Grabski Franciszek
Polish Naval Academy, Gdynia, Poland

Nonhomogeneous compound Poisson process application to modeling of random processes related to accidents in the Baltic Sea waters and ports

Keywords

nonhomogeneous Poisson process, nonhomogeneous compound Poisson process, safety characteristics

Abstract

A crucial role in construction of the models related to accidents on Baltic Sea water and ports play nonhomogeneous Poisson and nonhomogeneous compound Poisson process. The model of consequences and connected to it model of accidents number on Baltic sea waters and ports are here presented. Moreover some procedures of the models parameters identification are presented in the paper. Estimation of some model parameters was made based on data from reports of HELCOM [10, 11], Interreg project Baltic LINes [9] and EMSA [13].

1. Introduction

In the paper [7] the models of accidents number in the Baltic Sea waters and ports are presented. A crucial role in construction of the models plays a Poisson process and its extensions especially a nonhomogeneous Poisson process. Moreover some procedures of the model parameters identification are presented in the paper. Estimation of models parameters was made based on available data coming from reports of HELCOM [10, 11] and Interreg project Baltic LINes [9]. The models allow us to anticipate number of accidents on Baltic Sea waters and ports in future. The nonhomogeneous compound Poisson process as a model of the accidents consequences is also presented in this paper. Theoretical results [1], [2], [3], [4], [5] are applied for anticipation of the fatalities number , number of injured people and lost ships number in accidents at the Baltic Sea waters and ports in the specified time period.

2. Nonhomogeneous Poisson process

We will begin with a reminder of the concept of nonhomogeneous Poisson's process.

Let

$$\tau_0 = \vartheta_0 = 0$$

$$\tau_n = \vartheta_1 + \vartheta_2 + \cdots + \vartheta_n,\ n \in \mathbb{N}, \tag{1}$$

where $\vartheta_1,\ \vartheta_2, \dots, \vartheta_n$ are positive independent and idendical distributed random variables. Let

$$\tau_\infty = \lim_{n\to\infty} \tau_n = \sup\{\tau_n\colon\ n \in \mathbb{N}_0\}. \tag{2}$$

A stochastic process $\{N(t)\colon\ t \geq 0\}$ defined by the formula

$$N(t) = \sup\{n \in \mathbb{N}_0\colon\ \tau_n \leq t\} \tag{3}$$

is called a *counting process* corresponding to a random sequence $\{\tau_n : \in \mathbb{N}_0\}$.
Let $\{N(t)\colon\ t \geq 0\}$ be a stochastic process taking values on $S = \{0,1,2,\dots\}$, value of which represents the number of events in a time interval $[0,t]$.

A *counting process* $\{N(t)\colon\ t \geq 0\}$ is said to be *nonhomogeneous Poisson process* (NPP) with an intensity function $\lambda(t) \geq 0,\ \ t \geq 0$, if

1. $P(N(0) = 0) = 1$; (4)

2. *The process $\{N(t)\colon\ t \geq 0\}$ is the stochastic process with independent increments, the right continuous and piecewise constant trajectories*;

3. $P(N(t+h) - N(t) = k) =$

$$\frac{\left(\int_t^{t+h}\lambda(x)dx\right)^k}{k\,!}e^{-\int_t^{t+h}\lambda(x)dx}; \qquad (5)$$

From this definition it follows that the one dimensional distribution of NPP is given by the rule

$$P(N(t)=k)=\frac{\left(\int_0^t\lambda(x)dx\right)^k}{k\,!}e^{-\int_0^t\lambda(x)dx}, \qquad (6)$$

$k=0,1,2,\ldots$

The expectation and variance of NPP are the functions

$$\Lambda(t)=E[N(t)]=\int_0^t\lambda(x)dx\,, \qquad (7)$$

$$\mathrm{V(t)}=V[N(t)]=\int_0^t\lambda(x)dx\,,\quad t\geq 0. \qquad (8)$$

The corresponding standard deviation is

$$\mathrm{D(t)}=\sqrt{V[N(t)]}=\sqrt{\int_0^t\lambda(x)dx}\,,\quad t\geq 0. \qquad (9)$$

The expected value of the increment $N(t+h)-N(t)$ is

$$\Delta(t;h)=E\big(N(t+h)-N(t)\big)=\int_t^{t+h}\lambda(x)dx. \qquad (10)$$

The corresponding to it standard deviation is

$$\sigma(t;h)=\sqrt{\int_t^{t+h}\lambda(x)dx} \qquad (11)$$

An nonhomogeneous Poisson process with $\lambda(\,t)=\lambda$, $t\geq 0$ for each $t\geq 0$, is a regular Poisson process. The increments of an nonhomogeneous Poisson process are independent, but not necessarily stationary. A nonhomogeneous Poisson process is a Markov process.

3. Compound Poisson process

Let $\{N(t){:}\,t\geq 0\}$ be a Poisson proces with intensity $\lambda>0$ and $X_1,\ X_2,\ldots$ be sequence of independent and identically distributed (i.i.d.) random variables independent of $\{N(t){:}\,t\geq 0\}$. A stochastic process

$$X(t)=X_1+X_2+\cdots+X_{N(t)},\ \ t\geq 0 \qquad (12)$$

is called a *compound Poisson process* (CPP).

The probability discrete distribution function of $\{N(t){:}\,t\geq 0\}$ at k is

$$p(k;t)=P(N(t)=k)=\frac{(\lambda\,t)^k}{k\,!}e^{-\lambda\,t},$$

$k=0,1,2,\ldots$

We quote a well-known result.
If $E(X_1^2)<\infty$, then

1. $E[X(t)]=\lambda\,t\,E(X_1),$ (13)

2. $V[X(t)]=\lambda\,t\,E(X_1^2)\,.$ (14)

The concepts and facts can be generalized. We assume now that $\{N(t){:}\,t\geq 0\}$ is a *nonhomogeneous Poisson process* (NPP) with an intensity function $\lambda(\,t),\ t\geq 0$ such that $\lambda(\,t)\geq 0$ for $t\geq 0$, and $X_1,\ X_2,\ldots$ is a sequence of the independent and identically distributed (i.i.d.) random variables independent of $\{N(t){:}\,t\geq 0\}$. A stochastic process $\{X(t){:}\,t\geq 0\}$ determines by the formula

$$X(t)=X_1+X_2+\cdots+X_{N(t)},\ \ t\geq 0 \qquad (15)$$

is said to be a *nonhomogeneous compound Poisson process* (NCPP)

Proposition 1

If $\{N(t){:}\,t\geq 0\}$ is a *nonhomogeneous Poisson process* (NPP) with an intensity function $\lambda(\,t),\ t\geq 0$ such that $\lambda(\,t)\geq 0$ for $t\geq 0$ then cumulative distribution function (CDF) of the *nonhomogeneous compound Poisson process is given by the rule*

$$G(x,t)=I_{[0,\infty)}(x)e^{-\Lambda(\,t)}+\sum_{k=1}^{\infty}p(k;t)F_X^{(k)}(x), \qquad (16)$$

where

$F_X^{(k)}(x)$ denotes the *k*-fold convolution of CDF of the random variables X_i, *i*=1,2,… and

$$p(k;t)=\frac{(\Lambda(\,t))^k}{k\,!}e^{-\Lambda(\,t)},\ \ t\geq 0,\ \ k=0,1,\ldots, \qquad (17)$$

$$\Lambda(\,t)=E[N(t)]=\int_0^t\lambda(x)dx \qquad (18)$$

is discrete probability distribution of NPP.

Proof: Using total probability low we obtain cumulative distribution function (CDF) of NCPP.

$$G(x,t)=P(X(t)\leq x)=$$

$$=P\big(X_1+X_2+\cdots+X_{N(t)}\leq x\big)=$$

$$= \sum_{k=0}^{\infty} P\left(X_1 + \cdots + X_{N(t)} \le x | N(t) = k\right) \cdot$$

$$\cdot P(N(t) = k) = \sum_{k=0}^{\infty} p(k;t) F_X^{(k)}(x) =$$

$$= I_{[0,\infty)}(x) e^{-\Lambda(t)} + \sum_{k=1}^{\infty} p(k;t) F_X^{(k)}(x) .$$

Conclusion 1

If the random variables , i=1,2,… are absolutely continuous with density function $f_X(\cdot)$, then the density of NCPP is given by the rule

$$g(x,t) = \sum_{k=1}^{\infty} p(k;t) f_X^{(k)}(x), \ x \ne 0, \ t > 0, \quad (19)$$

where $f_X^{(k)}(x)$ denotes k –fold convolution of the density function $f_X(x)$.

Example 1
Let the random variables X_i, i=1,2,… have normal distrbution $N(m,\sigma)$. It means that a probabilility density function of $X_i = X$ is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \sigma > 0, \ m \in (-\infty,\infty), \quad (20)$$

$$x \in (-\infty,\infty).$$

The sum $X_1 + X_2 + \cdots + X_k$ has normal distribution $N(km, \sqrt{k}\sigma)$. Hence it's density is k –fold convolution of the density function $f_X(x)$ given by (20):

$$f_X^{(k)}(x) = \frac{1}{\sqrt{2\pi}\sqrt{k}\sigma} e^{-\frac{(x-km)^2}{2k\sigma^2}}$$

Therefore the density of NCPP given by (19) takes the form

$$g(x,t) = \frac{1}{\sqrt{2\pi}\,\sigma} \sum_{k=1}^{\infty} \frac{(\Lambda(t))^k}{\sqrt{k}\,k!} e^{-\Lambda(t)} e^{-\frac{(x-km)^2}{2k\sigma^2}},$$
$$x \ne 0, \ t > 0,$$

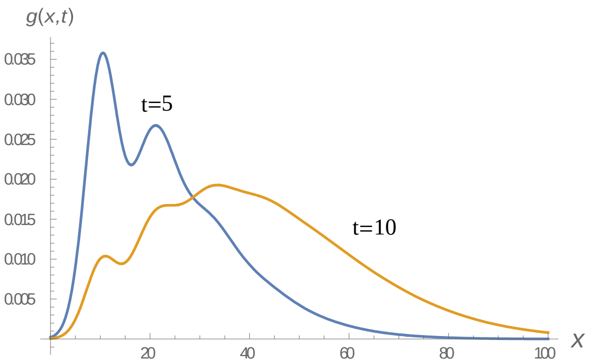

Figure 1. The density functions of CPP for
λ = 0.4, m = 10.2, σ = 3.2 and t = 5, t = 10

Conclusion 2
If the random variables X_i, i=1,2,… have a discrete probability function $p_X(x) = P(X = x)$, $x \in S$ then the discrete probability distribution of NCPP is given by the rule

$$g(x,t) = \sum_{k=1}^{\infty} p(k;t) p_X^{(k)}(x), \ t > 0 \quad (21)$$

where $p_X^{(k)}(x)$ denotes k –fold convolution of the discrete probability distribution $p_X(x)$.

Example 2
Assume that random variables $X_i, i = 1,2, ...$ have a Poisson distribution with parameter $\mu > 0$:

$$p_X(x) = \frac{\mu^x}{x!} e^{-\mu}, \ x = 0,1,2, ...$$

k –fold convolution of this discrete distribution functions is

$$p_X^{(k)}(x) = \frac{(k\mu)^x}{x!} e^{-k\mu}, \ x = 0,1,2, ... \ .$$

Then the rule (18) takes the form

$$g(x,t) = \sum_{k=1}^{\infty} \frac{(\Lambda(t))^k}{k!} e^{-\Lambda(t)} \frac{(k\mu)^x}{x!} e^{-k\mu}, \quad (22)$$
$$x = 0,1,2, ... \ , \ t > 0$$

Assuming $\Lambda(t) = \lambda t, \ t = 15, \ \lambda = 0.4, \ \mu = 0.1$ we have computed probabilties (22) The results are shown in *Table 1*.

Table 1. The values of the function (22)

x	0	1	2	3
$g(x,15)$	0,562495	0,306725	0,098596	0,023906
x	4	5	6	7
$g(x,15)$	0,004804	0,000840	0,000132	0,000018

Proposition 2
Let $\{X(t): t \ge 0\}$ be a *nonhomogeneous compound Poisson process* (NCPP).

If $E(X_1^2) < \infty$, then

1. $E[X(t)] = \Lambda(t)\, E(X_1)$ (23)

2. $V[X(t)] = \Lambda(t)\, E(X_1^2)$, (24)

Proof: Applying the property of conditional expectation

$$E[X(t)] = E[E(X(t)|N(t))]$$

we have

$$E[E(X(t)|N(t))] =$$

$$= E\left(E(X_1 + X_2 + \cdots + X_{N(t)})\middle|N(t)\right) =$$

$$= \sum_{n=0}^{\infty} E\left(X_1 + X_2 + \cdots + X_{N(t)}\middle|N(t) = n\right)P(N(t) = n) =$$

$$= \sum_{n=0}^{\infty} E(X_1 + X_2 + \cdots + X_n)\, P(N(t) = n)$$

$$= \sum_{n=0}^{\infty} E(X_1)\, n\, P(N(t) = n) = E(X_1)E\big(N(t)\big)$$

Using a formula

$$V[X(t)] = E[V(X(t)|N(t))] + V[E(X(t)|N(t))]$$

we get

$$= \sum_{n=0}^{\infty} V\left(X_1 + X_2 + \cdots + X_{N(t)}\middle|N(t) = n\right)P(N(t) = n)$$

$$= \sum_{n=0}^{\infty} V(X_1 + X_2 + \cdots + X_n)\, P(N(t) = n) =$$

$$= \sum_{n=0}^{\infty} V(X_1)\, n\, P(N(t) = n) = V(X_1)E\big(N(t)\big)$$

$$= V(X_1)\Lambda(t),$$

$$V\big[E(X(t)|N(t))\big] =$$

$$= V\left(E(X_1 + X_2 + \cdots + X_{N(t)})\middle|N(t)\right) =$$

$$= V(E(X_1)N(t)) = (E(X_1))^2 V\big(N(t)\big) =$$

$$= (E(X_1))^2 \Lambda(t).$$

Therefore

$$V[X(t)] = V(X_1)\Lambda(t) + (E(X_1))^2 =$$

$$= \Lambda(t)[E(X_1^2) - (E(X_1))^2 + (E(X_1))^2] =$$

$$= \Lambda(t)\, E(X_1^2).$$

Proposition 2

Let $\{X(t+h) - X(t): t \geq 0\}$ be an increament of *compound nonhomogeneous Poisson process* (CNPP).

If $E(X_1^2) < \infty$, then

$$E[X(t+h) - X(t)] = \Delta(t;h)\, E(X_1) \qquad (25)$$

$$D[X(t+h) - X(t)] = \sqrt{\Delta(t;h)\, E(X_1^2)}, \qquad (26)$$

where

$$\Delta(t;h) = \int_t^{t+h} \lambda(x)dx.$$

4. Corrected model of accidents number in Baltic Sea waters and ports

We will quote information from the paper [7], which is necessary for further consideration. Some mistakes in formulas (15) and (16) are noticed by author. Now this mistakes are corrected.
Assume that a stochastic process $\{N(t);\ t \geq 0\}$ taking values on $S = \{0,1,2,\dots\}$, represents the number of accidents in the Baltic Sea and Seaports in a time interval $[0,t)$. Due to the nature of these events, pre-assumption that it is a nonhomogeneous Poisson process with some parameter $\lambda(t) > 0$, seems to be justified. The expected value of increment of this process is given by (10) while its one dimensional distribution is determined by (5). We can use practically these rules if will know the intensity function $\lambda(t) > 0$. To define this function we utilize information presented in [5], [9], [10, 11] The statistical analysis of the data shows that the intensity function $\lambda(t)$ can be approximated by the linear function $\lambda(t) = at + b$.

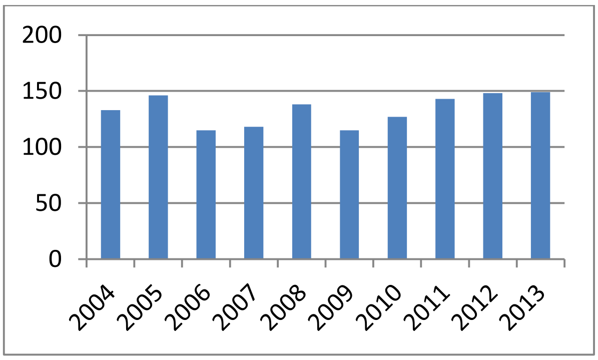

Figure 1. Total number of reporting ship accidents in the Baltic Sea during 2004-2013

4.1. Estimation of models parameters

Dividing the number of accidents in each year, by

365 or 366 we get the intensity in units of [1 / day]. The results are shown in *Table 2.* We approximate the empirical intensity by a linear regression function $y = ax + b$ that satisfied condition

$$S(a,b) = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2 \rightarrow min$$

Recall, that solution of above optimization problem leads to finding parameters a and b. The parameters are given by the rules:

$$a = \frac{\mu_{11}}{\mu_{20}}, \qquad b = m_{01} - am_{10} \qquad , \qquad (27)$$

$$\bar{x} = m_{10} = \frac{1}{n} \sum_{i=1}^{n} x_i, \qquad \bar{y} = m_{01} = \frac{1}{n} \sum_{i=1}^{n} y_i,$$

$$m_{11} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i, \qquad \mu_{11} = m_{11} - m_{10}\, m_{01},$$

$$m_{20} = \frac{1}{n} \sum_{i=1}^{n} x_i^2, \qquad \mu_{20} = m_{20} - {m_{10}}^2$$

Table 2. The empirical intensity of accidents in the Baltic Sea waters and ports

Year	Interval	Center of interval	Number of accidents	Intensity [1/day]
2004	[0, 366)	183	133	0,36338
2005	[366,731)	731,5	146	0,40000
2006	[731, 1096)	913,5	115	0,31506
2007	[1096, 1461)	1278,5	118	0,32328
2008	[1461, 1827)	1644	138	0,37704
2009	[1827, 2192)	2009,5	115	0,31506
2010	[2192, 2557)	2374,5	127	0,34794
2011	[2557, 2922)	2374,5	143	0,39178
2012	[2922, 3288)	3105	148	0,40437
2013	[3288, 3653)	3470,5	149	0,40821

Applaying the rules (27) for the data from *Table 2* and using Excel system we obtain

$$a = 0{,}000014756, \qquad b = 0{,}337925722. \qquad (28)$$

The linear intensity of accidents is

$$\lambda(x) = 0{,}000014756\, x + 0{,}337925722 \qquad (29)$$
$$x \geq 0 .$$

From (7) we have

$$\Lambda(t) = \int_0^t (0{,}000014756\, x + 0{,}337925722) dx.$$

Hence we obtain

$$\Lambda(t) = 0{,}0000073782\, t^2 + 0{,}337925722\, t, \qquad (30)$$
$$t \geq 0.$$

Therefore the one dimensional distribution of NPP is

$$P(N(t) = k) = \frac{(\Lambda(t))^k}{k!} e^{-\Lambda(t)}, \quad k = 0,1,2, \ldots, (31)$$

where $\Lambda(t)$ is given by (30).
Finnally we can say that the model of the accident number in the Baltic Sea waters and port is the nonhomogeneous Poisson process with the parameter $\Lambda(t), \ t \geq 0$ determines by (30).

5. Anticipattion of the accident number

From (5) and (11) we get

$$P(N(t+h) - N(t) = k) = \qquad (31)$$

$$= \frac{\Delta(t;h)}{k!} e^{-[\Delta(t;h)]}.$$

It means that we can anticipate number of accidents at any time interval with a length of h. The expected value of the increment $N(t+h) - N(t)$ is defined by (10). For the function

$$\Lambda(t) = a\frac{t^2}{2} + b\, t$$

we obtain the expeted value of the accidents at time interval $[t, t+h)$

$$\Delta(t;h) = h(\frac{a\,h}{2} + b + a\,t), \qquad (32)$$

The corresponding standard deviation is

$$\sigma(t;h) = \sqrt{h(\frac{a\,h}{2} + b + a\,t)}. \qquad (33)$$

Example 1
We want to predict the number of accidents from June 1 of 2017 to August 30 of 2017. We also want to calculate the probability of a given number of accidents.
First we have to determine parameters t and h. As extention of table 2 on year 2017 we obtain an interval [4749, 5114).

From January 1 of 2017 to June 1 of 2017 have pased 151 days. Hence $t = 4749 + 151 = 4900$.
From June 1 to August 31 have passed h =92 days. For these parameters using (32) and (33) we obtain

$$\Delta(t;h) = 34.45, \quad \sigma(t;h) = 5.87$$

This means that the average predicted number of accidents between June 1, 2017 and August 31, 2017 is about 34 with a standard deviation of about 6. Probability that the number of accidents at the Baltic Sea waters and ports in considered interval of time is not greater than ***d***=45 and not less that ***c***=25 is

$$P_{25\le k\le 45} = P(25 \le N(t+h) - N(t) \le 45) =$$

$$= \sum_{k=25}^{k=45} \frac{34.45^k}{k\,!} e^{-34.45};$$

Applying approximation by normal distribution we get

$$P_{25\le k\le 45} = \Phi\left(\frac{45-34.45}{5.87}\right) - \Phi\left(\frac{25-34.45}{5.87}\right) =$$
$$= \Phi(1.7972) - \Phi(-1.6098) = 0.910.$$

6. Models describing number of accidents at the Baltic ports

In the article [7] reasoned that, the intensity function of the process $N_1(\mathrm{t})$ describing number of accidents at the Baltic ports is given by

$$\lambda_1(x) = 0{,}44 \times \lambda(x). \qquad (34)$$

Because

$$a_1 = 0{,}44 \times 000014756 = 0{,}00000649264 \quad (35)$$

and

$$b_1 = 0{,}44 \times 0{,}337925722 = 0{,}14868731768 \quad (36)$$

then

$$\lambda_1(x) = 0{,}00000649264\ x + 0{,}14868731768 \quad (37)$$

The expected value and corresponding standard deviation of the accidents at time interval $[t, t+h)$ are

$$\Delta_1(t;h) = a_1\frac{h^2}{2} + b_1 h + 2a_1 t\, h, \qquad (38)$$

$$\sigma_1(t;h) = \sqrt{a_1\frac{h^2}{2} + b_1 h + 2\, a_1\, t}\ . \qquad (39)$$

Example 2
We want to anticipate the number of accidents in the ports of Baltic Sea from June 1 , 2017 to August 31, 2017. We calculate the probability of a given number of that kind of accidents. Parameters t and h are the same like in example l, parameters a_1 and b_1 are given by (36) and (37). From (39) and (40) we obtain the expected value and standard deviation of accidents in ports of Baltic Sea and in the time period $[t, t+h)$.

$$\Delta_1(t;h) = 13{,}77, \quad \sigma_1(t;h) = 3{,}71$$

For example, probability that the number of accidents in the Baltic Sea Ports in this time period is not greater than ***d***=20 and not less that ***c***=10 is approximately equal to

$$P_{10\le k\le 20} = \Phi\left(\frac{20-13{,}77}{3{,}71}\right) - \Phi\left(\frac{10-13{,}77}{3{,}71}\right) =$$

$$= \Phi(1{,}68) - \Phi(-1{,}02) = 0.799.$$

7. Anticipation of the accident consequences

Let $X_i,\ i = 1,2,\dots,N(t)$ denotes number of fatalities or injuried poeople or ships lost in ***i***-th accident. We suppose that the random variables $X_i, i = 1,2,\dots$ have the identical Poisson distribution with parameters

$$E(X_i) = V(X_i) = \mu, \quad i = 1,2,\dots,N(t).$$

The predicted number of fatalities in the time interval $[t, t+h)$ is described by the expectation of the increment $X(t+h) - X(t)$.
Recall that the expected value and standard deviation of the accidents number in the time interval $[t, t+h)$ are given by (10) and (11).

To calculate the expected number of fatalities in the considered time interval we apply **Proposition 2.**

Example 3
We want to anticipate the number of fatalities in accidents in the Baltic Sea waters and ports from June 1 , 2017 to August 31, 2017.

For the data from **Example 1** using (24) and (25) we obtain the expected value of fatalities in the time interval $[t, t+h)$:

$$\boldsymbol{EFN} = \Delta(t;h) \times \mu \qquad (40)$$

and the standard deviation

$$\boldsymbol{DFN} = \sqrt{\Delta(t;h) \times (\mu + \mu^2)}\ . \qquad (41)$$

We know that the average of the sample is an unbiased estimator of the expected value. Unfortunately, reliable data are not available for the moment. We roughly estimate this parameter using data presented in EMSA reports [12, 13] and paper [6]. These data's are only partially consistent with the previous ones. The approximate estimate of the parameter μ is the number

$$\boldsymbol{\mu = 0,056}\ .$$

Applying (40) and (41) we get

$$\boldsymbol{EFN = 1,9292} \quad \text{and} \quad \boldsymbol{DFN = 1,4273}\ .$$

In this case, the formula (19) takes the form

$$g(x,t;h) = \sum_{k=1}^{\infty} \frac{\Delta(t;h)^k}{k\,!} e^{-\Delta(t;h)} \frac{(k\mu)^x}{x!} e^{-k\mu}\ , \quad (42)$$
$$x = 0,1,2,\ldots\ \ ,\ t > 0.$$

For $t = 4900$, $h = 92$ we have $\Delta(t;h) = 34.45$. Using (43), for $\mu = 0{,}056$ we obtain a predicted dystribution of fatalities in accidents at the Baltic Sea and ports from June 1 , 2017 to August 31, 2017. *Table 3* and *Figure 2* show this distribution.
We can see that the most probable numbers of fatalities are 1 and 2. The probability that there will be no fatal accident is only about 15%.

Table 3. Distribution of fatalities number

x	0	1	2	3
$g(x)$	0,153175	**0,279411**	**0,262665**	0,169372
x	4	5	6	7
$g(x)$	0,0841492	0,0343135	0,0119478	0,0036499

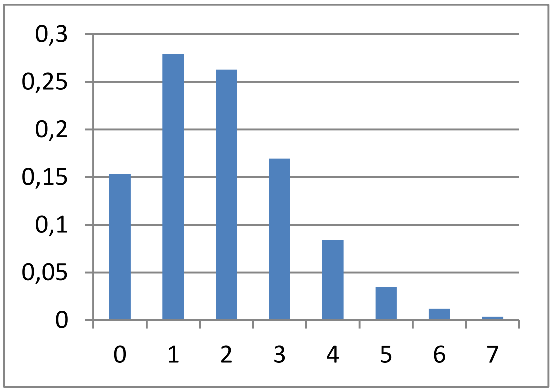

Figure 2. Distribution of fatalities number

We can see that the most probable numbers of fatalities are 1 and 2. The probability that there will be no fatal accident is only about 15%.

Example 4
The predicted number of injured person in accidents in the Baltic Sea and Ports from June 1 , 2017 to August 31, 2017 we will get in a similar way. In this case

$$\boldsymbol{\mu = 0,224.}$$

For the data from **Example 3** using (41) and (42) we obtain an expected value and a standard deviation of injured people number at considered period.

$$\boldsymbol{ENI} = 34{,}45 \times 0{,}224 = 7{,}7168$$

$$\boldsymbol{DNI} = \sqrt{34{,}45 \times (0{,}224 + 0{,}224^2)} = 3{,}0733$$

Equality (42) allows to compute predicted dystribution of the injured person number. The results are shown in *Table 4* and *Figure 3*.

Table 4. Distribution of injured person number

x	0	1	2	3
$g(x)$	0,0009942	0,0061322	0,019599 1	0,0431724
x	4	5	6	7
$g(x)$	0,0735834	0,103326	**0,124318**	**0,131639**
x	8	9	10	11
$g(x)$	**0,125075**	0,108205	0,086211	0,063839
x	12	13	14	15
$g(x)$	0,0442636	0,0289157	0,017890 2	0,0105299
x	16	17	18	19
$g(x)$	0,0059186	0,0031876	0,001649	0,0008226

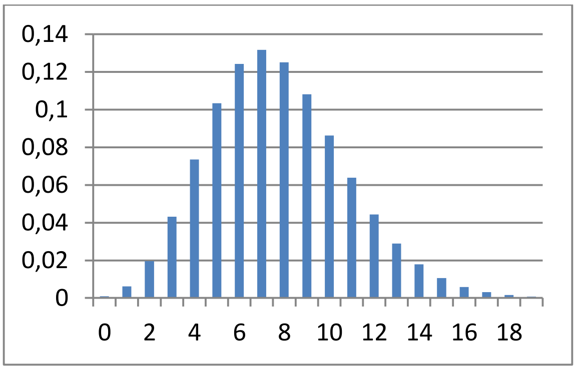

Figure 3. Distribution of injured person number

Example 5
For the ships lost number in accidents in the Baltic Sea and Sea Ports in considered time interval parameter

μ is

$$\boldsymbol{\mu} = \mathbf{0,016} .$$

For the data from **Example 3** using (40) and (41) we obtain an expected value and standard deviation of the ships lost number in considered period.

$$\boldsymbol{EIN} = 34{,}45 \times 0{,}016 = 0{,}5512$$

$$\boldsymbol{DIN} = \sqrt{34{,}45 \times (0{,}016 + 0{,}016^2)} = 0{,}74834$$

Equality (42) allows to compute predicted dystribution of injured person number. The results are shown in *Table 5.*

Table 5. Dystribution of ships lost number

x	0	1	2	3
g(*x*)	0,578791	0,313966	0,0876672	0,0167734
x	4	5	6	7
g(*x*)	0,0024704	0,00029833	0,00003074	0,0000003

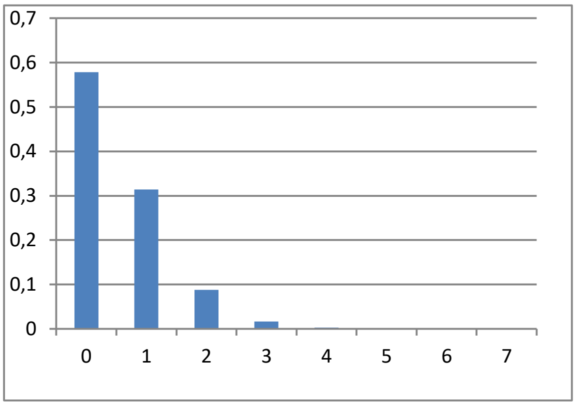

Figure 4. Distribution of ships lost number

We can notice that the most probable is no ships lost. This probability is about 58%.

8. Conclusions

The random processes theory deliver concepts and theorems that enable to construct stochastic models concerning accidents. The counting processes and processes with independent increaments are the most appropriate for modelling number of the accidents number in Baltic Sea waters and ports in specified period of time. A crucial role in the models construction plays a nonhomogeneous Poisson process and nonhomogeneous compound Poisson process. Based on the nonhomogeneous Poisson process the models of accidents number in the Baltic Sea waters and Seaports have been constructed. Moreover some procedures of the model parameters identification are presented in the paper. Estimation of model parameters was made based on data from reports of HELCOM (2014) and Interreg project Baltic LINes (2016-2019).

The nonhomogeneous compound Poisson process as a model of the accidents consequences is also presented in this paper. Theoretical results are applied for anticipation the number of fatalities, number injured people and number lost ships in accidents at the Baltic Sea waters and ports in specified period of time.

Acknowledgements

The paper presents the results developed in the scope of the HAZARD project titled “Mitigating the Effects of Emergencies in Baltic Sea Region Ports” that has received funding from the Interreg Baltic Sea Region Programme 2014-2020 under grant agreement No #R023. https://blogit.utu.fi/hazard/

References

[1] Di Crescenzo, A., Martinucci, B. & Zacks, S. Compound Poisson process with a Poisson subordinator. *Journal of Applied Probability* Vol. 52, No. 2, p. 360- 374**.**
[2] Fisz, M. (1969). *Probability and Mathematical* Statistics., Warsaw PWN; (in Polish).
[3] Grabski, F. (2002). *Semi-Markov models of reliability and operation.* Warsaw : IBS PAN (in Polish).
[4] Grabski, F. (2015). *Semi-Markov Processes: Applications in Systems Reliability and Maintenance.* Elsevier; Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Sydney.
[5] Grabski, F. (2017). Nonhomogeneous Poisson process application to modelling accidents number at Baltic waters and ports. *Journalof Polish Safety and Reliability Association;* volume 8, number 1, p. 39-46.
[6] Herdzik, J. (2016). Zdarzenia wypadkowe na morzu i ich głowne przyczyny (Accident events at sea and their main causes). *Autobusy* 10/16.p.33-38.
[7] Limnios, N. & Oprisan, G. (2001). *Semi-Markov Processes and Reliability.* Boston; Birkhauser.
[8] Shiryayev, A.N. (1984). *Probability*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo.

[9] Baltic LINes (2016). *Shipping in the Baltic Sea – Past,present and future developments relevant for Maritime Spatial Planning*. Project Report I. 35 p
[10] HELCOM (2013). *Annual report on shipping accidents in the Baltic Sea area during 2012*. 43 pp.
[11] HELCOM (2014). *Annual report on shipping accidents in the Baltic Sea in 2013*.
[12] EMSA (2015). Annual Overview of Marine Casualties and Incidents 2014
[13] EMSA (European Maritime Safety Agency) (2017),*ANNUAL overview of marine casualties and incidents 2016.* 119 pp.