ANADOLU ÜNİVERSİTESİ BİLİM VE TEKNOLOJİ DERGİSİ –B
Teorik Bilimler

ANADOLU UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY –B
Theoretical Sciences
Cilt/Vol.:1-Sayı/No: 2 : 91-104 (2011)

DISTRIBUTION FUNCTION OF FIRST EXIT TIME FOR A COMPOUND POISSON PROCESS

Gamze ÖZEL [1], Ceyhan İNAL

ABSTRACT

In this study, the first exit time of a compound Poisson process with positive jumps and an upper horizontal boundary is considered. An explicit formula is derived for the distribution function of the first exit time associated with the compound Poisson process. By means of a proposed algorithm, some numerical examples and an application on traffic accidents are also given to illustrate the usage of the distribution function of the first exit time and proposed algorithm.

Keywords: Compound poisson process, First exit time, Upper horizontal boundary, Oracle database.

BİRLEŞİK POISSON SÜRECİ İÇİN İLK AŞMA ZAMANININ DAĞILIM FONKSİYONU

ÖZ

Bu çalışmada, pozitif adımlı ve üst yatay sınırlı bir birleşik Poisson sürecinin ilk aşma zamanı göz önünde bulundurulmuş ve ilk aşma zamanının kesin formülüne ulaşılmıştır. İlk aşma zamanının dağılım fonksiyonunun ve önerilen algoritmanın uygulanabilirliğini göstermek için bazı sayısal örnekler ve trafik kazaları üzerine bir uygulama verilmiştir.

Anahtar Kelimeler: Birleşik poisson süreci, İlk aşma zamanı, Üst yatay sınır, Oracle veritabanı.

1. INTRODUCTION

Let $\{N_t, t \geq 0\}$ be a homogeneous (or nonhomogeneous) Poisson process and let Y_i, $i = 1, 2, 3, ...,$ be i.i.d. random variables, independent of the process $\{N_t, t \geq 0\}$. A stochastic process $\{X_t, t \geq 0\}$ is said to be a compound Poisson process if it can be represented as

$$X_t = \sum_{i=1}^{N_t} Y_i \tag{1}$$

If $E(Y_i) = \eta$, $V(Y_i) = \sigma^2$, $i = 1, 2, 3, ...,$ the expected value and variance of X_t are $E(X_t) = \lambda t \eta$, $V(X) = \lambda t(\sigma^2 + \eta^2)$, respectively. In particular, if Y_i, $i = 1, 2, 3, ...,$ are Poisson distributed in Eq. (1), $\{X_t, t \geq 0\}$ is called as a Neyman type A process and if Y_i, $i = 1, 2, 3, ...,$ are

[1,] Hacettepe University, Department of Statistics 06800 Beytepe, Ankara, Turkey.
Corresponding author: Gamze Özel (e-mail: gamzeozl@hacettepe.edu.tr)

Recieved: 10 December 2009; **Revised:** 22 March 2010; **Accepted:** 3 May 2010

distribute according to the binomial distribution, $\{X_t, t \geq 0\}$ is called as a Neyman type B process. Similarly, if Y_i, $i = 1, 2, 3, ...$, are geometric distributed, $\{X_t, t \geq 0\}$ is called as a Polya-Aeppli process (Ozel and Inal, 2008b).

The statistical significance of the compound Poisson process arises from its applicability in real life situations, where the researcher often observes only the total amount X_t, which is composed of an unknown random number N_t of random contributions Y_i, $i = 1, 2, 3, ...$ However, without knowledge of the probability function for X_t obstructs the usage of the compound Poisson distribution completely. Therefore, the explicit distribution function of the first exit time for the compound Poisson process has not been obtained yet (Ammussen, 2000).

Concerning applications, the first exit time can be defined as the length of the busy period for a M/G/1 queue in queuing theory (Rosencrantz, 1983). In risk analysis, the distribution function of the first exit time is that of the first time at which the accumulated total claims from an insurance company exceeds its capital. Besides these, the first exit time is defined as the first time until either outdating or total depletion of stock for perishable items in inventory theory (Dirickx and Koevoets, 1997).

Laplace transforms of the distribution function of the first exit time with two paralel boundaries were derived by Dvoretzky et al. (1953) for a homogeneous Poisson process. Then explicit formulas for the distribution function of the first exit time for the homogeneous Poisson process were obtained by Delucia and Poor (1997) and Rolski et al. (1998). The Laplace Stieltjes transforms of the distribution function of the first exit time with positive jumps were given by Bar-Lev et al. (1999) for the compound Poisson process where Y_i, $i = 1, 2, ...$, were continuous random variables. On the other hand, the explicit formula for the distribution function of the first exit time has not been obtained where Y_i, $i = 1, 2, ...$, are discrete random variables.

In this study it is assumed that $\{N_t, t \geq 0\}$ is a homogeneous Poisson process with parameter $\lambda > 0$ and Y_i, $i = 1, 2, ...$, are discrete random variables representing the positive integer-valued jump sizes. Then, the distribution function of the first exit time is obtained for the compound Poisson process with an upper horizontal boundary and positive integer-valued jump sizes. The paper is organized as follows. We start by introducing in Section 2 the probability function of X_t. This allows us in Section 3 to derive the distribution function of the first exit time for the compound Poisson process. Then, some numerical examples and an application to the traffic accidents are presented in Section 4 by means of the proposed algorithm in Oracle database. The conclusion is given in Section 5.

2. THE PROBABILITY FUNCTION OF X_t

Lemma 1. Let Y_i, $i = 1, 2, ...$, be discrete random variables and let $\{N_t, t \geq 0\}$ be a homogeneous Poisson process. Then the probability function of X_t is given by

$$p_{X_t}(k) = P(X_t = k) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} p_Y{}^{(n)}(k), k = 0, 1, 2, ... \tag{2}$$

where $p_Y{}^{(0)}(k) = I(k = 0)$ and $p_Y{}^{(n)}(k)$ is the n-fold convolution of $P(Y_i = j) = p_j$, $j = 1, 2, ...$ However, it is not easy to yield the explicit probabilities of X_t from Eq. (2) since it needs infinite sum. Furthermore it is very time consuming (Rolski et al., 1998).

Therefore, a recursive algorithm was derived by Panjer (1981) for X_t satisfying the relation

$$p_{N_t}(n) = \frac{\lambda t}{n} p_{N_t}(n-1), \quad n = 1, 2, \ldots \tag{3}$$

Then the following recursion holds for $p_{X_t}(k)$

$$p_{X_t}(0) = e^{-\lambda t(1-p_Y(0))},$$

$$p_{X_t}(k) = \lambda t \sum_{i=1}^{k} \frac{i}{k} p_Y(i) p_{X_t}(y-i), \tag{4}$$

where $p_Y(y)$ is the common probability function of Y_i, $i = 1, 2, \ldots$ The recursion in Eq. (4) starts with the calculated value of $P(X_t = 0) = e^{-\lambda t(1-p_Y(0))}$. As an example, in risk analysis for large insurance portfolios, this probability is very small, sometimes smaller than the smallest number that can be represented on the computer. When this occurs, this initial value is represented on the computer as zero and the recursion in Eq. (4) fails. Recently, the explicit probability function of X_t has been derived by Ozel and Inal (2008a) as in Theorem 1.

Theorem 1. Let $\{N_t, t \geq 0\}$ be a homogeneous Poisson process with parameter $\lambda > 0$ and $\lambda_j = \lambda p_j$, $j = 1, 2, \ldots, m$. If Y_i, $i = 1, 2, \ldots$, are discrete random variables with the probabilities $P(Y_i = j) = p_j$, $j = 0, 1, 2$, the explicit formula for the probability function of X_t is given by

$$\begin{aligned} p_{X_t}(k) &= e^{-\lambda t(1-p_0)}, \qquad k = 0 \\ &= e^{-\lambda t(1-p_0)} \sum_{i=0}^{[k/2]} \frac{(\lambda_1 t)^{k-2i} (\lambda_2 t)^i}{(k-2i)!\, i!}, \quad k = 1, 2, 3, \ldots \end{aligned} \tag{5}$$

where [ ] denotes the integer part of the number in the brackets.

If $m = 1$ and $j = 0, 1$, then for $i = 1, 2, \ldots$ $P(Y_i = 0) = p_0$, $P(Y_i = 1) = p_1 = 1 - p_0$. In this case, it is well known that X_t has the Poisson process with parameter $\lambda p_1 t$.

The probability function of X_t for $m > 2$ is given

$$P(X_t = 0) = e^{-\lambda t(1-p_0)},$$

$$P(X_t = 1) = e^{-\lambda t(1-p_0)} \frac{(\lambda_1 t)}{1!},$$

$$P(X_t = 2) = e^{-\lambda t(1-p_0)} \left[\frac{(\lambda_1 t)^2}{2!} + \frac{(\lambda_2 t)}{1!} \right],$$

$$P(X_t = 3) = e^{-\lambda t(1-p_0)} \left[\frac{(\lambda_1 t)^3}{3!} + \frac{(\lambda_1 t)(\lambda_2 t)}{1!1!} + \frac{(\lambda_3 t)}{1!} \right],$$

$$P(X_t = 4) = e^{-\lambda t(1-p_0)} \left[\frac{(\lambda_1 t)^4}{4!} + \frac{(\lambda_1 t)^2 (\lambda_2 t)}{2!1!} + \frac{(\lambda_1 t)(\lambda_3 t)}{1!1!} + \frac{(\lambda_2 t)^2}{2!} + \frac{(\lambda_4 t)}{1!} \right],$$

$$P(X_t = 5) = e^{-\lambda t(1-p_0)}\left[\frac{(\lambda_1 t)^5}{5!} + \frac{(\lambda_1 t)^3(\lambda_2 t)}{3!1!} + \frac{(\lambda_1 t)^2(\lambda_3 t)}{2!1!} + \frac{(\lambda_1 t)(\lambda_2 t)^2}{2!1!} + \frac{(\lambda_1 t)(\lambda_4 t)}{1!1!} + \frac{(\lambda_2 t)(\lambda_3 t)}{1!1!} + \frac{(\lambda_5 t)}{1!}\right], \tag{6}$$

$$P(X_t = 6) = e^{-\lambda t(1-p_0)}\left[\frac{(\lambda_1 t)^6}{6!} + \frac{(\lambda_1 t)^4(\lambda_2 t)}{4!1!} + \frac{(\lambda_1 t)^3(\lambda_3 t)}{3!1!} + \frac{(\lambda_1 t)^2(\lambda_2 t)^2}{2!2!} + \frac{(\lambda_1 t)^2(\lambda_4 t)}{2!1!} + \frac{(\lambda_1 t)(\lambda_2 t)(\lambda_3 t)}{1!1!1!} + \frac{(\lambda_1 t)(\lambda_5 t)}{1!1!} + \frac{(\lambda_2 t)^3}{3!} + \frac{(\lambda_2 t)(\lambda_4 t)}{1!1!} + \frac{(\lambda_3 t)^2}{2!} + \frac{(\lambda_6 t)}{1!}\right],$$

According to the probabilities $P(X_t = k)$, $k = 1, 2, 3, \ldots$, the right-hand side terms depend on how k can be partitioned into different forms with using integers $1, 2, \ldots, m$. Notice that it is the same manner with $m = 2$ and Eq. (5) can also be obtained taking $m = 2$ in Eq. (6). Furthermore, Eq. (6) can be used if Y_i, $i = 1, 2, 3, \ldots$, have infinite values $j = 1, 2, 3, \ldots$ and the probability $P(Y_i = j)$, $\lambda_j = \lambda p_j$ approaches zero for $j \to \infty$. The probabilities $P(X_t = k)$, $k = 1, 2, 3, \ldots$, can be computed directly from Eq. (6) since our formula is not recursive. It is not necessary to compute $P(X_t = \ell)$, $\ell = 0, 1, \ldots, k-1$ and this yields time advantage in computation. Moreover, Eq. (6) is independent from the type of probability function of Y_i, $i = 1, 2, 3, \ldots$

In this study, a new algorithm has been prepared in Oracle database for the probability function of X_t. The required algorithm can be obtained from author upon request. As an illustration of the proposed algorithm the probabilities $P(X_t = k)$, $k = 0, 1, 2, \ldots$, are obtained from Eq. (6). In all of these calculations it is assumed that $\{N_t, t \geq 0\}$ is a homogeneous Poisson process with different values of λ and $t = 5$ Figure 1 displays the probabilities $P(X_t = k)$, $k = 0, 1, 2, \ldots$, where Y_i, $i = 1, 2, 3, \ldots$, are Poisson distributed with parameter μ with λ chosen such as $E(X_t) = \lambda t \mu = 15$.

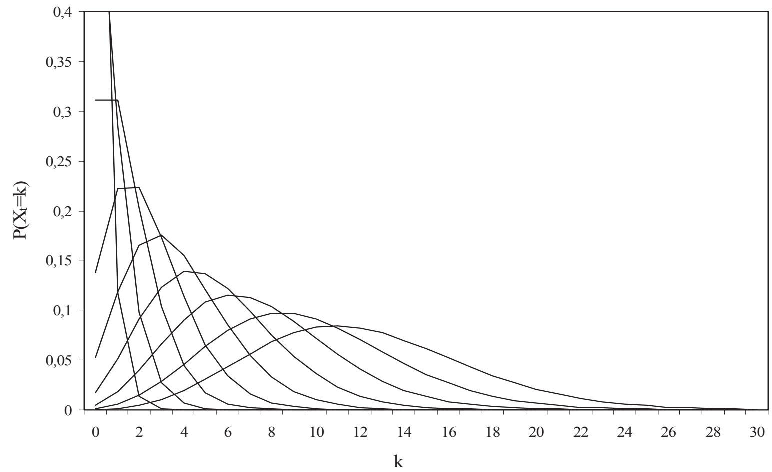

Figure 1. The probabilities $P(X_t = k)$, $k = 0, 1, 2, \ldots$, with $E(X_t) = \lambda t \mu = 15$ where Y_i, $i = 1, 2, 3, \ldots$, are Poisson distributed. From the top the bottom μ taking values 1.0, 0.9, 0.8, …, 0.1, respectively. Note that $\{X_t, t \geq 0\}$ is also called as a Neyman type A process.

Then, the probabilities $P(X_t = k)$, $k = 0, 1, 2, ...$, are shown in Figure 2 where Y_i, $i = 1, 2, 3, ...$, are binomial distributed with parameters (m, p) chosen such as $E(X) = \lambda tmp = 15$ and $m = 20$.

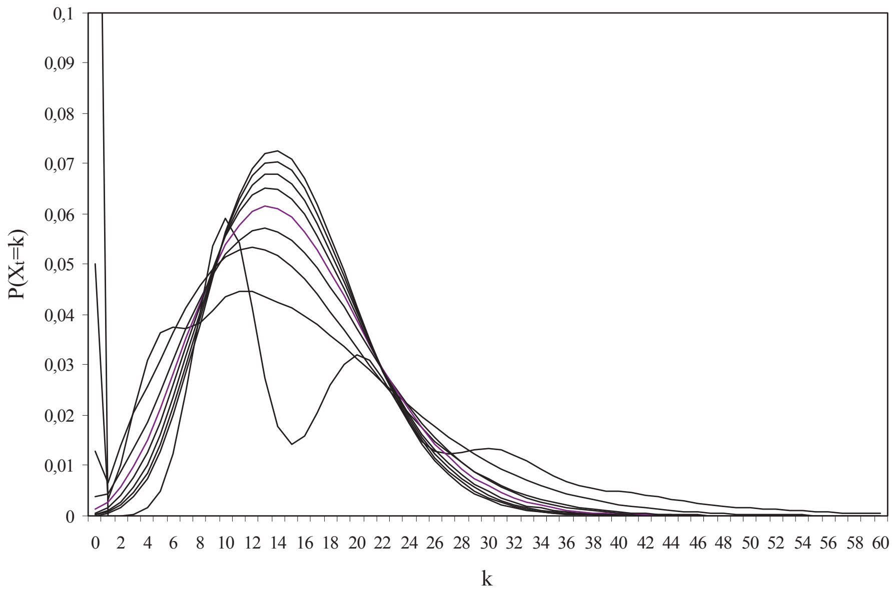

Figure 2. The probabilities $P(X_t = k)$, $k = 0, 1, 2, ...$, with $E(X_t) = \lambda tmp = 15$ and $m = 20$ where Y_i, $i = 1, 2, 3, ...$, are binomial distributed. From the top the bottom p taking values 0.05, 0.06, 0.07, 0.08, …, 0.5, respectively. Note that $\{X_t, t \geq 0\}$ is also called as a Neyman type B process.

Finally, the probabilities $P(X_t = k)$, $k = 0, 1, 2, ...$, are presented in Figure 3 where Y_i, $i = 1, 2, 3, ...$, are geometric distributed with parameter θ chosen such as $E(X_t) = \lambda / \theta = 15$.

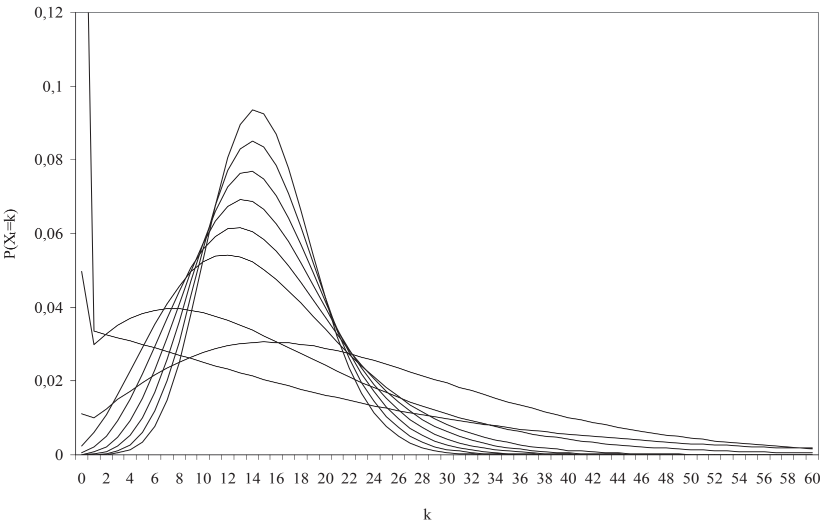

Figure 3. The probabilities $P(X_t = k)$, $k = 0, 1, 2, ...$, with $E(X_t) = \lambda t / \theta = 15$ Y_i, $i = 1, 2, 3, ...$, are geometric distributed. From the top the bottom θ taking values 1.0, 0.9, 0.8, …, 0.1, respectively. Note that $\{X_t, t \geq 0\}$ is also called as a Polya-Aeppli process.

3. DISTRIBUTION FUNCTION OF THE FIRST EXIT TIME FOR THE COMPOUND POISSON PROCESS

Now consider an upper horizontal boundary β for Eq. (7) then the first exit time T is defined as

$$T = \inf\{t : X_t \geq \beta\} \tag{7}$$

where $0 < t < \infty$ and $\beta > 0$. T can be described as the first instant at which a sample path crosses (jumps over) the boundary β.

Applications related with the distribution function of the first exit time are varied in probability and statistics, including financial mathematics, reliability, queues, inventory theory and sequential analysis; see, for instance, Gallot (1993), Ignatov and Kaishev (2004). In all of these studies Y_i, $i = 1, 2, 3, ...,$ were defined as continuous random variables. However, the distribution function of the first exit time is not considered for the compound Poisson process with an upper boundary where Y_i $i = 1, 2, ...,$ are discrete random variables.

In this study, the distribution function of T can be obtained in the following manner.

Theorem 2. Let $\{N_t, t \geq 0\}$ is a homogeneous Poisson process with parameter $\lambda > 0$ and Y_i, $i = 1, 2, ...,$ are discrete random variables with the probabilities $P(Y_i = j) = p_j$, $j = 0, 1, ..., m$. Let $\lambda_j = \lambda p_j$, $j = 1,2,...,m$ be the parameters. Then the distribution function of T for $m = 2$ is

$$F_T(t) = 1 - e^{-\lambda t(1-p_0)}\left[1 + \sum_{k=1}^{\beta-1}\sum_{i=0}^{[k/2]}\frac{(\lambda_1 t)^{k-2i}(\lambda_2 t)^i}{(k-2i)!\,i!}\right], \; t > 0 \tag{8}$$

Proof: Since $\{X_t, t \geq 0\}$ is an increasing process,

$$P(T > t) = F_{X_t}(\beta - 1)$$

Then the distribution function of T is given by

$$\begin{aligned} F_T(t) &= P(T \leq t) \\ &= 1 - P(X_t \leq \beta - 1) \\ &= 1 - \sum_{k=0}^{\beta-1} p_{X_t}(k) \\ &= 1 - F_{X_t}(\beta - 1) \end{aligned} \tag{9}$$

Substitution of Eq. (5) in Eq. (9) yields the formula in Eq. (8). The corresponding probability density function of T is

$$f_T(t) = e^{-\lambda t(1-p_0)}\left[\lambda t(1-p_0) + \lambda t(1-p_0)\sum_{k=2}^{\beta-1}\sum_{i=0}^{[k/2]}\frac{(\lambda_1 t)^{k-2i}(\lambda_2 t)^i}{(k-2i)!\,i!} + \sum_{k=1}^{\beta-2}\sum_{i=0}^{[k/2]}\frac{(k-i)\,(\lambda_1 t)^{k-2i}(\lambda_2 t)^i}{(k-2i)!\,i!}\right], \; t > 0 \tag{10}$$

For $m > 2$ the distribution function of T is obtained if the cumulative probabilities in Eq. (11) substituted into Eq. (9).

$$F_{X_t}(0) = e^{-\lambda t(1-p_0)},$$

$$F_{X_t}(1) = F_{X_t}(0) + F_{X_t}(0)\frac{(\lambda_1 t)}{1!},$$

$$F_{X_t}(3) = F_{X_t}(2) + \left[\frac{(\lambda_1 t)^3}{3!} + \frac{(\lambda_1 t)(\lambda_2 t)}{1!1!} + \frac{(\lambda_3 t)}{1!}\right], \tag{11}$$

$$F_{X_t}(4) = F_{X_t}(3) + \left[\frac{(\lambda_1 t)^4}{4!} + \frac{(\lambda_1 t)^2(\lambda_2 t)}{2!1!} + \frac{(\lambda_1 t)(\lambda_3 t)}{1!1!} + \frac{(\lambda_2 t)^2}{2!} + \frac{(\lambda_4 t)}{1!}\right]$$

$$F_{X_t}(5) = F_{X_t}(4) + \left[\frac{(\lambda_1 t)^5}{5!} + \frac{(\lambda_1 t)^3(\lambda_2 t)}{3!1!} + \frac{(\lambda_1 t)^2(\lambda_3 t)}{2!1!} + \frac{(\lambda_1 t)(\lambda_2 t)^2}{1!2!} + \frac{(\lambda_1 t)(\lambda_4 t)}{1!1!} \right.$$
$$\left. + \frac{(\lambda_2 t)(\lambda_3 t)}{1!1!} + \frac{(\lambda_5 t)}{1!}\right]$$

$$F_{X_t}(6) = F_{X_t}(5) + \left[\frac{(\lambda_1 t)^6}{6!} + \frac{(\lambda_1 t)^4(\lambda_2 t)}{4!1!} + \frac{(\lambda_1 t)^3(\lambda_3 t)}{3!1!} + \frac{(\lambda_1 t)^2(\lambda_2 t)^2}{2!2!} + \frac{(\lambda_1 t)^2(\lambda_4 t)}{2!1!} \right.$$
$$\left. + \frac{(\lambda_1 t)(\lambda_2 t)(\lambda_3 t)}{1!1!1!} + \frac{(\lambda_1 t)(\lambda_5 t)}{1!1!} + \frac{(\lambda_2 t)^3}{3!} + \frac{(\lambda_2 t)(\lambda_4 t)}{1!1!} + \frac{(\lambda_3 t)^2}{2!} + \frac{(\lambda_6 t)}{1!}\right]$$

According to the probabilities given above, a recursive formula is obtained for the distribution function of X_t with the following way: As an example, for $F_{X_t}(4)$, the terms in the square bracket of $F_{X_t}(3)$ is multiplied by $(\lambda_1 t)$, then the terms bigger than $(\lambda_1 t)$ in the square bracket of $F_{X_t}(2)$ is multiplied by $(\lambda_2 t)$ and the last term $(\lambda_4 t)$ is added. Finally, the denominators of the terms is arranged as suitable to the power of $(\lambda_j t)$'s.

An algorithm for $F_T(t)$ has been obtained in Oracle database using Eq. (9) and Eq. (11). In this algorithm, firstly for $F_{X_t}(k)$, the terms in the square bracket $F_{X_t}(k-1)$ is multiplied by $(\lambda_1 t)$, then the terms bigger than $(\lambda_1 t)$ in the square bracket of $F_{X_t}(k-2)$ is multiplied by $(\lambda_2 t)$ $(\lambda_2 t)$, …, the terms bigger than $(\lambda_{[k/2]-1} t)$ in the square bracket of $F_{X_t}\left[k - \left[\frac{k}{2}\right]\right]$ is multiplied by $(\lambda_{[k/2]} t)$ where [ ] denotes the integer part of the number in the brackets. Lastly the term $(\lambda_k t)$ is added and the denominators of the terms is arranged as suitable to the power of $(\lambda_j t)$'s. At the end of the algorithm, $F_T(t) = 1 - F_{X_t}(\beta - 1)$ is calculated according to the value of β.

Some numerical examples of $F_T(t)$ are presented in Figures 4-6 using the algorithm of $F_T(t)$. In these figures the upper horizontal boundary is taken as $\beta = 10$ and $\{N_t, t \geq 0\}$ is a homogeneous Poisson process with parameter λ taking several values. The distribution functions of T are given in Figure 4 where $i = 1, 2, ...,$ have Poisson distribution with parameter $\mu = 1.5$ and in Figure 5, the distribution functions of T are shown where Y_i, $i = 1, 2, ...,$ have binomial distribution with parameters ($p = 0.2$, $m = 5$) Finally, the distribution functions of T are presented in Figure 6 where Y_i $i = 1, 2, ...,$ have geometric distribution with parameter $\theta = 0.6$.

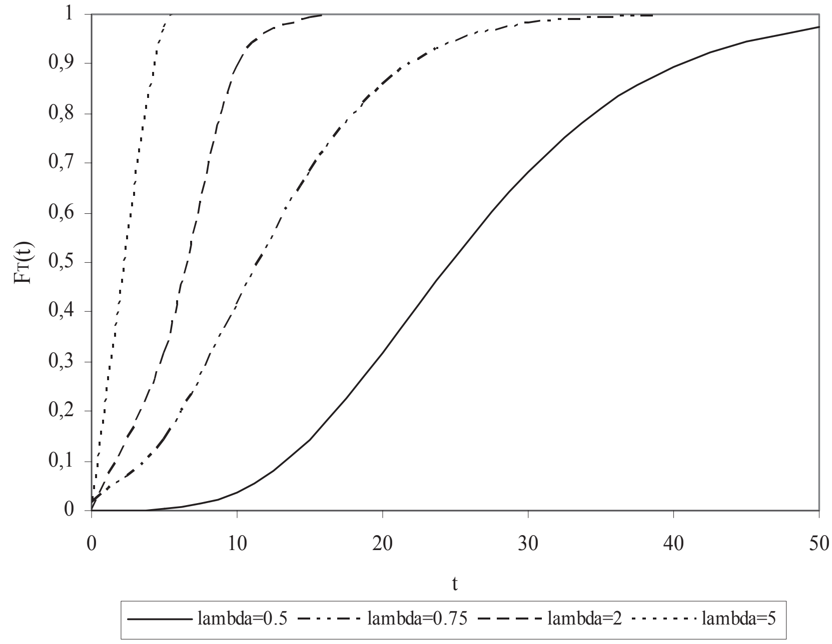

Figure 4. The distribution functions of T with $\beta = 10$ for several values of t where Y_i, $i = 1, 2, ...,$ have Poisson distribution with parameter $\mu = 1.5$. Note that $\{X_t, t \geq 0\}$ is a Neyman type A process.

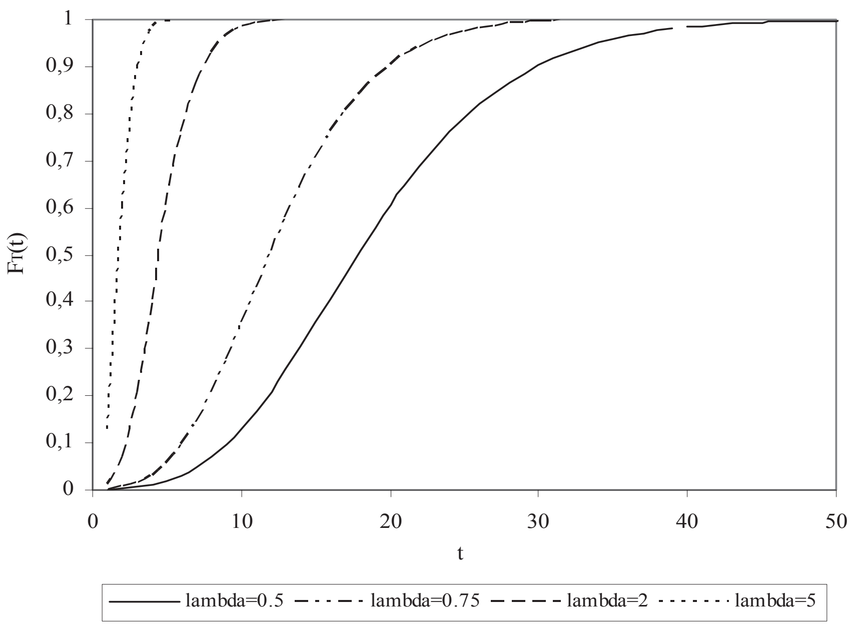

Figure 5. The distribution functions of T with $\beta = 10$ for several values of t where Y_i, $i = 1, 2, ...,$ have binomial distribution with parameters ($p = 0.2$, $m = 5$). Note that $\{X_t, t \geq 0\}$ is a Neyman type B process.

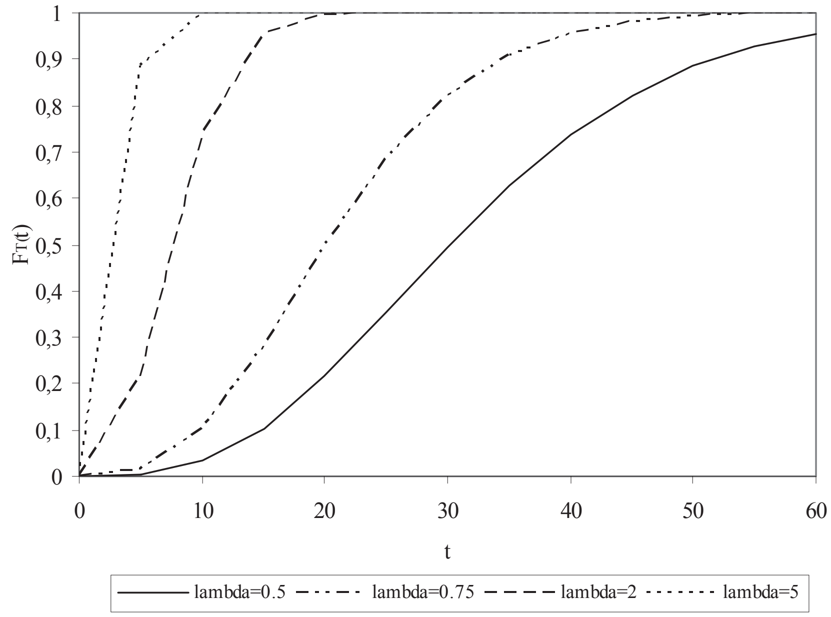

Figure 6. The distribution functions of T with $\beta = 10$ for several values of t where Y_i, $i = 1, 2, ...,$ have geometric distribution with parameter $\theta = 0.6$. Note that $\{X_t, t \geq 0\}$ is a Polya-Aeppli process.

4. APPLICATION OF THE COMPOUND POISSON PROCESS AND THE FIRST EXIT TIME TO THE TRAFFIC ACCIDENT DATA

Leiter and Hamdan (1973) suggested the joint distribution of the number of accidents and the number of fatal accidents can be expressed by a Poisson-Bernoulli model and the joint distribution of the number of accidents and the number of fatalities by a Neyman type A process. An alternative model with Neyman type B process proposed by Cacoullos and Papageorgiou (1980, 1982) to express relationship between the number of accidents and the number of fatalities. Van der Laan and Louther (1986) constructed a compound Poisson model for the costs of traffic accidents for Dutch motorists. Recently, the Neyman type A process has been used by Djauhari (2002) to explain traffic accidents and fatalities in Indonesia. Meintanis (2007) obtained a new goodness of fit test for certain bivariate distributions based on accident data and fatalities in The Netherlands. The data were obtained from the database of BRON of the Ministry of Transport, The Netherlands. In particular, total accidents and fatalities recorded on Sundays of each month over the period 1997-2004 in the region of Groningen are given in Table 1.

Table 1. Total Sunday accidents (left entry) and the corresponding number of fatalities (right entry) recorded in the region Groningen for each month during the years 1997-2004.

Month	**1997**		**1998**		**1999**		**2000**		**2001**		**2002**		**2003**		**2004**	
January	6	0	6	0	13	1	11	0	8	0	8	0	11	4	2	0
February	10	0	10	1	7	0	4	0	8	1	8	0	9	0	2	0
March	7	0	13	4	8	0	10	0	6	0	12	0	9	0	3	0
April	11	0	5	0	14	1	15	1	9	0	10	1	7	1	1	1
May	12	0	17	2	13	0	18	0	13	2	11	0	12	1	5	0
June	21	1	19	0	14	0	21	1	12	3	12	1	13	0	7	2
July	15	0	10	0	14	0	11	1	10	2	4	0	8	0	1	0
August	11	1	11	1	10	0	8	0	9	0	14	1	6	0	5	0
September	7	0	11	0	7	0	9	0	22	1	16	1	7	0	8	1
October	11	2	13	1	16	1	14	0	15	1	8	1	6	1	2	0
November	15	1	17	1	13	0	13	0	6	0	9	1	11	1	1	0
December	5	0	7	0	10	1	10	0	10	0	8	0	5	0	2	0

In this study the same data is used to show applicability of the compound Poisson process and the first exit time. The numbers of accidents and fatalities in Groningen between years 1997-2004 are presented in Figure 7.

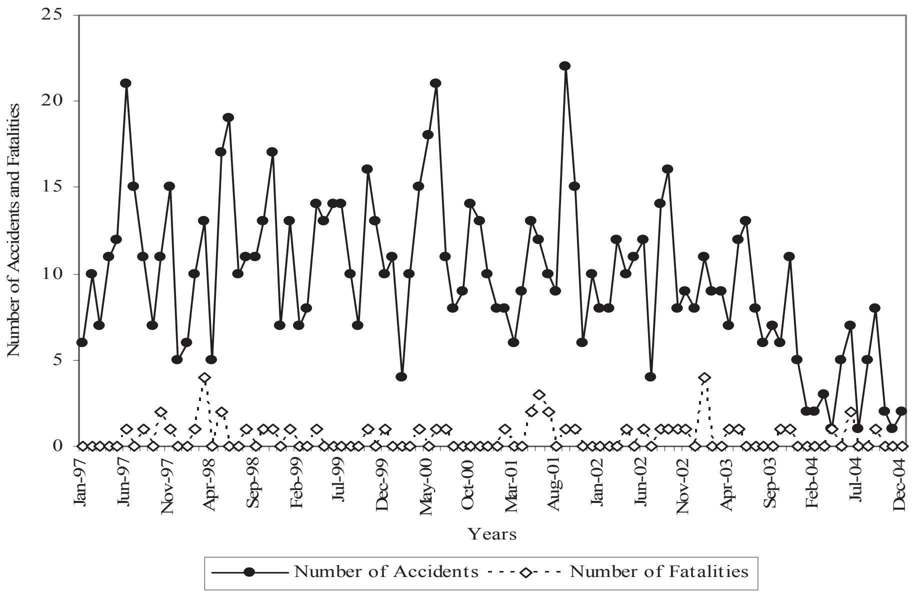

Figure 7. The numbers of accidents and fatalities by years.

For the construction of a model to explain the total number of fatalities from the accidents the following random variables are defined:

N_t : The number of accidents which occur in Groningen between years 1997-2004;

Y_i : The number of fatalities of i^{th} accident such that $i = 1, 2, 3, \ldots$;

X_t : The total number of fatalities in the time interval $(0, t]$.

We then get $X_t = \sum_{i=1}^{N_t} Y_i$ and we say that $\{X_t, t \geq 0\}$ is a compound Poisson process if the following conditions are hold:

Condition 1. Fit of the Homogeneous Poisson Process to Accidents: In the literature on accident statistics we mostly find the number of accidents is Poisson distributed during the time interval of length $t > 0$ (Djauhari, 2002; Shanmungan and Singh, 1980). Using goodness of fit test ($\chi^2 = 2.94$),we have seen that the number of the accidents which occur in Groningen between years 1997-2004, defined as $\{N_t, t \geq 0\}$, fit a homogeneous Poisson process with $\lambda = 9.84$ (month). It means that the expected number of traffic accidents approximately equals to 10 per month. The probabilities of the traffic accidents for $t = 1, 2, 3$ months are shown in Figure 8.

Condition 2. Independence Test of the Process $\{N_t, t \geq 0\}$ and Y_i, $i = 1, 2, 3, \ldots$: Now it must be shown the independence of Y_i, $i = 1, 2, \ldots$ and $\{N_t, t \geq 0\}$. According to Spearman's ρ test (Spearman's $\rho = 0.084$; $p = 0.432$), the independence is accepted.

Condition 3. Fit of the Poisson, Binomial or Geometric Distributions to Fatalities: To decide the best distribution between Poisson distribution, binomial distribution and geometric distribution for the number of fatalities, the goodness of fit tests were performed and the results are presented in Table 2. It is seen in Table 2 that the chi-square values are less than the critical table values for each distribution at the 5% level of significance. This means that the Poisson distribution with parameter $\mu = 0.53$, binomial distribution with parameters $(m = 4,\ p = 0.13)$ and geometric distribution with parameter $\theta = 0.62$ significantly fit the data.

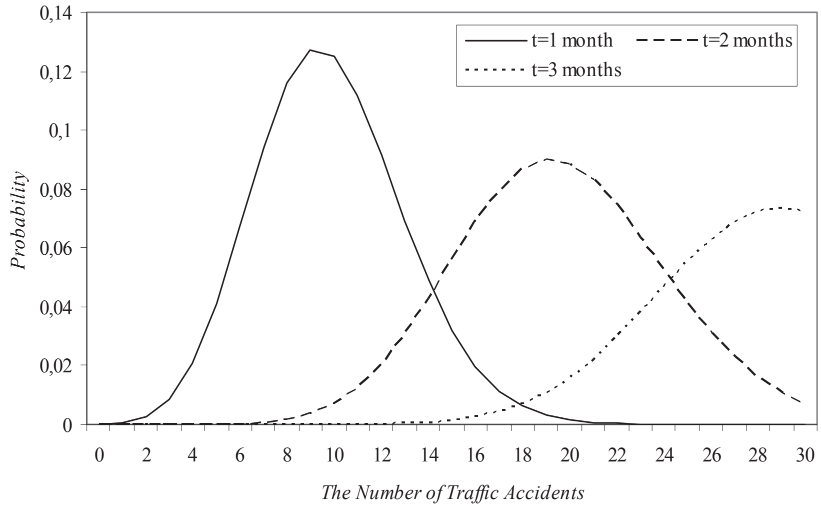

Figure 8. The occurrence probabilities of traffic accidents within $t = 1, 2, 3$ months.

Table 2. Comparison of fit of Poisson and binomial distributions to observed frequency for fatalities.

Number of Fatalities	**Observed Frequency**	**Expected Frequency**		
		Poisson	**Binomial**	**Geometric**
0	59	56.51	55.00	56.04
1	29	29.95	32.87	22.73
2	5	7.94	7.37	8.75
3	1	1.40	0.73	3.37
4	2	0.19	0.03	1.30
Total	96	96.00	96.00	95.18
Chi-Square		0.202	1.519	0.057
d.f.		2	3	2

For our data, binomial distribution is more appropriate than geometric and Poisson distributions for the random variables Y_i, $i = 1, 2, 3, \ldots$, since the number of fatalities are small. As mentioned in the study of Cacoullos and Papageorgiou (1980), Neyman type B process is almost as good as the Neyman type A process for small number of m. As m gets larger, the Neyman type B process approaches the Neyman type A process. This is expected since the values of p, the probability of a passenger being killed in a traffic accident, are small and so the binomial distribution can be approximated by the Poisson distribution. Naturally, the Neyman type B process gets closer to the Neyman type A process as m increases.

Since all conditions are hold, it can be said that $\{X_t, t \geq 0\}$ is a Neyman type B process. So, the probabilities $P(X_t = k)$, $k = 0, 1, 2, \ldots$, can be computed from Eq. (6) easily with our algorithm. The probabilities of the total number of fatalities for $t = 1, 2, 3$ months are given in Figure 9 and as seen in Figure 9, the probability of the total fatality number rapidly decreased after $k = 20$.

The distribution functions of T are presented in Figure 10 where the random variables Y_i, $i = 1, 2, ...,$ have binomial distribution with parameters $p = 0.13$ and $m = 4$. In Figure 10, $\{N_t, t \geq 0\}$ is a homogeneous Poisson process with $\lambda = 9.84$ and the upper horizontal boundaries are taken as $\beta = 5$ and $\beta = 10$.

5. CONCLUSION

We conclude with the comment that the probabilities $P(X_t = k)$, $k = 0, 1, 2, ...,$ can be computed and the distribution function of the first exit time can be obtained easily if $p_0 = P(Y_i = 0)$, $p_1 = P(Y_i = 1), ..., p_m = P(Y_i = m)$ are known.

In this study, a direct algorithm is proposed to compute the probabilities $P(X_t = k)$, $k = 0, 1, 2, ...$ and an efficient way is given to compute the distribution function and the probability density function of the first exit time for a compound Poisson process when the jump sizes are discrete random variables and the boundary is upper horizontal. Then, numerical examples and an application to traffic accident data have been presented to illustrate the usage of the probability function of the compound Poisson process and the distribution function of the first exit time using their algorithms in Oracle database.

Another important result of this study is that the probability function of the compound Poisson process and the distribution function of the first exit time can be obtained when Y_i, $i = 1, 2, 3, ...,$ have any discrete distribution and new algorithms are not required since our algorithms can be used for the general compound Poisson process.

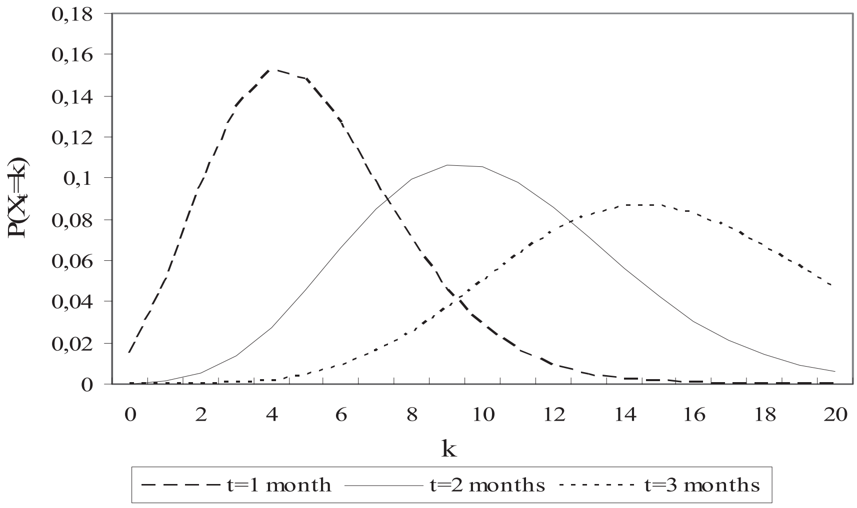

Figure 9. The probability of total fatality number which will occur within $t = 1, 2, 3$ months.

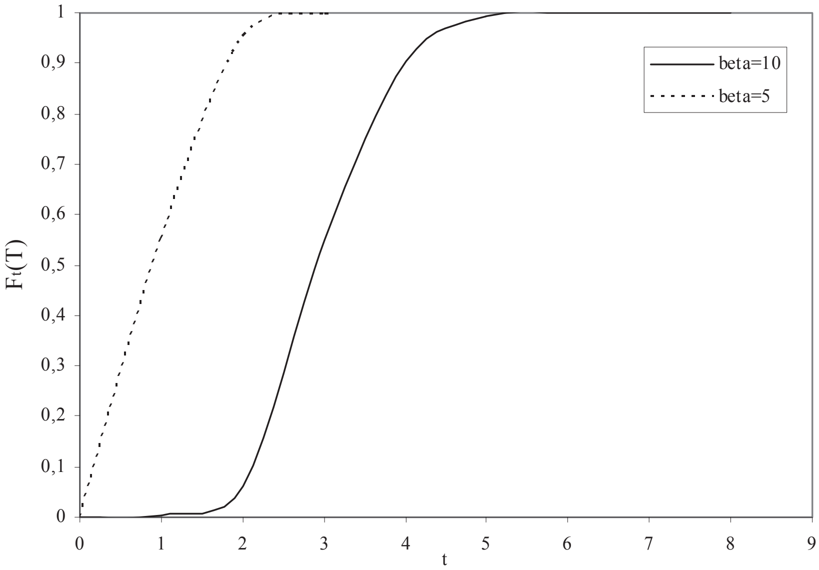

Figure 10. The distribution functions of T for $\beta = 5$ and $\beta = 10$ with several values of t where Y_i, $i = 1, 2, ...,$ have binomial distribution with parameters ($p = 0.13$, $m = 4$).

REFERENCES

Ammussen, S. (2000). *Ruin Probabilities*. World Scientific.

Bar-Lev, S., Bshouty, D., Perry, D. and Zacks, S. (1999). Distributions of stopping times for compound Poisson processes. *Stochastic Models* 15(1), 89-101.

Cacoullos, T. and Papageorgiou, H. (1980). On some bivariate probability models applicable to traffic accidents and fatalities. *International Statistical Review* 48, 345-356.

Cacoullos, T. and Papageorgiou, H. (1982). Bivariate negative binomial-Poisson and negative binomial-Bernoulli models with an application to accident data. *In Statistics and Probability: Essays in Honor of C. R. Rao, G. Kallianpur et al. (eds)* 155-168.

Delucia, J. and Poor, H.V. (1997). The moment generating function of the stopping time for a linearly stopped Poisson process. *Stochastic Models* 13, 275-292.

Dirickx, Y.M.I. and Koevoets, D. (1997). A continuous review inventory model with compound Poisson demand process and stochastic lead time. *Naval Research Logistics Quarterly* 24(4), 577.

Djauhari, M.A. (2002). Stochastic pattern of traffic accidents in Bandung city, *IATSS Research* 26(2), 85-91.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1953). Sequential decision problems for processes with continuous time parameter. *Annals of Mathematical Statistics* 24, 254-264.

Gallot, S.F.L. (1983). Absorption and first passage times for a compound process in a general upper boundary. *Journal of Applied Probability* 30, 835-850.

Ignatov, Z.G. and Kaishev, V.K. (2004). A finite-time ruin probability formula for continuous claim severites. *Journal of Applied Probability* 41(2), 570-578.

Klugman, S.A., Panjer, H.H. and Willmot, G.E. (1998). *Loss Models*. John Wiley & Sons.

Leiter, R.E. and Hamdan, H.A. (1973). Some bivariate probability models applicable to traffic accidents and fatalities. *International Statistical Review* 41, 87-100.

Meintanis, S.G. (2007). A new goodness of fit test for certain bivariate distributions applicable to traffic accidents. *Statistical Methodology* 4, 22-34.

Ozel, G. and Inal, C. (2008). The probability function of the compound Poisson process and an application to aftershock sequence in Turkey. *Environmetrics* 19(1), 79-85.

Ozel, G. and Inal, C. (2008). The probability function of the compound Poisson distribution using integer partitions and Ferrer's graph. *Bulletin of Statistics and Economics* 2, 70-79, (2008).

Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. *Astin Bulletin* 12, 22-26.

Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1998). *Stochastic Processes for Insurance and Finance*. John Wiley & Sons.

Rosencrantz, W.A. (1983). Calculation of the Laplace transform of the length of the busy period for the M/G/1 queue via martingales. *Annals of Probability* 11, 817-818.

Shanmungan, R. and Singh, J. (1980). Some bivarite probability models applicable to traffic accidents and fatalities. *Proceedings of the NATO Advanced Study Institute: Statistical Distributions in Scientific Work*, 6, 95-103.

Van Der Laan, B.S. and Louter, A.S. (1986). A statistical model for the costs of passenger car traffic accidents. *The Statistician* 35(2), 163-174.