

A NOTE ON DISCRIMINATING POISSON PROCESSES FROM OTHER POINT PROCESSES WITH STATIONARY INTER ARRIVAL TIMES

GUSZTÁV MORVAI AND BENJAMIN WEISS

We give a universal discrimination procedure for determining if a sample point drawn from an ergodic and stationary simple point process on the line with finite intensity comes from a homogeneous Poisson process with an unknown parameter. Presented with the sample on the interval $[0, t]$ the discrimination procedure g_t , which is a function of the finite subsets of $[0, t]$, will almost surely eventually stabilize on either POISSON or NOTPOISSON with the first alternative occurring if and only if the process is indeed homogeneous Poisson. The procedure is based on a universal discrimination procedure for the independence of a discrete time series based on the observation of a sequence of outputs of this time series.

Keywords: Point processes

Classification: 60G55

1. INTRODUCTION

The simplest point process is undoubtedly the homogeneous Poisson point process on the real line. It can be described as a random countable discrete subset of \mathbb{R} with the property that the random variables that count the number of points in disjoint intervals are independent with a Poisson distribution and parameter proportional to the lengths of the intervals. An alternate description is via the inter arrival process which in the case of a Poisson process consists of independent exponentially distributed random variables with a fixed parameter. There has been much work done in the statistical literature concerning tests to determine whether a given data set is best modeled by such processes. Our discrimination procedure will not be a statistical test in a sense that our result will be asymptotic, we will not give upper bounds for type I or type II errors.

We propose to study the identification problem from the point of view of the theory of universal discrimination procedures which are asymptotically point wise consistent. In an earlier paper we have given a series of procedures $g_n(X_1, X_2, \dots, X_n)$ which when presented with a sequence coming from a discrete time ergodic process will eventually almost surely stabilize on IND if the process is independent and NOTIND otherwise. We

will view a point process as a random discrete subset $\omega \subset \mathbb{R}$ and then provide a sequence of discrimination procedures $g_n(\omega \cap [0, n])$ which will eventually almost surely stabilize on POISSON if the process we are sampling is Poisson and on NOTPOISSON otherwise.

Here is a more formal description of the setup that we are considering. The random discrete subset can be described by random variables $\dots, R_{-1} < 0 \leq R_0 < R_1 \dots$ defined on a probability space $(\Omega, \Sigma, \mathbf{P})$ where the elements of Ω are discrete subsets of \mathbb{R} and the random variables $R_i(\omega)$ are the points of ω . The σ -algebra Σ is generated by functions that count the number of points in $\omega \cap [a, b]$ for arbitrary intervals $[a, b]$. There is a natural flow T_t defined on the space Ω that takes the element ω to $\omega - t$ and we will assume that \mathbf{P} is invariant and ergodic under this flow. In addition we will only consider those point processes with the property that the expected number of points in $\omega \cap [-N, N]$ is finite for all N . In this case it is well known that a probability space $(\Omega^0, \Sigma^0, \mathbf{P}^0)$ can be defined so that the inter arrival times, defined by $X_n = R_{n+1} - R_n$, form a stationary and ergodic stochastic process with $\mathbf{E}^0(X_n) < \infty$ and the probability that $X_i = 0$ equals zero. Now the σ -algebra Σ^0 is generated by the inter arrival times and \mathbf{P}^0 is what is called the Palm distribution in the literature of point processes.

Conversely, starting from any such ergodic stationary process one can define an ergodic stationary point process by suitably randomizing the position of zero. Essentially what is done is to take $(\Omega^0, \Sigma^0, \mathbf{P}^0)$ and form $(\Omega, \Sigma, \mathbf{P})$ by multiplying it by the unit interval with the uniform measure on $[0, 1]$. If U represents the new uniform random variable then the R_i are defined by setting $R_0 = UX_0$ and then defining the rest of the R_n 's by the inter arrival times. This is carried out in detail for example in (Thorisson [8], Chapter 8). In this duality ergodicity of the flow corresponds exactly to the ergodicity of the discrete stationary process (cf. 8.7 in Thorisson [8]).

As is well known the process is a Poisson point process with parameter λ if and only if the X_i 's are independent exponential random variables with parameter λ with respect to the Palm measure. It follows easily from the construction of the Palm measure that for processes with finite intensity, which are the only ones we are considering, almost sure results are the same for the two measures. Our estimation scheme will proceed in two parts. In the the first part we will probe the observed inter arrival times for independence. Whenever a positive answer is received we will further probe for the distribution being exponential. This second procedure will rely on the classical Dvoretzky–Kiefer–Wolfowitz inequality with the tight constant that was obtained by P. Massart [5]. In the first section we deal with the first part and basically rely on our earlier work [6]. In the second section, which stands entirely in the classical setting of determining an unknown distribution from a sequence of independent samples, we will give our procedure for discriminating the exponential family. There are of course many papers in the statistical literature which deal with the problem of testing for an exponential distribution (see for example Haywood and Khmaladze [2] and the references listed there). Since our asymptotic point wise setting is rather different we give a complete discussion of our procedure ab ovo.

2. DISCRIMINATING INDEPENDENT PROCESSES FROM OTHER STATIONARY PROCESSES

Let $\{X_n\}$ be a stationary and ergodic real valued process. When we observe a sample of the point process in the time interval $[0, T]$, i.e. when we observe $\omega \cap [0, T]$, we will observe a random number of inter arrival times $\{X_i | 0 \leq i \leq \tau_T\}$ where $\tau_T(\omega)$ is the greatest index n such that $R_{n+1} \leq T$. We shall probe this sequence for being derived from an independent process versus the alternative that this real valued process is not independent. Even though the sequence of X_i is not stationary with respect to the underlying sample space of the point process (the interval X_0 has a different distribution even when the inter arrival times are independent - as long as they are not constant) the Palm measure, with respect to which they do form a stationary process is equivalent to the probability measure of the process so that the notion of almost sure is the same for the two measures.

As usual, we will denote the sequence $\{X_1, X_2, \dots, X_n\}$ by X_1^n . Let $DIID_n(X_1^n)$ be an arbitrary discrimination procedure such that eventually almost surely

$$DIID_n(X_1^n) = \begin{cases} IND & \text{if the process is independent} \\ DEP & \text{otherwise.} \end{cases}$$

We gave an example of such a discrimination procedure in Morvai and Weiss [6]. (Cf. Ryabko and Astola [7] also.) We can now define a discrimination procedure $DIID_T(A)$ where A is a finite subset of $[0, T]$ with $A = \{r_0 < r_1 < \dots < r_{n+1}\}$ by

$$DIID_T(A) = DIID_n(r_1 - r_0, r_2 - r_1, \dots, r_{n+1} - r_n).$$

For a given ω , our procedure for discriminating independence becomes simply $DIID_T(\omega \cap [0, T])$.

3. DISCRIMINATING EXPONENTIALLY DISTRIBUTED PROCESSES FROM OTHER I.I.D. PROCESSES

Let X_1, X_2, \dots , be independent identically distributed real valued random variables. Define the empirical distribution $F_n(t)$ as

$$F_n(t) = \frac{|\{1 \leq i \leq n : X_i \leq t\}|}{n}.$$

Let $F^\lambda(t)$ denote the distribution function of the exponential distribution with parameter $\lambda > 0$, that is,

$$F^\lambda(t) = \begin{cases} 1 - e^{-\lambda t} & \text{if } t \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

We would like to compare $F_n(t)$ with $F^\lambda(t)$ but we do not know the value of λ . It turns out that given a certain level of error ϵ_n that we expect to see we can express the fact that there exists a positive number λ such that

$$\sup |F(t) - F^\lambda(t)| < \epsilon_n$$

as a finite number of inequalities. These are obtained in the following way. Let $Y_1 \leq Y_2 \leq \dots \leq Y_n$ be the ascending rearrangement of X_1, X_2, \dots, X_n . The empirical distribution function of the X_i takes the value i/n on the interval $[Y_i, Y_{i+1})$ and since distribution functions are monotone the maximal deviation between $F^\lambda(t)$ and $F(t)$ will take place at the endpoints of these intervals. We will take $\epsilon_n = \sqrt{\frac{\ln(n)}{n}}$ and this will give rise to the following possible intervals for λ :

$$I_i(X_1^n) = \left[\frac{-\ln\left(1 - \frac{i}{n} + \sqrt{\frac{\ln(n)}{n}}\right)}{Y_i}, \frac{-\ln\left(1 - \frac{i-1}{n} - \sqrt{\frac{\ln(n)}{n}}\right)}{Y_i} \right].$$

Define the intersection of the intervals I_1, \dots, I_n by J_n , that is,

$$J_n(X_1^n) = \bigcap_{i=1}^n I_i(X_1^n).$$

Finally we define the discrimination procedure $DEXP(X_1^n)$ as

$$DEXP_n(X_1^n) = \begin{cases} EXP & \text{if for } i \neq j : X_i \neq X_j \text{ and } J_n(X_1^n) \text{ is not empty} \\ NONEXP & \text{otherwise.} \end{cases}$$

Theorem 3.1. Let X_1, X_2, \dots , be independent identically distributed. Then $DEXP_n(X_1^n) = EXP$ eventually almost surely if X_i 's are exponentially distributed for some parameter λ and $NONEXP$ eventually almost surely otherwise.

Proof. The event that for some $i \neq j$ we will see $X_i = X_j$ can happen with positive probability only if the distribution of the X_i 's has an atom in which case it certainly can not have a density function not to speak of an exponential density function. If there is an atom then by ergodicity we will certainly see such events and so we can put $DEXP_n(X_1^n) = NONEXP$ whenever . for some $i \neq j : X_i = X_j$ and $n > \max(i, j)$. Thus we may assume that the distribution of the X_i 's is non atomic which implies that with probability one if $i \neq j$ then $X_i \neq X_j$. So we may assume that $Y_1 < Y_2 < \dots < Y_n$. Let $Y_0 = -\infty, Y_{n+1} = \infty$. Then for $-\infty < t < \infty$,

$$F_n(t) = \frac{i}{n} \text{ if } Y_i \leq t < Y_{i+1}.$$

Assume that the X_i 's are exponentially distributed with some parameter $\lambda > 0$. By the tight version of the Dvoretzky-Kiefer-Wolfowitz Inequality due to Massart [5]

$$P \left(\sup_t |F_n(t) - F^\lambda(t)| > \sqrt{\frac{\ln n}{n}} \right) \leq 2e^{-2 \ln n} = 2n^{-2}$$

which is summable. By the Borel–Cantelli lemma, eventually almost surely,

$$\sup_t |F_n(t) - F^\lambda(t)| \leq \sqrt{\frac{\ln n}{n}}.$$

Since $F^\lambda(t)$ is monotone increasing the maximum deviation of $F_n(t)$ from $F^\lambda(t)$ is at the points Y_i . Calculating the difference for $i = 1, 2, \dots, n$

$$\left| 1 - e^{-\lambda Y_i} - \frac{i}{n} \right| \leq \sqrt{\frac{\ln n}{n}}$$

and

$$\left| 1 - e^{-\lambda Y_i} - \frac{i-1}{n} \right| \leq \sqrt{\frac{\ln n}{n}}$$

eventually almost surely. Thus eventually almost surely, $\lambda \in I_i(X_1^n)$ for $i = 1, \dots, n$ and so $\lambda \in J_n(X_1^n)$ that is $J_n(X_1^n)$ is not empty and $DEXP(X_1^n) = EXP$ eventually almost surely.

Now assume that the X_i 's are not exponentially distributed and non atomic. Let $G(t) = P(X \leq t)$, the true unknown distribution function of the process. We again apply Massart's result [5] and conclude that:

$$P \left(\sup_t |F_n(t) - G(t)| > \sqrt{\frac{\ln n}{n}} \right) \leq 2e^{-2 \ln n} = 2n^{-2}$$

which is summable. By the Borel–Cantelli lemma almost surely there will be an N , such that for all $n \geq N$,

$$\sup_t |F_n(t) - G(t)| \leq \sqrt{\frac{\ln n}{n}}. \tag{1}$$

Since G is nonatomic there must exist $a < b$ such that $0 < G(a) < G(b) < 1$. We will use this to put a priori bounds on the possible values of the parameter λ in case the discrimination procedure will say EXP which will enable us to get a convergent sequence and contradict the assumption that $G(t)$ is not exponential.

Now we argue by contradiction. Assume that on a subsequence n_j , $DEXP(X_1, \dots, X_{n_j}) = EXP$. This means that there exists a sequence of λ_j such that $\lambda_j \in J_{n_j}(X_1, \dots, X_{n_j})$. In other words, for $i = 1, 2, \dots, n_j$

$$\left| 1 - e^{-\lambda_j Y_i} - \frac{i}{n_j} \right| \leq \sqrt{\frac{\ln n_j}{n_j}}$$

and

$$\left| 1 - e^{-\lambda_j Y_i} - \frac{i-1}{n_j} \right| \leq \sqrt{\frac{\ln n_j}{n_j}}.$$

Since $F^\lambda(t)$ is monotone increasing the maximum deviation of $F_{n_j}(t)$ from $F^{\lambda_j}(t)$ is at the points Y_i and the deviations at the points Y_i are

$$\left| 1 - e^{-\lambda_j Y_i} - \frac{i}{n_j} \right|$$

and

$$\left| 1 - e^{-\lambda_j Y_i} - \frac{i-1}{n_j} \right|$$

thus

$$\sup_t |F_{n_j}(t) - F^{\lambda_j}(t)| \leq \sqrt{\frac{\ln n_j}{n_j}}. \tag{2}$$

By (1), (2) and the triangle inequality, for all $n_j > N$,

$$\sup_t |G(t) - F^{\lambda_j}(t)| \leq 2\sqrt{\frac{\ln n_j}{n_j}}. \tag{3}$$

Since $0 < G(a) < G(b) < 1$ implies that for some $0 < L < M < \infty$ for all $n_j > N$, $\lambda_j \in [L, M]$. Since λ_j is from a bounded interval $[L, M]$, it has a convergent subsequence with some limit $\lambda > 0$. Thus

$$\sup_t |F^\lambda(t) - F^{\lambda_j}(t)| \rightarrow 0. \tag{4}$$

By(3)

$$\sup_t |G(t) - F^{\lambda_j}(t)| \rightarrow 0. \tag{5}$$

But $G \neq F^\lambda$. So (4) and (5) can not hold at the same time. This is a contradiction. The proof of Theorem 3.1 is complete. \square

4. DISCRIMINATING POISSON PROCESSES FROM OTHER POINT PROCESSES WITH STATIONARY INTER ARRIVAL TIMES

Let $0 \leq R_1 < R_2 < \dots$ be the arrival times of the ergodic point process. The inter arrival times $X_n = R_{n+1} - R_n$ form a stationary and ergodic real valued process under the Palm measure which has the same null sets as the probability measure of our process. This means that we can assume that we are sampling the X_i from a stationary ergodic process. To simplify the notation we will now write our discrimination procedure in terms of these inter arrival times rather than in terms of the point process ω . Define

$$\begin{aligned} &DPOISSON(X_1^n) \\ &= \begin{cases} POISSON & \text{if } DIID_n(X_1^n) = IID \text{ and } DEXP_n(X_1^n) = EXP \\ NONPOISSON & \text{otherwise.} \end{cases} \end{aligned}$$

Remark 4.1. One has to calculate $DEXP(X_1^n)$ only if $DIID(X_1^n) = IID$.

Theorem 4.2. Assume that the point process $\{R_i\}$ has stationary and ergodic inter arrival times $\{X_n\}$. Then $DPOISSON(X_1^n) = POISSON$ eventually almost surely if the point process $\{R_i\}$ is a Poisson process and $NONPOISSON$ eventually almost surely otherwise.

Proof. Since a point process is Poisson if and only if it has inter arrival times which are independent, identically distributed with exponential distribution for some $\lambda > 0$ the proof of Theorem 4.2 is complete. \square

ACKNOWLEDGEMENT

This first author was supported by Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, 13-15 Reáltanoda utca, H-1053, Budapest, Hungary.

(Received November 29, 2018)

REFERENCES

-
- [1] D. J. Daley and D. Vere-Jones: An introduction to the theory of point processes. Vol. II. General theory and structure. Second edition. In: Probability and its Applications. Springer, New York 2008. DOI:10.1007/978-0-387-49835-5
 - [2] J. Haywood and E. Khmaladze: On distribution-free goodness-of-fit testing of exponentiality. *J. Econometr.* 143 (2008), 5–18. DOI:10.1016/j.jeconom.2007.08.005
 - [3] O. Kallenberg: Foundations of modern probability. Second edition. In: Probability and its Applications. Springer-Verlag, New York 2002. DOI:10.1007/978-1-4757-4015-8
 - [4] P. A. W. Lewis: Some results on tests for Poisson processes. *Biometrika* 52 (1965), 1 and 2, 67–77. DOI:10.1093/biomet/52.1-2.67
 - [5] P. Massart: The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *Ann. Probab.* 18 (1990), 3, 1269–1283. DOI:10.1214/aop/1176990746
 - [6] G. Morvai and B. Weiss: Testing stationary processes for independence. *Ann. Inst. H. Poincaré Probab. Statist.* 47 (2011), 4, 1219–1225. DOI:10.1214/11-aihp426
 - [7] B. Ryabko and J. Astola: Universal codes as a basis for time series testing. *Statist. Methodol.* 3 (2006), 375–397. DOI:10.1016/j.stamet.2005.10.004
 - [8] H. Thorisson: Coupling, stationarity, and regeneration. In: Probability and its Applications. Springer-Verlag, New York 2000. DOI:10.1007/978-1-4612-1236-2

Gusztáv Morvai, Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, 13-15 Reáltanoda utca, H-1053, Budapest, Hungary and MTA-BME Stochastics Research Group, 1 Egy József utca, Building H, Budapest, 1111. Hungary.
e-mail: morvai@math.bme.hu

Benjamin Weiss, Hebrew University of Jerusalem, Jerusalem 91904. Israel.
e-mail: weiss@math.huji.ac.il