



Algorithms for Point Processes Analysis

Bernard Picinbono

► To cite this version:

Bernard Picinbono. Algorithms for Point Processes Analysis. Communications in Statistics - Simulation and Computation, 2013, 42 (1), pp.90-106. 10.1080/03610918.2011.633197 . hal-00793385

HAL Id: hal-00793385

<https://centralesupelec.hal.science/hal-00793385>

Submitted on 22 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

In order to obtain a reprint of the published version of this paper, ask to the
author:

bernard.picinbono@lss.supelec.fr

Algorithms for point processes analysis

B. Picinbono

Laboratoire des Signaux et Systèmes ¹,
Supélec, Plateau de Moulon, 91190 Gif-sur-Yvette, France

Tel: (33)-(1)-69.85.17.40, Email : bernard.picinbono@lss.supelec.fr
July 2011

Abstract

A time point process can be defined either by the statistical properties of the time intervals between successive points or by those of the number of points in arbitrary time intervals. There are mathematical expressions to link up these two points of view, but they are in many cases too complicated to be used in practice. In this paper we present an algorithmic procedure to obtain the number of points of a stationary point process recorded in some time intervals by processing the values of the distances between successive points. We present some results concerning the statistical analysis of these numbers of points and when analytical calculations are possible the experimental results obtained with our algorithms are in excellent agreement with those predicted by the theory. Some properties of point processes in which theoretical calculations are almost impossible are also presented.

Some key words. Point processes, renewal processes, lifetime, counting.

¹Unité mixte de recherche (UMR 8506) du Centre national de la recherche scientifique (CNRS), de l'École supérieure d'électricité (Suplec) et de l'Université de Paris-Sud 11 (UPS).

1 Introduction

Point processes (PP) play an important role in probability theory (Cox and Isham, 1980), physics (Lowen and Teich, 2005), or signal processing (Snyder and Miller, 1991). In the one-dimensional case, and especially for time PPs, there are two main methods to describe their properties. The first is by using the statistical properties of the numbers of points of the process appearing in non-overlapping time intervals. These numbers are nonnegative discrete valued random variables (RV). The second is by using the distances between successive points of the process, sometimes called lifetimes. These quantities are in general continuous positive RVs. This approach is especially well-suited to the case of renewal PPs defined by the fact that these lifetimes are independent and identically distributed (IID) RVs, in such a way that a stationary renewal PP is completely defined by the probability distribution common to all these lifetimes. The relations between these two approaches are the subject of number of papers (McFadden, 1962; Picinbono, 2004). Even if there exist a number of theoretical results on this subject, they are often difficult to be used in practice because of the complexity of the analytical expressions. For example, even in the simplest case of renewal PPs, it is often almost impossible to obtain explicitly the analytical expression of the probabilities of the numbers of points recorded in some intervals in terms of the distribution function (DF) of the lifetime which completely defines the process.

One of the objectives of this paper is to overcome this difficulty by introducing algorithmic procedures allowing the transformation of some properties of the lifetimes into properties of counting in some intervals. These algorithms can be used either with simulated data generated by computer or with real data coming from physical experiments. In the cases of PPs for which the theoretical calculations yield relatively simple analytical results, we show that the experimental results obtained with our method on computer simulated data are in excellent agreement with the theory, which is a test of the validity of the algorithms presented in the paper.

It is clear that the problem analyzed hereafter is not new in its principle. The two possible methods of definitions for PPs appear in almost all the books on this subject. However the calculations that yield the counting properties in terms of those of intervals between points are rarely indicated, except in the

especially important case of renewal processes (Cox and Isham, 1980). The use of computer methods to study this kind of problem appear in a sequence of papers of Lewis and the most important references are indicated in (Cox and Lewis, 1966). However the specific problem presented in this paper was not analyzed previously and the recent advances in numerical simulation allow us to introduce procedures that are rapid and precise. Similar procedures have been already used for the analysis of dead-time problems in PPs (Picinbono, 2007; Picinbono, 2009).

The paper is organized as follows. Section 2 is mainly devoted to counting analysis. We first present an overview of the known theoretical results concerning the calculations of the counting probabilities of a PP in terms of the statistical properties of the intervals between points. We note that, even for the simplest PPs such as renewal process the calculations are often rather complicated. After a short review of the possible experimental methods for the determination of the counting statistics, we present the principles of an algorithmic method allowing this determination from a sequence of measurements of successive intervals between points of the PP. In Section 3 we consider the same kind of problems appearing when using residual lifetimes. The theoretical calculations are in general almost impossible, which justifies the simulation approach presented in this section. In the last section (4) various experimental results obtained with the algorithms previously described are presented. In the cases where theoretical results are possible there is a very good agreement between theory and simulation experiment.

2 Counting Analysis

2.1 Theoretical Considerations

Let X_k be a sequence of N positive random variables and T_i the sequence deduced from the X_k s by $T_1 = X_1$ and $T_k = T_{k-1} + X_k$, $1 < k \leq N$. These T_k constitute an increasing sequence of values which are the random time instants of a time PP defined by the X_k s. The RVs X_k are thus the distances between successive points called in what follows lifetimes of order one. The lifetime of order n is the interval between T_k and T_{k+n} which is clearly equal to $S_n[k] = X_{k+1} + X_{k+2} + \dots + X_{k+n}$ often written simply S_n when no confusion

is possible. Note that the recursion defining the instants T_i yields $T_n = S_n[0]$.

In *counting analysis* one is interested in the probability distribution of the number of points appearing in some intervals. There are two main procedures for counting analysis. In the first we consider time intervals of duration D beginning at a point of the PP and we study the statistical properties of the number of points of the PP appearing in such intervals. The condition on the beginning of the interval leads to the term of *triggered counting*. If no such condition is introduced, or if the beginning of the counting interval has no reason to be a point of the PP, we use the term of relaxed counting.

Consider first the problem of the calculation of the triggered counting probability. Let N be the number of points of the PP in an interval such $[t_i, t_i + D]$, where t_i is a point of the PP and D is the counting duration. This number is a discrete valued RV and we want to calculate its distribution function from some properties of the RVs X_i .

Let \mathcal{E}_n be the event that there are n points of the PP in the interval $[t_i, t_i + D]$. This event can be written as $\mathcal{E}_n = (S_n \leq D) \cap (S_{n+1} > D)$, where S_n and S_{n+1} are written for $S_n[i]$ and $S_{n+1}[i]$ respectively. Its probability $q_n = \Pr[\mathcal{E}_n]$ can be expressed in terms of the DFs $F_n(x)$ of S_n . Indeed it results from the relation

$$\Pr[(S_n \leq D)] = \Pr[(S_n \leq D) \cap (S_{n+1} \leq D)] + \Pr[(S_n \leq D) \cap (S_{n+1} > D)] \quad (1)$$

and from the fact that $\Pr[(S_n \leq D) \cap (S_{n+1} \leq D)] = \Pr[(S_{n+1} \leq D)]$ because $S_n < S_{n+1}$ that

$$q_n(D) = F_n(D) - F_{n+1}(D), \quad (2)$$

with also $S_0 = 0$, which yields $q_0 = 1 - F_1(D)$, where $F_1(x)$ is the DF of the lifetime of order one X_k . This expression is given for renewal processes on p. 52 of (Cox and Isham, 1980).

Let us now analyze a slightly more complicated case useful for what follows. Consider two non-overlapping intervals $[0, a]$ and $[a, b]$, with $b = a + D$ and let $p_{m,n}(a, b)$ be the probability that these two intervals contain m and n points respectively. A calculation similar to the previous yields

$$p_{m,n}(a, b) = F_{m,m+n}(a, b) - F_{m+1,m+n}(a, b) - F_{m,m+n+1}(a, b) + F_{m+1,m+n+1}(a, b), \quad (3)$$

where the functions $F(.,.)$ are the DFs defined by

$$F_{m,n}(x, y) = \Pr[(S_m \leq x) \cap (S_n \leq y)] , \quad (4)$$

with $S_0 = 0$.

From these probabilities it is possible to study the relaxed counting statistics which, as seen above, correspond to the random number of points of the PP in an interval $[a, b[$ without specifying that a is a point of the PP. The corresponding probability is

$$p_n(a, b) = \sum_{m=0}^{\infty} p_{m,n}(a, b), \quad (5)$$

where $p_{m,n}(a, b)$ is given by (3). When the PP is stationary and under general ergodicity conditions, $p_n(a, b)$ tends to a limit when a tends to infinity in such a way that

$$p_n(D) = \lim_{a \rightarrow \infty} p_n(a, D). \quad (6)$$

This relation can also be interpreted in another way. Instead of making $a \rightarrow \infty$, it is possible to take $a \rightarrow -\infty$ while b remains finite. Physically this means that the origin of the PP is removed to the infinity in the past, which corresponds to the idea of stationarity. Equation (6) means that the effect of the initial state at the infinity in the past does not have any effect on the present.

The triggered and relaxed counting probabilities $q_n(D)$ and $p_n(D)$ given by (2) and (6) respectively are quite different. We shall verify this point in the forthcoming computer experiments. Furthermore it is clear that the mean value of the random number of points in relaxed counting is equal to λT , where λ is the density of the PP, while this expression has no reason to be valid for the triggered counting. We shall also verify this point experimentally.

The practical use of the previous expressions is not always simple. The main problem comes from the fact that the analytic expressions of the various DFs appearing in (2) or (3) are rarely known. Consider the simplest case of a renewal PP. This means that RVs X_k are IID. Thus the PP is completely defined by the DF $F(x)$ common to all these variables. Even with this assumption, the calculation of p_n or q_n can remain rather complicated. The DFs $F_n(x)$ appearing in (2) are the DFs of sums of IID random variables. The probability density functions (PDF) of such sums are a multiple convolutions of the PDFs common to all the RVs X_k and can in some cases have a rather

simple expression. But this is not always so and for example in the very simple situation where the RVs X_k are uniformly distributed in a finite interval, the analytical expression of the convolutions becomes extremely complicated for $n > 5$. The problem becomes still more difficult for the calculation of the relaxed counting probabilities even with renewal PPs. Indeed it is necessary to calculate the DF $F_{m,n}(x, y)$ defined by (4). It results from the definition of the sums S_n that $S_{m+n} = S_m + \bar{S}_{n-m}$ where \bar{S}_{n-m} is still of a sum of RVs X_k . If the PP is a renewal PP the RVs S_m and \bar{S}_{n-m} are independent, which strongly simplifies the calculation of the DF $F_{m,n}(x, y)$. The problem is then reduced to the calculation of the DF of IID random variables and the difficulty is the same as for the triggered counting probabilities.

When the PP is no longer a renewal PP it is almost impossible to use the previous equations because the calculation of the DF of sums of correlated RVs is in general very tedious. This explains why an experimental approach seems of great interest.

2.2 Experimental Methods for Counting Statistics

The first experimental method for counting analysis of a PP is by using a counting linear filter. To any time PP, or random sequence of time instants T_i , we can associate a signal $S(t)$ called *shot noise of the PP* and defined by

$$S(t) = \sum_i h(t - T_i), \quad (7)$$

where $h(t)$ is an arbitrary impulse response of a stable linear filter. If $h(t)$ is a rectangular function equal to 1 for $0 < t < D$ and 0 otherwise, the signal $S(t)$ is equal to the number of points of the PP appearing in the interval $(t - D, t)$. Note that $S(t)$ is a continuous time signal taking only integer values. The statistical analysis of $S(t)$ can in particular yield the relaxed counting probabilities p_n defined above. For obtaining the triggered counting probabilities we must take into consideration the values of $S(t)$ at the times instants $T_k + D$. Note that the values of $S(t)$ at times t and $t + \tau$ are in general correlated. For example in the case where the T_i s are points of a stationary Poisson process, the covariance function of $S(t)$ is a triangular function in the interval $[-D, +D]$ in such a way that $S(t)$ and $S(t + \tau)$ are independent when $\tau > D$. This means that if one wants to make a statistical analysis of $S(t)$

with independent samples $S(\theta_i)$ it is necessary to take $|\theta_i - \theta_j| > D$.

The second method consists in the use of electronic counters, as described in p. 101 of (Cox and Isham, 1980). In such systems a window is open at time instants θ_i and remains open till $\theta_i + D$ and the counter counts the number of points T_i appearing in this interval. During the interval of counting $[\theta_i, \theta_i + D]$ the counter cannot open a new window, which introduces thinning effect analyzed in (Cox and Isham, 1980). When the number of points used for the statistical analysis of the counts is very great, this thinning effect does not yield important effect on the precision of the results.

2.3 Algorithms for Counting Statistics

Starting from a realization x_k , $1 \leq k \leq N$, of a sequence of positive random variables X_i , considered as lifetimes of a PP, we want to deduce an estimation of the counting probabilities previously introduced. We assume that X_i is stationary and ergodic, in such a way that expectation values can be estimated by averaging with respect to the variable k corresponding to the discrete time. The terms x_k of the starting sequence can be real data obtained from a physical experiment or simulated data obtained from a computer program. The number N which plays a central role in the precision of the results must be as large as possible and in most of our experiments it is of the order of 10^7 . When analytical results are available, the comparison between theory and experiment yields an idea of the precision of the methods introduced.

Let us begin with *triggered counting* which requires relatively simple methods. The problem is to estimate the terms such as $F_n(T)$ appearing in (2), where $F_n(\cdot)$ is the DF of the random sums S_n introduced above. The solution appears in two steps: calculation of the sum and estimation of the probability that $S_n \leq T$. For the calculation of the sum it is appropriate to use a recursive method. Indeed we have $s_n[i] = x_{i+1} + x_{i+2} + \dots + x_{i+n}$ valid for $1 \leq i \leq N - n$. This can be written recursively in the form $s_n[i] = s_{n-1}[i] + x_{i+n}$. From this term we introduce the quantity $y_n[i] = u\{T - s_n[i]\}$, where $u(x)$ is the unit step function equal to 1 if $x \geq 0$ and to 0 otherwise. It is clear that $y_n[i]$ is a realization of a Bernoulli RV $Y_n[i]$ taking only the values 0 and 1 and such that $\Pr\{Y_n[i] = 1\} = \Pr\{S_n[i] \leq T\} = F_n(T)$. As for a Bernoulli RV we have $E\{Y_n[i]\} = \Pr\{Y_n[i] = 1\}$, we therefore deduce that $F_n(T)$ is simply the mean

value of the RVs $Y_n[i]$ directly deduced from the sums $s_n[i]$ by time averaging and thus from the data x_i . This operation must be realized for the values of n satisfying $1 \leq n \leq Q$, where Q corresponds to the number of probabilities q_n that we want to estimate.

Consider now the case of *relaxed counting*. The sequence of algorithms that follow is a numerical realization of the procedures used in some electronic counters discussed above. We start again from a realization $\{x_k\}$ of the sequence of RVs X_k . It could be possible to introduce a method similar to the one used just above in order now to determine the DFs appearing in (3). The problem, however, is much more complicated and furthermore the use of (5) would require the calculation of an infinite number of terms. It is then more appropriate to avoid this step by using a method directly introducing the calculation of the number of points of the PP used for obtaining the counting probability. For this purpose we shall use a method similar to those introduced in (Picinbono, 2007) and (Picinbono, 2009) for the study of dead time effects in PPs.

The aim of the algorithm is to associate with a sequence of values x_k defining the set of points t_k of the PP by the same procedure as the one introduced above for passing from X_i to T_i , another sequence n_i of nonnegative numbers which are the numbers of points of the PP appearing in intervals such that $[t_i + a, t_i + a + D[$, where the t_i s are some points of the PP. For this we introduce an intermediary signal σ_j defined by the recursion

$$\sigma_i = \sigma_{i-1} + x_i, \quad (8)$$

with $\sigma_1 = x_1$.

This signal is always increasing because the x_i are positive, and increases from x_1 to $\sum_i x_i$. It is clearly related with the number of points of the PP in some time intervals. For example, the number of points t_k appearing in the interval $[\alpha, \beta[$ is equal to the number of values of σ_i satisfying $\alpha \leq \sigma_i < \beta$.

As we are interested in intervals such as $[t_i + a, t_i + b[$, with $b = a + D$, we do not take into account the values of σ_i greater than b and it is then convenient to limit the possible values of σ_i to the interval $[0, b[$. For this purpose we replace this signal by

$$r_i = u(b - r_{i-1})(r_{i-1} + x_i), \quad (9)$$

where $u(\cdot)$ is again the unit step function and with the initial condition $r_1 = x_1$. This signal is always increasing except when $r_{i-1} > b$ in which case it comes back to the value 0 at the index (or time) i .

There are other values of r_i that are without interest in our counting problem. There are those corresponding to $r_i < a$. This leads to introduce the signal y_i defined by

$$y_i = u(r_i - a). \quad (10)$$

It is clear that y_i is a Bernoulli signal taking only the values 0 or 1. Furthermore, it contains all the information necessary to count the number of points t_i appearing in the interval $[t_j + a, t_j + b[$. It remains to extract this information.

For this it should be noted that if the signal y_i takes successively n times the values 1 before coming back to the value 0, this means that there are $n - 1$ points of the PP defined by the x_i s in the interval $[t_j + a, t_j + b[$. In order to count this number it suffices to calculate the sum of the successive values of y_i , which is realized by the recursive algorithm

$$z_{i+1} = (z_i + y_{i+1})y_{i+1}. \quad (11)$$

This signal is equal to 0 when $y_i = 0$ or increases by steps of value 1 when $y_i = 1$, and comes back to 0 with the signal y_i . The number of points recorded in the interval under consideration is thus the maximum value of the signal ζ_i which is obtained by the recursion

$$\zeta_i = z_i u(z_i - z_{i+1}). \quad (12)$$

This signal is a sequence of zeros and nonzero integers. It is clear that if $\zeta_i = n$, $n > 0$, then there are $n - 1$ points in the interval of measurements.

In order to understand more precisely the behaviour of this set of algorithms, we present in Table 1 an example of processing of a sequence of 40 successive values x_i of a random signal X_k . The parameters of the algorithms are $a = 5$ and $D = 1$, which yields $b = 6$.

TABLE I. PRINCIPLE OF THE WORKING OF THE ALGORITHMS.

Successive values of the signal x_i

i	1	2	3	4	5	6	7	8	9	10
x_i	0.2938	0.8587	0.7187	1.9908	0.8075	0.1889	0.5914	0.4089	1.7414	0.7652
i	11	12	13	14	15	16	17	18	19	20
x_i	0.6689	0.8693	1.2778	0.7848	0.9106	3.4896	0.3908	2.5792	0.9050	1.7570
i	21	22	23	24	25	26	27	28	29	30
x_i	1.0001	0.3945	0.2926	1.3201	1.0460	0.6134	0.5438	1.3064	0.2787	0.3313
i	31	32	33	34	35	36	37	38	39	40
x_i	0.3396	0.2055	1.3729	0.7775	1.8647	0.4704	0.7418	0.6884	1.7139	1.9120

Successive values of the signal r_i

i	1	2	3	4	5	6	7	8	9	10
s_i	0.2938	1.1526	1.8713	3.8621	4.6696	4.8585	5.4500	5.8589	7.6002	0
i	11	12	13	14	15	16	17	18	19	20
s_i	0.6689	1.5383	2.8161	3.6009	4.5115	8.0010	0	2.5792	3.4841	5.2412
i	21	22	23	24	25	26	27	28	29	30
s_i	6.2412	0	0.2926	1.6127	2.6587	3.2720	3.8158	5.1222	5.4009	5.7322
i	31	32	33	34	35	36	37	38	39	40
s_i	6.0717	0	1.3729	2.1504	4.0151	4.4855	5.2273	5.9157	7.6296	0

Successive values of the signal ζ_i

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
t_i	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	0
i	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
t_i	2	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	3	0

We have indicated in the same table the values of r_i and those between 5 and 6 are written in boldface. The number of such values is the number of points of the PP generated by the x_i of the first table in intervals $[t_i + a, t_i + a + T[$. We see that these numbers are 2, 0, 1, 3 and 2. It is exactly what appears from the table giving the values of the signal ζ_i by using the relation $n_i = \zeta_i - 1$.

It is important to note that in this operation we obtain 5 values of the numbers of points in intervals from a sequence of 40 values of the initial signal x_i . There is thus an important loss of information, of the order of 10 %, but when we are using, as in our experiments, a set of approximately 10^7 values of x_i , there remains a set of approximately 10^6 values of the numbers of points in intervals of measurements, which is quite sufficient to reach results with a fairly good degree of precision.

More generally the signal ζ_i is a sequence of N nonnegative values. Those equal to zero have no interest for our analysis. It is appropriate to delete all these values; this yields a sequence of N' positive values m_i . The number N' is not known in advance and depends on the sequence of values x_i analyzed. The numbers n_i of points of the PP interesting for our statistical analysis is thus $n_i = m_i - 1$.

Note that the procedure for passing from the lifetimes x_i to the number of points n_j is similar to the one used in electronic counters where the thinning effect yields a loss of points for the statistical analysis. On the other hand this introduces an advantage indicated above which is the diminution of the correlation between samples of points analyzed. For example in the case where the RVs X_i are exponentially distributed, which implies that the corresponding PP is a Poisson process, then the values N_j deduced from the X_i s are independent RVs, which is an advantage for the statistical analysis.

Indeed the last step is to evaluate some statistical properties of the random numbers n_i obtained by this procedure. The simplest result that can easily be obtained is the mean value and the variance. For some experiments it is the only interesting feature that can be extracted from the counting experiment. In some more advanced studies the interest lies in counting probabilities $p_n(a, b)$ defined by (5). For this purpose it is sufficient to use a normalized histogram of the experimental numbers n_i . Note that the algorithms presented above avoid the step involving the calculation of the series appearing in (5).

3 Residual Lifetime

3.1 Theoretical Considerations

Consider again a PP defined from an arbitrary origin t_0 by a sequence of stationary positive RVs X_i , $i \geq 1$, such that the distance between successive random points T_i is $T_i - T_{i-1} = X_i$. Let a be a positive number. Suppose that the interval $[0, a[$ contains n points of the process and let $T_{n+1}^{[a]}$ be the first point of the PP posterior to a . Let $Z_c^{(a,n)}$ be the RV defined by $Z_c^{(a,n)} = T_{n+1}^{[a]} - a$. This is called the constrained residual lifetime of the PP. The expression “constrained” comes from the fact that there are n points in $[0, a[$. On the other hand, the residual lifetime $Z^{(a)}$ is simply the distance between a and the first point of the PP posterior to a . The calculation of the DFs of $Z_c^{(a,n)}$ and of $Z^{(a)}$ is very tedious and their expressions in terms of the DFs of the lifetimes introduced above are very complicated for finite values of a .

There is, however, a simple and interesting result concerning the asymptotic situation when a tends to infinity. Indeed it is shown on pps. 8 and 54 of (Cox and Isham, 1980) that, under very general conditions, the PDF of the residual lifetime $Z^{(a)}$ takes for $a \rightarrow \infty$ the form

$$f(z) = \frac{1}{m_X} [1 - F_1(z)], \quad (13)$$

where $F_1(z)$ is the DF of the RVs X_i already used previously and m_X its mean value.

3.2 Algorithms for Lifetimes Statistics

The numerical methods for lifetimes statistics are a direct extension of those used for counting analysis. We start from a realization x_i of a sequence of

positive RVs X_i . The main idea is to calculate, as previously, a recursive sum of these data x_i , but returning to zero when this sum becomes greater than a . This is given by

$$\bar{s}_{i+1} = u(a - \bar{s}_i)(\bar{s}_i + x_{i+1}), \quad (14)$$

analog to (9) used above for counting analysis. We must now eliminate all the values smaller than a and preserve the one which is immediately greater than a . This is obtained by the signal \bar{y}_i defined by

$$\bar{y}_i = u(\bar{s}_i - a)\bar{s}_i \quad (15)$$

It is clear that \bar{y}_i is always equal to 0 except for the indices i specifying that the sum becomes greater than a . Indeed let i be the first index such that $\bar{y}_i > 0$ after a sequence of zeros. This means that $\bar{s}_i > a$. According to (14) this implies that $\bar{s}_{i+1} = 0$, which generates a new sequence of \bar{y}_i equal to 0, and so on. After elimination of the samples \bar{y}_i equal to 0 we get a sequence of nonzero \bar{y}_i and the quantity $\bar{q}_i = \bar{y}_i - a$ yields samples of residual lifetime that can be analyzed.

The simplest results that can be obtained is the mean and the variance of the residual lifetime. But it is also of great interest to study the PDF of this residual lifetime and especially to see how this PDF tends to the limit given by (13) when a increases. For this it suffices to use normalized histograms of the values \bar{q}_i .

4 Experimental Results

4.1 Erlang Point Processes

These PPs are renewal processes characterized by the fact that the RVs X_i describing the distance between successive points are IID and defined by the PDF $p(x) = 4\lambda^2 x \exp(-2\lambda x)$. The corresponding mean value is $1/\lambda$ and λ is then the density. They can be defined from Poisson processes of density $\mu = 2\lambda$ in which one point out of two has been regularly deleted. From this property it is possible to calculate all the quantities introduced above in the theoretical considerations.

According to a method presented in the Appendix, the counting probabilities q_n of triggered counting in an interval $[0, D]$ and defined by (2) are given

by

$$q_n = \exp(-\mu D)[1 + \mu D/(2m + 1)](\mu D)^{2n}/(2n)!. \quad (16)$$

The corresponding mean value is

$$M_d = \lambda D - (1/4)(1 - e^{-4\lambda D}). \quad (17)$$

This mean value is always smaller than λD because of the fact that the counting interval is open by a point of the PP. It tends however to λD when $D \rightarrow \infty$.

On the other hand the relaxed counting probabilities are given by more complicated expressions. In particular we have

$$p_{m0}(a, a + D) = e^{-\mu(a+D)} \left[\frac{(\mu a)^{2m}}{(2m)!} (1 + \mu D) + \frac{(\mu a)^{2m+1}}{(2m+1)!} \right] \quad (18)$$

and after a sum on the variable m we get

$$p_0(a, a + D) = e^{-\mu(a+D)} [(1 + \mu D) \cosh(\mu a) + \sinh(\mu a)]. \quad (19)$$

When $a \rightarrow \infty$ we obtain

$$p_0(D) = e^{-\mu D} (1 + \mu D/2) = e^{-2\lambda D} (1 + \lambda D). \quad (20)$$

Similarly, for $n > 0$, we obtain

$$p_n(D) = (1/2)e^{-\mu D} \frac{(\mu D)^{2n-1}}{(2n-1)!} \left[1 + \frac{\mu D}{n} + \frac{(\mu D)^2}{2n(2n+1)} \right]. \quad (21)$$

It is easy to verify that $\sum p_n(D) = 1$ and that the corresponding mean value is $\mu D/2 = \lambda D$. There is thus a clear difference with respect to the case of triggered counting given by (17).

The simulation of RVs X_i corresponding to an Erlang processes is easy. Indeed it is clear that these lifetimes can be considered as a sum of two IID positive RVs with an exponential PDF, which corresponds to the construction from a Poisson process. In the experiments presented below we assume that $\mu = 2$, which yields $\lambda = 1$.

In Table 2 we present experimental measurements of triggered counting in an interval of duration $D = 2$. The line “TH” indicates the theoretical values given by (16) with $\lambda = 1$ and $D = 2$. The line “EXP” presents experimental results and we observe an excellent agreement between these two lines.

TABLE 2. TRIGGERED COUNTING PROBABILITIES.

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	Sum	Mean
TH	0.0916	0.3419	0.3517	0.1637	0.0430	0.0072	0.0008	0.0001	0.0000	1	1.7501
EXP	0.0916	0.3416	0.3521	0.1637	0.0429	0.0072	0.0008	0.0001	0.0000	1.0000	1.7499

In Table 3 we present theoretical and experimental results concerning relaxed counting probabilities with $\lambda = 1$, $a = 10$, and $D = 2$. Here also the agreement between the two lines is excellent. We have verified that the results remain unchanged when $a > 5$. On the other hand for smaller values of this parameter a we are no longer in the asymptotic situation and other theoretical and more complicated expressions must be used

TABLE 3. RELAXED COUNTING PROBABILITIES.

	p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	Sum	Mean
TH	0.0549	0.2808	0.3712	0.2121	0.0662	0.0129	0.0017	0.0002	0.0000	1	2
EXP	0.0549	0.2805	0.3717	0.2122	0.0661	0.0128	0.0017	0.0001	0.0000	1.0000	1.9998

These results in a case where the theory is available can be considered as a test of the good performance of the algorithms previously introduced and analyzed.

Let us now present results concerning lifetimes measurements. They appear in Fig. 1. As indicated above the PDF of the lifetime of an Erlang Process of density λ is $p(x) = 4\lambda^2 \exp(-2\lambda x)$. Applying (13) we deduce that the PDF of the residual lifetime is $p(z) = \lambda(1 + 2\lambda z) \exp(-2\lambda z)$. In the framework of our algorithms the first is obtained with $a = 0$ and the latter for $a \rightarrow \infty$. We have verified that the measurements of PDFs yield the same result as soon as $a > 3$. In Fig. 1 we present in the continuous line the curves representing these two PDFs and the points correspond to experimental results obtained by using the algorithms of the previous section with $a = 0$ and $a = 5$ respectively. We observe that these points are located with a very good degree of precision on the theoretical curves, which is an illustration of the correct behaviour of our analysis method.

4.2 Renewal Process with Uniformly Distributed Lifetime

Consider now a renewal PP such that the lifetimes are IID RVs uniformly distributed between 0 and 2. Their mean value is 1, which is also the value of

the density λ of the PP. In Table 4 we present results concerning probability of counting measurements in an interval of duration $D = 2$. It is interesting to compare these results with those obtained for a Poisson process with the same density. The theoretical values of these probabilities are given in the first line of the table. Experimental values of triggered (q_n) and relaxed (p_n) counting probabilities are given on lines 2 and 3. The values of a and D are the same as in the experiments presented in the previous section ($a = 10$, $D = 2$). As expected, we see that these probabilities are almost the same and correspond with a very good degree of approximation to those appearing in line 1. Similar results for the renewal PP with uniformly distributed lifetime appear in the last two lines. In this case there is a clear difference between triggered and relaxed counting probabilities, which also yields a clear difference between the mean values and the variance. Furthermore the mean value of the relaxed counting is, as expected, equal to $\lambda D = 2$.

TABLE 4. TRIGGERED AND RELAXED COUNTING PROBABILITIES.

n	0	1	2	3	4	5	6	7	8			
Poisson												
TH p_n	0.1353	0.2707	0.2707	0.1804	0.0902	0.0361	0.0120	0.0034	0.0009			
EXP q_n	0.1351	0.2709	0.2710	0.1801	0.0903	0.0361	0.0119	0.0035	0.0008			
EXP p_n	0.1353	0.2712	0.2705	0.1798	0.0904	0.0363	0.0119	0.0035	0.0009			
Uniform										Sum	Mean	Variance
q_n	0.0000	0.4980	0.3347	0.1255	0.0334	0.0069	0.0012	0.0002	0.0000	1.0000	1.7209	0.7655
p_n	0	0.3330	0.4174	0.1828	0.0525	0.0118	0.0021	0.0003	0.0000	1.0000	2.0003	0.8742

Results of residual lifetime measurements appear in Fig. 2. They are presented as in Fig. 1. The lifetime PDF defining the renewal process is equal to $1/2$ in the interval $[0, 2]$ and to 0 otherwise. It is obtained for $a = 0$ and this clearly appears in the figure where the points (\cdot or $+$) represents experimental results. In the asymptotic situation, for $a \rightarrow \infty$, Eq. (13) shows that the PDF $f(z)$ of the residual lifetime is a triangular function in $[0, 2]$, or equal to $1 - z/2$ for $0 \leq z \leq 2$ and 0 otherwise. We observe that the experimental points obtained for $a = 5$ correspond perfectly to this equation. Finally we observe for $a = 0.5$ and $a = 1$ how the PDF of the residual lifetime tends to its asymptotic value.

4.3 Point Processes with Correlated Lifetimes

For some applications the assumption of independence between the successive lifetimes X_k , which is the basis of the concept of renewal processes, is

too restrictive and it is interesting to present statistical models in which this assumption is not introduced. Among the various possible solutions, one of the simplest is the exponential autoregressive model of order one [EAR(1)] introduced by (Jacobs and Lewis, 1977) and discussed in details in (Cox and Isham, 1980). It is a natural extension of Poisson processes in which the lifetimes X_k have an exponential PDF, as for Poisson processes, but are no longer independent.

In this model the RVs X_k satisfy the recursion

$$X_k = \alpha X_{k-1} + B_k Z_k, \quad (22)$$

where B_k and Z_k are two independent sequences of IID RVs. Furthermore B_k is a Bernoulli RV taking only the values 0 or 1 with probabilities α and $1 - \alpha$ respectively, while Z_k is a positive RV with an exponential PDF. It is then easy to verify that X_k and B_k have the same exponential PDF. If $\alpha = 0$, $X_k = Z_k$, and the corresponding PP becomes a Poisson process. This is the reason why the PP defined by the X_k of (22) is considered as an extension of Poisson processes as indicated in p. 62 of (Cox and Isham, 1980). It is not Poisson because the X_k are correlated and the corresponding covariance function is $\gamma_p = \sigma_X^2 \alpha^{|p|}$. where σ_X^2 is the variance of X_k .

The computer simulation of the sequence X_k given by (22) is especially simple. There are however no experimental results published on the PPs associated with these lifetimes and our purpose is then to use the algorithms introduced above to study some of their properties and in particular to clarify their difference with Poisson processes. It is clear that the theoretical calculations of counting probabilities or of residual lifetime are almost impossible, which justifies a simulation approach.

The first point however is to verify the correlation properties of the X_k generated by the recursion (22). For this purpose we generate by computer a sequence of approximately 10^7 values x_i that are outcomes of positive RVs X_i given by (22) and with the PDF $p_X(x) = \exp(-x)$. By using standard procedures of measurements of the covariance functions we obtain the first ten values $\hat{\gamma}_i$ of the covariance function γ_i of X_k for various values of the parameter α of (22). The results are displayed in Table 5. The values chosen for α are 0, 0.3, 0.5, and 0.7. They appear in the table for γ_1 since $\gamma_1 = \alpha$. We see in line 2 that the samples x_i for $\alpha = 0$ are uncorrelated with a very good precision. For

the other values of α the experimental values $\hat{\gamma}_i$ are in quite good agreement with their theoretical values γ_i .

TABLE 5. THEORETICAL AND EXPERIMENTAL VALUES OF THE COVARIANCE FUNCTION OF X_k .

i	0	1	2	3	4	5	6	7	8	9
γ_i	1	0	0	0	0	0	0	0	0	0
$\hat{\gamma}_i$	1.0007	0.0001	0.0002	0.0002	-0.0004	-0.0008	0.0003	-0.0006	0.0003	-0.0001
γ_i	1.0000	0.3000	0.0900	0.0270	0.0081	0.0024	0.0007	0.0002	0.0001	0.0000
$\hat{\gamma}_i$	0.9998	0.3004	0.0905	0.0268	0.0082	0.0031	0.0010	0.0005	0.0002	0.0002
γ_i	1.0000	0.5000	0.2500	0.1250	0.0625	0.0313	0.0156	0.0078	0.0039	0.0020
$\hat{\gamma}_i$	1.0014	0.5013	0.2512	0.1263	0.0640	0.0326	0.0169	0.0094	0.0053	0.0032
γ_i	1.0000	0.7000	0.4900	0.3430	0.2401	0.1681	0.1176	0.0824	0.0576	0.0404
$\hat{\gamma}_i$	1.0001	0.6998	0.4902	0.3435	0.2407	0.1688	0.1183	0.0825	0.0578	0.0407

Let us now consider counting experiments for which we have no theoretical results except for $\alpha = 0$ where the PP defined by (22) becomes a Poisson process.

Some results are displayed in Table 6, where we list measurements of triggered and relaxed counting probabilities q_n and p_n for various values of α . The part A corresponds to $0 \leq n \leq 9$ and the part B to $10 \leq n \leq 15$. We also present the sum of these measured values and the corresponding mean value. The values of the other parameters are $\lambda = 1$, $D = 1$.

For $\alpha = 0$ the PP becomes a Poisson process for which $q_n = p_n = \exp(-m)m^n/n!$ with $m = \lambda D$. These theoretical values appear on the first line of the table and the corresponding experimental values of q_n and p_n appear on the second and third line. The agreement between these three lines is quite good. Furthermore we verify with a high degree of precision that the mean value is equal to $\lambda D = 1$.

When α increases, *i.e.* when the correlation between successive intervals between points increases, we observe that q_0 increases while p_0 decreases, and the inverse situation appears for q_1 and p_1 . Furthermore the mean value associated with q_n varies, while the one corresponding to p_n remains approximately equal to 1, according to the theoretical considerations.

TABLE 6. TRIGGERED AND RELAXED COUNTING PROBABILITIES
FOR THE EAR(1) POINT PROCESS.

A										
n	0	1	2	3	4	5	6	7	8	9
$\alpha = 0$										
TH	0.3679	0.3679	0.1839	0.0613	0.0153	0.0031	0.0005	0.0001	0	0
q_n	0.3678	0.3679	0.1841	0.0612	0.0153	0.0031	0.0005	0.0001	0.0000	0.0000
p_n	0.3679	0.3686	0.1829	0.0621	0.0147	0.0032	0.0005	0.0001	0.0000	0.0000
$\alpha = 0.3$										
q_n	0.3847	0.3209	0.1574	0.0746	0.0344	0.0158	0.0070	0.0030	0.0013	0.0005
p_n	0.3686	0.4124	0.1322	0.0503	0.0210	0.0091	0.0037	0.0015	0.0007	0.0003
$\alpha = 0.5$										
q_n	0.4184	0.3031	0.1332	0.0657	0.0348	0.0193	0.0109	0.0062	0.0036	0.0020
p_n	0.3684	0.4369	0.1123	0.0408	0.0193	0.0101	0.0052	0.0030	0.0017	0.0010
$\alpha = 0.7$										
q_n	0.4657	0.2994	0.1061	0.0435	0.0274	0.0181	0.0127	0.0082	0.0056	0.0038
p_n	0.3655	0.4565	0.1015	0.0355	0.0164	0.0085	0.0051	0.0037	0.0023	0.0015
B										
n	10	11	12	13	14	15	Sum	Mean		
$\alpha = 0$										
TH	0	0	0	0	0	0	1	1		
q_n	0	0	0	0	0	0	1	1		
p_n	0	0	0	0	0	0	1	0.9991		
$\alpha = 0.3$										
q_n	0.0002	0.0001	0	0	0	0	1	1.1571		
p_n	0.0001	0.0001	0	0	0	0	1.0000	0.9999		
$\alpha = 0.5$										
q_n	0.0012	0.0007	0.0004	0.0002	0.0001	0	0.9999	1.1865		
p_n	0.0006	0.0003	0.0002	0.0001	0.0001	0	0.9999	1.0000		
$\alpha = 0.7$										
q_n	0.0027	0.0019	0.0013	0.0010	0.0007	0	0.9992	1.1419		
p_n	0.0010	0.0007	0.0004	0.0003	0.0003	0.0002	0.9996	0.9972		

It remains now to present measurements of residual lifetimes. If $\alpha = 0$, which means in the case of a stationary Poisson process of density λ , we have verified that the PDF of the residual lifetime is $\lambda \exp(-\lambda z)$ irrespective of the value of the parameter a introduced in the theoretical considerations. This corresponds to the properties of a stationary Poisson process characterized by the fact that the past and the future are independent. The situation is quite different when $\alpha \neq 0$. In this case the PDF of the residual lifetime is still exponential for $a = 0$, which is simply the lifetime X_k , and we find again the same situation when $a \rightarrow \infty$, according to (13). The situation, however, is rather different from the experiments presented in Section IV. Indeed, because of the correlation between successive values of X_k , it is necessary to take greater values of a in order to reach the asymptotic form of the PDF given by (13). For example when $\alpha = 0.6$, this asymptotic structure appears only for $a > 10$, while $a > 5$ was sufficient for renewal processes, as seen in Figs. 1 and 2.

Results of computer experiments on life time measurements are reported in

Figures 4 and 5. In Fig. 4 we study the influence of the parameter a introduced above on the PDF of the residual life time. As $p_X(x) = \exp(-x)$, it results from (13) that the PDF $f(z)$ is also $\exp(-z)$. This function is indicated in continuous line in the figures 3A, 3B, and 3C. These three figures are obtained for $\alpha = 0.7$ and $a = 0$ for 3A, $a = 1$ for 3B, and $a = 20$ for 3C. The points represent the experimental results of PDF measurements. As expected these points are very precisely located on the exponential curve for the extreme cases corresponding to figures 3A and 3C. On the other hand the experimental values of the PDF of the residual waiting time for $a = 1$, which is almost impossible to calculate theoretically, is clearly not exponential, as it appears on Fig. 3B.

This figure also indicates that the corresponding PDF $f(z)$ presents a crossing point with the curve $\exp(-z)$ obtained in the two extreme cases of Figs. 3A and 3C. This crossing effect appears for $z = 0.7$, value of the product $a\alpha$. Various other experiments confirm this fact. For a better understanding of this point we present in Fig. 4 results of measurements of PDFs of residual lifetime for $a\alpha = 1/2$ and $a = 0.7$, $a = 1$, and $a = 1.5$ for curves 1, 2, and 3 respectively. We observe that the crossing with the curve $\exp(-z)$ appears precisely for $z = 1/2$. We have up to now no way to find a proof of this experimental result.

Appendix : Calculations concerning Erlang processes

In order to calculate the counting probabilities of an Erlang process of density λ we start from a Poisson process of density $\mu = 2\lambda$ in which one point out of two has been regularly deleted (periodic deterministic thinning). Consider the interval $[0, D]$ and suppose that there is a non-erased point of the Poisson process at the time 0. The event that there are n points of the Erlang process in the interval $[0, D[$ is realized if there are either $2n$ or $2n + 1$ points of the Poisson process in this interval. The probability of this event is clearly (16) valid for any integer $n \geq 0$. After simple algebra this yields the mean value (17). The same procedure is applied to obtain Eqs. (18) to (21)

BIBLIOGRAPHY

- Cox D. and Isham V. (1980). *Point Processes*. London: Chapman and Hall.
- Cox D. and Lewis P. (1966). *The Statistical Analysis of Series of Events*. London: Wiley.
- Jacobs P. and Lewis P. (1977). “A mixed autoregressive-moving average exponential sequence and point process (EARMA 1,1) ,” *Adv. Appl. Prob.*, **9**, pp. 87-104.
- Lowen S. and Teich M. (2005). *Fractal-based Point Processes*. Hoboken (N. J.): Wiley.
- McFadden, J. (1962). On the lengths of intervals in a stationary point process. *J. Roy. Statist. Soc.*, **B 24**, 364–382.
- Picinbono B. (2004). Time intervals and counting in point processes. *IEEE Trans. Inform. Theory*, **IT 50**, 1336–1340.
- Picinbono B. (2007). Temps mort d’entrée dans les processus ponctuels. *Traitement du signal*, **24**, 411–428.
- Picinbono B. (2009). Output dead-time in point processes. *Communications in Statistics, Simulation and Computation*, **38**, 2198-2213.
- Snyder D. and Miller M. (1991). *Random Point Processes in Time and Space*. New York: Springer-Verlag.

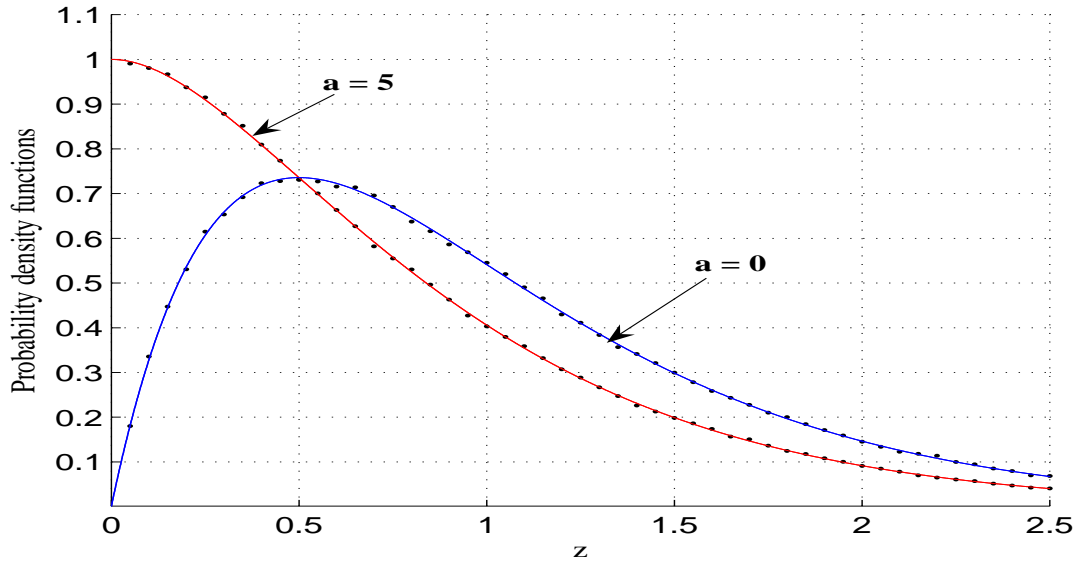


Figure 1: Probability density functions $p(z)$ of lifetime ($a = 0$) and residual lifetime ($a = 5$) for an Erlang process of density $\lambda = 1$. Points: experiment, continuous curves: theory.

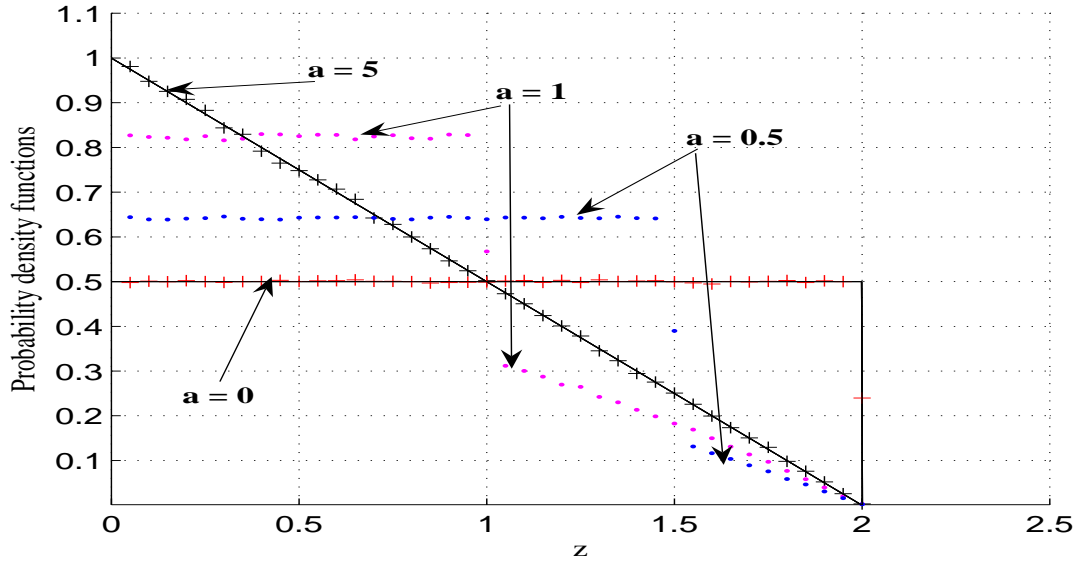


Figure 2: Probability density functions $p(z)$ of residual lifetime of a renewal process of uniform lifetime and unit density λ for various values of a . Points: experiment, continuous curves: theory.

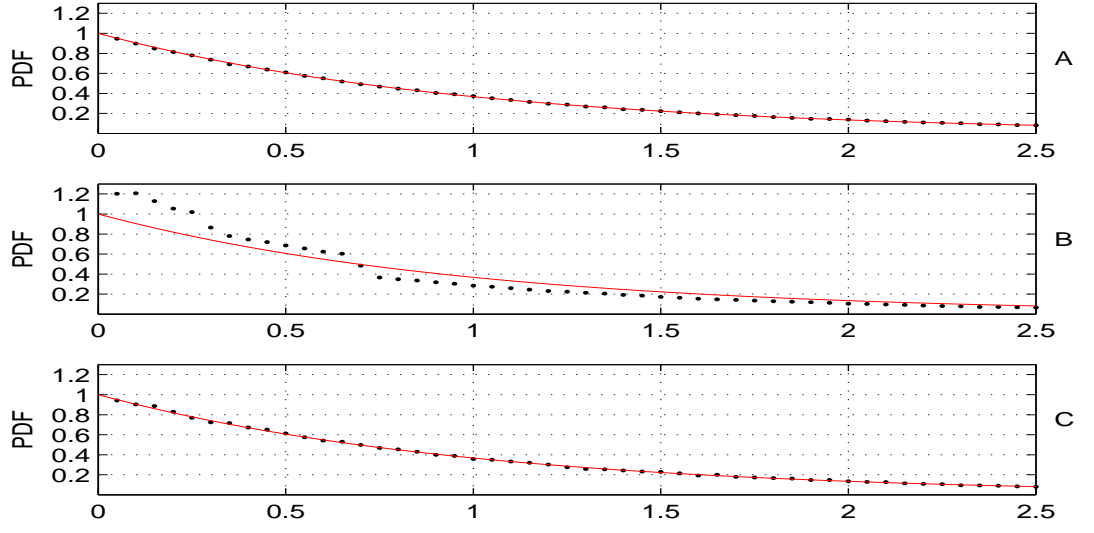


Figure 3: Probability density functions of residual lifetime for EAR(1) process and various values of a : A: $a = 0$, B: $a = 1$, and C: $a = 20$.

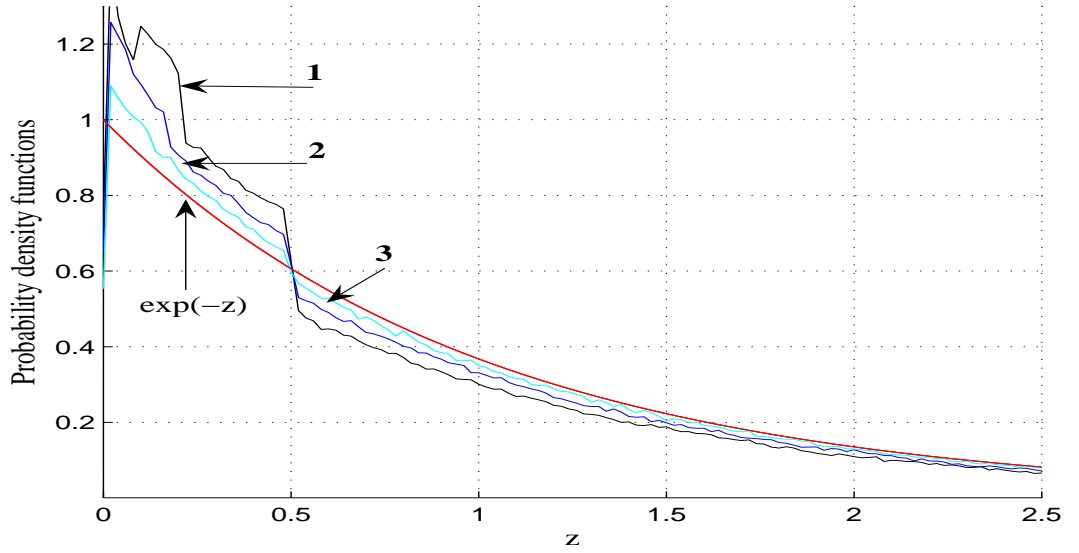


Figure 4: Probability density functions of residual lifetime for EAR(1) process for $\alpha a = 1/2$. 1: $a = 0.7$, 2: $a = 1$, 3: $a = 1.5$.