function of the gamma distribution (full line) (a) and cumulative distribution function of the gamma distribution (full line) (b) and with the results of the Kolmogorov-Smirnov test inserted

There are no bases for rejecting the null hypothesis that the empirical distribution of the time between failures τ_n of the marine power plant fuel system complies with the gamma distribution.

The Kolmogorov-Smirnov test applied to the log-normal distribution has resulted in its following parameters: number of categories: 13; lower limit: 0; upper limit: 130; mean (M):

2,0908; variance: 1,6149; observed mean: 14,1979; observed variance: 247,5795.

Histogram and cumulative histogram of the time between the failures of marine power plants fuel oil systems with the probability density function of the log-normal distribution and the cumulative distribution function of the log-normal distribution and with the Kolmogorov-Smirnov test results, inserted, have been shown in fig. 7.



a) *b)*

Fig. 7. Histogram (a) and cumulative histogram (b) of the time between failures of marine power plants fuel oil systems with probability density function of the log-normal distribution (full line) (a) and cumulative distribution function of the log-normal distribution (full line) (b) and with the results of the Kolmogorov-Smirnov test inserted

There are no bases for rejecting the null hypothesis that the empirical distribution of the time between failures (τ_n) of the marine power plants fuel oil systems complies with the log-normal distribution.

To define the goodness of fit of the obtained empirical distribution of the time between failures in the marine power plants fuel oil systems (τ_n) with the Weibull distribution, a graphical method (the window *Graphs 2D>Histograms 2D* of the program *STATISTICA 8.0*) have been used. Parameters of the Weibull distribution α, β have been estimated by the Maximum Likelihood test. The obtained parameters for the distribution are as follows: number of categories: 13; lower limit: 0; upper limit: 130; α=0,9928; β=14,1522.

Histogram and cumulative histogram of the time between failures (τ_n) of marine power plants fuel oil systems and the probability density function of the Weibull distribution and the cumulative distribution function of the Weibull distribution have been shown in fig. 8.

a) *b)*

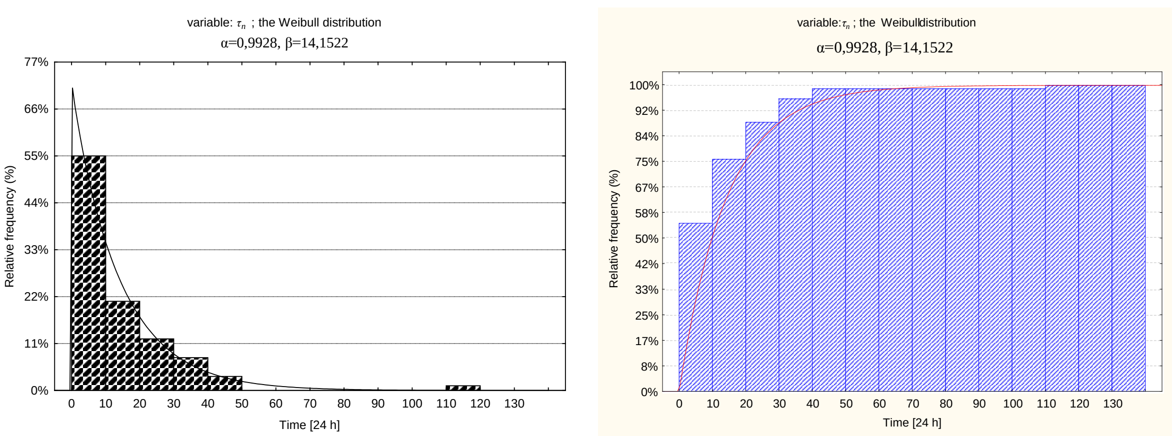


Fig. 8. Histogram (a) and cumulative histogram (b) of the time between failures of marine power plants fuel oil systems with probability density function of the Weibull distribution (full line) (a) and cumulative distribution function of the Weibull distribution (full line) (b) with the distribution parameters: α=0,9928, β=14,1522

5. Final remarks

In case of the consecutive failure moments of the marine power plants fuel oil systems, non of the discussed theoretical distributions turns out a sufficiently appropriate model for the obtained empirical distribution.

However, for the time between failures of the marine power plants fuel oil systems, there are no bases for rejecting the null hypothesis assuming the obtained empirical distribution comply with each of the considered theoretical distributions, that is, the exponential, gamma, log-normal and Weibull.

Program *STATISTICA 8.0* enabled the fit of the theoretical distribution and carrying out the tests of goodness of fit in case of the assumption that exponential, gamma and log-normal distributions become the models. But in reference to the Weibull distribution, due to the Maximum Likelihood test, only the distribution parameters could be estimated. The analysis carried out by means of a statistical package *STATGRAPHICS*, presented in [2], in reference to the Weibull distribution, brought slightly different, more unique results; but it was made by a much older statistical package than *STATISTICA 8.0*, and for the space reason it was not possible to discuss in details the comparison of the test results obtained by means of both statistical packages.

References

[1] Chybowski L., Matuszak Z., *Porównanie kompleksowych pakietów oprogramowania do prowadzenia wieloaspektowej analizy niezawodności.* Sbornik naucznych trudow mieżdunarodnoj baltijskoj associacii maszinostroitielej BALTTECHMASZ -2006, Izdatelstwo KGTU, Kaliningrad 2006, pp. 277-289.

[2] Matuszak Z., *Modeli otkazow i prinadleżnost dannych ob otkazach k generalnoj sowokupnosti na primierie sudowych energeticzeskich ustanowok.* Kaliningradskij gosudarstwiennyj techniczeskij uniwersytet, Kaliningrad 2002, Monografia.

[3] Matuszak Z., *Kompozicii raspredelenij charakteristik nadiożnosti i modeli otkazow sistem sudowych energeticzeskich ustanowok.* Kaliningradskij gosudarstwiennyj techniczeskij uniwersytet, Kaliningrad 2003, Monografia.

[4] Stanisz A., *Przystępny kurs statystyki. Tom 1: Statystyki podstawowe*. StatSoft, Kraków 2006.

[5] StatSoft – Statistica 8.0, *Podręcznik elektroniczny STATISTICA*.