Scientific Journals
of the Maritime University of Szczecin

Zeszyty Naukowe
Akademii Morskiej w Szczecinie

2016, 47 (119), 80–84
ISSN 1733-8670 (Printed)
ISSN 2392-0378 (Online)
DOI: 10.17402/152

Received: *29.04.2016*
Accepted: *15.09.2016*
Published: *20.09.2016*

Extreme value distributions in the analysis of traffic time on the Świnoujście–Szczecin fairway

Lech Kasyk

Maritime University of Szczecin
1–2 Wały Chrobrego St., 70-500 Szczecin, Poland, e-mail: l.kasyk@am.szczecin.pl

Key words: traffic time, fairway, generalized extreme value distributions, Frechet distribution, crossing time, simulated models

Abstract

The present article addresses the issue of crossing time on the fairway, modeling in restricted areas, where vessel traffic flow is disturbed. Data of movement time on the Świnoujście–Szczecin fairway was grouped according to ship type. The probability distributions describing the crossing time of different ship groups were analyzed. Using the Pearson chi-square goodness-of-fit and Cramer–von Mises tests it has been shown that the best distributions describing traffic time of all ship groups are the generalized extreme value distributions.

Introduction

To provide solutions to the many problems of marine traffic engineering, it is first necessary to know the probability distributions of paricular random variables. This is very important, especially in simulated models (Gucma, Gucma & Zalewski, 2008). Recently, only certain standard distributions (e.g. normal, lognormal, uniform, exponential and Poisson) have been used for analysis of marine traffic (Kasyk, 2012).

The generalized extreme value distribution is a family of continuous probability distributions used in extreme value theory, especially in the economic and social sciences. Belonging to this family, among others, are: a Weibull distribution, a Gumbel distribution, an extreme value distribution and a Frechet distribution. Applications of Weibull distributions in different transport problems have been known for a long time (Gucma & Jankowski, 2001; Curbach, Gucma & Proske, 2005; Gucma, 2005; Smolarek, 2005). However, recently, distributions from this family have been used in the analysis of vessel traffic flows. The vessel speed in restricted areas, where a speed limit exists, may be described and modeled by Gumbel distribution (Kasyk & Kijewska, 2014). The speed limit means that we deal with a distribution of extreme values. It is similar in the case of traffic time in restricted areas, where a speed limit exists.

The probability density function of the Weibull distribution is as follows:

$$f(x)=\frac{\lambda}{\beta}\exp\left(-\left(\frac{x-\alpha}{\beta}\right)^{\lambda}\right)\left(\frac{x-\alpha}{\beta}\right)^{\lambda-1} \tag{1}$$

The probability density function of the Gumbel distribution is as follows:

$$f(x)=\frac{\exp\left(\frac{x-\alpha}{\beta}-\exp\left(\frac{x-\alpha}{\beta}\right)\right)}{\beta} \tag{2}$$

The probability density function of the Frechet distribution is as follows:

$$f(x)=\frac{\lambda}{\beta}\exp\left(-\left(\frac{x-\alpha}{\beta}\right)^{-\lambda}\right)\left(\frac{x-\alpha}{\beta}\right)^{-1-\lambda} \tag{3}$$

The probability density function of the extreme value distribution is as follows:

$$f(x) = \frac{\exp\left(\frac{\alpha - x}{\beta} - \exp\left(\frac{\alpha - x}{\beta}\right)\right)}{\beta} \tag{4}$$

In all above formulas, α is a location parameter, β is a scale parameter and λ is a shape parameter.

Below, in Table 1, formulas for means and variances of the aforementioned distributions are presented.

Table 1. Means and variances of extreme value distributions

Distribution	Mean	Variance
Weibull	$\alpha + \beta \cdot \Gamma\left(1 + \frac{1}{\lambda}\right)$	$\beta^2\left(\Gamma\left(1 + \frac{2}{\lambda}\right) - \Gamma\left(1 + \frac{1}{\lambda}\right)^2\right)$
Gumbel	$\alpha - \gamma \cdot \beta$	$\pi^2\beta^2/6$
Frechet	$\alpha + \beta \cdot \Gamma\left(1 - \frac{1}{\lambda}\right)$	$\beta^2\left(\Gamma\left(1 - \frac{2}{\lambda}\right) - \Gamma\left(1 - \frac{1}{\lambda}\right)^2\right)$
Extreme value distribution	$\alpha + \gamma \cdot \beta$	$\pi^2\beta^2/6$

where:

γ – Euler's constant, with numerical value ≈ 0.577216;

$\Gamma(z)$ – the Euler gamma function, which satisfies $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t$.

The vessel traffic flow is not homogeneous. Different groups of ships have their own distributions. Consequently, in this article the data set of traffic time has been divided into seven groups, for the following types:

- barges,
- tankers,
- containers,
- cargo,
- general cargo,
- carriers,
- other ships.

For particular ship groups, a traffic time, from point Karsibor to point Dok5 in Szczecin harbor has been analyzed. Data from VTS Szczecin have been used. In hypothesis tests, a small p-value suggests that it is unlikely that the data came from the considered distribution (significance level is 0.05).

Distribution of traffic time on the Świnoujście–Szczecin fairway

The *Karsibor – Dok5* section is a part of the Świnoujście–Szczecin fairway, which is 53.7 km long. In the first half of 2009, in the north-south direction, 739 ships were registered on this section.

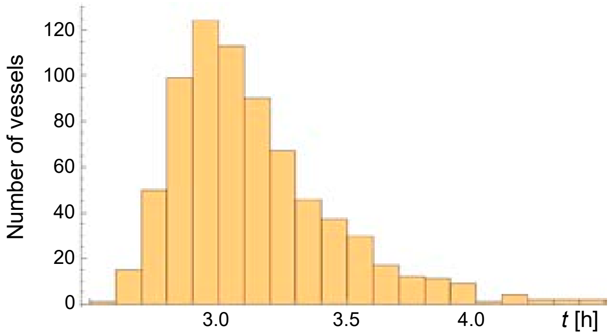

Figure 1. Histogram of traffic time on the Karsibor – Dok5 section

The minimal traffic time on this section is equal to 2.585 h, and the maximal traffic time is equal to 7.25 h. The mean time for all ships amounts to 3.16 h. And the standard deviation is equal to 0.41 h.

Table 2 presents the division of this set into seven groups.

Table 2. The number of different types of vessels on the Karsibor – Dok5 section

No.	Vessel type	Number of vessels
1.	Barges	27
2.	Tankers	67
3.	Containers	69
4.	Cargo	366
5.	General cargo	134
6.	Carriers	49
7.	Other ships	27

Traffic time distribution for barges

Figure 2 presents the frequency histogram and the graph of the probability density function of the extreme value distribution, fitted to data for barges.

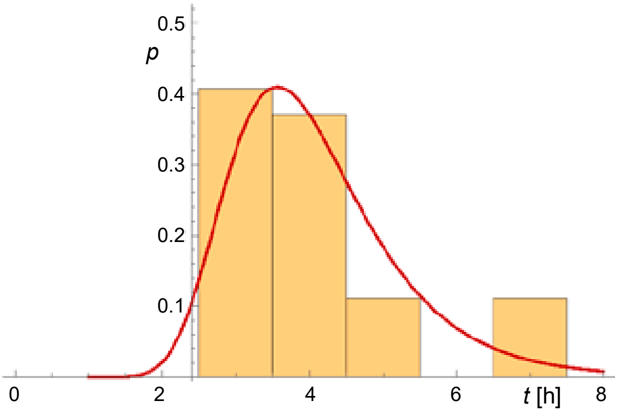

Figure 2. Frequency histogram of traffic time for barges

In the Pearson goodness-of-fit test chi-square test (Sobczyk, 2004; Rogowski, 2015) $p = 0.624$. In the

Cramér–von Mises goodness-of-fit test (Rogowski, 2015) $p = 0.805$. We are therefore unable to reject the hypothesis that the traffic time for barges between points 11 km and I Brama Torowa has an extreme value distribution. In this case, location parameter $\alpha = 3.566$ and scale parameter $\beta = 0.898$. The Empirical mean of data is equal to 4.078 h and theoretical mean from the extreme value distribution is equal to 4.084 h. Relative difference between these values is equal to 0.2%. Empirical variance of data is equal to 1.469 and theoretical variance from the extreme value distribution is equal to 1.326. Relative difference between these values is equal to 10%.

Traffic time distribution for tankers

Figure 3 presents the frequency histogram of data connected with tankers' traffic time.

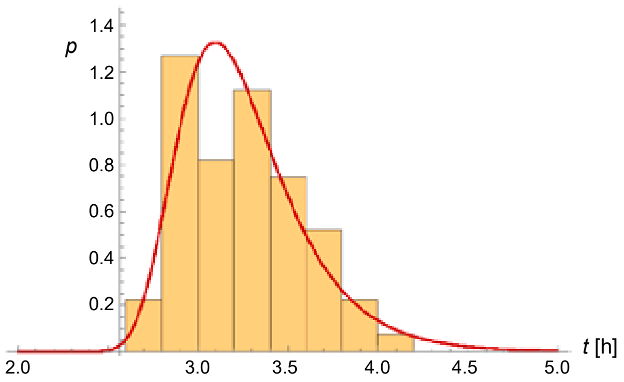

Figure 3. Frequency histogram of traffic time for tankers

Using the chi-square goodness-of-fit test we find that the $p = 0.47$. In the Cramér–von Mises goodness-of-fit test $p = 0.929$. So we can say that the traffic time for tankers has an extreme value distribution. In this case, location parameter $\alpha = 3.097$ and scale parameter $\beta = 0.277$.

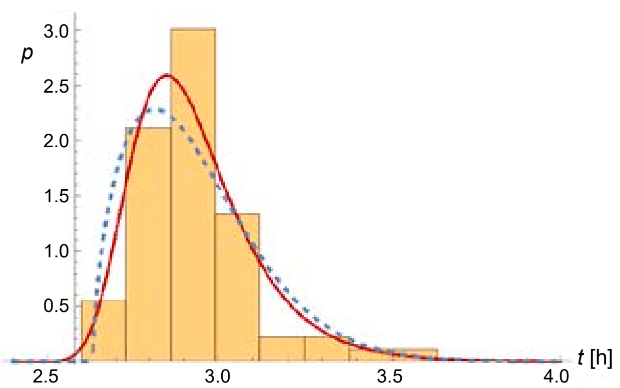

Figure 4. Frequency histogram of traffic time for containers

Traffic time distribution for containers

Figure 4 presents the histogram and the graph of extreme value distribution and Weibull (dashed) probability density function, fitted to data for containers.

The Pearson chi-square test and Cramér–von Mises test have shown that two distributions from the family of generalized extreme value distribution tests adequately describe the traffic time of containers between reporting points 11 km and I Brama Torowa. There are: the extreme value distribution with parameters $\alpha = 2.848$ and $\beta = 0.142$, and the Weibull distribution with parameters $\alpha = 2.633$, $\beta = 0.333$ and $\gamma = 1.602$.

Traffic time distribution for cargo ships

Based on the Pearson chi-square test and Cramér–von Mises test, we can accept that the extreme value distribution with parameters $\alpha = 3.0$ and $\beta = 0.23$, is well fitted to the data of traffic time for containers. Test probability $p = 0.09$ in the Pearson chi-square goodness-of-fit test and 0.33 in the Cramér–von Mises test.

Figure 5 presents the frequency histogram of data connected with cargo ship traffic time and probability density function of the fitted distribution.

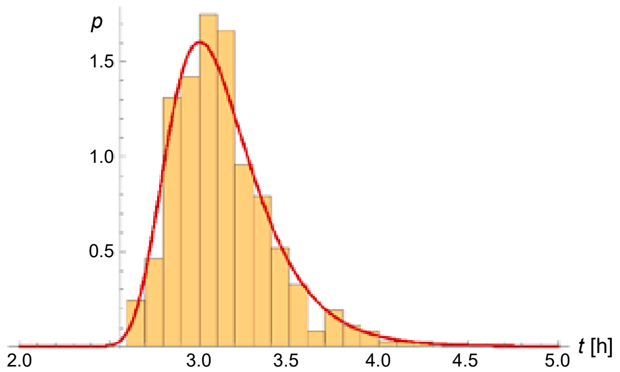

Figure 5. Frequency histogram of traffic time for cargo ships

Traffic time distribution for general cargo ships

Figure 6 presents the frequency histogram and the graph of extreme value distribution probability density function, applied to data for general cargo ships.

Using the chi-square goodness-of-fit test, we found $p = 0.43$. In the Cramér–von Mises goodness-of-fit test $p = 0.97$. Consequently, we are unable to reject the hypothesis that the traffic time of general cargo ships between reporting points 11 km and

I Brama Torowa has an extreme value distribution. In this case, location parameter $\alpha = 2.96$ and scale parameter $\beta = 0.219$.

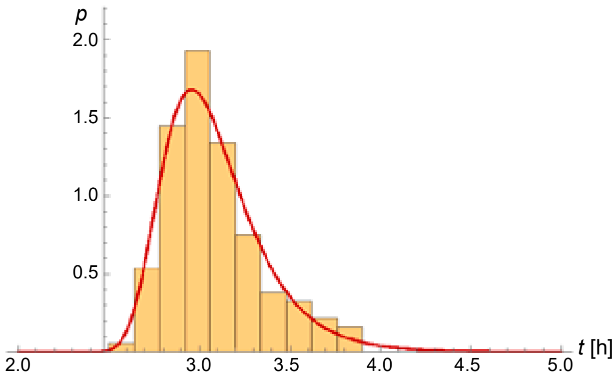

Figure 6. Frequency histogram of traffic time for general cargo ships

Traffic time distribution for carriers

Figure 7 presents the histogram and the graph of probability density function of distributions applied to data for carriers.

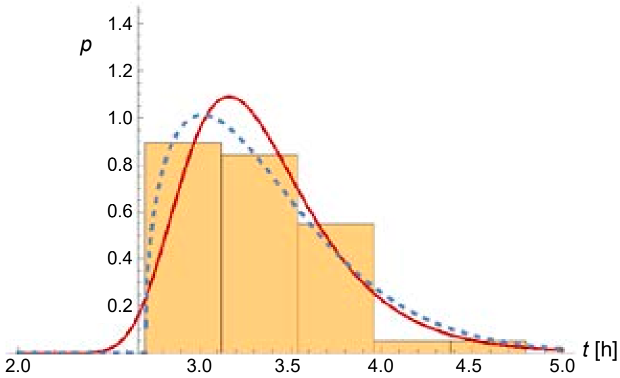

Figure 7. Frequency histogram of traffic time for carriers

On the basis of the Pearson chi-square test and Cramér–von Mises test, we can accept that the extreme value distribution with parameters $\alpha = 3.16$ and $\beta = 0.338$, and the Weibull distribution with parameters $\alpha = 2.7$, $\beta = 0.729$ and $\gamma = 1.429$ are well fitted to the data of traffic time for carriers.

Traffic time distribution for other ships

The group called "other ships" is comprised of: tugs, factory trawlers, research/survey vessels, suction dredgers, diving support vessels, ro-ro/passenger ships, fishing vessels and offshore supply ships.

Based on the Pearson chi-square test and Cramér–von Mises test, we can accept that the extreme value distribution with parameters $\alpha = 2.95$ and $\beta = 0.288$,

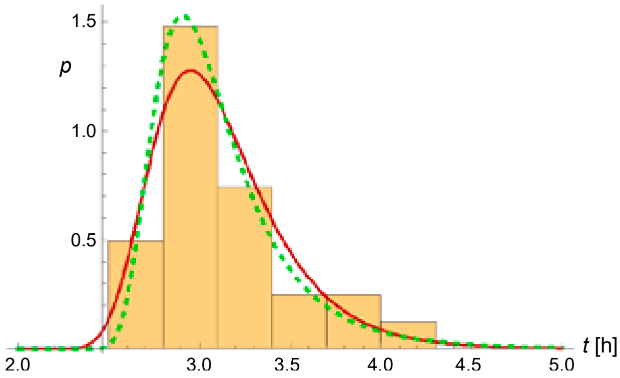

Figure 8. Frequency histogram of traffic time for other ships

and the Frechet distribution with parameters $\alpha = 1.09$, $\beta = 1.84$ and $\gamma = 7.59$ are well fitted to the data of traffic time for other ships.

Conclusions

All frequency histograms are asymmetric and all have negative coefficients of skewness. Different regulations (especially the speed limit) necessitate the use of extreme values in the analysis of vessel traffic flows. All presented vessel traffic time distributions are in accordance with an extreme value distribution. Furthermore, to some of them, a Weibull or a Frechet distribution fits well. Relative differences between empirical and theoretical means are less than 1%. Relative differences between empirical and theoretical variances are less than 20%. Using new distributions (especially Frechet) for the analysis of vessel traffic flows allows the building of better simulation models of vessel traffic. Two statistical goodness-of-fit tests have been used: the most universal and popular Pearson chi-square test, and the definitely stronger, but rarely used Cramér–von Mises test. Using the Cramér–von Mises test, we have a stronger basis for the application of verified distributions.

Acknowledgments

This research outcome has been achieved under the research project No. 4/S/INM/15 financed from a subsidy of the Ministry of Science and Higher Education for statutory activities of Maritime University of Szczecin.

References

1. Curbach, M., Gucma, L. & Proske, D. (2005) *Complex method of bridge safety assessment in respect to ship collision*. Monographs, 2nd International Congress "Seas and Oceans 2005". Volume 1. Szczecin: Wydawnictwo AM.

2. Gucma, L. (2005) *Probability method of the ship's underkeel clearance determination*. Monographs, 2nd International Congress "Seas and Oceans 2005". Volume 1. Szczecin: Wydawnictwo AM.
3. Gucma, S., Gucma, L. & Zalewski, P. (2008) *Symulacyjne metody badań w inżynierii ruchu morskiego*. Szczecin: Wydawnictwo Naukowe Akademii Morskiej w Szczecinie.
4. Gucma, L. & Jankowski, S. (2001) *Method of determining probabilistic models of propeller streams speed at the bottom of manoeuvring ships*. Proceedings of the IX International Scientific and Technical Conference on MTE. Szczecin: Wydawnictwo AM.
5. Kasyk, L. (2012) *Probabilistyczne metody modelowania parametrów strumienia ruchu statków na akwenach ograniczonych*. Radom: Wydawnictwo UTH.
6. Kasyk, L. & Kijewska, M. (2014) Gumbel Distribution in Analysis of Vessel Speed on the Świnoujście–Szczecin Fairway. *Zeszyty Naukowe Akademii Morskiej w Szczecinie* 2.
7. Rogowski, A. (2015) Testowanie hipotezy o rozkładzie Poissona w oparciu o statystykę Cramera–von Misesa. *Logistyka* 3.
8. Smolarek, L. (2005) *The application of computer simulation in life raft safety parameter estimation*. Monographs, 2nd International Congress "Seas and Oceans 2005". Volume 1. Szczecin: Wydawnictwo AM.
9. Sobczyk, M. (2004) *Statystyka*. Warszawa: Wydawnictwo Naukowe PWN.