A note on the history of the Poisson process.

Tomasz Rolski

September 23, 2019

1 Introduction

This note reviews how the contemporary concept of the Poisson process (sometimes called the Poisson random measure) evolved through out the time up to recent times. That is from F. Lundberg and W. Feller up to A. Renyi paper [30] from 1967 and finally, a construction of Poisson processes by binomial (or Bernoulli) processes, which can be found in J.E. Moyal [24], J. Mecke [23] and J.F. Kingman [16] papers. Also in C. Ryll-Nardzewski [32] this construction was used in a proof of Theorem for homogeneous Poisson processes.

Let us begin with a quotation from Harald Cramér [2]. "It was not until the 1930's that the foundation of a general and rigorous theory of stochastic processes were laid by Kolmogorov, Khintchine, Lévy, Feller, and the other authors. In this connections it became clear that the Brownian movement process, and the Lundberg risk process, introduced many years earlier by the pioneers, form the main building stones of an important general class of stochastic process, known as processes with independent increments." Notice that a special case of the Lundberg risk process,with all the sums at risk equal to 1, leads to what we now understand by a Poisson process. In general by a Lundberg risk process Cramér understood what we recently mean by a compound Poisson process.

One has to distinguish results obtained in a less mathematical formal way from today standards, that is before a foundation of probability theory and stochastic processes were laid. Because of this Filip Lundberg works for many years were rather unnoticed by mathematicians, as well his papers were published in Swedish. The main ideas, which later became a part of the so called collective risk theory (which is a part of actuarial mathematics), where included in his doctoral thesis [20]. He introduced therein the compound Poisson process and involved work on the central limit theorem.

From another side, this time applied to a telecommunication problem, A.K. Erlang considered a Poisson process as an input to an automatic telephone exchanges; [7], although the first ideas which lead to a prototype of the Poisson processes appeared in [6]. Again the concepts were difficult to understand for mathematicians, and besides, another obstacle for international community was the papers in Danish.

A contemporary reader has available monographs on point processes, and in particular on Poisson processes (Kallenberg [14], a two volume Daley and Vere-Jones [4, 5]). A modern approach to point processes and Poisson processes defined on abstract spaces can be found in a recent book by Last and Penrose [19], wherein the reader can find a section on the history of Poisson processes. One can also recommend as an introductory text the book by Kingman [17].

1.1 The name.

We first recall the origin of names "point process" and "Poisson process". Thus in the dissertation of Conny Palm in 1943, there exists the first known recorded use of the term point process as *Punktprozesse* in German; see [13]. It is believed, see Guttorp and Thorarinsdottir [12], that William Feller [9] was the first in print to refer to it as the Poisson process in a 1940 paper. Although the Swedish statistician Ove Lundberg [21] used the term Poisson process in his 1940 PhD dissertation.

As Guttorp and Thorarinsdottir write, it is likely that the term was adopted by most participants of the Berkeley Symposium on Probability and Mathematical Statistics in 1945 and from now on was commonly used in the literature.

1.2 Early works

Guttorp and Thorarinsdottir [12] mentioned about Filip Lundberg result from 1903, where by a rather heuristic reasoning he derives for probability function $f(x)$ equation

$$\frac{\partial f(x, P)}{\partial P} = f(x-1, P) - f(x, P),$$

which is the forward equation for the Poisson process and it was later derived by Kolmogorov [18] for Markov processes. Clearly the solution of this equation is $f(x, P) = P^x e^{-x}/x!$. In his memoirs [3], Cramér mentioned about Lundberg risk process in the context of early work on Poisson process. He also mentioned the work by Erlang from 1909 in the context of telephone traffic problem and of Rutherford and Geiger from 1908 in the analysis of radioactive disintegration. Cramér also recall his early work (in Swedish) from 1919 on the Poisson process. Finally, let us point out a rather overlooked and forgotten pioneering discovery by Norbert Wiener [34], that the assumption of independence of the number of points in disjoint subsets of an Euclidean space leads to Poisson distribution of a number of points in a subset. The author of this note thanks to professor Guenter Last for pointing me out this reference.

In the paper by Feller from 1949 [10], when surveying some processes he writes: *The Poisson process.*– This is quite familiar to physicists, who often refer to it as "random events" and occasionally call the Poisson distribution after Bateman.

1.3 Poisson process as a counting process

By a counting process we mean a stochastic process $(X(t))_{t\geq 0}$ with non-decreasing right continuous and $\mathbb{Z}_+$-valued. Following Palm one can interpret the increment $X(t+s)-X(t)$ as a number of points in interval $(t,t+s]$. By a Poisson process with parameter λ it was meant (see e.g. Feller [10]): "events occurring in time, such as telephone calls, radioactive disintegrations, impact of particles (cosmic rays), and the like. Let it be assumed that (i) the probability of an event in any time interval of length dt, is asymptotically, $\lambda\, dt$, where λ is positive constant; (ii) the probability of more than one event in a time interval dt is of smaller order of magnitude than dt, in symbols $o(dt)$; (iii) the number of events in non-overlapping intervals represent independent random variables." If $P_n(t)$ was denoting the probability of having exactly n events in a time interval of length t, than Feller showed $P_n(t) = (\lambda t)^n e^{-\lambda t}/n!$. In modern language, Poisson process $N(t)_{t\geq 0}$ is a stochastic process, with (iii) independent increments, (ii) with stationary increments and (i) orderly. The name of orderly comes from Russian *ordinarnii* (see e.g. Khintchin [?]). In Daley and Vere-Jones [4] in Theorem 2.2.III it was recalled, that $N(t)$ with stationary and independent increments, which is also orderly is a Poisson process, that is $\mathbb{P}(N(b)-N(a)=n) = (\lambda(b-a))^n e^{-\lambda(b-a)}/n!$.

1.4 Poisson processes

In this section we recall what today is understood by a Poisson process (or Poisson point process or Poisson random measure). We follow Daley and Vere-Jones [4], section 2.4, who consider E a complete separable metric space, $\mathcal{E}$ the Borel σ-field generated by open spheres of E and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Remark that in the book of Last and Penrose [] there is presented a definition of Poisson process on a general state space E, however then one must overcome difficulties of the lack of local finitness.

A contemporary approach to point process is to define N as a measurable mapping $(\Omega, \mathcal{F})$ to $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, where $\mathcal{N}$ is a space of discrete measures locally finite on $(E, \mathcal{E})$, and $\mathcal{N}$ is endowed with the vague topology. If N assumes values in $\mathcal{N}_0$ of measure with unit atoms, then N is said to be a simple point process (or without multiple points). In other words we have a family of $\mathbb{Z}_+$-valued random variables $N(E))_{E\mathcal{E}}$, such that for each $\omega \in \Omega$, $N(\cdot)$ is a discrete measure.

We now state what one recently means by a Poisson process on $(E, \mathcal{E})$. Suppose Λ, sometimes called a parameter measure, is a boundedly finite measure on E. One says that a point process N is a Poisson process with parameter measure Λ if

(PP1) for every finite family of disjoint bounded Borel sets $\{B_i, i = 1, \ldots, k\}$ $N(B_1), \ldots, N(B_k)$ are independent,

(PP2) for bounded Borel set B

$$P(N(B) = n) = \frac{[\Lambda(B)]^n}{n!} e^{-\Lambda(B)}.$$

Sometimes N is said to be a *Poisson random measure.* In other words we have a family of discrete random variables $(N(B))B \in \mathcal{B}$, such that for each $\mathbb{P}$-each ω, $\mathbb{P}(\cdot)$ is a discrete measure. Remark that there are theories that no topological assumptions are required; see eg. the book of Last and Penrose [19].

One of problems is to demonstrate the existence of the Poisson measure. In the literature it is considered that the positive answer is due to Kingman [16] or Mecke[23], where the construction was via the so called binomial process; i.e. in set B there is $N(B)$ points, which are thrown independently in B according to $m(dx)/m(B)$ distribution; see Subsection 1.7. Similar construction can be found in Moyal [24], however the for stationary Poisson process on real line it is given in the Lemma by Ryll-Nardzewski [32]. We will come back to the problem of existence later.

1.5 Marczewski and Ryll-Nardzewski works in Wrocław

It was Edward Marczewski (see [8]), with collaborators Kazimierz Florek and Ryll-Nardzewski, who initiated in early 50-ties a thorough study of Poisson processes. At this time still $E = [0, \infty)$ and a Poisson process was an integer valued, right-continuous and non-decreasing stochastic process having a property of independent stationary increments. In this paper it was proposed to consider Ω as a functional space of integral valued function on $[0, \infty)$, non-decreasing and right-continuous, $\mathcal{F}$ the σ-field generated by $\{\omega(t) < y\}$. A probability measure P on $(\Omega, \mathcal{F})$ establishes a counting process. By $\Omega_1 \subset \Omega$ one denotes functions with unit jumps only. A stochastic process $\omega(t) \in \Omega$ is Poisson if $\mu(\Omega_1) = 1$, it has independent increments and the distribution of $\omega((t, t+y])$ does not depend on t.

Cramér [1] and Marczewski in [22] on the base of counting process $(\Omega, \mathcal{F}, P)$ extended the concept of number of points in intervals to number of points $N(B)$ in Borel set B via $\#\{t \in B : \omega(t) - \omega(t-) > 0\}$. Then in [22] it was noticed that for the Poisson process, $N(B)$ is a random variable.

In what follows, Ryll-Nardzewski [31] approached to the notion of point processes in a fashion similar to what we understand now, however he, similarly as others, overlooked the paper by Wiener [34]. In this paper Wiener introduced the name *Poisson chaos* or *discrete chaos*. He considered a fixed Borel subset E of finite dimensional Euclidean space and $\mathcal{E}_0$ a denumerable field of Borel subsets generating the σ-field $\mathcal{E}$ of Borel subsets of E. Then Ω it was the space of finite real valued set functions on $\mathcal{E}_0$ which are σ-additive, and $\mathbb{P}$ is a probability measure defined on a σ-field of subsets of Ω. He supposed that

(1^o) $\mathcal{E}_0 \ni B \to \omega(B)$ is $\mathbb{P}$-measurable and

(2^o) $\omega(B_1), \ldots, \omega(B_k)$ are independent, whenever B_j are disjoint sets belonging to $\mathcal{E}_0$.

Then he proved that the above conditions are fulfilled for a σ-additive set functions on $\mathcal{E}$. A point $x_0 \in E$ is called *singular* if

$$\mathbb{P}(\omega(\{x_0\}) \neq 0) > 0.$$

The number of singular points can be at most denumerable. Define $m(B) = \int \omega(B)\, d\mathbb{P}$, which today is said to be an intensity (or mean) measure. It was assumed that the point process is simple, that is $\mathbb{P}(\Omega_1) = 1$, where $\Omega_1 \subset \Omega$ consisting ω having unit atoms only. The following theorem was stated. *Suppose that m is a finite measure on $\mathcal{E}$. If $x_0 \in E$ is an atom of m (or, in other words, if x_0 is singular), then $\omega(\{x_0\})$ assumes the value 1 or 0 with probability $m(\{x_0\})$ or $1 - m(\{x_0\})$ respectively. If B contains no* *singular points, then $\omega(B)$ has the Poisson distribution*

$$\mathbb{P}(\omega(B) = k) = \frac{m(B)^k}{k!} e^{-m(B)}.$$

In the remaining part of the paper the theory was extended for the case when m is σ–finite.

Although in the paper Ryll-Nardzewski writes "process" but the term *non-homogeneous Poisson process* appears in the title.

In the paper Ryll-Nardzewski [32] by a process (i.e. point process) it was meant a random denumerable set $\boldsymbol{X}$ of real numbers, defined by a sequence $\{X_j\}$ of random variables. Introduce random variable

$$N(I) = \text{the number of indices } j \text{ such that } X_j \in I$$

for all intervals $I \subset \mathbb{R}$. Furthermore, it was assumed

(i) $\mathbb{P}(X_j \neq X_k) = 1$ for $j \neq k$,

(ii) $\mathbb{P}(\lim_j |X_j| = \infty) = 1$,

(iii) $\mathbb{P}(X_j = a) = 0$ for every j and real a.

It was said that $\boldsymbol{X}$ is a Poisson process when random variables $N(I)$ have the Poisson distribution, and if for any disjoint intervals $I_1, \ldots, I_k$, random variables $N(I_1), \ldots, N(I_k)$ were stochastically independent. It was mentioned that in [31] it was shown that

$$\mathbb{P}(N(I) = k) = \frac{m(B)^k}{k!} e^{-m(I)},$$

for some σ-finite measure m on the line. Now it was assumed that $m(I) = \lambda|I|$, that is in modern language that the Poisson process is homogeneous. It is interesting to note that Ryll-Nardzewski discovered, what Kingman [17] called later, a characterization of Poisson processes by binomial (or Bernoulli) processes. *Let us fix an interval I and denote by Q_k the event $\boldsymbol{n}(I) = k$.*

Then under the assumption Q_k, the conditional distribution of the set of k points of $\boldsymbol{X}$ belonging to I is the same as the uniform distribution of k independent point belonging to I.

Next Ryll-Nardzewski introduced the following transform, which one could call by a characteristic functional: for a complex valued function of real variable f such that $\int_{-\infty}^{\infty} |f(t)|\, dt < \infty$ he defined $\phi(f) = \mathbb{E} \prod_j (1 + f(X_j))$ and then under the assumption that $\boldsymbol{X}$ is a homogeneous Poisson process, with the use of Lemma above he computed characteristic functional

$$\phi(f) = \exp(\lambda \int_{-\infty}^{\infty} f(t)\, dt).$$

Compare with a probability generating functional defined as $\mathbb{E} \prod_j h(x_j)$ for some suitable test functions h (see [4], p. 15). With this tool Ryll-Nardzewski was ready to prove an invariance theorem of the homogeneous Poisson process.

For a sequence $\{Y_i\}$ of random variables are equivalent in the sense of de Finetti (that is if the distribution of any finite system $Y_{i_1}, Y_{i_2}, \ldots, Y_{i_k}$, where the indices i_j are different, depends only on k. the process with point $\{X_j + Y_j\}$ is again homogeneous Poisson with the same parameter λ.

1.6 Budapest probability school

Although neither mathematicians from Wrocław nor from Budapest could contant at that time with colleagues from the West, they could meet themselfes and their research on Poisson process was surely influenced by among them. For example Andraás Prekopa in [28], although did not quote Ryll-Nardzewski paper [31], in the introduction he writes: "The idea of our general method was suggested by a lecture of C. RYLL-NARDZEWSKI, who gave an elegant solution of a telephone-problem", with a footnote "This lecture was held in Wroclaw at the Colloquium on Stochastic Processes in 1953." In Prekopa paper in 1957 [28] there appeared already a contemporary definition of Poisson process on a general space E.

In Ryll-Nardzewski papers the matter of existence of a point process was not explicitely considered. This problem was taken up later by Prekopa in [26], where he considered the so called by him a stochastic set function $\xi(B)$, where B belongs to a family of subsets of space E. It is required that $\xi(B)$ has the following finite additivity property: $\mathbb{P}(\xi(A) = \bigcup_{k=1}^{l} \xi(A_k)) = 1$, whenever $A_1, \ldots, A_l$ are disjoint subsets from $\mathcal{B}_0$. In this paper, Prekopa was interested in extension of process $(\xi(B))_{B\in\mathcal{B}_0}$, where $\mathcal{B}_0$ is a field, to $(\xi(B))_{B\in\mathcal{B}}$, where $\mathcal{B}$ is the σ-field generated by $\mathcal{B}_0$ and $\xi(B)$ has the following countable additivity property: $\mathbb{P}(\xi(A) = \bigcup_{k=1}^{\infty} \xi(A_k)) = 1$, whenever $A_1, \ldots$ are disjoint subsets from $\mathcal{B}$.

It was a natural question about relations between conditions PP1 and

PP2. For example suppose that

$$\mathbb{P}(N(B) = k) = \frac{\Lambda(B)^k}{k!} e^{-\Lambda(B)},$$

is fulfilled for a family $\mathcal{E}$ of subsets of E. Is it true that $(N(B))_{B\in\mathcal{B}}$ is a Poisson process with parameter measure Λ. Alfred Renyi [30] gave the following answer to the problem for $E = \mathbb{R}$.

Suppose that Λ is a locally finite non-atomic measure on $\mathbb{R}$ and $\mathcal{B}_0$ be a field generated by finite intervals $[s, t)$. Suppose that N is a random additive set function on $\mathcal{B}_0$ such that $N(B)$ has Poisson distribution with mean $\Lambda(B)$ for all $B \in \mathcal{B}_0$. Then $N(B_1), \ldots, N(B_k)$ are independent for disjoint $B_j \in \mathcal{B}$, that is N is a Poisson process with parameter measure Λ.

The second important result from [30] deals with the question whether the so called avoidance function $B \to \mathbb{P}(N(B) = 0)$ defines the process.

In the setting of the previous theorem, if

$$\mathbb{P}(N(B) = 0) = e^{-\lambda(B)}$$

for all $B \in \mathcal{B}_0$ and if N is without multiple points, then N is a Poisson process.

It was also a question, whether it suffices to assume that E is a family of intervals $(s, t]$. The counterexample was provided by L. Shepp, and is enclosed in Goldman [11]. A recent version of this theorem in a more general setting can be found in Theorem 3.3. from Kallenberg [14].

1.7 Construction via binomial process

The following construction proves in a simple and useful way the existence of the Poisson random measure. Consider a measurable space $(E, \mathcal{E})$ with a σ-finite measure λ. Consider first the case when $\lambda(E) < \infty$.

Let $Z, X_1, X_2 \ldots$ be independent random variables on $(\Omega, calF, \mathbb{P})$, Z with Poisson distribution with mean $\Lambda(E)$ and $X_1, X_2, \ldots$ i.i.d. with the common distribution $\lambda(\cdot)/\lambda(E)$. Then a point process N defined by

$$N(B) = \sum_{j=1}^{Z} \mathbf{1}(X_j \in B),$$

is a Poisson process on E with intensity measure Λ, that is fulfilling the axioms (PP1) and (PP2) from Subsection 1.4 of the Poisson random measure on E.

In the case of not finite measure λ on E, one constructs Poisson processes $N_1, N_2, \ldots$ on $E_1, E_2, \ldots$ respectively, where $\lambda(E_j) < \infty$. Then

$$N(B) = \sum_j N_j(B \cap E_j)$$

is a Poisson process on E. fulfils the axioms (PP1) and (PP2) from Subsection 1.4

The above construction is due to Kingman [16]; see also Mecke [23], although it appeared earlier in [24]. However for a special case of a homogeneous Poisson process on $\mathbb{R}$ it is implicite in Ryll-Nardzewski [32].

Ackonwledgement The author benefited from the conversation with professor Guenter Last from Karlsruhe Institute of Technology.

References

[1] Cramér, H. (1951) A contribution to the theory of stochastic processes. *Proceedings of the second Berkeley Symposium*, 329–340.

[2] Cramér, H. (1969) Historical review of Filip Lundberg's work on risk theory. *Skand. Aktuarietidskr. Suplement*, **52**, 6–12. In Harald Cramer, Collected Works II, Springer 1994.

[3] Cramér (1994) H. Half a century with probability theory. Some personal recollections. *Ann. Probab.* **4**, 509–546. In Harald Cramer, Collected Works II, Springer.

[4] Daley, D.J. and Vere–Jones, D. (2003) An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods. Springer.

[5] Daley, D.J. and Vere–Jones, D. (2008) An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure. Springer.

[6] Erlang, A.K. (1909). Sandsynlighedsregning og Telefonsamtaler. *Nyt Tidsskrift for Matematik.* **20** (B), 33—39.

[7] Erlang, A.K. (1917) Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges (in Danish), *Electroteknikeren* **13**, 5—13.

[8] Florek, K. ; Marczewski, E. ; Ryll-Nardzewski, C. (1953) Remarks on the Poisson stochastic process. I. *Studia Math.* **13**, 122–129.

[9] Feller, W. (1940) On the integro-differential equations of purely discontinuous Markov processes. *T. Am. Math. Soc.* **48**, 488–515.

[10] Feller, W. (1949) On the theory of stochastic processes, with particular reference to applications. *Proceedings of the Berkeley Symposium*, 403–432.

[11] Goldman, J.R. (1967) Stochastic point processes: Limit theorems. *Ann. Math. Statist.* **38**, 771–779.

[12] Guttorp, P. and Thorarinsdottir, T.L. What happened to discrete chaos, the Quenouille process, and the sharp Markov property? Some history of stochastic point processes. MS 2012.

[13] Haugen, R.B. (1995) The life and work of Conny Palm – some personal comments and experiences, In VTT Symposium, volume 154, pages 207-207. Valtion Teknillien Tutkimuskeskus.

[14] Kallenberg, O. *Random Measures*, 3th ed., Akademie-Verlag, 1983.

[15] Khinchin, A. Ya. (1955) *Mathematical methods of the theory of mass service.* Trudy Mat. Inst. Steklov., 49, Acad. Sci. USSR, Moscow, 3–122.

[16] Kingman, J.F.C. Completely random measures. *Pacific J. Math.* **21**, 1967, 59–78.

[17] Kingman, J.F.C. (1993) *Poisson processes*, Claredon Press, Oxford.

[18] Kolmogoroff, A. (1931) Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung *Mathematische Annalen*, **104.1**, 415–458.

[19] Last, G., & Penrose, M. (2017). *Lectures on the Poisson Process* (Institute of Mathematical Statistics Textbooks). Cambridge University Press, 2017, doi:10.1017/9781316104477.

[20] , Lundberg, F. (1903) Approximations of the Probability Function/Reinsurance of Collective Risks. Doctoral thesis.

[21] Lundberg, O. *On random processes and their application. to sickness and accident statistics.* Uppsala: Almqvist & Wiksell, 1940.

[22] Marczewski, E. (1953) Remarks on the Poisson stochastic process (II) *Studia Mathematica* **13**, 130–136.

[23] Mecke, J. (1967) Stationäre zufällige Masse auf lokalkompakten Ableschen Gruppen. Z. Wahrscheinlichkeitheorie verw. Geb. **9**, 36–58

[24] J.E. Moyal (1962) The general theory of stochastic population processes *Acta mathematica*, **108**, 1–31.

[25] C. Palm, (1943) Intensitätsschwankungen im Fernsprechverkehr. *Ericsson Techniks*, No. 44 MR 11402

[26] Prékopa, A. (1956) On stochastic set functions I, *Acta Math. Acad. Sci. Hung.* **7**, 215–263.

[27] Prékopa, A. (1957) On Poisson and composed Poisson stochastic set functions. *Studia Math.* **16**, 142–155.

[28] Prékopa, A. (1958) A secondary processes generated by a random point distribution of Poisson type. *Ann. Univ. Sci. Budapest Sectio Math.* **1**, 153–170.

[29] Renyi, A. (1956) A characterization of Poisson processes - Magyar Tud. Akad. Mat. Kutató Int. Közl.

[30] Renyi, A. (1967) Remarks on the Poisson Process. *Studia Sci. Math. Hungar.* **2** 119–123.

[31] Ryll-Nardzewski, C. On the non-homogeneous Poisson process. I. *Studia Math.* **14** (1953), 124–128.

[32] Ryll-Nardzewski, C. (1954) Remarks on the Poisson stochastic process. III. (On a property of the homogeneous Poisson process). *Studia Math.* **14**, 314–318.

[33] Ryll-Nardzewski, C.. Remarks on processes of calls. 1961 Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II pp. 455–465 Univ. California Press, Berkeley, Calif.

[34] Wiener, N. (1938) The homogeneous chaos. *American Journal of Mathematics*, **60**, 897–936.