Periodic Non-stationary Arrival Processes in Queueing Networks and their Characterization*

B.Dimitrov,[1] D.Green, Jr,[1] V.Rykov,[2] Z.Krougly [3]

[1] Department of Mathematics & Sciences, Kettering University, Flint, Michigan, USA
1700 West Third Avenue, Flint, Vichigan 48504-4898, USA

[2] Department of Applied Mathematics & Computer Modelling
Faculty of Automatization & Computer Sciences,
Russian State Oil and Gas University,
Leninsky prospect 65, Moscow 117917, Russia

[3] Department of Statistical and Actuarial Sciences,
University of Western Ontario, London. Ont., Canada N6A5B9

Abstract. Some properties and statistical analysis of periodic Non-stationary Poisson Processes (NPP) via an almost-lack-of-memory (ALM) distributions are presented. Problems of parameters estimation and testing ALM distributions versus other distributions are considered. A procedure for calculating critical level and power of likelihood ratio test, based on Monte-Carlo simulation method is shown.

1 Introduction

It is well known that the traffic through a network depends on the time. There are intervals of higher activity alternating with others of low activity, which could be days, weeks, or seasons. There is a natural reason for such phenomena, due to the environment where a network and its users are surrounded. Many authors have noticed some periodicity in the intensities of emergency calls, phone calls, service needs, accidents, other casualty activities, etc.

The times $\{T_n\}$ when a need occurs form the points of a *point process* [12]. The expected number $\Lambda(t)$ of points occurred on an interval $[0, t)$ is called *hazard*

* This research was partially supported by RFFI Grant No. 01-05-90259.

rate of the point process. Its derivative (supposed to exist) $\lambda(t) = \Lambda'(t)$ is called *intensity*, or *rate* of the process.

When the *inter arrival times* (i.a.t.) $X_n = T_{n+1} - T_n$ between two consecutive points are independent identically distributed (i.i.d.) random variables (r.v.'s), the process is called *renewal process*, and these r.v.'s, vice versa, generate appropriate process in the following way

$$N(t) = \max\{n : T_n \leq t\}, \quad \text{with } T_n = \sum_{1 \leq i \leq n} X_i. \tag{1}$$

Usually, most of results in Queueing Theory and Queueing Networks are obtained under this assumption. Since it is fulfilled

$$\lim_{t\to\infty} \frac{\Lambda(t)}{t\mathbf{E}[X_n]} = 1 \quad \text{and} \quad \lim_{t\to\infty} \lambda(t) = \frac{1}{\mathbf{E}[X_n]},$$

most of respective results are converged to the stationary regime, i.e. some characteristics "beyond the horizon", in infinite time.

For the traffic in finite time there are not many of analytical results. Here we focus on non-stationarity, namely, on periodic non-stationarity. There are cases where the process intensity function is a periodic function. There exists number $c > 0$ such that $\lambda(t)$ satisfies the equation

$$\lambda(t + c) = \lambda(t) \qquad \text{for any} \quad t \geq 0. \tag{2}$$

This number c is called *period* (or *cycle*) of the process. We use some more general and rarely noticed results from probability theory in order to introduce the concept of an underlying r.v. X, generating a NPP, and vice versa. This relationship is then used to characterize a periodic NPP via the underlying r.v. X with ALM properties, and some statistical estimations of parameters on the ALM distributions based on observations of the associated process are discussed.

2 A characterization of a NPP

Let $\{N(t),\ t \leq 0\}$ be a NPP with the *leading function* (for definition see [11])

$$\Lambda(t) = \int_0^t \lambda(u)du, \tag{3}$$

which represents the expected number of points on $[0, t)$. It is assumed that

$$\Lambda(t) < \infty \quad \text{for any} \ \ t < \infty \qquad \text{and} \ \ \Lambda(t) \to \infty \quad \text{for} \ \ t \to \infty. \tag{4}$$

Special case of this process is the Stationary Poisson Process (SPP) with $\Lambda(t) = \lambda t$.

Notice that the i.a.t.'s for the NPP are not independent in general. For example, their joint multi-dimensional p.d.f. is

$$f_{X_1,X_2,\dots,X_n}(x_1,x_2,\dots,x_n) =$$
$$= \lambda(x_1)\lambda(x_1+x_2)\cdots\lambda(x_1+x_2+\cdots+x_n)e^{-\Lambda(x_1+x_2+\cdots+x_n)}.$$

It can not be represented in product form except the case of SPP, when $\Lambda(t) = \lambda t$, and $\lambda(t) = \lambda$. This means that the points of NPP can not be represented as sum of i.i.d. r.v. Nevertheless, there exists some r.v. X, generated by any NPP and procedure, such that the process is uniquely specified by a sequence of i.i.d. r.v., same as X. To introduce such a procedure and the appropriate r.v., we need the notion of *records*.

Definition 1. Let $\{X_n,\ n = 1,2,\dots\}$ be a sequence of i.i.d. r.v. The sequence of r.v. $\{T_n,\ n = 1,2,\dots\}$, where

$$T_1 = X_1,\ T_n = X_{\nu_n} \quad \text{with} \ \ \nu_n = \min\{k : X_k > T_{n-1} \ \text{for} \ \ k > \nu_{n-1}\}. \qquad (5)$$

is called the *sequence of records* of $\{X_n\}$. □

Assume, that $\lambda(t) \geq 0$ is known, and $\Lambda(t)$ is defined by (3) and has the properties (3). Introduce the function

$$F(t) = 1 - \exp\{-\Lambda(t)\} = 1 - \exp\{-\int_0^t \lambda(u)du\}, \quad t \geq 0. \qquad (6)$$

The function $F(t)$ possesses the properties of a cumulative distribution function (c.d.f.) defined as $F(t) = 0$ for $t \leq 0$. Therefore, there exists a r.v. X with c.d.f $F(t) = \mathbf{P}\{X \leq t\}$, which we call *associated with the NPP* $N(t)$.

Reversely: Let X be a continuous r.v. representing the time to arrival of a call in the system, with a c.d.f. $F(t)$, and p.d.f. $f(t)$. Consider the sequence of i.i.d. r.v. $\{X_1, X_2, \dots\}$ with c.d.f. $F(t)$. Define the records T_n of $\{X_n\}$ as points of process, and $N(t)$ similar to (1)

$$N(t) = \max\{n : T_n \leq t\}, \qquad (7)$$

with T_n defined by (5), instead of given there. The following theorem (see Feller [9], Karlin and Taylor [10]) shows that this procedure leads to NPP with leading function $\Lambda(t)$ and intensity function $\lambda(t)$, given by

$$\Lambda(t) = -\ln(1 - F(t)), \qquad \text{and} \qquad \lambda(t) = \frac{f(t)}{1 - F(t)}. \qquad (8)$$

Theorem 1. *The points of any NPP with leading function $\Lambda(t)$ are records of the sequence of i.i.d. r.v. with c.d.f. $F(x)$ defined by (6). Inversely, any sequence of i.i.d. r.v. with common c.d.f. $F(x)$ generates by (7) a NPP $N(t)$ with leading and intensity functions (8), whose points are the records of $\{X_n\}$, given by (5).*

Notice that this theorem provides some new interpretation of the nature of input processes in queueing networks.

3 Periodic NPP's and the ALM distributions

In this section we show how periodic NPP's is connected with recently introduced ALM distributed r.v. (see [1, 2, 6, 7]). We start with the general concept of LM property for r.v. The proofs of the following properties could be found in referred papers.

Definition 2. A non-negative r.v. X non-degenerating at zero with $0 < \mathbf{P}\{X \geq c\} < 1$ is said to have the *LM* property at the point $c > 0$ if and only if the equality

$$\mathbf{P}\{X \geq c + x \mid X \geq c\} = \mathbf{P}\{X \geq x\} \tag{9}$$

holds for any $x \geq 0$. We call c *point of regeneration* (r.p.) for the r.v. X and for its distribution. □

Notice that if this property holds for all points c it implies that X has either the exponential distribution (if X is continuous) or the geometric distribution (if X is discrete). However, if X has the LM-property at a point c it also has the LM-property at any point from the sequence $\{c_m = mc\}_{m=0}^{\infty}$. It holds

Theorem 2. *Let the distribution of a r.v. X has two r.p. a and b. Then*

(i) if a and b are commensurable, then there exists a number c such that X has LM property with respect to the sequence of r.p. $c_m = mc$, to which the numbers a and b belong;

(ii) if a and b are incommensurable, then the r.v X has exponential distribution.

These properties hint to give another definition

Definition 3.

A non - negative r.v. X has the ALM property if there exists an infinite sequence of distinct constants $\{c_m = mc\}_{m=0}^{\infty}$ such that the equality

$$\mathbf{P}(X \geq c_m + x \mid X \geq c_m) = P(X \geq x) \tag{10}$$

holds for any c_m and any $x \geq 0$.

The following theorem ([5]) gives a characterization of the class of ALM distributions with $[x]$ is the largest integer less then x.

Theorem 3. *A r.v. X has the ALM property over the sequence $\{c_m = mc\}_{m=0}^{\infty}$ if and only if*
(i) in continuous case its p.d.f. $f_X(x)$, $x \geq 0$ has the form

$$f_X(x) = (1-a)a^{[x/c]} f_Y(x - [x/c]c), \tag{11}$$

where $a = \mathbf{P}\{X \geq c\}$, and $f_Y(.)$ is the p.d.f. of a continuous r.v. Y with support $[0, c)$;
(ii) in discrete case its p.m.f. $p_X(x)$ has the form

$$p_X(x) = (1-a)a^{[x/c]} p_Y(x - [x/c]c) \tag{12}$$

where $a = \mathbf{P}\{X \geq\}$ and $p_Y(.)$ is the p.m.f. of a discrete r.v. Y with support $\{0, 1, \ldots, c-1\}$;
(iii) moreover in any case it can be represented as

$$X = Z_c + cK \tag{13}$$

where Z_c and K are independent r.v.'s, Z_c is concentrated on $[0, c)$ and K has a geometric distribution with parameter a and p.m.f. $p_K(k) = (1-a)a^k$, $k = 0, 1, \ldots$.

The next theorem shows the connection between NPP and family of ALM distributions.

Theorem 4. *The r.v. X associated with any periodic NPP possesses the ALM property, and, on reverse, any NPP is generated by some ALM distributed r.v. X in the sense of the Theorem 1.*

Consider now some statistical properties of ALM distributions.

4 Parameter Estimation

For parameter estimation one can use both maximum likelihood (ML) Method and/or the method of moments and will get the same results. Let us show how moment's method works in the case, observation on X available.

It is well known that for any sample $X_1, \dots, X_n$ the sample mean

$$\bar{X} = \frac{1}{n} \sum_{1 \le i \le n} X_i$$

is the best estimator for expected value. From (13) one can find that

$$\mu_X = \mathbf{E}X = \mathbf{E}Z + c\,\mathbf{E}K = \mu_Z + c\frac{a}{1-a}.$$

By changing expectations with sample means, which are their estimations one can obtained that

$$\frac{\bar{X} - \bar{Z}}{c} = \frac{\hat{a}}{1-\hat{a}}, \qquad \text{or} \quad \hat{a} = \frac{\bar{X} - \bar{Z}}{c\left(1 + \frac{\bar{X} - \bar{Z}}{c}\right)} = \frac{\bar{X} - \bar{Z}}{c + \bar{X} - \bar{Z}}.$$

Here

$$\bar{Z} = \frac{1}{n} \sum_{1 \le i \le n} Z_i, \qquad \text{with} \quad Z_i = X_i - \left[\frac{X_i}{c}\right] \quad \text{for} \quad mc \le X_i < (m+1)c.$$

Analogous estimators was given in ([4]) by the ML method.

5 Likelihood Ratio Test

It is intuitively clear (and proved in [8]), that when $c \to 0$, the class of ALM distributions should tend to the exponential. This remark shows that it is of interest to consider the exponential class as a competing class of distributions versus the ALM-property with any sequence c_m. On the other hand, the parameter c usually may be known from practice reason. Therefore it is reasonable to compare two ALM distributions with the same value of parameter c and different p.d.f. on the cycle.

Let $X_1, \dots, X_n$ be an independent random sample from unknown distribution with p.d.f. $f(x)$. We are interested in testing the null hypothesis

$$H_0 : f(x) = f_0(x)$$

versus the alternative

$$H_1 : f(x) = f_1(x).$$

According to Neumann-Pearson theorem, the most powerful test for testing H_0 versus H_1 is the likelihood ratio test. For this test, due to independence of observations, the critical region can be represented in the form

$$W = \left\{ (x_1, \ldots, x_n) : \frac{l_1(x_1, \ldots, x_n)}{l_0(x_1, \ldots, x_n)} \prod_{1 \le i \le n} \frac{f_1(x_i)}{f_0(x_i)} > t \right\}, \quad t > 0. \tag{14}$$

The significance level of the test is

$$\alpha = \mathbf{P}_{H_0}\{W\} = \mathbf{P}_{H_0}\{(X_1, \ldots X_n) \in W\},$$

and the power of the test is

$$1 - \beta = P_{H_1}\{W\}.$$

In calculating the critical value t_α for given significance level α, and power $1 - \beta$ of the test we use the Monte Carlo method. From numerical point of view instead of the product in formula (14) it would have been better to use its logarithm. For this reason we consider a natural logarithm of likelihood ratio. For convenience we will refer to the statistic

$$w = \ln \prod_{1 \le i \le n} \frac{f_1(x_i)}{f_0(x_i)} = \sum_{1 \le i \le n} (\ln f_1(x_i) - \ln f_0(x_i)) \tag{15}$$

as to a test's statistic.

In cases of large sized samples the statistics (15) has approximately normal distribution. It allows us to limit ourselves in calculation of only its two moments and then calculate appropriate significance level and power of the test with the use of respective normal approximation.

To show how it works let us denote by U and V the r.v.'s

$$U = \ln f_1(X) - \ln f_0(X), \qquad V = \ln f_1(Y) - \ln f_0(Y),$$

where X and Y are taken from distributions with densities $f_0(.)$ and $f_1(.)$ corresponding to hypotheses H_0 and H_1. Denote by μ_U, μ_V and σ_U^2, σ_V^2 their expectations and variances respectively. Then, for large samples the test's statistics

2. B. Dimitrov, and Z. Khalil (1992) A class of new probability distributions for modeling environmental evolution with periodic behavior. *Environmetrics*, **3** (4), 447–464.
3. Chukova, S., Dimitrov, B. and Khalil, Z. (1993) A characterization of probability distributions similar to the exponential, *The Canadian Journal of Statistics*, **21** (3), 269-276.
4. Dimitrov, B., Khalil, Z., and Le Normand, L. (1993) A statistical study of periodic environmental processes, in *Proceedings of the International Congress on Modelling and Simulation*, Dec. 6-10, 1993, The Univ. of Western Australia, Perth, v. 41, pp. 169 - 174.
5. B. Dimitrov, S. Chukova, and D. Green Jr. (1997) Distributions in Periodic Random Environment and Their Applications. *SIAM J. on Appl. Math.* **57** (2), 501-517.
6. Dimitrov B., Chukova S. and Garrido J.(2000) Compound counting processes in periodic random environment, *J. of Statist. Research*, **34 (2)**, 99-111.
7. B. Dimitrov, Z. Khalil, M. Ghitany, and V. Rykov (2001) Likelihood Ratio Test for Almost Lack of Memory Distributions. *Technical Report No. 1/01, November 2001* Concordia University, 2001.
8. B. Dimitrov, V. Rykov, Z. Krougly, M. Ghitany (2003) On Properties and Statisticac Estimation of ALM Distributions. In: *Proceedings of Hawaii International Conference on Statistics and Related Fields*, Hawaii, Honolulu, June 4-8, 2003.
9. W. Feller (1966) *An Introduction to Probability Theory and its Applications*, Vol. 1, Wiley, N.-Y.
10. S. Karlin and H. Taylor (1975) *First Course in Stochastic Processes*, 2d ed., Academic Press, New York.
11. A. Khinchin (1963) *The works on mathematical theory of queuing.* FIZMATGIZ, M., 235 p. (In Russian)
12. D. Koenig, V. Rykov, V. Schmidt (1980) Stationary queuing systems with dependencies. In: *Itogi nauki i techniki, Teor. Ver. Mat. Stat. Teoretich Cybern.* V. 18, PP. 95-186, Moscow, VINITI (In Russian). Translated in English in: *Journal of Soviet. Math.*, V. 21 (1983), No. 6, pp. 938-994.