Simulating of poisson point process using conditional intensity function (Hazard function)

Behrouz Fathi-Vajargah*, Hassan Khoshkar-Foshtomi

Department of Statistics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran
*Corresponding author E-mail: fathi@guilan.ac.ir

Abstract

In this article we first study linear point process then based on introducing conditional intensity function (Hazard function), we present an algorithm for simulating Monte Carlo point process and try to test and study its behaviors on some individuals of time points.

Keywords: *Conditional Intensity Function, Monte Carlo Simulation, Point Process, Time Points.*

1. Introduction

A temporal point pattern is basically a list of times of events. Many real phenomena produce data that can be represented as a temporal point pattern; the Events, Earthquakes, Arrivals at a server and Accidents shows a few examples. Common to these examples is that we do not know how many events will occur, or at what times they will occur. Usually complex mechanisms are behind these seemingly random times, for example earthquakes cause new earthquakes in the form of aftershocks. An essential tool for dealing with these mechanisms, for example in predicting future events, is a stochastic process modelling the point patterns: a temporal point process.

In this note, familiarity with the Poisson process on the line as well as basic probability theory and statistics is assumed. On the other hand, measure theory is not assumed; for a much more thorough treatment with all the measure theoretical details, see [9], [10].

2. Evolutionary point processes

There are many ways of treating temporal point processes. In this note we will explore one approach based on the so-called conditional intensity function. To understand what this is, we first have to understand the concept of evolutionarily [36].

2.1. Evolutionarily

Usually we think of time as having an evolutionary character: what happens now may depend on what happened in the past, but not on what is going to happen in the future. This order of time is also a natural starting point for defining practically useful temporal point processes. Roughly speaking, we can define a point process by specifying a stochastic model for the time of the next event given we know all the times of previous events. The term evolutionary point process is used for processes defined in this way.

The past in a point process is captured by the concept of the history of the process. If we consider the time t, then the history H_t is the list of times of events $(...,t_1,t_2,...,t_n)$ up to but not including time t. Note that theoretically the point process may extend infinitely far back in time, but it does not have to do this.

Note also that we assume that we have a simple point process, i.e. a point process where no points coincide, such that the points can be strictly ordered in time.

2.2. Inter event times

When specifying a temporal point process we can use many different approaches. In this note, we will consider two, where we specify the distribution of
   I.   The time lengths between subsequent events, or
   II.   The number of events occurring in an arbitrary time-interval.
We will start with (I).
The lengths of the time intervals between subsequent events are known as inter event times. We can define a temporal point process by specifying the distributions of these. Let $f^*(t) = f(t|H_t)$ be the conditional density function of the time of the next event t_{n+1} given the history of previous events $(...,t_{n-1},t_n)$.

Here we use the notation $*$ from [19] to remind ourselves that this density is conditional on the past, rather than writing explicitly that a function depends on the history.

Note that the density function $f^*(t)$ specifies the distribution of all inter event times, one by one, starting in the past, and thus the distribution of all events is given by the joint density

$$f(...,t_1,t_2,...) = \prod_i f(t_i |...,t_{i-2},t_{i-1}) = \prod_i f^*(t_i) \tag{1}$$

Let us consider a simple example of a point process specified by specifying $f^*(t)$:

Example 2.1: (Renewal process and Wold process)
The simplest process we can define by specifying the distribution of the inter event times is the renewal process. This process is defined by letting the inter event times be i.i.d. stochastic variables, i.e. $f^*(t_n) = f(t_n - t_{n-1})$ where f is a density function for a distribution on $(0,\infty)$.

A special case of this is the homogeneous Poisson process with intensity λ , where f is the density of the exponential distribution with inverse mean λ . Fig 1 shows generations of three different renewal processes. A generalization of the renewal process is the Wold process where, rather than being independent, the inter event times may depend on the previous inter event time (i.e. the inter event times are a first-order Markov chain) [29].
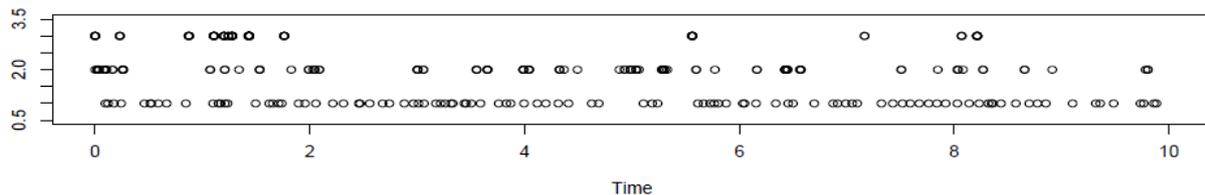
Fig. 1: Three Generations of Renewal Processes with Different Inter Event Time Distributions: Gamma (0.02, 0.2) (Upper), Gamma (0.1, 1) (Middle), Gamma (2, 20) (Lower). Note How the Upper Case is clustered and the Lower Case is regular Compared to the Middle Case (Which Is a Poisson Process). Also Note That All the Simulations Have Roughly 100 Points for Easy Comparison (They are Very Densely Packed Together for the Upper Case).

The above examples show cases where the placement of t_n depends on t_{n-1} and, in the case of the wold process, t_{n-2} .

However, in general it may depend on the whole history, and it turns out that $f^*(t)$ is not the most convenient or intuitive way of specifying the general case.

2.3. Conditional intensity function

The conditional intensity function is a convenient and intuitive way of specifying how the present depends on the past in an evolutionary point process. Consider the conditional density f^* and its corresponding cumulative distribution function F^* . Then the conditional intensity function (or Hazard function) is defined by

$$\lambda^*(t) = \frac{f^*(t)}{1 - F^*(t)} . \tag{2}$$

The conditional intensity function can be interpreted heuristically in the following way: consider an infinitesimal interval around t , say dt , then

$$\lambda^*(t)dt = \frac{f^*(t)dt}{1-F^*(t)}$$

$$= \frac{P(\text{point in } dt | H_t)}{P(\text{point not before } t | H_t)}$$

$$= \frac{P(\text{point in } dt, \text{ point not before } t | H_t)}{P(\text{point not before } t | H_t)} \qquad (3)$$

$$= P(\text{point in } dt | \text{point not before } t, H_t)$$

$$= P(\text{point in } dt | H_t)$$

$$= E(N(dt)|H_t).$$

Here $N(A)$ denotes the number of points falling in an interval, and the last equality follows from the assumption that no points coincide, so that there is either zero or one point in an infinitesimal interval. In other words, the conditional intensity function specifies the mean number of events in a region conditional on the past [33], [34], [36].

We consider an example of point process where the conditional intensity has particular functional form:

Example 2.2: (Poisson process)

The (inhomogeneous) Poisson process is among other things characterized by the number of points in disjoint sets being independent. The conditional intensity function inherits this independence. The Poisson process is quite simply the point process where the conditional intensity function is independent of the past, i.e. the conditional intensity function is equal to the intensity function of the Poisson process, $\lambda^*(t) = \lambda(t)$.

A little creativity and common sense can be used to define many new models using the conditional intensity function. This, of course, depends on the fact that the conditional intensity function uniquely defines a point process. To prove this we first need to note that the definition of the conditional intensity function can also be reversed such that an expression for the density or cumulative distribution functions of the inter event times can be obtained:

Theorem 2.1: *The reverse relation of (2) is given by*

$$f^*(t) = \lambda^*(t)\exp\left(-\int_{t_n}^{t} \lambda^*(s)\,ds\right), \qquad (4)$$

Or

$$F^*(t) = 1 - \exp\left(-\int_{t_n}^{t} \lambda^*(s)\,ds\right), \qquad (5)$$

Where t_n is the last point before t.

Proof: By (2), we get that

$$\lambda^*(t) = \frac{f^*(t)}{1-F^*(t)} = \frac{\frac{d}{dt}F^*(t)}{1-F^*(t)} = -\frac{d}{dt}\log(1-F^*(t)). \qquad (6)$$

Integrating both sides, we get by the fundamental theorem of calculus that

$$\int_{t_n}^{t} \lambda^*(s)\,ds = -(\log(1-F^*(t)) - \log(1-F^*(t_n))) = -\log(1-F^*(t)) \qquad (7)$$

Since $F^*(t_n) = 0$ here (point t_{n+1} falls on top of t_n with probability zero, since the point process is simple [9]).

Isolating $F^*(t)$ we get (5), and (4) then follows by differentiating $F^*(t)$ with respect to t, again using the fundamental theorem of calculus.

Theorem 2.2: *A conditional intensity function $\lambda^*(t)$ uniquely defines a point process if it satisfies the following conditions for all t and all possible point patterns before t:*

1. $\lambda^*(t)$ *is well-defined and non-negative,*

2.   *the integral $\int_{t_n}^{t} \lambda^*(s)\,ds$ is well-defined,*

3. $\int\limits_{t_n}^{t} \lambda^*(s)ds \to \infty \;\; For\, t \to \infty$.

Proof: Distribution of the point process is well-defined, if all inter event times have well-defined densities, i.e. $f^*(t)$ should be a density function for all t , or equivalently $F^*(t)$ should be a cumulative distribution function. From the three assumptions and (5) it follows that

• $0 \le F^*(t) \le 1$,

• $F^*(t)$ is a non-decreasing function of t ,

• $F^*(t) \to 1$ For $t \to \infty$,

Which means that the distributions of the inter event times are well-defined. Uniqueness follows from Theorem 2.1, since $F^*(t)$ is uniquely obtained from $\lambda^*(t)$ using (5).

Note that item 3. In Theorem 2.2 implies that the point process continuous forever, a property which is often not desirable for practical use luckily we can get rid of this assumption. If we remove this, the proof still holds except that item 3. In the proof has to be removed. Now $F^*(t) \to p$ for some probability $p < 1$, so we have to understand what it means when the cumulative distribution function for the inter event time does not tend to one when time tends to infinity. Basically this means that there is only probability p of having one (or more) points in the rest of the process, and with probability $1-p$ the process terminates with no more points [13], [36].

Example 2.3: (Two terminating point processes)

Consider a unit rate Poisson process on $[0,1]$ .This has conditional intensity function $\lambda^*(t) = 1[t \in [0,1]]$. Thus starting at zero (with no points so far), we get that

$$F^*(t) = 1 - \exp(-\int\limits_0^t 1[s \in [0,1]]ds) = 1 - \exp(-\min\{t,1\}) \tag{8}$$

Where $1[.]$ denotes the indicator function. For $t > 1$, this equals $1 - \exp(-1) \approx 0.63$, so there is a probability of a bout 0.37 of having no points at all. If we do get a point there is an even smaller chance of getting another point after the first one. Another terminating unit rate process could be a process that stops after getting n points. In this case

$$F^*(t) = (1 - \exp(-t))\, 1\, [N\, ([0,t) < n] . \tag{9}$$

Both these examples illustrate that assumption 3. In Theorem 2.2 is not necessary to get well-defined point processes [18].

3. Simulation

Suppose that, starting at 0 say, we trace $\mathbb{R}_+$ in such a way that at the time we are passing position t our speed is $\dfrac{1}{\lambda^*(t)}$, which can be ∞ . (The value $\lambda^*(t)$ is determined by the past, i.e. by what happened up to t .) Then the time instants at which we shall meet all the points in $\mathbb{R}_+$ of the process form a homogeneous Poisson process.

In other language, the random time transformation $\tau = \Lambda^*(t) = \int\limits_0^t \lambda^*(u)du$ takes the point process with conditional intensity function $\lambda^*(t)$ into a unit rate Poisson process [27].

Lemma 3.1: *Let X be a random variable with continuous distribution function $F^*(.)$ and integrated Hazard function $H(x) = -\log[1 - F^*(x)]$. Then $Y = H(X)$ has a unit exponential distribution (i.e. with unit mean). Conversely, if Y is a random variable with unit exponential distribution, then $X = H^{-1}(Y)$ has distribution function $F^*(.)$.*

Therefore, we have a sequence of interval lengths $X_1, X_2,...$ with continuous distributions $F_1^*(t), F_2^*(t),...$; the corresponding sequence of transformed random variables $Y_1 = H_1(X_1), Y_2 = H_2(X_2),...$ is a sequence of unit exponential random variables.

If the intervals $X_1, X_2,...$ represent the sequence of intervals for a point process with conditional intensity function $\lambda^*(t)$ that can be represented in terms of integrated hazard functions as above, then the joint distribution of any finite sequence of these intervals is the product of the distribution functions $F_i^*(t)$, and the joint distribution of the corresponding transformed random variables $H_1(X_1), H_2(X_2),...$ is the product of unit exponential distributions and therefore represents the joint distribution of a set of i.i.d. unit exponential random variables. But such a point process is just a unit rate Poisson process.

The requirement that the compensator $\Lambda^*(t)$ should increase without bound ensures that there is no last point in the process. The basic result remains valid without it, except insofar as the final interval is then infinite and so cannot belong to a unit rate Poisson process. The extreme case in this example makes the point [9], [10].

Example 3.1: (One point process)

Let a point process on $(0,\infty)$ have exactly one point, at t_1, say, where $P\{t_1 \leq x\} = F^*(x)$, and we assume that the D.F F^* is continuous. Then

$$\Lambda^*(t) = \int_0^{\min(t,t_1)} \frac{dF^*(u)}{1 - F^*(u)} = -\log(1 - F^*[\min(t,t_1)]).\tag{10}$$

The initial interval transforms to an interval with unit exponential distribution; the transformed process then terminates. The converse part of Lemma 3.1 contains within it the basis for one general approach to simulating point processes. Using the notation $X_1, X_2,...$, $F_1^*(.), F_2^*(.),...$ and $H_1(.), H_2(.),...$ as in the Lemma 3.1, it may be summarized as follows.

Algorithm 3.1: (Simulation of point processes by conditional intensity function (Hazard function))
1. Simulate a sequence $Y_1, Y_2,...$ of unit exponential random variables (respectively, a sequence $U_1, U_2,...$ of uniform $U(0,1)$ random variables).
2. Transform to the sequence of successive interval lengths $X_1 = H_1^{-1}(Y_1), X_2 = H_2^{-1}(Y_2),...$ (respectively, the sequence $F_1^{*-1}(U_1), F_2^{*-1}(U_2),...$).
3. Form the point process $(t_1, t_2,...)$ by setting $t_1 = X_1, t_2 = X_1 + X_2,...$.

The use of exponential or uniform random variables to initiate the algorithm is immaterial in that both lead to point processes with identical properties. The use of the exponential variates shows more clearly the relation to the Poisson process and may be marginally more convenient when the process is specified through its conditional intensity function because $t_1, t_2,...$ then solve the successive equations:

$$\int_0^{t_1} \lambda^*(u)du = Y_1 \ ,\tag{11}$$

$$\int_{t_1}^{t_2} \lambda^*(u)du = Y_2 \ ,\tag{12}$$

And so on. The main constraint in the use of this algorithm is the common need to introduce an iterative numerical method to find the inverse of the integrated Hazard or distribution function.

In principle, the method may be extended to situations where the interval distributions are conditioned by external as well as internal variables, provided that all the relevant conditioning information is available at the beginning of each new interval [5], [7], [8] and [9].

4. Numerical results

Example 4.1: Suppose we have

$$F^*(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^2 + x}{2} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \qquad F^{*-1}(x) = \begin{cases} 0 & x < 0 \\ -\dfrac{1}{2} + \sqrt{2x + \dfrac{1}{4}} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}\tag{13}$$

Then, we have:

$$f^*(x) = \frac{2x + 1}{2} \qquad 0 \leq x < 1\tag{14}$$

Now, according to relations (2) and (3), we have:

$$\lambda^*(u) = \frac{f^*(u)}{1-F^*(u)} = \frac{2u+1}{2-u^2-u} \qquad 0 \le u < 1 \tag{15}$$

In which case according to Algorithm 3.1, we have:

Table 1: Simulation of Point Process for n=10 (Elapsed Time=0.926694)

| Interval lengths (X)	Point process (t)
1.0000	1.0000
1.0000	2.0000
1.0000	3.0000
1.0000	4.0000
0.9111	4.9111
0.7026	5.6137
0.5474	6.1612
0.4449	6.6061
0.2768	6.8828
0.1222	7.0051

Table 2: Simulation of Point Process for n=70 (Elapsed Time=0.261724)

| Interval lengths (X)	Point process (t)
1.0000	1.0000
1.0000	2.0000
1.0000	3.0000
.	.
.	.
.	.
0.0846	47.8957
0.0723	47.9680
0.0454	48.0134
0.0088	48.0223

Table 3: Simulation of Point Process for n=400 (Elapsed Time=0.369915)

Interval lengths (X)	Point process (t)
1.0000	1.0000
1.0000	2.0000
1.0000	3.0000
.	.
.	.
.	.
0.0153	274.6100
0.0130	274.6230
0.0072	274.6302
0.0018	274.6320

Also according to the conditional intensity function and high tables for the sightly distribution function, we have:
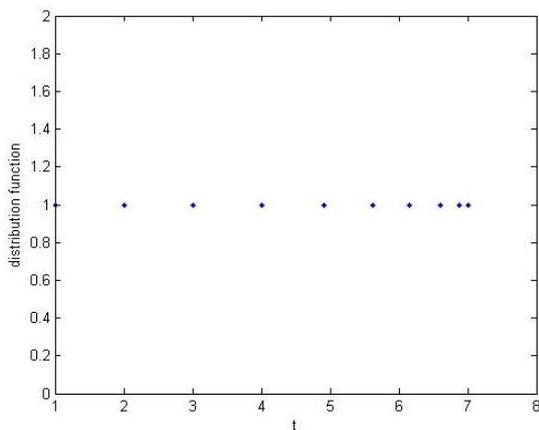
Fig. 2: Monte Carlo Simulation of Table1
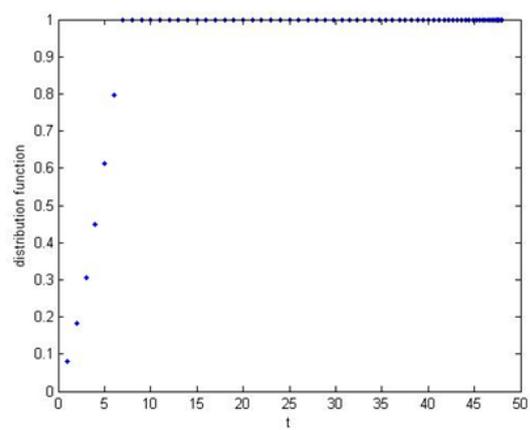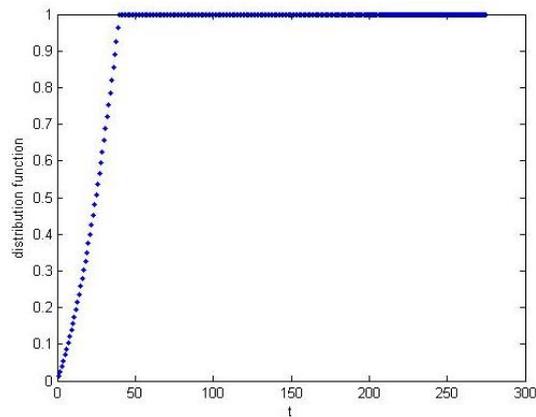
Fig. 3: Monte Carlo Simulation of Table2

Fig. 4: Monte Carlo Simulation of Table3.

As we can see, the horizontal axis represents time points examined and the vertical axis shows the cumulate distribution function values. Since, the conditional intensity function is non-negative (because the distribution function is non-negative), with the rise over time, accumulate points at a distance less than the desired distribution function, the smoother the point on desire areas will happen. We also note that no matter how much time will the contrary, the time not only will not increase but the reduction will also be happened. It would be desirable to increase the convergence speed of the process.

Also, another point to note here is that in Algorithm 1.3 to obtain the results we employed partitioned rand-library function in Matlab software and step length as $X = \dfrac{t}{(\frac{n}{10})}$, and it will improve the speed computations time points.

5.   Conclusion

The aim of this paper was to review the long term time points of a point process with a given conditional intensity function $\lambda^*(t)$. We see upside of these areas by increasing the number of points in time when the recovery was smooth and shapes of the distribution function converge to the desired time point.

References

[1]   Badar, M. G., and Priest, A. M., "Statistical aspects of fiber and bundle strength in hybrid composites", Progress in Science and Engineering Composites, Hayashi, T., Kawata, K. and Umekawa, S. (Eds.), ICCM- IV, Tokyo, (1982) 1129-1136.

[2]   Bennett, S., "Logistic regression model for survival data", *Applied Statistics*, 32, (1983) 165-171.

[3]   Bjerkedal, T., "Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli", *American Journal of Hygiene*, 72, (1960) 130-148.

[4]   Bremaud, P. and Massoulie, L., "Hawkes branching point processes without ancestors", J. Appl. Probab, 38, (2001) 122–135.

[5]   Brown, T. M. and Nair, M., "A simple proof of the multivariate random time change theorem for point processes", J. Appl. Probab, 25, (1988) 210–214.

[6]   Carlsson, H. and Nerman, O., "An alternative proof of Lorden's renewal inequality", Adv. Appl. Probab, 18, (1986) 1015–1016.

[7]   Chung, K. L., "Crudely stationary point processes", Amer. Math. Monthly, 79, (1972) 867–877.

[8]   Copson, E. C., *An introduction to the theory of functions of a complex variable*, Oxford University Press, Oxford, (1935).

[9]   Daley, D. J. and Vere-Jones, D., *An introduction to the theory of point processes,* Volume I: Elementary Theory and Methods. Springer, New York, 2nd edition, (2003).

[10]   Daley, D. J. and Vere-Jones, D., *An introduction to the theory of point processes,* Volume II: General Theory and Structure. Springer, New York, 2nd edition, (2008).

[11]   Daley, D. J., "On a class of renewal functions", Proc. Cambridge Philos. Soc. 61, (1965) 519–526.

[12]   Daley, D.J. and Vere-Jones, D., "A summary of the theory of point processes", In Lewis, (1972) 299–383.

[13]   Davies, R.B., "Testing the hypothesis that a point process is Poisson", Adv. Appl. Probab, 9, (1977) 724–746.

[14]   Desmond, A. F., "Stochastic models of failure in random environments", *Canadian Journal of Statistics*, 13, (1985) 171-183.

[15]   Diggle, P.J., Gratton, R.J., "Monte Carlo methods of inference for implicit statistical models (with discussion)", J. Royal Statist. Soc. Ser. B, 46, (1984) 193–227.

[16]   Doob, J.L., "Renewal theory from the point of view of the theory of probability", Trans. Amer. Math. Soc. 63, (1948) 422–438.

[17]   Glaser, R. A., "Bathtub and related failure rate characterization", *Journal of the American Statistical Association*, 75, (1980) 667-672.

[18]   Gupta, R. C., Kannan, N. and RayChoudhuri, A., "Analysis of log-normal survival data", *Mathematical Biosciences*, 139, (1997) 103-115.

[19]   Hawkes, A. G., "Point spectra of some mutually exciting point processes", *Journal of the Royal Statistical Society*, Series B, 33, (1971) 438–443.

[20]   Hawkes, A. G., "Spectra of some self-exciting and mutually exciting point processes", *Journal of Biometrika*, 58(1), (1971) 83–90.

[21]   Hawkes, A. G., "Spectra of some mutually exciting point processes with associated variables", In P. A. W. Lewis, editor, *Stochastic Point Processes*, Wiley, New York, (1972) 261–271.

[22]   Hawkes, A. G. and Oakes, D., "A cluster representation of a self-exciting process", *Journal of Applied Probability*, 11, (1974) 493–503.

[23]   Isham, V. and Westcott, M., "A self-correcting point process", Stoch. Proc. Appl., 8, (1979) 335–347.

[24]   Ludwig, J.A., Reynolds, J.F., *Statistical ecology: A primer on methods and computing*, John Wiley & Sons, New York, (1988).

[25] Mann, N. R., Schafer, R. E., and Singpurwalla, N. D., *Methods for statistical analysis of reliability and life data*, John Wiley & Sons, New York, (1974).

[26] Møller, J. and Rasmussen, J. G., "Perfect simulation of hawkes processes", Adv. in Appl. Probab., 37(3), (2005) 629–646.

[27] Møller, J. and Rasmussen, J. G., "Approximate simulation of hawkes processes", Methodol. Comput. Appl. Probab., 8, (2006) 53–65.

[28] Møller, J. and Waagepetersen, R. P., *Statistical inference and simulation for spatial point processes*. Chapman & Hall, Boca Raton, Florida, (2004).

[29] Ogata, Y., "On Lewis' simulation method for point processes", IEEE Trans- actions on Information Theory, IT-27(1), (1981) 23–31.

[30] Ogata, Y., "Statistical models for earthquake occurrences and residual analysis for point processes", *Journal of the American Statistical Association*, 83(401), (1988) 9–27.

[31] Ogata, Y., "Space-time point-process models for earthquake occurrences", *Annals of the Institute of Statistical Mathematics*, 50(2), (1998) 379–402.

[32] Parzen, E., *Stochastic processes*, Holden-Day, San Francisco, (1962).

[33] Picinbono B., "Time intervals and counting in point processes", IEEE Trans. Inform. Theory, 50, (2004) 1336-1340.

[34] Pinelis, I., "Monotonicity properties of the relative error of a Pade approximation for Mills' ratio", *Journal of Inequalities in Pure and Applied Mathematics*, 3, 20, (2002).

[35] Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T., *Numerical Recipes*: *The Art of Scientific Computing*, Cambridge University Press, Cambridge, U.K. (1991).

[36] Rasmussen, J. G., "Temporal point processes the conditional intensity function", January 24, (2011).

[37] Ross, Sh. M., *Introduction to probability models*, Academic Press, San Diego, 8nd edition, (2003).

[38] Rubenstein, R. Y., *Simulation and the Monte Carlo method*, Wiley, New York, (1981).

[39] Schmeiser, B. W., "Random variate generation a survey", Proc, Winter Simulation Conf., Orlando, FL, (1980) 79-104.

[40] Xie, F.C. and Wei, B.C., "Diagnostics analysis for log-Birnbaum-Saunders re- gression models", *Computational Statistics & Data Analysis*, 51, (2007) 4692-4706.