

Poisson processes, ordinary and compound

Onno J. Boxma and Uri Yechiali

Abstract: The Poisson process is a stochastic counting process that arises naturally in a large variety of daily-life situations. We present a few definitions of the Poisson process and discuss several properties as well as relations to some well-known probability distributions. We further briefly discuss the compound Poisson process.

Keywords: Counting process; Poisson process; exponential distribution; uniform distribution; binomial distribution; queueing system; PASTA; uniformization; compound Poisson process.

1 Introduction

Many processes in everyday life that "count" events up to a particular point in time can be accurately described by the so-called Poisson process, named after the French scientist Siméon Poisson (1781-1840; appointed as full professor at the Ecole Polytechnique, Paris, in 1806 as a successor of Fourier). An (ordinary) Poisson process is a special Markov process [*ref. to Stadge in this volume*], in continuous time, in which the only possible jumps are to the next higher state. A Poisson process may also be viewed as a counting process that has particular, desirable, properties. In this text we shall first give a few equivalent definitions of the Poisson process (Section 2). Subsequently, we describe the relation between the Poisson process and the (negative) exponential distribution (Section 3); we then show relations between the Poisson process and the uniform distribution (Section 4), and between the Poisson process and the binomial distribution (Section 5). In Section 6 we discuss a few "Poisson conservation" results that are extremely useful in, e.g., analyzing queueing networks in which customers arrive according to a Poisson process. Section 7 is devoted to the uniformization principle, a useful principle in studying the transient behavior of Markov processes. Finally, Section 8 considers a generalisation of the Poisson process, viz., the Compound Poisson process.

2 Definitions of the Poisson process

A counting process $\{C(t), t \geq 0\}$ is a stochastic process that keeps count of the number of events that have occurred up to time t . Obviously, $C(t)$

is non-negative and integer-valued for all $t \geq 0$. Furthermore, $C(t)$ is non-decreasing in t . $C(t) - C(s)$ equals the number of events in the time interval $(s, t]$, $s < t$.

$C(t)$ could, e.g., denote the number of arrivals of customers at a railway station in $(0, t]$, or the number of accidents on a particular highway in that time interval, or the number of births of animals in a particular zoo in $(0, t]$, or the number of calls to a telephone call-center during that period. A Poisson process is a counting process that has the desirable additional properties that the number of events in disjoint intervals are independent ("*independent increments*") and that the number of events in any given interval depends only on the length of that interval, and not on its particular position in time ("*stationary increments*"). In the case of the arrivals at the railway station, the stationarity assumption is clearly not fulfilled; there will be many more arrivals between 5 P.M. and 6 P.M. than between, say, 5 A.M. and 6 A.M. Still, one might wish to study the arrival process at the railway station during the rush hour. Restricting oneself to subsequent working days between 5 P.M. and 6 P.M. does allow one to use the stationary increments assumption.

Similarly, the independent increments assumption may be violated in the zoo example, but it will be a reasonably accurate representation of reality in many cases. From the viewpoint of mathematical tractability, these two properties are extremely important. We refer to Feller [3] for an extensive and lucid discussion of stochastic processes with stationary and independent increments. In the sequel we make an additional assumption, which reduces counting processes with stationary and independent increments to Poisson processes.

Definition 2.1

A Poisson process $\{N(t), t \geq 0\}$ is a counting process with the following additional properties:

- (i) $N(0) = 0$.
- (ii) The process has stationary and independent increments.
- (iii) $\mathbb{P}(N(h) = 1) = \lambda h + o(h)$ and $\mathbb{P}(N(h) \geq 2) = o(h)$, $h \downarrow 0$, for some $\lambda > 0$.

Above, the $o(h)$ symbol indicates that the ratio $\frac{\mathbb{P}(N(h) \geq 2)}{h}$ tends to zero for $h \downarrow 0$.

An equivalent definition is:

Definition 2.2

A Poisson process $\{N(t), t \geq 0\}$ is a counting process with the following additional properties:

- (i) $N(0) = 0$.
- (ii) The process has stationary and independent increments.
- (iii) $\mathbb{P}(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$, $n = 0, 1, 2, \dots$

Of course, the last property states that the number of events in any interval of length t is Poisson distributed with mean λt . λ is called the rate of the Poisson process. It readily follows that the Probability Generating Function of $N(t)$ is given by $\mathbb{E}[z^{N(t)}] = \sum_{n=0}^{\infty} z^n \mathbb{P}(N(t) = n) = e^{-\lambda(1-z)t}$. Differentiation yields $\mathbb{E}[N(t)] = \lambda t$, $\mathbb{E}[N(t)(N(t) - 1)] = (\lambda t)^2$, and hence $\text{Var}(N(t)) = \lambda t$.

The last property of Definition 2.1 may look awkward at first sight, but is insightful. It states that having two or more events in a small time interval is extremely unlikely, while the probability of a single event is approximately proportional to the length of that small interval.

Remark 2.1

A Poisson process arises naturally in large populations. Consider such a population of size n , and observe the number of events of a certain type, occurring during a unit time interval. For example, if the probability of an individual calling a call-center between, say, 9.00 and 9.01 is p , then the total number of calls during that minute is binomially distributed with parameters n and p . If n becomes large and p gets small such that $np \rightarrow \lambda > 0$, the result is a Poisson process with rate λ of calls per minute.

We close this section by giving yet another, equivalent, definition of the Poisson process.

Definition 2.3

A Poisson process $\{N(t), t \geq 0\}$ is a counting process with the following additional properties:

- (i) $N(0) = 0$.
- (ii) The only changes in the process are unit jumps upward. The intervals between jumps are independent, exponentially distributed random variables with mean $1/\lambda$, $\lambda > 0$.

Notice that Part (iii) of Definition 2.2 indeed implies that $\mathbb{P}(N(t) = 0) = e^{-\lambda t}$,

i.e., that the interval until the first event is exponentially distributed with mean $1/\lambda$, denoted $\exp(\lambda)$. The exponential distribution has the memoryless property (uniquely among all continuous distributions): If $T \sim \exp(\lambda)$, then $\mathbb{P}(T > s + t | T > s) = \mathbb{P}(T > t)$. It should also be noted that the memoryless property of the exponential distribution and the independence of successive jump intervals indeed imply that increments are stationary and independent.

Each of the above equivalent definitions has features that make it useful for deriving particular properties, as we shall see in the sequel.

3 Relation between the Poisson process and the exponential distribution

There is an intimate relation between the Poisson process and the exponential distribution, as is already being revealed by Definition 2.3. In this section we go somewhat deeper into this relation.

Let T_1 denote the time of the first event of a Poisson process and let T_n denote the time between the $(n - 1)$ st and n th event, $n = 2, 3, \dots$. Let $S_n = \sum_{i=1}^n T_i$ denote the instant of the n th event. An important observation is that

$$N(t) \geq n \iff S_n \leq t.$$

That is, the number of events during $(0, t]$ is at least n iff the time until the n th event is no larger than t . Hence

$$\mathbb{P}(N(t) \geq n) = \mathbb{P}(S_n \leq t), \quad n = 1, 2, \dots, \quad t \geq 0.$$

Definition 2.2 implies that $\mathbb{P}(N(t) \geq n) = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$, and hence

$$\mathbb{P}(S_n \leq t) = 1 - \sum_{j=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!}, \quad n = 1, 2, \dots, \quad t \geq 0. \quad (1)$$

This is the well-known result that the sum of n independent, $\exp(\lambda)$ distributed, random variables is Erlang($n; \lambda$) distributed. In particular, $T_1 = S_1$ is $\exp(\lambda)$. The density of this Erlang distribution equals $\lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$, $n = 1, 2, \dots, t > 0$, which also follows from the reasoning below:

$$\begin{aligned} \mathbb{P}(t < S_n \leq t + h) &= \mathbb{P}(N(t) = n - 1, \text{one event in } (t, t + h]) + o(h) \\ &= \mathbb{P}(N(t) = n - 1) \mathbb{P}(\text{one event in } (t, t + h]) + o(h) \\ &= e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda h + o(h). \end{aligned} \quad (2)$$

Now divide by h and let $h \downarrow 0$.

A well-known property of the exponential distribution is that it has constant failure rate equal to λ : if T is $\exp(\lambda)$ distributed, then

$$\mathbb{P}(T \leq t + h | T \geq t) = \lambda h + o(h), \quad h \downarrow 0,$$

implying that $\lim_{h \downarrow 0} \mathbb{P}(T \leq t + h | T \geq t)/h = \lambda$. We have also seen this property in Definition 2.1 of the Poisson process.

To illustrate the concept of failure rate, consider a probability distribution function $F(\cdot)$ describing the life time of a device, and denote its density by $f(\cdot)$. Then the failure rate of $F(\cdot)$ is defined as $r(t) = f(t)/[1 - F(t)]$; $r(t)dt$ gives the probability that the device fails during $(t, t + dt]$, given that it is ‘alive’ at time t .

4 Relations between the Poisson process and the uniform distribution

In this section we discuss a property of the Poisson process that often is very useful in applications. If exactly one event of a Poisson process has occurred in $(0, t]$, then the time of that occurrence is uniformly distributed on $(0, t)$. The informal explanation is that, because of the stationary and independent increments, each subinterval of equal length in $(0, t)$ has the same probability to contain that event. The formal derivation is:

$$\begin{aligned} & \mathbb{P}(T_1 \leq s | N(t) = 1) \\ = & \frac{\mathbb{P}(\text{one event in } (0, s], \text{ no event in } (s, t])}{\mathbb{P}(N(t) = 1)} \\ = & \frac{\mathbb{P}(N(s) = 1)\mathbb{P}(N(t-s) = 0)}{\mathbb{P}(N(t) = 1)} = \frac{s}{t}, \quad 0 \leq s \leq t. \end{aligned} \quad (3)$$

More generally, the following can be proved for a Poisson process: If $N(t) = n$, then the event times S_1, \dots, S_n are distributed like the order statistics of n independent random variables that are uniformly distributed on $(0, t)$.

The property that a Poisson arrival “is just as likely to occur in any interval” has proved to be extremely useful in, e.g., queueing theory. Firstly, it forms the basis of the well-known PASTA property: *Poisson Arrivals See Time Averages*. In queueing terms this property states that an outside observer, arriving to a queue according to a Poisson process, sees the system as if it were in steady state. Consider, e.g., an $M/G/s$ queue. This is a

queueing model in which arrivals occur according to a Poisson process (M), requiring a generally distributed (G) service time at one of s servers. The PASTA property implies for the $M/G/s$ queue that the number of customers seen by an arriving customer has the same distribution as the steady-state number of customers. The PASTA property was first rigorously proved by R.W. Wolff [11].

Another well-known application of the property that Poisson arrivals are uniformly distributed over an interval is provided by the $M/G/\infty$ queue [9]. This is an $M/G/s$ queue with an ample ($s = \infty$) number of servers. The $M/G/\infty$ system may be used, e.g., to model certain production or transportation systems; the customers might then be pallets moving on a conveyor belt, or cars on a road. Given that n arrivals have occurred in the $M/G/\infty$ system up to t , the above property specifies the distribution of the arrival epochs. One can then easily determine the probability that such a customer, who has arrived in $(0, t]$, is still present at t . Thus one can obtain the distribution of the number of customers $L(t)$ in the $M/G/\infty$ system at t , starting from an empty system at 0 (cf. Chapter 3 of Takács [9]). It turns out that $L(t)$ has a Poisson distribution with time-dependent rate $\lambda \int_0^t \mathbb{P}(B > y) dy$, B denoting service time. When $t \rightarrow \infty$, one gets the so-called stationary distribution of $L(t)$, which is Poissonian with rate $\lambda \mathbb{E}B$. Furthermore, the number of customers served in $(0, t]$ is also Poisson distributed, with rate $\lambda \int_0^t \mathbb{P}(B \leq y) dy$, and it is statistically independent of the process $\{L(t), t \geq 0\}$.

5 Relations between the Poisson process and the binomial distribution

Two important consequences of the stationary and independent increments properties are the following.

(i) For any $t > u \geq 0$ and integers $n \geq k$,

$$\mathbb{P}(N(u) = k | N(t) = n) = \binom{n}{k} \left(\frac{u}{t}\right)^k \left(1 - \frac{u}{t}\right)^{n-k}, \quad k = 0, 1, \dots, n.$$

That is, given that n events occurred in $(0, t]$, the probability that k of them occurred in $(0, u]$ is given by the binomial distribution with parameters n and ‘success’ probability $p = u/t$.

(ii) For two independent Poisson processes $N_1(t)$ and $N_2(t)$, having rates λ_1

and λ_2 , respectively,

$$\mathbb{P}(N_1(t) = k | N_1(t) + N_2(t) = n) = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k}, \quad k = 0, 1, \dots, n.$$

This result implies, in particular, that if two such processes ‘compete’ on which one will be the first to occur, the probability that $N_i(t)$ is ‘quicker’ is given by $\frac{\lambda_i}{\lambda_1 + \lambda_2}$, $i = 1, 2$.

See Remark 2.1 for yet another relation between the Poisson process and the binomial distribution.

6 Conservation properties

The Poisson process satisfies some conservation properties that greatly enhance its applicability. In particular, random splitting of a Poisson process results in independent Poisson processes; similarly, merging independent Poisson processes again produces a Poisson process. We now formalize these statements, without proof.

Proposition 5.1

Suppose that events occur according to a Poisson process $\{N(t), t \geq 0\}$ of rate λ . Further, suppose that each event is classified as a type i event with probability $p_i \in (0, 1)$, $i = 1, \dots, K$. Let $N_i(t)$ denote the number of type- i events in $(0, t]$. Then $\{N_1(t), t \geq 0\}, \dots, \{N_K(t), t \geq 0\}$ are independent Poisson processes with corresponding rates $\lambda p_1, \dots, \lambda p_K$.

Indeed, it is easy to see that an arbitrary interval of length h will contain a type- i event with probability $\lambda p_i h + o(h)$, $h \downarrow 0$. The independence of the various Poisson processes is less obvious. But remember that the knowledge that, say, $N_1(t) = j$ just implies that j out of the t/h intervals of length h contain a type-1 event; it does not really imply anything about the occurrence of events of other types.

Now consider the dual situation, in which there are K independent Poisson processes $\{N_1(t), t \geq 0\}, \dots, \{N_K(t), t \geq 0\}$ with corresponding rates $\lambda_1, \dots, \lambda_K$, and in which we count together all events of all these processes.

Proposition 5.2

The sum process $\{N(t), t \geq 0\}$ of K independent Poisson processes $\{N_i(t), t \geq 0\}$, $i = 1, \dots, K$, with $N(t) = \sum_{i=1}^K N_i(t)$, $t \geq 0$, is a Poisson process with rate equal to the sum $\sum_{i=1}^K \lambda_i$ of the rates of the individual processes.

This proposition can be easily proved using any of the three Definitions 2.1, 2.2 or 2.3. E.g., it is straightforward (and intuitive) to verify the following property of Definition 2.1: $\mathbb{P}(N(h) = 1) = \lambda h + o(h)$, from the similar property of the individual Poisson processes. The proposition also follows using the multiplication property of Probability Generating Functions of independent random variables and using the fact that the PGF of $N_i(t)$ equals $e^{-\lambda_i(1-z)t}$. Alternatively, Definition 2.3 also works well here; use the fact that the minimum of K independent $\exp(\lambda_i)$ distributed random variables is $\exp(\sum_{i=1}^K \lambda_i)$ distributed, combined with the memoryless property of the exponential distribution.

Networks of queues

In 1956 Paul Burke [2] proved the Output Theorem, a result that plays a crucial role in the study of networks of queues. This theorem states the following: Consider an $M/M/s$ queue, viz., a queueing system with s servers, First-Come-First-Served service operation, a Poisson(λ) arrival process of customers, and independent $\exp(\mu)$ distributed service times. Assume that $\lambda < s\mu$, implying that the process of number of customers in the system reaches an equilibrium distribution. Then the output process of the $M/M/s$ queue, counting the number of departures in $(0, t]$, is a Poisson process with rate λ . It can be shown that, when $s < \infty$, the above "conservation of Poisson flows" property only holds for the multiserver queue when service times are exponential. However, when $s = \infty$, it holds for general service times.

In 1957 Reich [7] provided an extremely elegant proof of the Output Theorem. He observed that the process constituted by the number of customers in the $M/M/s$ system is a reversible stochastic process. Viewed backward in time, this process is statistically indistinguishable from the original process. The output process of the reversed-in-time $M/M/s$ queue coincides with the Poisson input process of the original $M/M/s$ queue, and hence is also Poissonian. Using once more the reversibility property leads to the conclusion that the output process of the original $M/M/s$ queue is a Poisson process.

The above discussed properties of conservation of Poisson flows under merging, splitting and passing through an $M/M/s$ system allows one to analyze the queue length process in a network of queues Q_1, \dots, Q_K , all having external Poisson arrival processes, exponential service times, possibly multiple servers, and with Markovian routing of customers from queue to queue. Indeed, the output of an $M/M/s_i$ queue Q_i is Poisson. If a fraction p_{ij} is routed to Q_j , then the resulting flow is again Poisson; and if it merges with another,

independent, Poisson process, the resulting process is also Poisson. This is only part of the explanation why such networks of exponential multiserver queues have a joint queue length distribution that exhibits a product-form, the i th term of the product corresponding to an $M/M/s_i$ queue Q_i in isolation. For an excellent exposition of the theory of product-form queueing networks we refer to Kelly [5].

7 Uniformization

In many application areas it is important to determine the *transient* behavior of a Markov process. Consider a Markov process $\{X(t), t \geq 0\}$, with transition probabilities P_{ij} and visit times to all states being exponentially distributed with the *same* mean $1/\lambda$. Hence the number of transitions in $(0, t]$ is Poisson distributed with mean λt . Denoting the corresponding Poisson process by $\{N(t), t \geq 0\}$ we have:

$$\mathbb{P}(X(t) = j | X(0) = i) = \sum_{n=0}^{\infty} P_{ij}^{(n)} e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad (4)$$

where $P_{ij}^{(n)}$ denote the n -step transition probabilities between the states; put differently, these are the n -step transition probabilities of the discrete-time Markov chain underlying the Markov process. The above formula is derived by conditioning on the number of transitions in $(0, t]$.

Formula (4) is computationally advantageous, as the infinite sum can be truncated while the $P_{ij}^{(n)}$ can be evaluated efficiently. An interesting idea, apparently due to A. Jensen [4], allows extension of this principle to the case of *unequal* mean visit times. Let us assume that the mean visit time to state i is $1/\lambda_i$. Let λ be such that $\lambda \geq \lambda_i$ for all i . Consider a Poisson process with rate λ . Now consider a Markov process that spends $\exp(\lambda)$ in any state i and then jumps with probability $\hat{P}_{ij} = \frac{\lambda_i}{\lambda} P_{ij}$ to j and with probability $\hat{P}_{ii} = 1 - \frac{\lambda_i}{\lambda}$ back to i ; one might call this a fictitious transition. Remember that a sum of a geometrically distributed number of independent exponentially distributed random variables is again exponentially distributed; hence the sum of consecutive visit times to state i , before another state is visited, is $\exp(\lambda_i)$. A little thought now shows that this new Markov process is really the same as the original Markov process. The transient behavior of the original Markov process can hence be inferred from that of the new Markov process, by using Formula (4), with $P_{ij}^{(n)}$ being replaced by the n -step transition probabilities of the new discrete-time Markov chain that has transition probabilities \hat{P}_{ij} .

As an application of this uniformization technique, consider a system that alternates between up and down periods, up (down) periods being exponentially distributed with parameter μ_u (μ_d). We leave it to the reader to verify that by taking $\lambda = \mu_u + \mu_d$ as uniformization parameter, one gets new transition probabilities $\hat{P}_{uu} = \hat{P}_{du} = \frac{\mu_d}{\mu_u + \mu_d}$, and hence $\hat{P}_{uu}^{(n)} = \hat{P}_{du}^{(n)} = \frac{\mu_d}{\mu_u + \mu_d}$, $n = 1, 2, \dots$, leading to

$$P(X(t) = u | X(0) = u) = \frac{\mu_d}{\mu_u + \mu_d} + \frac{\mu_u}{\mu_u + \mu_d} e^{-(\mu_u + \mu_d)t},$$

and by symmetry,

$$P(X(t) = d | X(0) = d) = \frac{\mu_u}{\mu_u + \mu_d} + \frac{\mu_d}{\mu_u + \mu_d} e^{-(\mu_u + \mu_d)t}.$$

There is a large variety of other applications of the uniformization property in stochastic analysis and simulation, that may be found in various textbooks; the present example is taken from Ross [8], Section 5.8.

8 The compound Poisson process

A limitation of the Poisson process is that the jumps are always of unit size. A stochastic process $\{X(t), t \geq 0\}$ is called a *Compound Poisson Process* if it can be represented by

$$X(t) = \sum_{i=0}^{N(t)} Y_i, \quad t \geq 0,$$

where $\{N(t), t \geq 0\}$ is a Poisson process and Y_1, Y_2, \dots are independent, identically distributed random variables that are also independent of $\{N(t), t \geq 0\}$. $X(t)$ could, e.g. represent the accumulated workload input into a queueing system in $(0, t]$: Customers arrive according to a Poisson process $\{N(t), t \geq 0\}$, and the i th customer requires a service time of length Y_i . Alternatively, the Poisson arrival process might represent the number of insurance claims in $(0, t]$, while the Y_i represent independent claim sizes. $X(t)$ is then the total amount of monetary claims up to time t .

As another example one can envision a device subject to a series of independent random shocks. If the damage caused by the i -th shock is Y_i and the number of shocks in $(0, t]$ is $N(t)$, then the total accumulated damage up to time t is given by $X(t)$.

It is easily seen that, if the rate of the Poisson process equals λ and the Y_i have a common Laplace-Stieltjes transform $\beta(s) = \mathbb{E}[e^{-sY_i}]$, then

$$\mathbb{E}[e^{-sX(t)}] = e^{-\lambda(1-\beta(s))t}.$$

Differentiation then readily yields that $\mathbb{E}[X(t)] = \lambda t \mathbb{E}Y_1$ and $\text{Var}[X(t)] = \lambda t \mathbb{E}[Y_1^2]$.

Compound Poisson processes are an important subclass of Lévy processes [ref. to Kella contribution in this volume]. We refer to Bertoin [1] for a detailed account of the theory of Lévy processes.

Epilogue

The Poisson process is a stochastic counting process that arises naturally in daily-life situations in which there is a large population of individuals who, more or less independently of each other, have a small probability of contributing to the count in the next small time interval. The (compound) Poisson process has beautiful mathematical properties which make it a very powerful tool for stochastic modeling and analysis. We'd like to refer the interested reader to the monograph of Kingman [6]. Excellent textbook treatments can be found in, a.o., the books of Ross [8] and Tijms [10].

References

- [1] J. Bertoin (1996). *Lévy Processes*. Cambridge University Press.
- [2] P.J. Burke (1956). The output of a queuing system. *Oper. Res.* 4, 699-704.
- [3] W. Feller (1966). *An Introduction to Probability Theory and its Applications, Vol. II*. Wiley, New York.
- [4] A. Jensen (1953). Markoff chains as an aid in the study of Markoff processes. *Skand. Aktuarietidskr.* 36, 87-91.
- [5] F.P. Kelly (1979). *Reversibility and Stochastic Networks*. Wiley, New York.
- [6] J.F.C. Kingman (1993). *Poisson Processes*. Oxford Studies in Probability, 3. The Clarendon Press, Oxford University Press, New York.
- [7] E. Reich (1957). Waiting times when queues are in tandem. *Ann. Math. Statist.* 28, 768-773.

- [8] S.M. Ross (1983). *Stochastic Processes*. Wiley, New York.
- [9] L. Takács (1962). *Introduction to the Theory of Queues*. Oxford University Press, Oxford.
- [10] H.C. Tijms (1994). *Stochastic Models. An Algorithmic Approach*. Wiley, Chichester.
- [11] R.W. Wolff (1982). Poisson arrivals see time averages. *Oper. Res.* 30, 223-231.