Automation and Remote Control, Vol. 65, No. 10, 2004, pp. 1597–1610. Translated from Avtomatika i Telemekhanika, No. 10, 2004, pp. 85–100.

Periodic Poisson Processes and Almost-lack-of-memory Distributions[1]

B. N. Dimitrov*, V. V. Rykov, and Z. L. Krougly*****

*Kettering University, Flint, Michigan, USA

**Gubkin State University of Oil and Gas, Moscow, Russia

***University of Western Ontario, London, Ontario, Canada

Received April 15, 2004

Abstract—Certain characterization properties of time-varying periodic Poisson flows are studied in terms of almost-lack-of-memory (ALM) distributions. Parameter estimation formulas are derived. A method for verifying the hypothesis on the membership of a sample to the class of ALM-distributions is developed. Algorithms for computing critical levels and power of the likelihood ratio test by the Monte Carlo method are designed.

1. INTRODUCTION

Many natural phenomena, including flows in data transfer networks, communication systems, reliability models, ecological data descriptions, etc., as is known, depend on time. Several researchers have noticed the periodic nature of the intensity of first-aid calls, telephone calls, inquires at different services, accidents, natural collisions etc. High intensity intervals alternate with relatively low loads. Such intervals may be days, weeks, or even seasons. There are obvious reasons for the phenomena due to the environmental medium and users of the network.

It is convenient to describe event flows in terms of point processes [1, 2]. A *point process* is an ordered sequence of instants $\{T_n : \ T_n < T_{n+1}, \ n = 1, 2, \dots\}$ of occurrence of events. The process

$$N(t) = \max\{n : T_n < t\}, \tag{1}$$

reckoning the number of events occurring in the interval $[0, t)$ is called the *counter* of the point process. The mean number $\Lambda(t) = \mathbf{E}N(t)$ of events in an interval $[0, t)$ is called the *intensity function* of the point process. Its derivative $\lambda(t) = \Lambda'(t)$ (which is assumed to exist) is called the *intensity density*, or simply the *intensity* of the process.

If the *intervals between instants of occurrence of events* $X_n = T_n - T_{n-1}$, $n = 1, 2, \dots, T_0 = 0$, form a sequence of identically distributed independent random variables, the process is said to be a *recurrent* or *renewal* process. Such a process and its counter $N(t)$ are uniquely generated by the respective sequence of identically distributed independent random variables by the relation

$$T_n = \sum_{1 \le i \le n} X_i, \quad T_0 = 0.$$

Most results in queueing and stochastic network theory have been derived under this assumption and pertain to the stationary operation of systems, i.e., characterize the performance of the system "at infinity" as $t \to \infty$ by the relations

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

[1] This work was supported in part by the Russian Foundation for Basic Research, projects nos. 01-07-90259 and 04-0790115-v.

Analytical results for flows on a finite time interval are not many in number. Practical needs are ensured by numerical methods and imitation modeling. These approaches may aid in finding an approximate solution to any practical problem, but are not adequate to meet the requirements of in-depth theoretical analysis.

In this paper, we study time-varying, viz., periodic time-varying flows, confining ourselves to the class of periodic time-varying Poisson flows. Consequently, the event flow intensity is a periodic function, i.e., there exists a number $c > 0$ such that the intensity $\lambda(t)$ satisfies the relation

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

The least of these numbers c is called the *period* or *cycle* of the process.

In Section 2, we use the well-known results of probability theory to introduce a procedure and a random variable X generated by a time-varying Poisson flow such that the corresponding process can be uniquely estimated by this procedure from a sequence of identically distributed independent random variables having the same distribution as X. This relationship is used in Section 3 to characterize a periodic time-varying Poisson flow in terms of its generating random variable X, which, incidentally, has the property of partial lack of memory. In Sections 4 and 5, we examine certain statistical properties of parameter estimates and testing of hypothesis for these distributions based on observations on the corresponding process. In Sections 6 and 7, we describe algorithms and computer programs for computing critical values and power of hypothesis testing criteria for a wide class of ALM-distributions and give examples to illustrate the performance of these programs in Section 8.

2. CHARACTERIZATION OF TIME-VARYING POISSON FLOWS

Let us consider a time-varying Poisson flow $\{N(t),\ t \leq 0\}$ defined by the following properties [3].

(i) **Independence**: numbers of points (events) in any disjoint intervals are independent.

(ii) **Ordinariness**: any interval $[t, t + \Delta t)$ of small length Δt contains not more than one point

$$\mathbf{P}\{N(t + \Delta t) - N(t) = 1\} = \lambda(t)\Delta t + \bar{\bar{o}}(\Delta t), \quad \text{where} \quad \lambda(t) \geq 0, \quad \text{and}$$

$$\mathbf{P}\{N(t + \Delta t) - N(t) > 1\} = \bar{\bar{o}}(\Delta t) \quad \text{as} \quad \Delta t \to 0.$$

Under these assumptions, the number of points in any interval $[t, t + s)$ is Poisson distributed

$$\mathbf{P}\{N(t + s) - N(t) = n\} = \frac{[\Lambda(t + s) - \Lambda(t)]^n}{n!} \exp\{-[\Lambda(t + s) - \Lambda(t)]\}, \quad t, s \geq 0, \quad n = 0, 1, \ldots.$$

Furthermore, the function

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

called the *leading function* of the process [3], is equal to the mean number of points of the process in the interval $[0, t)$, i.e., the intensity function of the corresponding point process. We take

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

A particular case of this process is the stationary Poisson process for which $\Lambda(t) = \lambda t$.

Note that the event inter-occurrence intervals of a time-varying Poisson flow are not independent random variables. For example, their multidimensional distribution density is of the form

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \lambda(x_1)e^{-\Lambda(x_1)}\lambda(x_1+x_2)e^{-[\Lambda(x_1+x_2)-\Lambda(x_1)]}$$
$$\times\cdots\times\lambda(x_1+x_2+\cdots+x_n)e^{-[\Lambda(x_1+x_2+\cdots+x_n)-\Lambda(x_1+\cdots+x_{n-1})]}$$
$$= \lambda(x_1)\cdots\lambda(x_1+x_2+\cdots+x_n)e^{-\Lambda(x_1+x_2+\cdots+x_n)}$$

and can be expressed as the product of distribution densities of intervals, except for the one-dimensional case, for which $\Lambda(t) = \lambda t$ and $\lambda(t) = \lambda$.

Therefore, points of a time-varying Poisson flow cannot be expressed as the sum of identically distributed independent random variables. Nevertheless, there is a random variable X generated by a time-varying Poisson flow $N(t)$ and a procedure for which the time-varying Poisson flow in turn is uniquely renewed by a sequence of identically distributed independent random variables having the same distribution as X. We refer to such a random variable as the *generating* random variable of the time-varying Poisson flow. To introduce the corresponding procedure and random variable, we require the concept of *records*.

Definition 1. For a sequence of identically distributed independent random variables $\{X_n,\ n = 1, 2, \ldots\}$, the recursive sequence of the random variable $\{T_n,\ n = 1, 2, \ldots\}$ for which

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

is called the sequence of records of the sequence $\{X_n,\ n = 1, 2, \ldots\}$.

The main assertion of this section is the existence of a random variable X generated by a time-varying Poisson flow $N(t)$ such that this time-varying Poisson flow with the properties described above is renewed by the record sequence defined by a sequence of identically distributed independent random variables having the same distribution as X.

Let us assume that $\lambda(t) \geq 0$ is known and $\Lambda(t)$ is defined by relation (3) and has properties (4). Let us introduce a function

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

If supplemented with $F(t) = 0$, $t \leq 0$, this function has all properties of a distribution function. Consequently, there exists a random variable X having this distribution function, $F(t) = \mathbf{P}\{X \leq t\}$, called the *generating random variable of the time-varying Poisson flow* $N(t)$. Note that this random variable X is continuous and has a distribution density, equal to

$$f_X(t) = F'(t) = \lambda(t)\exp\{-\Lambda(t)\}, \quad t \geq 0, \tag{7}$$

with carrier in $[0, \infty)$. Consequently, the random variable X can be used to define the *time to the occurrence of some event* (or the *time to the occurrence of some call*) (in the terminology of reliability theory, it can also be called the continuous operation time). In particular, the function

$$\lambda_X(t) = \frac{f_X(t)}{1 - F_X(t)} = \lambda(t) \tag{8}$$

can be regarded as the *occurrence intensity* of an event (*arrival intensity* of a call or *failure hazard*) with accumulated intensity function

$$\Lambda_X(t) = \int_0^t \lambda_X(u)du = \Lambda(t).$$

Thus there is a complete correspondence between the main characteristics of an individual call in a packet of statistically equivalent calls related to some time-varying Poisson flow and the corresponding properties of the time-varying Poisson flow itself.

Conversely, let X be a continuous random variable, which defines the time to the occurrence of an event of a flow with distribution function $F(t)$ and distribution density $f(t)$. Consider a sequence of random variables $\{X_1, X_2, \dots\}$ with distribution function $F(t)$. Let us find a point process with the help of the "records" of this sequence, i.e., the points of this process are constructed according to relation (5). We can show (see [4, 5]) that this procedure yields a time-varying Poisson flow with leading function $\Lambda(t)$ and intensity $\lambda(t)$.

Theorem 1. *The instants of occurrence of events of any time-varying Poisson flow with leading function $\Lambda(t)$ are the records of a sequence of identically distributed independent random variables $\{X\} = \{X_n,\ n = 1, 2, \dots\}$ with distribution function $F(x)$ (6). Conversely, any such sequence $\{X\}$ with distribution function $F(x)$ generates a time-varying Poisson flow defined by the relation*

$$N(t) = \max\{n :\ T_n \le t\},$$

where $T = \{T_n,\ n = 1, 2, \dots\}$ is the sequence of records of the initial sequence $\{X\}$ (5).

Note that this theorem gives a new interpretation to the nature of the input flow of queueing systems and networks. This procedure can be interpreted as follows: all calls arrive at the system concurrently, but each with its own service commencement instant X_n. The dispatcher receives the inquiry from the first call and admits it to the system at instant X_1, while all other calls with an earlier service commencement instant are lost. The customer with the first service commencement instant greater than X_1 is taken next for service, etc.

3. PERIODIC TIME-VARYING POISSON FLOWS AND ALM-DISTRIBUTIONS

In this section, we show the relationship between *periodic* time-varying Poisson flows and distributions with almost lack of memory (see [6–11]) (ALM-distributions), see [6–11]. We begin with the general description of the lack of memory property of a random variable. Let us assume that a random variable describes the inter-occurrence intervals of events at some outer medium, which periodically changes with interval $c > 0$. Then the time axis can be subdivided into intervals $[0, c), [c, 2c), \dots, [mc, (m+1)c), \dots$. We assume that the conditions of formation on each of the intervals (cycles) are of identical randomness. Let a random variable Y_m define the time interval up to the occurrence of an event inside a given interval $[mc, (m+1)c)$, $m = 0, 1, \dots$. Note that this event may not occur in a given interval with some probability $0 < a < 1$. Thus, the waiting time X for the occurrence of the first event in a long time interval in the periodic phenomenon described above is a random number of cycles on which the event had not occurred plus the time up to the occurrence of this event in a separate cycle. Formally, the property of partial lack of memory of a random variable is formulated by

Definition 2. A nonnegative nonvanishing random variable X is said to lack memory at a point $c > 0$ if

$$\mathbf{P}\{X \ge c + x | X \ge c\} = \mathbf{P}\{X \ge x\} \quad \text{for all} \quad x \ge 0. \tag{9}$$

The point c is called the regeneration point of the random variable X and its distribution.

Remarks

(1) Obviously, this concept is meaningful only if $0 < \mathbf{P}\{X \ge c\} = a < 1$.

(2) If a random variable X lacks memory at all positive points c, then it has either an exponential (in the continuous case) or a geometric (in the discrete case) distribution. Moreover, we have

Lemma 1. *If a random variable X lacks memory at a point $c > 0$, then it lacks memory at all points of the sequence* $\{c_m = mc\}_{m=0}^{\infty}$.

Proof. By the condition of the lemma, the random variable X lacks memory at a point $c > 0$. Assume that it lacks memory at all points $c_k = kc$, $k \le m$. Then

$$\mathbf{P}\{X > c_{m+1} + x | X > c_{m+1}\} = \frac{\mathbf{P}\{X > c_{m+1} + x,\ X > c_{m+1}\}}{\mathbf{P}\{X > c_{m+1}\}}$$

$$= \frac{\mathbf{P}\{X > c + c_m + x | X > c\}\mathbf{P}\{X > c\}}{\mathbf{P}\{X > c_{m+1} | X > c\}\mathbf{P}\{X > c\}}$$

$$= \frac{\mathbf{P}\{X > c_m + x|\}}{\mathbf{P}\{X > c_m\}} = \mathbf{P}\{X > c_m + x | X > c_m\}. \quad \square$$

Furthermore, every random variable X has not more than one sequence for which the lack of memory property holds, or it has an exponential distribution.

Lemma 2. *A random variable X lacking memory at two points $a > 0$ and $b > 0$ also lacks memory at all positive points $c_{kl} = kb - la$, where k and l are (positive or negative) integers.*

Proof. Assume that $a < b$. By Definition (7),

$$\mathbf{P}\{X > x\} = \mathbf{P}\{X > b + x | X > b\} = \frac{\mathbf{P}\{X > b + x\}}{\mathbf{P}\{X > b\}}$$

$$= \frac{\mathbf{P}\{X > b + x,\ X > a\}}{\mathbf{P}\{X > b\}} + \frac{\mathbf{P}\{X > b + x,\ X \le a\}}{\mathbf{P}\{X > b\}}$$

$$= \frac{\mathbf{P}\{X > a + b - a + x | X > a\}\mathbf{P}\{X > a\}}{\mathbf{P}\{X > a + b - a | X > a\}\mathbf{P}\{X > a\}}$$

$$= \frac{\mathbf{P}\{X > b - a + x\}}{\mathbf{P}\{X > b - a\}} = \mathbf{P}\{X > b - a + x | X > b - a\}.$$

In the second equality, the probability of the second term is zero since the events within braces are incompatible. The last inequality implies that the point $b - a$ is also a regeneration point for the random variable X. By Lemma 1, all positive numbers $(kb - la)$ are also generation points. $\square$

This lemma implies

Theorem 2. *For a random variable X having two regeneration points a and b,*

(i) if a and b commensurable, there exists a number c such that X lacks memory at the points of a sequence $c_m = mc$ containing the points a and b,

(ii) if a and b are incommensurable, then the random variable X has an exponential distribution (or a geometrical distribution in the discrete case).

Proof. If the numbers a and b are commensurable, assume that $a < b$, i.e., $\frac{a}{b} = r = \frac{p}{q} < 1$, where r is a rational number and p and q are irreducible integers. Repeatedly applying the procedure used in the proof of Lemma 2, we can show that the number $c = \frac{b}{q}$ can be expressed as $(kb - la)$ with integral k and l, i.e., is a regeneration point by this lemma. Hence all positive numbers of the type $(kb - la)$ can be expressed as $\frac{nb}{q}$. Consequently, the numbers $a = pc$ and $b = qc$ also belong to this sequence.

If the numbers a and b are incommensurable, then any number c can be expressed with any degree of accuracy as a number of the type $(kb - la)$. Therefore, every number is a regeneration point of the random variable X and, by Remark 2, has an exponential distribution. $\square$

This property prompts us to generalize Definition 2.

Definition 3. The distribution of a nonnegative nonvanishing random variable X is said to be a distribution with almost lack of memory (ALM-distribution) if there exists an infinite sequence of numbers $\{c_m = mc\}_{m=0}^{\infty}$ such that

$$\mathbf{P}\{X \geq c_m + x | X \geq c_m\} = \mathbf{P}\{X \geq x\} \quad \text{for all} \quad c_m \quad \text{and any} \quad x \geq 0. \tag{10}$$

Theorem 3 describes the main characterization properties of the class of ALM-distributions. Its proof is given in [6–8]. In the formulation given below, $[x]$ denotes the largest integer not greater than x and the parameter a here and in what follows has nothing common with the letter a used in Lemma 2 and Theorem 2 to denote the regeneration point.

Theorem 3. *A nonnegative nonvanishing random variable X (interpreted as the time up to the occurrence of an event) has an ALM-distribution for the sequence $\{c_m = mc\}_{m=0}^{\infty}$ if and only if any one of the following assertions holds:*

(i) The distribution function of the random variable X is of the form

$$F_X(x) = 1 - a^{[x/c]} \left(1 - (1-a)F_Y(x - [x/c]c)\right), \tag{11}$$

where $a = \mathbf{P}\{X \geq c\}$ and $F_Y(.)$ is the distribution function of a random variable Y concentrated on the interval $[0, c)$.

(ii) The distribution density $f_X(x)$, $x \geq 0$, of the continuous random variable X is of the form

$$f_X(x) = (1-a)a^{[x/c]} f_Y(x - [x/c]c), \tag{12}$$

where $a = \mathbf{P}\{X \geq c\}$ and $f_Y(.)$ is the distribution density of a continuous random variable Y with carrier $[0, c)$.

The distribution $f_X(x)$ of a discrete random variable X is defined by the same formula, where $f_Y(.)$ is the distribution of a discrete random variable with carrier $\{0, 1, \ldots, c-1\}$.

(iii) The event occurrence intensity

$$\lambda_X(x) = \frac{(1-a)\, f_Y(x - [x/c]c)}{1 - (1-a)F_Y(x - [x/c]c)} \tag{13}$$

for the random variable X is a periodic function with period c.

(iv) The random variable X can be expressed as

$$X = Z_c + cK, \tag{14}$$

where the independent random variables Z_c is concentrated on the interval $[0, c)$ and the random variable K has a geometric distribution with parameter a, $p_K(k) = (1-a)a^k$, $k = 0, 1, \ldots$.

Note that the fourth property contains an equivalent representation for the class of ALM-distributions, which is useful in modeling these distributions.

Finally, let us formulate a theorem establishing a relationship between periodic time-varying Poisson flows and the class of ALM-distributions.

Theorem 4. *The random variable generating a time-varying Poisson flow has a distribution with lack of memory. Conversely, any random variable having an ALM-distribution generates a periodic time-varying Poisson flow by the record procedure.*

Proof. By Theorem 1 on the relationship between time-varying Poisson flows and records of the sequence of its generating random variable, it suffices to show that the random variable generating the periodic time-varying Poisson flow has one of the properties (i)–(iv) stated in Theorem 3. Indeed, this is true since the intensity of failures of a periodic time-varying Poisson flow is a periodic function $\lambda(t+c) = \lambda(t)$ and, by virtue of relation (6), its generating random variable X has a distribution, which for $mc \le t < (m+1)c$ can be expressed as

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(u)du\right\} = 1 - \left[\exp\left\{-\int_0^c \lambda(u)du\right\}\right]^m \exp\left\{-\int_0^{t-mc} \lambda(u)du\right\}$$
$$= 1 - a^m(1 - F(t-mc)),$$

which for $a = \exp\{-\int_0^c \lambda(u)du\}$ is the same as (11).

On the other hand, according to (13), the intensity of occurrence of an event of the random variable X with an ALM-distribution is a periodic function and, as a consequence of relation (8), the corresponding time-varying Poisson flow is periodic. □

Now we describe certain statistical properties of ALM-distributions.

4. PARAMETER ESTIMATION

The parameters of an ALM-distribution can be estimated either by the maximum likelihood or the moment method; both yield the same results [9–12]. We could not succeed in estimating the parameter c. In applications, it is usually judged from "physical" considerations. Therefore, assuming that it is known, we shall show that a simple estimate for the parameter a can be obtained by the method of moments. As is known, the sample mean

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

of a sample $X_1, \ldots, X_n$ is the best mean of the expectation of the random variable. From relation (14) we obtain

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

In this relation, replacing the theoretical means by their sample analogs, which are their estimates, we obtain

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

where

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

A similar estimate is obtained in [9] by the maximum likelihood method. Let $X_{(1)}, \ldots, X_{(n)}$ be the ordered statistics of a random sample $X_1, \ldots, X_n$ from a general population with distribution density (12). Then the likelihood function takes the form

$$l(x_1, \ldots, x_n) = (1-a)^n a^{n\bar{k}} \prod_{0\le m\le r}\ \prod_{1\le j\le k_{m+1}} f_Y(x_{(K_m+j)} - mc), \tag{15}$$

where k_m is the number of terms in the sample in the interval $[mc, (m+1)c)$, $K_0 = 0$, $K_m = \sum_{j<m} k_j$, $\bar{k} = n^{-1}\sum_m mk_m$, and $r = [X_{(n)}/c]$. In [9], estimates of other parameters of ALM-distributions are found by the maximum likelihood method.

5. TESTING THE HYPOTHESIS ON ALM-DISTRIBUTIONS

Intuitively, it is clear that ALM-distributions must tend to an exponential distribution as $c \to 0$. We now give a rigorous proof of this assertion.

Theorem 5. *If $c \to 0$ and $a \to 1$ so that $1 - a \approx \lambda c$, where λ is a positive constant, then the limit of an ALM-distribution is an exponential distribution with parameter λ.*

Proof. By Theorem 4, the Laplace–Stieltjes transform of an ALM-distribution is of the form

$$\varphi_X(s) = \varphi_Y(s) \frac{1-a}{1 - a\,exp\{-cs\}}. \tag{16}$$

Since the carrier of the random variable Y is the interval $[0, c]$, we find that $\varphi_Y(s) \to 1$ as $c \to 0$. The second factor in expression (16) for $\varphi_X(s)$ satisfies the relation

$$\frac{1-a}{1 - aexp\{-cs\}} = \frac{1-a}{1 - a(1 - cs + (cs)^2/2 - \cdots}) = \frac{1}{1 + \frac{cs}{1-a} + o(c)} \to \frac{1}{1 + \frac{s}{\lambda}} = \frac{\lambda}{\lambda + s}.$$

Hence taking the limit as $c \to 0$ in the last expression, since $\dfrac{1-a}{c} \to \lambda$, we arrive at the assertion of the theorem. This assertion is also implied by the theorem on the continuous dependence of Laplace–Stieltjes transforms and their originals. □

According to this theorem, it is worthwhile to use exponential distributions as a competitive hypothesis in studying the class of ALM-distributions with a sequence $c_m = mc$. On the other hand, the parameter c is often known from the nature of the investigated phenomenon. Therefore there is justification for comparing different ALM-distributions for the same value of the parameter c and different distribution densities on a cycle. Preliminary results of such an investigation are reported in [13–15]. Below we state their generalizations and give examples.

Let $x_1, \ldots, x_n$ be a user sample of a general population with unknown distribution density $f(x)$. Consider the problem of testing the zero hypothesis H_0: $f(x) = f_0(x)$ against the alternative hypothesis H_1: $f(x) = f_1(x)$.

According to the Neumann-Pearson theorem, the most powerful criterion for testing a simple hypothesis H_0 against an alternative H_1 is the likelihood ratio test. Since observations are independent, the critical domain for this test can expressed as

$$\mathbf{W} = \left\{ (x_1, \ldots, x_n) : \ \frac{f_1(x_1, \ldots, x_n)}{f_0(x_1, \ldots, x_n)} = \prod_{1 \le i \le n} \frac{f_1(x_i)}{f_0(x_i)} \ > t \right\}, \quad t > 0. \tag{17}$$

The significance level α and power π of this test are

$$\alpha = \mathbf{P}_{H_0}\{\mathbf{W}\} = \mathbf{P}_{H_0}\{(X_1, \ldots, X_n) \in \mathbf{W}\},$$
$$\pi = \mathbf{P}_{H_1}\{\mathbf{W}\} = \mathbf{P}_{H_1}\{(X_1, \ldots, X_n) \in \mathbf{W}\}.$$

In computations, it is more convenient to use the natural logarithm of the product in (17). For this purpose, let us introduce the natural algorithm of the likelihood ratio, which, for the sake of brevity, is called the *statistic of the test*

$$W = \ln \prod_{1 \le i \le n} \frac{f_1(x_i)}{f_0(x_i)} = \sum_{1 \le i \le n} \left(\ln f_1(x_i) - \ln f_0(x_i) \right). \tag{18}$$

Applying the Monte Carlo method, we compute the critical value t_α of the significance level α and power π of the test. A suitable algorithm is described in the next section.

For a large sample, the distribution of the statistic W is asymptotically normal. Therefore, to compute the significance level and power of the test, we can use a normal approximation and restrict to estimation of the first two moments of the corresponding distributions. To illustrate the performance of this approach, let us denote the random variables

$$U = \ln f_1(X) - \ln f_0(X), \quad V = \ln f_1(Y) - \ln f_0(Y), \tag{19}$$

by U and V, where X and Y are random variables with distribution density $f_0(.)$ and $f_1(.)$ for the hypotheses H_0 and H_1, respectively. Let μ_U, μ_V, and σ_U^2, σ_V^2 denote their mean and variance, respectively. For large samples, the statistic W (18) under zero and alternative hypotheses has a normal distribution with parameters $n\mu_U$, $n\sigma_U^2$, and $n\mu_V$, $n\sigma_V^2$ respectively. Therefore, the critical value t_α of the test statistic for a given significance level α can be found from the equation

$$\alpha = \mathbf{P}_{H_0}\{W > t_\alpha\} = \mathbf{P}_{H_0}\left\{\frac{W - n\mu_U}{\sigma_U\sqrt{n}} > t_\alpha - \frac{n\mu_U}{\sigma_U\sqrt{n}}\right\} = 1 - \Phi\left(t_\alpha - \frac{n\mu_U}{\sigma_U\sqrt{n}}\right)$$

or

$$t_\alpha - \frac{n\mu_U}{\sigma_U\sqrt{n}} = z_{1-\alpha},$$

where $z_{1-\alpha}$ is the $(1-\alpha)$th quantile of the standard normal distribution. Hence the critical value t_α for a given critical level α is

$$t_\alpha = n\mu_U + z_{1-\alpha}\sigma_U\sqrt{n}. \tag{20}$$

Accordingly, the power of the test is

$$\begin{aligned}\pi = \mathbf{P}_{H_1}\{W > t_\alpha\} &= \mathbf{P}_{H_1}\left\{\frac{W - n\mu_V}{\sigma_V\sqrt{n}} > \frac{t_\alpha - n\mu_V}{\sigma_V\sqrt{n}}\right\} \\ = 1 - \Phi\left(\frac{t_\alpha - n\mu_V}{\sigma_V\sqrt{n}}\right) &= 1 - \Phi\left(\frac{\mu_U - \mu_V}{\sigma_V}\sqrt{n} + z_{1-\alpha}\frac{\sigma_U}{\sigma_V}\right).\end{aligned} \tag{21}$$

This relation shows that the power of the test largely depends on the difference between the expectations of the random variables U and V.

The parameters μ_U, μ_V, and σ_U^2, σ_V^2 for certain particular cases can be computed in closed form. In the general case, they can also be estimated by the Monte Carlo method and estimates, instead of their exact values, can be used. Suitable algorithms for computing the tails of (additional) empirical distribution functions under the zero H_0 and alternative H_1 hypotheses for both cases are described below.

6. ALGORITHMS

In this section, we describe algorithms for computing the distributions of the test statistic W under the zero and alternative hypotheses. The first algorithm can be used for a sample of any size. The second algorithm is applicable only to sufficiently large samples (for example, $n > 30$). Both algorithms use the Monte Carlo method.

Algorithm 1.

Beginning. Choose the distribution densities $f_0(.)$ and $f_1(.)$ for the zero and alternative hypothesis, and sample size n.

Step 1. Generate N independent random samples $(x_1^{(j)}, \ldots, x_n^{(j)})$, $j = 1, 2, \ldots, N$, of size n with independent elements from a general population with distribution density $f_0(.)$ and compute the values of the test statistic

$$w_0^{(j)} = w\left(x_1^{(j)}, \ldots, x_n^{(j)}\right) = \sum_{1 \le i \le n} \left(\ln f_1(x_i^{(j)}) - \ln f_0(x_i^{(j)})\right). \tag{22}$$

Step 2. Compute the tail of the empirical distribution function

$$\bar{F}_{0,N}(t) = \frac{1}{N}\left\{\text{number of quantities } \; w_0^{(j)} > t\right\}, \quad t > 0. \tag{23}$$

Step 3. Generate N independent random samples $(y_1^{(j)}, \ldots, y_n^{(j)})$, $j = 1, 2, \ldots, N$, of size n with independent elements from a general population with distribution density $f_1(.)$ and compute the values of the test statistic $w_1^{(j)}$ by formula (22), using $y_i^{(j)}$ for $x_i^{(j)}$.

Step 4. Compute the tail of the empirical distribution function $\bar{F}_{1,N}(t)$ by (23), replacing $w_0^{(j)}$ by $w_1^{(j)}$.

Step 5. For a given user sample $(x_1, \ldots, x_n)$, compute the value w of the test statistic (18)

$$w = w(x_1, \ldots, x_n) = \sum_{1 \le i \le n} (\ln f_1(x_i) - \ln f_0(x_i)).$$

Compute the significance level of the user sample (p-value) $\alpha = \alpha(w)$ for rejecting the zero hypothesis H_0 in favor of the alternative H_1 with function $\bar{F}_{0,N}(x)$ in the form

$$\alpha(w) = \bar{F}_{0,N}(w).$$

Compute the critical value t_α and power π of the test with functions $\bar{F}_{0,N}(x)$ and $\bar{F}_{1,N}(x)$ in the form

$$t_\alpha = \bar{F}_{0,N}^{(-1)}(\alpha), \qquad \pi = \bar{F}_{1,N}(t_\alpha).$$

Step 6. Print the results:

• the critical value t_α and significance level $\alpha(w) = \bar{F}_{0,N}(w)$ of the user sample for rejecting the zero hypothesis H_0 in favor of the alternative hypothesis H_1,

• the test power $\pi = \bar{F}_{1,N}(t_\alpha)$,

• the curve of the significance level function $\alpha(t) = \bar{F}_{0,N}(t)$, and

• the curve of the test power function $\pi(t) = \bar{F}_{1,N}(t)$.

End.

Decision on the rejection of the zero hypothesis in favor of the alternative hypothesis is taken by comparing the p-value of the user statistic of significance level α.

This algorithm generates desired results due to the strong law of large numbers, which asserts that the empirical distribution function converges to the theoretical distribution with probability 1 with the growth of the observation number N, which in this algorithm can chosen as large as we please.

For user samples of a sufficiently large size, we can apply the central limiting theorem to simplify the algorithm perceptibly. According to this theorem, the sum of identically distributed independent random variables tends to a normal distribution with the growth of the number of terms. Therefore, to compute the distributions of the test statistic under zero and alternative hypotheses for large user samples, we can use a normal approximation and restrict to estimation of its first two moments. A suitable algorithm is described below.

Algorithm 2.

Beginning. Choose the distribution densities $f_0(.)$ and $f_1(.)$ for the zero and alternative hypotheses and sample size n.

Step 1. Generate N independent random variables $x_1, \ldots, x_N$ from a general population with distribution density $f_0(.)$ and compute the value u_j of the statistic U (19)

$$u_j = u(x_j) = \ln f_1(x_j) - \ln f_0(x_j) \tag{24}$$

and the corresponding sample mean $\bar{u}$ and variance S_u^2

$$\bar{u} = \frac{1}{N} \sum_{1 \le j \le N} u_j, \quad S_u^2 = \frac{1}{N} \sum_{1 \le j \le N} u_j^2 - (\bar{u})^2 = \bar{u^2} - (\bar{u})^2. \tag{25}$$

Step 2. Generate N independent random variables $y_1, \ldots, y_N$ from a general population with distribution density $f_1(.)$, compute the value v_j of the statistic V (19) by formula (24), replacing x_j by y_j and compute the corresponding sample mean $\bar{v}$ and variance S_v^2 by formulas similar to (26).

Step 3. For a given significance level α, compute by (19) the critical value t_α for rejecting the zero hypothesis H_0 in favor of the alternative hypothesis H_1

$$t_\alpha = n\bar{u} + z_{1-\alpha} S_u \sqrt{n},$$

where $z_{1-\alpha}$ is the $(1-\alpha)$th quantile of the standard normal distribution.

Step 4. Compute the power π of the test by (20)

$$\pi = 1 - \Phi\left(\frac{\bar{u} - \bar{v}}{S_v}\sqrt{n} + z_{1-\alpha}\frac{S_u}{S_v}\right).$$

Step 5. For a given user sample $(x_1, \ldots, x_n)$, compute the value of the test statistic

$$w = w(x_1, \ldots, x_n) = \sum_{1 \le i \le n} (\ln f_1(x_i) - \ln f_0(x_i)).$$

Compute the significance level of the user sample (p-value) $\alpha = \alpha(w)$ for rejecting the zero hypothesis H_0 in favor of the alternative hypothesis H_1:

$$\alpha(w) = \mathbf{P}_{H_0}\{W > w\} = 1 - \Phi\left(\frac{w - n\bar{u}}{S_u^2\sqrt{n}}\right).$$

Step 6. Print the results:

• the critical value t_α and significance level of the user sample $\alpha = \alpha(w)$ for rejecting the zero hypothesis H_0 in favor of the alternative hypothesis H_1,

• the power π of the test,

• the function of the significance level $\alpha(t) = 1 - \Phi\left(\frac{t-\bar{u}}{S_u\sqrt{n}}\right)$, and

• the test power function $\pi(t) = 1 - \Phi\left(\frac{t-n\bar{v}}{S_v\sqrt{n}}\right)$.

End.

For supporting the decision of the problem on hypothesis testing for ALM-distributions, we developed an applied program packet ALM-soft in C++ using the WINDOWS tools. It has a user graphic interface for input and output of results and routines for modeling different distributions with or without the ALM property (uniform, exponential, normal, Gamma, Weilbull, logarithmic normal, and other distributions), computing the corresponding statistics, and plotting their curves. The packet can applied for a single computation or parametric computation (in a cycle). Results can be displayed as tables or curves, depending on the variations of parameters (in particular, a, α, etc.), and their ranges.

This packet was used in the examples given below.

7. EXAMPLES

We give three model examples to illustrate the application of our algorithms and their software.

Example 1. Testing the hypothesis on the membership of a sample to an exponential distribution on a line against the hypothesis of its membership to an ALM-distribution with an exponential distribution on a cycle.

It is readily seen that if the parameters of the density of the ALM-exponential distribution $f_0(x)$ are chosen such that

$$c > 0, \quad a = e^{-\lambda c}, \quad f_Y(x) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda c}}, \tag{26}$$

then the random variable X with ALM-distribution (12) has an exponential distribution on the whole line with distribution density $\lambda e^{-\lambda x}$. Any other choice for the parameter $a \neq e^{-\lambda c}$ would yield an ALM-distribution $f_1(x)$ different from the exponential distribution on the whole line.

For this example, let us test the hypothesis H_0 on the membership of a sample to an exponential distribution on the whole line

$$H_0: \ f(x) = f_0(x) = \lambda e^{-\lambda x}$$

against the alternative hypothesis H_1 of its membership to an ALM-distribution with an exponential distribution on a cycle

$$H_1: \ f(x) = f_1(x) = (1-a)a^{[\frac{x}{c}]}\lambda e^{-\lambda(x - [\frac{x}{c}]c)}.$$

The parameters were chosen as follows: $\lambda = c = 1$ and $a_0 = e^{-1}$ for the distribution density $f_0(x)$ and the same parameters $\lambda = c = 1$ with a variable parameter a_1 for the alternative hypothesis with $f_1(x)$. Algorithms 1 and 2 were used in experiments. The results are shown in Figs. 1a–1c. Figure 1a shows the tails of the empirical distribution functions $\bar{F}_0(t)$ (bottom curve) and $\bar{F}_1(t)$ (top curve) for $a_0 = e^{-1}$ and $a_1 = 0.5$ and small samples. Figures 1b and 1c show the test power function as a function of the parameters $a = a_1$ and $\pi = \pi(a)$ for small and large samples computed by algorithms 1 and 2, respectively.

Example 2. Comparison of two ALM-distributions with uniform distribution on a cycle.

In this example, we tested the zero hypothesis on the membership of a sample to an ALM-distribution with uniform distribution on a cycle $H_0: \ f(x) = f_0(x) = (1-a_0)a_0^{[\frac{x}{c}]}\,\frac{x}{c}$ against the alternative $H_1: \ f(x) = f_1(x)$ with the same distribution, but with a variable parameter $a = a_1$.

The distribution parameters for this example were chosen as follows: $c = 1$ and $a_0 = 0.5$. The results are shown in Figs. 2a–2c, which are similar to Figs. 1a–1c. Figure 2a shows the tails of empirical distribution functions $\bar{F}_0(t)$ (bottom curve) and $\bar{F}_1(t)$ (top curve) for $a_0 = 0.5$ and $a_1 = 0.3$ and small samples. Figures 2b and 2c show the test power function as a function of $a = a_1$ and $\pi = \pi(a)$ for small and large samples computed by algorithms 1 and 2.

Example 3. Test the hypothesis on the membership of an ALM-distribution sample with uniform distribution on a cycle against the hypothesis of its membership to an ALM-distribution with arcsin-distribution on a cycle.

For this example, we tested the zero hypothesis H_0 on the membership of an ALM-distribution sample with uniform distribution on a cycle $H_0: \ f(x) = f_0(x) = (1-a_0)a_0^{[\frac{x}{c}]}\,\frac{x}{c}$ against the

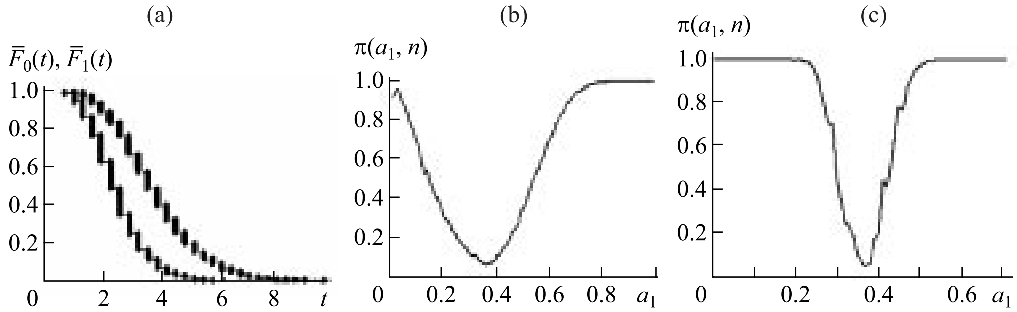

Fig. 1. Tails of empirical distribution functions and test power function of Example 1.

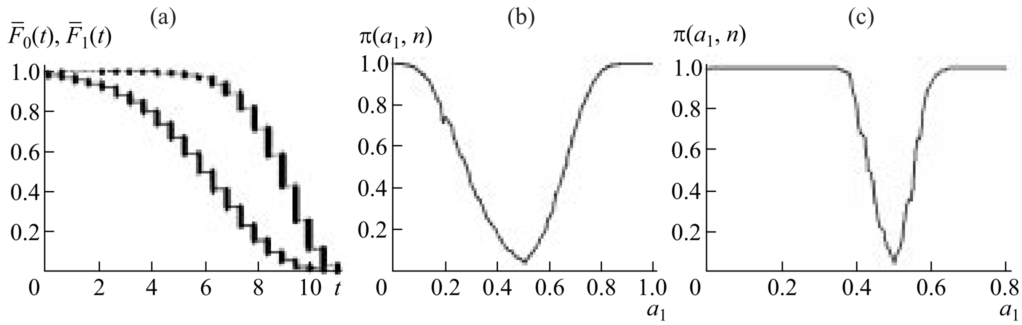

Fig. 2. Tails of empirical distribution functions and test power function of Example 2.

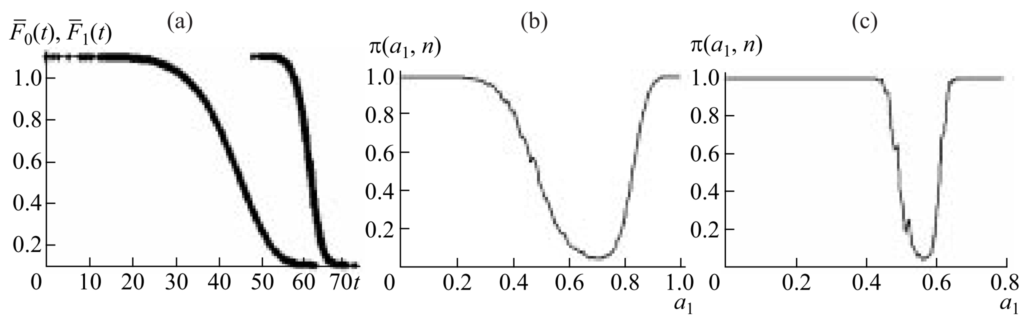

Fig. 3. Tails of empirical distribution functions and test power function of Example 3.

alternative hypothesis H_1 on its membership to an ALM-distribution with arcsin distribution on a cycle with variable parameter $a = a_1$. The density of this distribution is $f_1(x) = \frac{1}{\pi\sqrt{x(1-x)}}$, $0 \le x \le 1$. The distribution parameters for this example were chosen as before $c = 1$ and $a_0 = 0.7$. The results are shown in Figs. 3a–3c, which are similar to the previous figures. Figure 3a shows the tails of the empirical distribution functions $\bar{F}_0(t)$ and $\bar{F}_1(t)$ for $a_0 = 0.7$ and $a_1 = 0.3$ and small samples. Figures 3b and 3c show the test power function as a function of $a = a_1$ and $\pi = \pi(a)$ for small and large samples generated by algorithms 1 and 2.

8. CONCLUSIONS

Time-varying periodic flows of events occur in numerous applications, particularly in data transfer networks, communication systems, reliability models, ecological data descriptions, etc. A new characterization for a periodic time-varying Poisson flow of events is elaborated in terms of distributions with almost lack of memory property. Statistical parameter estimation and testing of hypothesis for such distributions are studied. Algorithms and their software realizations are described. Examples are given to illustrate the application of the developed program software.

REFERENCES

1. Kening, D., Rykov, V., and Shmidt, F., Stationary Queue Systems with Dependencies, in *Itogi nauki i tekhniki. Teoriya veroyatn. Mat. statist. Teor. Kibern.*, (Advances in Science and Technology. Probability Theory, Mathematical Statistics, Theoretical Cybernetics), Moscow: VINITI, 1981, vol. 18, pp. 95–196.
2. Snyder, D.L. and Miller, V.I., *Random Point Processes in Time and Space*, New York: Springer, 1995.
3. Khinchin, A.Ya., *Raboty po matematicheskoi teorii massovogo obsluzhivaniya* (Collected Papers on Queueing Theory), Moscow: Fizmatgiz, 1963.
4. Feller, W., *An Introduction to Probability Theory and Its Applications*, New York: Wiley, 1971, vol. 2. Translated under the title *Vvedenie v teoriyu veroyatnostei i ee prilozheniya*, Moscow: Mir, 1984.
5. Karlin, S., *A First Course in Stochastic Processes,* New York: Academic, 1975. Translated under the title *Osnovy teorii sluchainykh protsessov*, Moscow: Mir, 1971.
6. Chukova, S., Dimitrov, B., and Khalil, Z., A Characterization of Probability Distributions Similar to the Exponential, *Can. J. Statist.*, 1992, vol. 21, no. 3, pp. 269–276.
7. Dimitrov, B. and Khalil, Z., A Class of New Probability Distributions for Modeling Environmental Evolution with Periodic Behavior, *Environmetrics*, 1992, vol. 3, no. 4, pp. 447–464.
8. Chukova, S. and Dimitrov, B., On Distributions having the "Almost Lack of Memory Property," *J. Appl. Probability*, 1993, vol. 29, no. 3, pp. 691–698.
9. Dimitrov, B., Khalil, Z., and Le Normand, L., A Statistical Study of Periodic Environmental Processes, *Proc. Int. Congress Modelling and Simulation*, Perth, 1993, vol. 1, pp. 169–174.
10. Dimitrov, B., Chukova, S., and Green, D., Jr., Distributions in Periodic Random Environment and Their Applications, *SIAM J. Appl. Math.,* 1997, vol. 57, no. 2, pp. 501–517.
11. Dimitrov, B., Chukova, S., and Garrido, J., Compound Counting Processes in a Periodic Random Environment, *J. Statist. Research*, 2000, vol. 34, no. 2, pp. 99–111.
12. Dimitrov, B., Krougly, Z., Ghitany, M., and Rykov, V., On Properties and Statistical Estimation of ALM distributions, *Proc. Hawaii Int. Conf. Statistics and Related Fields*, Honolulu, 2003 (on a CD).
13. Dimitrov, B., Khalil, Z., Ghitany, M., and Rykov, V., *Likelihood Ratio Test for Almost Lack of Memory Distributions*, Technical Report No. 1/01, Concordia University, 2001.
14. Dimitrov, B., Green, D., Jr., Rykov, V., and Stanchev, P., On Statistical Hypotheses Testing via Simulation Method, *Proc. 28th Int. Conf. Information and Communication Technologies and Programming*, Varna, 2003, pp. 55–72.
15. Dimitrov, B., Krougly, Z., Green, D., Jr., and Rykov, V., Periodic Nonstationary Arrival Process in Queueing Networks and Their Characterization, *Proc. Int. Workshop. Distributed Computer and Communication Networks. Stochastic Modelling and Optimization*, Moscow: TECHNOSPHERA, 2003, pp. 64–72.

This paper was recommended for publication by V.M. Vishnevskii, a member of the Editorial Board