Poisson Processes

Stochastics

Illés Horváth

2021/09/29

Introduction

The Poisson distribution is often used to model the number of random events during a given time period.

Introduction

The Poisson distribution is often used to model the number of random events during a given time period.

Some examples:

- the number of fires in a city in a given year;
- the number of packages arriving to an internet server in a given time interval;
- the number of cars passing on the street in a given minute;
- the number of customers entering a shop in a given time interval;
- the number of errors in a book on a given page.

Introduction

The Poisson distribution is often used to model the number of random events during a given time period.

Some examples:

- the number of fires in a city in a given year;
- the number of packages arriving to an internet server in a given time interval;
- the number of cars passing on the street in a given minute;
- the number of customers entering a shop in a given time interval;
- the number of errors in a book on a given page.

Now we are interested in not just the total number, but also the entire process.

Point processes in general

A *point process* in general is a subset of $\mathbb{R}$ (or just $[0, \infty)$). The points are often referred to as *arrivals*.

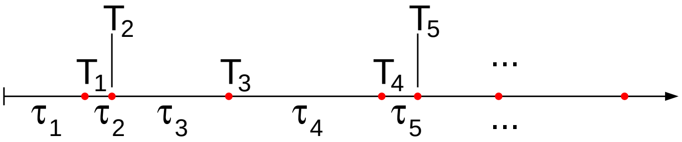
Point processes in general

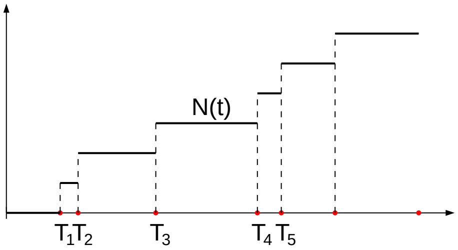
A *point process* in general is a subset of $\mathbb{R}$ (or just $[0, \infty)$). The points are often referred to as *arrivals*. A point process on $[0, \infty)$ can be described in several equivalent ways:

- by the *times of arrivals* $T_1, T_2, \ldots$, or
- by the *interarrival times* $\tau_1, \tau_2, \ldots$, or
- by the *counting process* $N(t)$, defined as the number of arrivals up to time t.

A *point process* in general is a subset of $\mathbb{R}$ (or just $[0, \infty)$). The points are often referred to as *arrivals*. A point process on $[0, \infty)$ can be described in several equivalent ways:

- by the *times of arrivals* $T_1, T_2, \ldots$, or
- by the *interarrival times* $\tau_1, \tau_2, \ldots$, or
- by the *counting process* $N(t)$, defined as the number of arrivals up to time t.





Relations between T_n, τ_n and $N(t)$ are as follows:

$$T_n = \tau_1 + \cdots + \tau_n, \qquad \tau_n = T_n - T_{n-1},$$
$$N(t) = \max\{n : T_n < t\}, \qquad T_n = \max\{t : N(t) \leq n-1\}.$$

Memoryless property

We look to define a *random point process* to model the earlier example scenarios.

We look to define a *random point process* to model the earlier example scenarios.

First question: how long do we need to wait for the first event to occur? That is, if $T_1 = \tau_1$ denotes the time we need to wait for the first event, what is the distribution of T_1?

Memoryless property

We look to define a *random point process* to model the earlier example scenarios.

First question: how long do we need to wait for the first event to occur? That is, if $T_1 = \tau_1$ denotes the time we need to wait for the first event, what is the distribution of T_1?

Consider the following. Assuming that some time t has passed *without an event*, did we get “closer” to the first event?

Memoryless property

We look to define a *random point process* to model the earlier example scenarios.

First question: how long do we need to wait for the first event to occur? That is, if $T_1 = \tau_1$ denotes the time we need to wait for the first event, what is the distribution of T_1?

Consider the following. Assuming that some time t has passed *without an event*, did we get "closer" to the first event?

For example, if we start counting cars on the street at 9:00 (corresponding to $t = 0$), and no cars pass for 1 minute, then the time we need to wait after that for the first car has the same distribution as T_1 originally, because car drivers don't care when we start counting.

Memoryless property

This can be formalized as

$$\mathbb{P}(T_1 < t + s | T_1 > t) = \mathbb{P}(T_1 < s) \qquad \forall s, t > 0,$$

or, equivalently,

$$\mathbb{P}(T_1 > t + s | T_1 > t) = \mathbb{P}(T_1 > s) \qquad \forall s, t > 0.$$

This is known as the *memoryless* property of T_1.

Memoryless property

Theorem

a) *The exponential distribution is memoryless. That is, if* $T \sim EXP(\lambda)$ *for some* $\lambda > 0$*, then*

$$\mathbb{P}(T > t + s | T > t) = \mathbb{P}(T > s) \qquad \forall s, t > 0.$$

b) *The only memoryless continuous distribution is the exponential distribution. That is, if*

$$\mathbb{P}(T > t + s | T > t) = \mathbb{P}(T > s) \qquad \forall s, t > 0$$

holds for a continuous random variable T*, then* $T \sim EXP(\lambda)$ *for some* $\lambda > 0$*.*

Memoryless property

Proof.

(a) Let $T \sim \mathrm{EXP}(\lambda)$ for some $\lambda > 0$. Its cdf is

$$F(x) = \mathbb{P}(T < x) = 1 - e^{-\lambda x},$$

so

$$\mathbb{P}(T > t) = e^{-\lambda t}.$$

Proof.

(a) Let $T \sim \mathrm{EXP}(\lambda)$ for some $\lambda > 0$. Its cdf is

$$F(x) = \mathbb{P}(T < x) = 1 - e^{-\lambda x},$$

so

$$\mathbb{P}(T > t) = e^{-\lambda t}.$$

Then

$$\mathbb{P}(T > t + s | T > t) = \frac{\mathbb{P}(T > t + s)}{\mathbb{P}(T > t)} = $$
$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbb{P}(T > s).$$

Memoryless property

(b) Proof (sketch). If T satisfies the memoryless property, then

$$\frac{\mathbb{P}(T > t + s)}{\mathbb{P}(T > t)} = \mathbb{P}(T > s).$$

Let

$$g(t) = \log \mathbb{P}(T > t).$$

(b) Proof (sketch). If T satisfies the memoryless property, then

$$\frac{\mathbb{P}(T > t + s)}{\mathbb{P}(T > t)} = \mathbb{P}(T > s).$$

Let

$$g(t) = \log \mathbb{P}(T > t).$$

Then $g(t)$ satisfies Cauchy's functional equation

$$g(t + s) = g(t) + g(s) \qquad \forall t, s > 0$$

and is bounded from above, so its solution is linear:

$$g(t) = -\lambda t, \quad \text{and} \quad \mathbb{P}(T > t) = e^{-\lambda t}.$$

Construction of the process

So now we know that $T_1 = \tau_1$ (the waiting time until the first arrival) must have exponential distribution.

Construction of the process

So now we know that $T_1 = \tau_1$ (the waiting time until the first arrival) must have exponential distribution.

After the first arrival, the time between the first and second arrival τ_2 is also memoryless, so it must also have exponential distribution, and is independent from the waiting time before the first event, and so on.

So now we know that $T_1 = \tau_1$ (the waiting time until the first arrival) must have exponential distribution.

After the first arrival, the time between the first and second arrival τ_2 is also memoryless, so it must also have exponential distribution, and is independent from the waiting time before the first event, and so on.

This gives the following definition. The *Poisson point process* (or simply Poisson process) is a random point process on $[0, \infty]$ where the $\tau_1, \tau_2, \ldots$ interarrival times are iid random variables with distribution $\mathrm{EXP}(\lambda)$ for some $\lambda > 0$.

Construction of the process

So now we know that $T_1 = \tau_1$ (the waiting time until the first arrival) must have exponential distribution.

After the first arrival, the time between the first and second arrival τ_2 is also memoryless, so it must also have exponential distribution, and is independent from the waiting time before the first event, and so on.

This gives the following definition. The *Poisson point process* (or simply Poisson process) is a random point process on $[0, \infty]$ where the $\tau_1, \tau_2, \dots$ interarrival times are iid random variables with distribution $\text{EXP}(\lambda)$ for some $\lambda > 0$.

λ is the *rate parameter* of the process. We will also use the notation $\text{PPP}(\lambda)$ for the process.

Characterization

Why is it called Poisson?

Characterization

Why is it called Poisson?

Theorem (Poisson point process characterization)

(a) *For PPP(λ), the following properties hold:*
- *the number of arrivals in an interval $[a, b]$ has distribution POI($\lambda(b-a)$), and*
- *the number of arrivals in disjoint intervals are independent random variables.*

(b) *The above two properties uniquely determine PPP(λ). That is, any point process that satisfies the above two properties has the same distribution as PPP(λ).*

Characterization

Why is it called Poisson?

Theorem (Poisson point process characterization)

(a) *For PPP(λ), the following properties hold:*
   - *the number of arrivals in an interval $[a, b]$ has distribution POI($\lambda(b-a)$), and*
   - *the number of arrivals in disjoint intervals are independent random variables.*

(b) *The above two properties uniquely determine PPP(λ). That is, any point process that satisfies the above two properties has the same distribution as PPP(λ).*

No proof.

The rate parameter

The characterization theorem states that the number of arrivals X in any interval $[a, b]$, which can be computed as $X = N(b) - N(a)$, has Poisson distribution $X \sim \mathrm{POI}(\lambda(b - a))$.

The rate parameter

The characterization theorem states that the number of arrivals X in any interval $[a, b]$, which can be computed as $X = N(b) - N(a)$, has Poisson distribution $X \sim \mathrm{POI}(\lambda(b - a))$.

The parameter $\lambda(b - a)$ is proportional both to the length of the interval $b - a$ and also to the rate parameter λ.

The rate parameter

The characterization theorem states that the number of arrivals X in any interval $[a, b]$, which can be computed as $X = N(b) - N(a)$, has Poisson distribution $X \sim \mathrm{POI}(\lambda(b-a))$.

The parameter $\lambda(b-a)$ is proportional both to the length of the interval $b-a$ and also to the rate parameter λ.

The rate λ is a density type of parameter; the process has more arrivals on average for higher values of λ. This is consistent with

$$\mathbb{E}(\text{number of arrivals in an interval of length 1}) = \lambda,$$
$$\mathbb{E}(\text{interarrival time}) = \frac{1}{\lambda}.$$

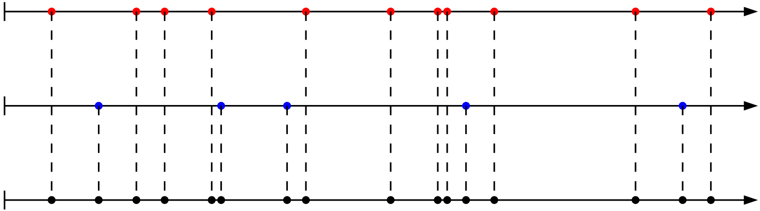
Union

Theorem (Union or superposition)

The union of an independent $PPP(\lambda_1)$ and $PPP(\lambda_2)$ is a $PPP(\lambda_1 + \lambda_2)$.

Theorem (Union or superposition)

The union of an independent PPP(λ_1) and PPP(λ_2) is a PPP($\lambda_1 + \lambda_2$).



Union

Proof. Due to the characterization theorem, the following lemma is sufficient.

Lemma

Let $X_1 \sim POI(\mu_1)$ *and* $X_2 \sim POI(\mu_2)$ *be independent. Then* $X_1 + X_2 \sim POI(\mu_1 + \mu_2)$.

Proof. Due to the characterization theorem, the following lemma is sufficient.

Lemma

Let $X_1 \sim POI(\mu_1)$ and $X_2 \sim POI(\mu_2)$ be independent. Then $X_1 + X_2 \sim POI(\mu_1 + \mu_2)$.

Proof. The probability generating function of POI(μ) is

$$G(z) = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} z^k = e^{-\mu} \sum_{k=0}^{\infty} \frac{(\mu z)^k}{k!} = e^{-\mu} e^{\mu z} = e^{\mu(z-1)},$$

so the probability generating function of $X_1 + X_2$ is

$$e^{\mu_1(z-1)} \cdot e^{\mu_2(z-1)} = e^{(\mu_1+\mu_2)(z-1)}.$$

Thinning

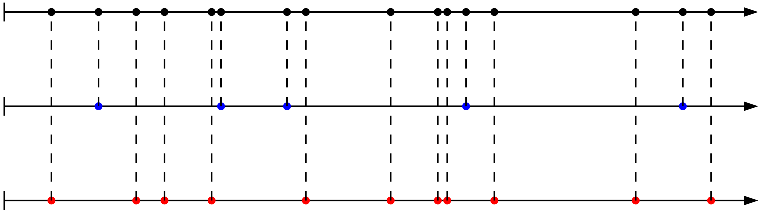
Theorem (Thinning)

For a PPP(λ), we color each arrival red with probability p and blue with probability $1 - p$, independently from other arrivals. Then the red arrivals form a PPP($p\lambda$), the blue arrivals form a PPP($(1 - p)\lambda$), and the two processes are independent from each other.

Thinning

Theorem (Thinning)

For a PPP(λ), we color each arrival red with probability p and blue with probability $1 - p$, independently from other arrivals. Then the red arrivals form a PPP($p\lambda$), the blue arrivals form a PPP($(1-p)\lambda$), and the two processes are independent from each other.



Thinning

Proof. Due to the characterization theorem, the following lemma is sufficient.

Lemma

Let $X \sim POI(\mu)$. For each of X arrivals, we color each arrival red with probability p and blue with probability $1 - p$ independently. Then the red arrivals have distribution $POI(p\mu)$, the blue arrivals have distribution $POI((1 - p)\mu)$, and the number of red and number of blue arrivals are independent.

Thinning

Proof. Due to the characterization theorem, the following lemma is sufficient.

Lemma

Let $X \sim POI(\mu)$. For each of X arrivals, we color each arrival red with probability p and blue with probability $1-p$ independently. Then the red arrivals have distribution POI($p\mu$), the blue arrivals have distribution POI($(1-p)\mu$), and the number of red and number of blue arrivals are independent.

Proof. The number of red arrivals can be obtained as a sum of X Bernoulli random variables with probability generating function $(1-p)+pz$, so the probability generating function of red arrivals is

$$e^{\mu((1-p)+pz-1)} = e^{p\mu(z-1)}.$$

Thinning

Proof. Due to the characterization theorem, the following lemma is sufficient.

Lemma

Let $X \sim POI(\mu)$. For each of X arrivals, we color each arrival red with probability p and blue with probability $1 - p$ independently. Then the red arrivals have distribution POI($p\mu$), the blue arrivals have distribution POI($(1 - p)\mu$), and the number of red and number of blue arrivals are independent.

Proof. The number of red arrivals can be obtained as a sum of X Bernoulli random variables with probability generating function $(1 - p) + pz$, so the probability generating function of red arrivals is

$$e^{\mu((1-p)+pz-1)} = e^{p\mu(z-1)}.$$

We omit the proof of independence.

Example

For example, consider a road where only cars and motorcycles pass. Let's assume that we are given the following information:

- On average 4 vehicles per minute pass, and on average 3/4 of the vehicles are cars and 1/4 of the vehicles are motorcycles.

Example

For example, consider a road where only cars and motorcycles pass. Let's assume that we are given the following information:

- On average 4 vehicles per minute pass, and on average $3/4$ of the vehicles are cars and $1/4$ of the vehicles are motorcycles.

According to the union and thinning theorem, the above information is equivalent to the following:

Example

For example, consider a road where only cars and motorcycles pass. Let's assume that we are given the following information:

- On average 4 vehicles per minute pass, and on average $3/4$ of the vehicles are cars and $1/4$ of the vehicles are motorcycles.

According to the union and thinning theorem, the above information is equivalent to the following:

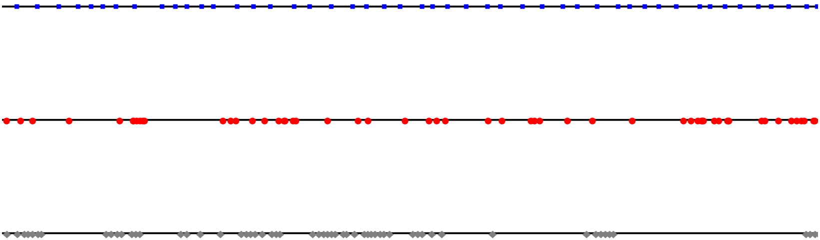
- On average 3 cars per minute pass, and on average 1 motorcycle per minute passes, and cars and motorcycles are independent.

Example

Out of the following three point processes, only one is a Poisson point process.

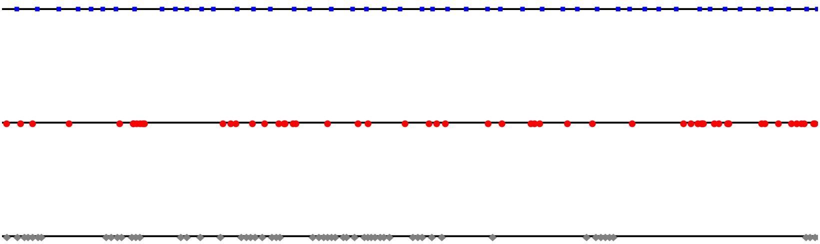
Example

Out of the following three point processes, only one is a Poisson point process.



Example

Out of the following three point processes, only one is a Poisson point process.



The red middle one is the Poisson process. For the first, blue process, interarrival times are too even, while for the last, gray process, the arrivals are too clustered but there are also some large gaps (actually, interarrival time has Pareto distribution).

Position within an interval

Let's assume that last year, only a single fire occurred in a city. What is the distribution of the time of that fire within the year?

Position within an interval

Let's assume that last year, only a single fire occurred in a city. What is the distribution of the time of that fire within the year?

Based on the above information, it could have been at any time during the year, so its distribution should be uniform within the year.

Position within an interval

Let's assume that last year, only a single fire occurred in a city. What is the distribution of the time of that fire within the year?

Based on the above information, it could have been at any time during the year, so its distribution should be uniform within the year. Indeed this is the case:

Lemma

Assuming a PPP(λ) only has a single arrival in an interval $[a, b]$, the distribution of that arrival is $U([a, b])$.

Position within an interval

Proof. Let $c \in [a, b]$ and define the following events:

- A: there is exactly one arrival in $[a, c]$,
- B: there is exactly one arrival in $[a, b]$.

Position within an interval

Proof. Let $c \in [a, b]$ and define the following events:

- A: there is exactly one arrival in $[a, c]$,
- B: there is exactly one arrival in $[a, b]$.

We want to compute $\mathbb{P}(A|B)$. We know

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)};$$

Position within an interval

Proof. Let $c \in [a, b]$ and define the following events:

- A: there is exactly one arrival in $[a, c]$,
- B: there is exactly one arrival in $[a, b]$.

We want to compute $\mathbb{P}(A|B)$. We know

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)};$$

first we compute $\mathbb{P}(B)$:

$$\mathbb{P}(B) = \mathbb{P}(1 \text{ arrival in } [a, b]) = \frac{(\lambda(b-a))^1}{1!} e^{-\lambda(b-a)}$$

since the number of arrivals in $[a, b]$ has distribution $\mathrm{POI}(\lambda(b-a))$.

Position within an interval

Next we compute $\mathbb{P}(A \text{ and } B)$:

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(1 \text{ arrival in } [a, c] \text{ and } 1 \text{ arrival in } [a, b]).$$

Position within an interval

Next we compute $\mathbb{P}(A \text{ and } B)$:

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(1 \text{ arrival in } [a, c] \text{ and } 1 \text{ arrival in } [a, b]).$$

The number of arrivals in $[a, c]$ and the number of arrivals in $[a, b]$ are not independent, because the intervals are not disjoint. However, the event that there is 1 arrival in $[a, c]$ and 1 arrival in $[a, b]$ is equivalent to saying that there is 1 arrival in $[a, c]$ and 0 arrivals in $[c, b]$, which are disjoint.

Next we compute $\mathbb{P}(A \text{ and } B)$:

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(1 \text{ arrival in } [a,c] \text{ and } 1 \text{ arrival in } [a,b]).$$

The number of arrivals in $[a,c]$ and the number of arrivals in $[a,b]$ are not independent, because the intervals are not disjoint. However, the event that there is 1 arrival in $[a,c]$ and 1 arrival in $[a,b]$ is equivalent to saying that there is 1 arrival in $[a,c]$ and 0 arrivals in $[c,b]$, which are disjoint. So

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(1 \text{ arrival in } [a,c] \text{ and } 1 \text{ arrival in } [a,b]) =$$
$$\mathbb{P}(1 \text{ arrival in } [a,c] \text{ and } 0 \text{ arrival in } [c,b]) =$$
$$\mathbb{P}(1 \text{ arrival in } [a,c]) \cdot \mathbb{P}(0 \text{ arrival in } [c,b]) =$$
$$\frac{(\lambda(c-a))^1}{1!} e^{-\lambda(c-a)} \cdot \frac{(\lambda(b-c))^0}{0!} e^{-\lambda(b-c)} = \lambda(c-a)e^{-\lambda(b-a)}.$$

Position within an interval

So we have

$$\mathbb{P}(B) = \lambda(b-a)e^{-\lambda(b-a)},$$
$$\mathbb{P}(A \text{ and } B) = \lambda(c-a)e^{-\lambda(b-a)}.$$

Position within an interval

So we have

$$\mathbb{P}(B) = \lambda(b-a)e^{-\lambda(b-a)},$$
$$\mathbb{P}(A \text{ and } B) = \lambda(c-a)e^{-\lambda(b-a)}.$$

Putting it together,

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)} = \frac{\lambda(c-a)e^{-\lambda(b-a)}}{\lambda(b-a)e^{-\lambda(b-a)}} = \frac{c-a}{b-a},$$

which is exactly equal to the probability that a $U([a,b])$ variable falls within $[a,c]$.

Actually, a more general statement is true:

Theorem

Assuming a PPP(λ) has k arrivals in an interval $[a, b]$, their joint distribution is the same as the joint distribution of k independent $U([a, b])$ random variables.

Actually, a more general statement is true:

Theorem

Assuming a PPP(λ) has k arrivals in an interval $[a, b]$, their joint distribution is the same as the joint distribution of k independent $U([a, b])$ random variables.

No proof (but similar to the previous lemma).

Summary (so far)

For a PPP(λ), the following properties hold:

- interarrival times are EXP(λ) and independent;
- the number of arrivals in an interval $[a, b]$ has distribution POI($\lambda(b - a)$);
- the number of arrivals in disjoint intervals is independent;
- the union of independent PPP's is also a PPP with the rates summed;
- a PPP can be divided into two independent PPP by flipping a coin for each arrival;
- assuming the number of points in an interval is k, their position within the interval is independent uniform.

Extension to the negative half-line

Assume we have a PPP(λ) on $[0, \infty)$ with interarrival times $\tau_1, \tau_2, \dots$.

Extension to the negative half-line

Assume we have a PPP(λ) on $[0, \infty)$ with interarrival times $\tau_1, \tau_2, \ldots$.

It can be extended to the entire real line the following way. Let $\tau_{-1}, \tau_{-2}, \ldots$ be iid EXP(λ) variables, independent from each other and from $\tau_1, \tau_2, \ldots$.
Arrivals in $[-\infty, 0]$ occur at

$$\begin{aligned} T_{-1} &= -\tau_{-1}, \\ T_{-2} &= -(\tau_{-1} + \tau_{-2}), \\ T_{-3} &= -(\tau_{-1} + \tau_{-2} + \tau_{-3}), \\ &\vdots \end{aligned}$$

Extension to the negative half-line

The arrivals cut up the real line into intervals of random length. Calculate the expected length of the interval containing 0.

Extension to the negative half-line

The arrivals cut up the real line into intervals of random length. Calculate the expected length of the interval containing 0.

The interval containing 0 is $[-\tau_{-1}, \tau_1]$, so its length is $\tau_{-1} + \tau_1$, which has expected length

$$\mathbb{E}(\tau_{-1} + \tau_1) = 2\mathbb{E}(\tau_1) = \frac{2}{\lambda}.$$

Extension to the negative half-line

The arrivals cut up the real line into intervals of random length. Calculate the expected length of the interval containing 0.

The interval containing 0 is $[-\tau_{-1}, \tau_1]$, so its length is $\tau_{-1} + \tau_1$, which has expected length

$$\mathbb{E}(\tau_{-1} + \tau_1) = 2\mathbb{E}(\tau_1) = \frac{2}{\lambda}.$$

On the other hand, for all of the other intervals, their length is simply equal to the corresponding waiting time with expectation

$$\mathbb{E}(\tau_i) = \mathbb{E}(\tau_{-i}) = \frac{1}{\lambda}.$$

Extension to the negative half-line

The arrivals cut up the real line into intervals of random length. Calculate the expected length of the interval containing 0.

The interval containing 0 is $[-\tau_{-1}, \tau_1]$, so its length is $\tau_{-1} + \tau_1$, which has expected length

$$\mathbb{E}(\tau_{-1} + \tau_1) = 2\mathbb{E}(\tau_1) = \frac{2}{\lambda}.$$

On the other hand, for all of the other intervals, their length is simply equal to the corresponding waiting time with expectation

$$\mathbb{E}(\tau_i) = \mathbb{E}(\tau_{-i}) = \frac{1}{\lambda}.$$

Is the interval containing 0 really twice as long on average than all the other intervals?

Size bias

The answer is yes! The reason is that 0 is more likely to fall into a longer interval (simply because it is longer). In other words, *the length of the interval containing 0* and *the length of an interval* have different distributions.

Size bias

The answer is yes! The reason is that 0 is more likely to fall into a longer interval (simply because it is longer). In other words, *the length of the interval containing 0* and *the length of an interval* have different distributions.

The probability that a point falls into an interval is proportional to the length of the interval.

Size bias

The answer is yes! The reason is that 0 is more likely to fall into a longer interval (simply because it is longer). In other words, *the length of the interval containing 0* and *the length of an interval* have different distributions.

The probability that a point falls into an interval is proportional to the length of the interval. So if the interarrival time τ_1 has pdf $f(x)$, then the length of the interval containing 0 (or any other given point) has pdf

$$\tilde{f}(x) = \frac{xf(x)}{\mathbb{E}(\tau_1)}.$$

(The division by $\mathbb{E}(\tau_1)$ is normalization.)

The pdf $\tilde{f}(x)$ is the *size-biased version* of $f(x)$.

Detour: size bias

Specifically for $\tau_1 \sim \mathrm{EXP}(\lambda)$,

$$f(x) = \lambda e^{-\lambda x}, \qquad \tilde{f}(x) = \lambda^2 x e^{-\lambda x}.$$

The distribution with pdf $\tilde{f}(x)$ is known as the Erlang(2,λ) distribution. Apart from being the size-biased version of EXP(λ), it also satisfies

$$\tau_1 + \tau_{-1} \sim \mathrm{Erlang}(2, \lambda)$$

for τ_1, τ_{-1} iid EXP(λ).

Examples for size bias

The next example is known as the *waiting time paradox*.

Examples for size bias

The next example is known as the *waiting time paradox*.

Buses arrive on average every 10 minutes. We go to the bus stop at a random time. On average, how long do we have to wait for the bus to arrive?

Examples for size bias

The next example is known as the *waiting time paradox*.

Buses arrive on average every 10 minutes. We go to the bus stop at a random time. On average, how long do we have to wait for the bus to arrive?

The natural answer would be 5 minutes, but this is only true if the buses arrive exactly 10 minutes after each other.

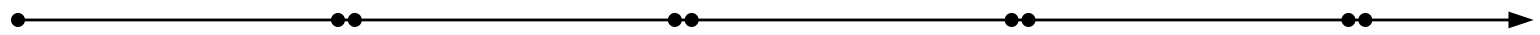
Examples for size bias

The next example is known as the *waiting time paradox*.

Buses arrive on average every 10 minutes. We go to the bus stop at a random time. On average, how long do we have to wait for the bus to arrive?

The natural answer would be 5 minutes, but this is only true if the buses arrive exactly 10 minutes after each other. If the interarrival times are random, there are shorter and longer intervals, and the time when we go to the bus stop will fall within a longer interval with higher probability.

For an even more specific example, assume that two buses arrive 1 minute after each other, then there is a 19 minute break, and this pattern is repeated:



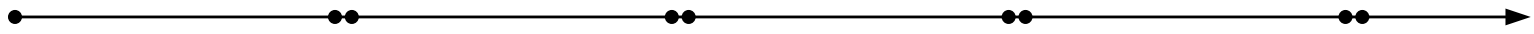
Further examples for size bias

For an even more specific example, assume that two buses arrive 1 minute after each other, then there is a 19 minute break, and this pattern is repeated:

In this setup, we will arrive at the long break with 95% probability, and have to wait 9.5 minutes on average!

Further examples for size bias

For an even more specific example, assume that two buses arrive 1 minute after each other, then there is a 19 minute break, and this pattern is repeated:



In this setup, we will arrive at the long break with 95% probability, and have to wait 9.5 minutes on average! Including the short intervals, the average (according to total expectation) is

$$0.95 \times 9.5 + 0.05 \times 0.5 = 9.05,$$

significantly more than 5 minutes.

Further examples for size bias

We also examine the time spent waiting in queues: at a bank, in a shop etc. (This is another version of the waiting time paradox.)

Further examples for size bias

We also examine the time spent waiting in queues: at a bank, in a shop etc. (This is another version of the waiting time paradox.)

It often seems like the queue you are standing in is slower than the other queues.

Further examples for size bias

We also examine the time spent waiting in queues: at a bank, in a shop etc. (This is another version of the waiting time paradox.)

It often seems like the queue you are standing in is slower than the other queues.

This is indeed the case: on average, you spend more time in slow queues than in fast queues, simply because they are slow.

Further examples for size bias

We also examine the time spent waiting in queues: at a bank, in a shop etc. (This is another version of the waiting time paradox.)

It often seems like the queue you are standing in is slower than the other queues.

This is indeed the case: on average, you spend more time in slow queues than in fast queues, simply because they are slow.

This phenomenon can also be interpreted the following way: the speed of a queue *as observed from outside* is different than the speed when observed *from within the queue*.

Further examples for size bias

We also examine the time spent waiting in queues: at a bank, in a shop etc. (This is another version of the waiting time paradox.)

It often seems like the queue you are standing in is slower than the other queues.

This is indeed the case: on average, you spend more time in slow queues than in fast queues, simply because they are slow.

This phenomenon can also be interpreted the following way: the speed of a queue *as observed from outside* is different than the speed when observed *from within the queue*.

One more example: when hiking, even if a hiker covers *the same distance* going uphill and downhill, he will spend *more time going uphill* because it is slower.

A natural way to simulate a Poisson process on $[0, t]$ is by generating the interarrival times $\tau_1, \tau_2, \ldots$ as independent EXP(λ) variables. Then the arrivals occur at

$$T_1 = \tau_1,$$
$$T_2 = \tau_1 + \tau_2,$$
$$\vdots$$

It is sufficient to generate the τ_n's until we reach $T_n \geq t$.

Simulation with POI variables

The disadvantage of the previous approach is that if we want to generate a PPP only on a small interval far away from 0, we need to generate a lot of the τ_n's just to get there.

The disadvantage of the previous approach is that if we want to generate a PPP only on a small interval far away from 0, we need to generate a lot of the τ_n's just to get there.

To generate a PPP only on an interval $[a, b]$, another option is to first generate the number of points X in $[a, b]$ according to $X \sim \mathrm{POI}(\lambda(b-a))$.

Simulation with POI variables

The disadvantage of the previous approach is that if we want to generate a PPP only on a small interval far away from 0, we need to generate a lot of the τ_n's just to get there.

To generate a PPP only on an interval $[a, b]$, another option is to first generate the number of points X in $[a, b]$ according to $X \sim \mathrm{POI}(\lambda(b - a))$.

Then we generate X points within the interval according to $\mathrm{U}([a, b])$ independently.

When is using PPP justified?

Generally, using Poisson point process to model an arrival process is justified when:

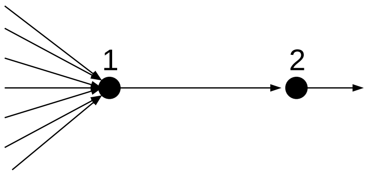
- arrivals are coming from many different sources, AND
- each individual source has a small contribution.

When is using PPP justified?

Generally, using Poisson point process to model an arrival process is justified when:

- arrivals are coming from many different sources, AND
- each individual source has a small contribution.

Example. In the network below, modeling the input of server 1 with PPP is justified, but modeling the input of server 2 with PPP is not justified.



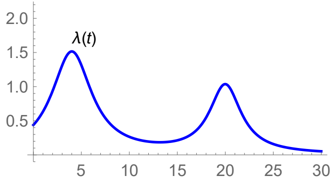
Inhomogeneous Poisson point process

In some cases, the rate of arrivals is not constant, but changes over time. One example for this is traffic: during rush hour, the arrival rate is higher, while outside rush hour it is smaller.

Inhomogeneous Poisson point process

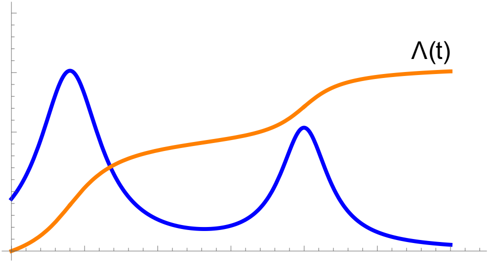
In some cases, the rate of arrivals is not constant, but changes over time. One example for this is traffic: during rush hour, the arrival rate is higher, while outside rush hour it is smaller.

We want to define a Poisson process where the arrivals are according to some nonnegative rate function $\lambda(t)$ which is changing over time.

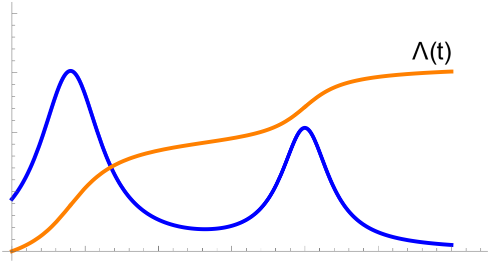


This is carried out the following way. Let

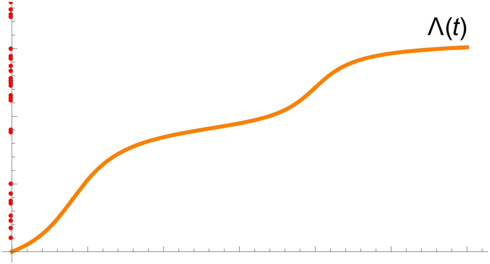
$$\Lambda(t) = \int_0^t \lambda(s)\mathrm{d}s.$$



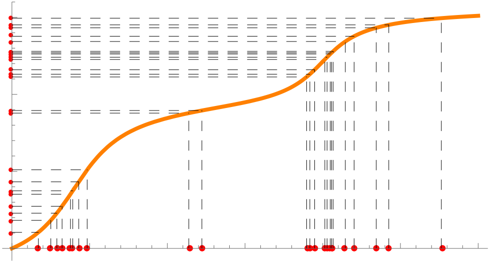
$\Lambda(t)$ is an increasing function starting at 0; when $\lambda(t)$ is large, $\Lambda(t)$ is increasing faster, and when $\lambda(t)$ is small, $\Lambda(t)$ is increasing slower. We will work with $\Lambda(t)$ from now on.



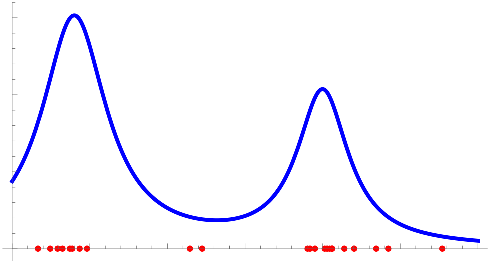
Let $Y_1, Y_2, \ldots$ be arrival times of a PPP(1) on $[0, \infty]$. (On the plot, they are shown along the Y axis.)



Then the points $\Lambda^{-1}(Y_1), \Lambda^{-1}(Y_2), \ldots$ form an inhomogeneous Poisson process. (They are mapped to the X axis via the graph of $\Lambda(t)$.)



Generally, the resulting process has more arrivals around points where $\lambda(t)$ is large and fewer arrivals around points where $\lambda(t)$ is small.



Properties of the inhomogeneous Poisson process:

Properties of the inhomogeneous Poisson process:

- ✓ the number of arrivals in $[a, b]$ has distribution $\mathrm{POI}\left(\int_a^b \lambda(t)\mathrm{d}t\right)$;

Inhomogeneous Poisson point process

Properties of the inhomogeneous Poisson process:

- ✓ the number of arrivals in $[a, b]$ has distribution $\mathrm{POI}\left(\int_a^b \lambda(t)\mathrm{d}t\right)$;
- ✓ the number of arrivals in disjoint intervals is independent;

Inhomogeneous Poisson point process

Properties of the inhomogeneous Poisson process:

- ✓ the number of arrivals in $[a, b]$ has distribution POI$\left(\int_a^b \lambda(t)\mathrm{d}t\right)$;
- ✓ the number of arrivals in disjoint intervals is independent;
- × interarrival times are no longer EXP distribution;

Inhomogeneous Poisson point process

Properties of the inhomogeneous Poisson process:

- ✓ the number of arrivals in $[a, b]$ has distribution $\mathrm{POI}\left(\int_a^b \lambda(t)\mathrm{d}t\right)$;
- ✓ the number of arrivals in disjoint intervals is independent;
- × interarrival times are no longer EXP distribution;
- × interarrival times are no longer independent.

2-dimensional PPP

The Poisson process can be generalized for higher dimensions.

2-dimensional PPP

The Poisson process can be generalized for higher dimensions.

Example. A 2-dimensional Poisson process can be used to model the position of trees within a forest.

2-dimensional PPP

The Poisson process can be generalized for higher dimensions.

Example. A 2-dimensional Poisson process can be used to model the position of trees within a forest.

In a 2-dimensional Poisson process, arrivals (points) are in $\mathbb{R}^2$. The rate of the process corresponds to the density of points (e.g. average number of points per unit area). Interarrival times no longer make sense, but simulation with POI variables still works:

2-dimensional PPP

The Poisson process can be generalized for higher dimensions.

Example. A 2-dimensional Poisson process can be used to model the position of trees within a forest.

In a 2-dimensional Poisson process, arrivals (points) are in $\mathbb{R}^2$. The rate of the process corresponds to the density of points (e.g. average number of points per unit area). Interarrival times no longer make sense, but simulation with POI variables still works:

- in a planar shape of area A, the number of points has distribution POI(λA);

2-dimensional PPP

The Poisson process can be generalized for higher dimensions.

Example. A 2-dimensional Poisson process can be used to model the position of trees within a forest.

In a 2-dimensional Poisson process, arrivals (points) are in $\mathbb{R}^2$. The rate of the process corresponds to the density of points (e.g. average number of points per unit area). Interarrival times no longer make sense, but simulation with POI variables still works:

- in a planar shape of area A, the number of points has distribution POI(λA);
- assuming k points, their joint distribution is independent uniforms;

2-dimensional PPP

The Poisson process can be generalized for higher dimensions.

Example. A 2-dimensional Poisson process can be used to model the position of trees within a forest.

In a 2-dimensional Poisson process, arrivals (points) are in $\mathbb{R}^2$. The rate of the process corresponds to the density of points (e.g. average number of points per unit area). Interarrival times no longer make sense, but simulation with POI variables still works:

- in a planar shape of area A, the number of points has distribution POI(λA);
- assuming k points, their joint distribution is independent uniforms;
- uniform distribution is easy to simulate e.g. in rectangles.

Problem 1

A call center receives an average of 8 local and 2 long-distance calls during 5 minutes.

a. What is the probability that during 2 minutes, they receive exactly 1 long-distance call?
b. What is the probability that during 2 minutes, they receive at most 3 calls in total?
c. What is the conditional probability that during 2 minutes, they receive exactly 1 long-distance call, assuming that during the same period of time, they receive at most 3 calls in total?
d. We start logging calls at $t = 0$. What is the distribution and the mean of the time of the first local call?
e. What is the distribution and the mean of the time of the first call (of any type)?
f. What is the probability that the next call is local?

Problem 1

Solution.

(a) The long-distance calls form a Poisson process with rate $2/5 = 0.4$ (calls per minute), so for the number of long-distance calls X in a 2 minute interval, we have

$$X \sim \mathrm{POI}(0.4 \times 2) = \mathrm{POI}(0.8),$$

Problem 1

Solution.

(a) The long-distance calls form a Poisson process with rate $2/5 = 0.4$ (calls per minute), so for the number of long-distance calls X in a 2 minute interval, we have

$$X \sim \mathrm{POI}(0.4 \times 2) = \mathrm{POI}(0.8),$$

and

$$\mathbb{P}(X = 1) = \frac{0.8^1}{1!} e^{-0.8} \approx 0.359.$$

Problem 1

(b) The process of all calls form a Poisson process with rate $10/5 = 2$ (calls per minute), so for the number of calls Y in a 2 minute interval, we have

$$Y \sim \mathrm{POI}(2 \times 2) = \mathrm{POI}(4),$$

Problem 1

(b) The process of all calls form a Poisson process with rate $10/5 = 2$ (calls per minute), so for the number of calls Y in a 2 minute interval, we have

$$Y \sim \mathrm{POI}(2 \times 2) = \mathrm{POI}(4),$$

and

$$\mathbb{P}(Y \leq 3) = \mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2) + \mathbb{P}(Y = 3) =$$
$$\frac{4^0}{0!}e^{-4} + \frac{4^1}{1!}e^{-4} + \frac{4^2}{2!}e^{-4} + \frac{4^3}{3!}e^{-4} \approx 0.433.$$

Problem 1

(c) Using the same notation, we need to compute

$$\mathbb{P}(X = 1 | Y \leq 3) = \frac{\mathbb{P}(X = 1, Y \leq 3)}{\mathbb{P}(Y \leq 3)}.$$

Problem 1

(c) Using the same notation, we need to compute

$$\mathbb{P}(X = 1 | Y \leq 3) = \frac{\mathbb{P}(X = 1, Y \leq 3)}{\mathbb{P}(Y \leq 3)}.$$

$\mathbb{P}(Y \leq 3)$ was already computed in part (b).

(c) Using the same notation, we need to compute

$$\mathbb{P}(X = 1 | Y \leq 3) = \frac{\mathbb{P}(X = 1, Y \leq 3)}{\mathbb{P}(Y \leq 3)}.$$

$\mathbb{P}(Y \leq 3)$ was already computed in part (b).

For $\mathbb{P}(X = 1, Y \leq 3)$, note first that X and Y are not independent, since Y also includes the X long-distance calls.

(c) Using the same notation, we need to compute

$$\mathbb{P}(X = 1 | Y \leq 3) = \frac{\mathbb{P}(X = 1, Y \leq 3)}{\mathbb{P}(Y \leq 3)}.$$

$\mathbb{P}(Y \leq 3)$ was already computed in part (b).

For $\mathbb{P}(X = 1, Y \leq 3)$, note first that X and Y are not independent, since Y also includes the X long-distance calls. Write

$$Y = X + Z,$$

where Z is the number of local calls during the same 2 minute interval. $Z \sim \mathrm{POI}(8/5 \times 2 = 3.2)$, and is independent from X.

(c) We have

$$\mathbb{P}(X=1, Y\le 3) = \mathbb{P}(X=1, Z\le 2) = \mathbb{P}(X=1)\mathbb{P}(Z\le 2) =$$
$$\frac{0.8^1}{1!}e^{-0.8}\cdot\left(\frac{3.2^0}{0!}e^{-3.2}+\frac{3.2^1}{1!}e^{-3.2}+\frac{3.2^2}{2!}e^{-3.2}\right)\approx 0.137,$$

and

$$\mathbb{P}(X=1|Y\le 3) = \frac{\mathbb{P}(X=1, Y\le 3)}{\mathbb{P}(Y\le 3)} =$$
$$\frac{\frac{0.8^1}{1!}e^{-0.8}\cdot\left(\frac{3.2^0}{0!}e^{-3.2}+\frac{3.2^1}{1!}e^{-3.2}+\frac{3.2^2}{2!}e^{-3.2}\right)}{\frac{4^0}{0!}e^{-4}+\frac{4^1}{1!}e^{-4}+\frac{4^2}{2!}e^{-4}+\frac{4^3}{3!}e^{-4}}\approx 0.315.$$

Problem 1

(d) Local calls form a PPP(8/5), so interarrival times have distribution EXP(8/5), which has mean 5/8, so the average time we need to wait for a local call is 5/8 minutes.

Problem 1

(d) Local calls form a PPP(8/5), so interarrival times have distribution EXP(8/5), which has mean 5/8, so the average time we need to wait for a local call is 5/8 minutes.

(e) The process of all calls is a PPP(10/5), so interarrival times have distribution EXP(10/5), which has mean 5/10, so the average time we need to wait for a call is 1/2 minutes.

(d) Local calls form a PPP(8/5), so interarrival times have distribution EXP(8/5), which has mean 5/8, so the average time we need to wait for a local call is 5/8 minutes.

(e) The process of all calls is a PPP(10/5), so interarrival times have distribution EXP(10/5), which has mean 5/10, so the average time we need to wait for a call is 1/2 minutes.

(f) According to the thinning theorem, each call is local with probability 8/10, independent from other calls.

Problem 2

A certain type of cookie contains on average 3 chocolate chips per cookie and 2 raisins per cookie.

a) What is the probability that a random cookie will contain exactly 2 chocolate chips?
b) What is the probability that a random cookie will contain no raisins?
c) Assuming that a cookie contains a total of 2 pieces (of either chocolate chips or raisins), what is the conditional probability that both of them are chocolate chips?
d) Joe eats half of a cookie. What is the probability that it contains at least 1 raisin?
e) Joe eats the second half of the cookie too. What is the conditional probability that the entire cookie contains at least 2 raisins, assuming that the first half contained at least 1 raisin?

Problem 2

Solution.

(a) The chocolate chips are a PPP with rate 3 (per cookie), so if X denotes the number of chocolate chips in a random cookie, then $X \sim POI(3 \times 1)$, and

$$\mathbb{P}(X = 2) = \frac{3^2}{2!}e^{-3} \approx 0.224.$$

Problem 2

Solution.

(a) The chocolate chips are a PPP with rate 3 (per cookie), so if X denotes the number of chocolate chips in a random cookie, then $X \sim POI(3 \times 1)$, and

$$\mathbb{P}(X = 2) = \frac{3^2}{2!} e^{-3} \approx 0.224.$$

(b) The raisins are a PPP with rate 2 (per cookie), so if Y denotes the number of raisins in a random cookie, then $Y \sim POI(2 \times 1)$, and

$$\mathbb{P}(Y = 0) = \frac{2^0}{0!} e^{-2} \approx 0.135.$$

(c) X is the number of chocolate chips and Y is the number of raisins in the cookie. Then we need to compute

$$\mathbb{P}(X = 2 | X + Y = 2) = \frac{\mathbb{P}(X = 2, X + Y = 2)}{\mathbb{P}(X + Y = 2)}.$$

Problem 2

(c) X is the number of chocolate chips and Y is the number of raisins in the cookie. Then we need to compute

$$\mathbb{P}(X = 2 | X + Y = 2) = \frac{\mathbb{P}(X = 2, X + Y = 2)}{\mathbb{P}(X + Y = 2)}.$$

Since the chocolate chips and raisins are independent Poisson processes, so their union is a PPP with rate $2 + 3 = 5$, so $X + Y \sim \mathrm{POI}(5)$, and

$$\mathbb{P}(X + Y = 2) = \frac{5^2}{2!} e^{-5} \approx 0.0842.$$

(c) X is the number of chocolate chips and Y is the number of raisins in the cookie. Then we need to compute

$$\mathbb{P}(X = 2 | X + Y = 2) = \frac{\mathbb{P}(X = 2, X + Y = 2)}{\mathbb{P}(X + Y = 2)}.$$

Since the chocolate chips and raisins are independent Poisson processes, so their union is a PPP with rate $2 + 3 = 5$, so $X + Y \sim \mathrm{POI}(5)$, and

$$\mathbb{P}(X + Y = 2) = \frac{5^2}{2!} e^{-5} \approx 0.0842.$$

Also,

$$\mathbb{P}(X = 2, X + Y = 2) = \mathbb{P}(X = 2, Y = 0) =$$
$$\mathbb{P}(X = 2) \cdot \mathbb{P}(Y = 0) = \frac{3^2}{2!} e^{-3} \cdot \frac{2^0}{0!} e^{-2} \approx 0.0303.$$

Problem 2

(c) Putting it together, we have

$$\mathbb{P}(X = 2 | X + Y = 2) = \frac{\frac{3^2}{2!} e^{-3} \cdot \frac{2^0}{0!} e^{-2}}{\frac{5^2}{2!} e^{-5}} = \frac{9}{25} = 0.36.$$

(c) Putting it together, we have

$$\mathbb{P}(X = 2 | X + Y = 2) = \frac{\frac{3^2}{2!} e^{-3} \cdot \frac{2^0}{0!} e^{-2}}{\frac{5^2}{2!} e^{-5}} = \frac{9}{25} = 0.36.$$

Second solution. Each piece is a chocolate chip with probability 3/5 and a raisin with probability 2/5, independently from other pieces.

(c) Putting it together, we have

$$\mathbb{P}(X = 2|X + Y = 2) = \frac{\frac{3^2}{2!}e^{-3} \cdot \frac{2^0}{0!}e^{-2}}{\frac{5^2}{2!}e^{-5}} = \frac{9}{25} = 0.36.$$

Second solution. Each piece is a chocolate chip with probability $3/5$ and a raisin with probability $2/5$, independently from other pieces.

Assuming we have 2 pieces total in a cookie, the conditional distribution of the number of chocolate chips is BIN$(2, 3/5)$, and for $X \sim \mathrm{BIN}(2, 3/5)$,

$$\mathbb{P}(X = 2) = \binom{2}{2}\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)^0 = \frac{9}{25} = 0.36.$$

(d) In a half cookie, the number of raisins is

$$X \sim \mathrm{POI}(2 \times 1/2 = 1),$$

so

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{1^0}{0!} e^{-1} \approx 0.632.$$

Problem 2

(d) In a half cookie, the number of raisins is

$$X \sim \mathrm{POI}(2 \times 1/2 = 1),$$

so

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{1^0}{0!}e^{-1} \approx 0.632.$$

(e) Define the following events:

- A: the entire cookie contains at least 2 raisins,
- B: the first half contains at least 1 raisin.

Problem 2

(d) In a half cookie, the number of raisins is

$$X \sim \mathrm{POI}(2 \times 1/2 = 1),$$

so

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{1^0}{0!}e^{-1} \approx 0.632.$$

(e) Define the following events:

- A: the entire cookie contains at least 2 raisins,
- B: the first half contains at least 1 raisin.

Then the question is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)}.$$

(d) In a half cookie, the number of raisins is

$$X \sim \mathrm{POI}(2 \times 1/2 = 1),$$

so

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{1^0}{0!} e^{-1} \approx 0.632.$$

(e) Define the following events:

- A: the entire cookie contains at least 2 raisins,
- B: the first half contains at least 1 raisin.

Then the question is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)}.$$

We have already computed $\mathbb{P}(B)$ in part (d).

Problem 2

(e) Let X_1 denote the number of raisins in the first half and X_2 denote the number of raisins in the second half. Then

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(X_1 \geq 1, X_1 + X_2 \geq 2).$$

(e) Let X_1 denote the number of raisins in the first half and X_2 denote the number of raisins in the second half. Then

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(X_1 \geq 1, X_1 + X_2 \geq 2).$$

X_1 and X_2 are independent POI(1/2). The event A and B can be divided into 2 disjoint events:

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(X_1 \geq 1, X_2 \geq 2) =$$
$$\mathbb{P}(X_1 = 1 \text{ and } X_2 \geq 1) + \mathbb{P}(X_1 \geq 2) =$$
$$\mathbb{P}(X_1 = 1)\mathbb{P}(X_2 \geq 1) + \mathbb{P}(X_1 \geq 2).$$

Problem 2

(e)

$$\mathbb{P}(X_1 = 1)\mathbb{P}(X_2 \geq 1) + \mathbb{P}(X_1 \geq 2) =$$
$$\frac{(1/2)^1}{1!} e^{-1/2} \left(1 - \frac{(1/2)^0}{0!} e^{-1/2}\right) +$$
$$\left(1 - \frac{(1/2)^0}{0!} e^{-1/2} - \frac{(1/2)^1}{1!} e^{-1/2}\right) \approx 0.210,$$

and

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)} \approx 0.331.$$

Problem 4

Two types of jobs arrive at a server: type A and type B. On average, the arrival rates are 1 job/second for type A and 2 jobs/second for type B.

a) What is the probability that the first job arriving is of type A?
b) What is the distribution of the waiting time before the first arrival?
c) What is the distribution of the number of type B jobs that arrive before the first type A job?

Problem 4

Two types of jobs arrive at a server: type A and type B. On average, the arrival rates are 1 job/second for type A and 2 jobs/second for type B.

a. What is the probability that the first job arriving is of type A?
b. What is the distribution of the waiting time before the first arrival?
c. What is the distribution of the number of type B jobs that arrive before the first type A job?

Solution.

(a) 1/3.

Problem 4

(b) Since all jobs together form a PPP with rate 3 jobs/second, the waiting time before the first arrival has distribution EXP(3).

Problem 4

(b) Since all jobs together form a PPP with rate 3 jobs/second, the waiting time before the first arrival has distribution EXP(3).

(c) Consider the following: each job will be type A with probability 1/3 and type B with probability 2/3. So when counting the number of type A jobs before the first type B job, we are basically counting the number of unsuccessful trials (A) before the first success (B).

Problem 4

(b) Since all jobs together form a PPP with rate 3 jobs/second, the waiting time before the first arrival has distribution EXP(3).

(c) Consider the following: each job will be type A with probability 1/3 and type B with probability 2/3. So when counting the number of type A jobs before the first type B job, we are basically counting the number of unsuccessful trials (A) before the first success (B).

This is exactly how PGEO(1/3) is defined.

Problem 6

A 200 page manuscript contains on average 3 typos (errors) per page. During proofreading, 90% of the typos are found and removed.

1. The manuscript is 200 pages long. What is the average number of typos remaining in the manuscript *after* proofreading?
2. What is the probability that a page originally contained 2 typos, and both are found during proofreading?
3. What is the probability that on the first page, all typos are found during proofreading?

Problem 6

Solution.

(a) The total number of errors is a PPP with rate 3 (errors per page), while the remaining errors are a PPP with rate $3 \times 0.1 = 0.3$ due to thinning.

Problem 6

Solution.

(a) The total number of errors is a PPP with rate 3 (errors per page), while the remaining errors are a PPP with rate $3 \times 0.1 = 0.3$ due to thinning.

Then the total number of remaining erros on 200 pages has distribution $\mathrm{POI}(200 \times 0.3 = 60)$, which has mean 60.

Problem 6

Solution.

(a) The total number of errors is a PPP with rate 3 (errors per page), while the remaining errors are a PPP with rate $3 \times 0.1 = 0.3$ due to thinning.

Then the total number of remaining erros on 200 pages has distribution POI$(200 \times 0.3 = 60)$, which has mean 60.

(b) Let Z denote the total number of errors on a page;

$$Z = X + Y,$$

where X is the number of found errors and Y is the number of remaining errors on that page.

Solution.

(a) The total number of errors is a PPP with rate 3 (errors per page), while the remaining errors are a PPP with rate $3 \times 0.1 = 0.3$ due to thinning.

Then the total number of remaining erros on 200 pages has distribution POI($200 \times 0.3 = 60$), which has mean 60.

(b) Let Z denote the total number of errors on a page;

$$Z = X + Y,$$

where X is the number of found errors and Y is the number of remaining errors on that page. Then X and Y are independent, $X \sim \mathrm{POI}(2.7)$ and $Y \sim \mathrm{POI}(0.3)$.

Problem 6

(b) The question is $\mathbb{P}(X = Z = 2)$.

Mind the wording; this is not a conditional probability.

(b) The question is $\mathbb{P}(X = Z = 2)$.

Mind the wording; this is not a conditional probability. It can be written equivalently as

$$\mathbb{P}(X = Z = 2) = \mathbb{P}(X = 2, Y = 0) = \mathbb{P}(X = 2)\mathbb{P}(Y = 0) = \frac{2.7^2}{2!}e^{-2.7} \cdot \frac{0.3^0}{0!}e^{-0.3} \approx 0.181.$$

Problem 6

(b) The question is $\mathbb{P}(X = Z = 2)$.

Mind the wording; this is not a conditional probability. It can be written equivalently as

$$\mathbb{P}(X = Z = 2) = \mathbb{P}(X = 2, Y = 0) = \mathbb{P}(X = 2)\mathbb{P}(Y = 0) = \frac{2.7^2}{2!}e^{-2.7} \cdot \frac{0.3^0}{0!}e^{-0.3} \approx 0.181.$$

(c) All typos found is equivalent to 0 errors remaining. The number of remaining errors is $Y \sim \mathrm{POI}(0.3)$, and

$$\mathbb{P}(Y = 0) = \frac{0.3^0}{0!}e^{-0.3} \approx 0.741.$$

Problem 8

In a forest, there are on average 10 trees per $100\mathrm{m}^2$. Let us assume that each tree has diameter 20 cm on the ground level. (Ignore the possibility that they may overlap.)

- (a) What is the probability that there are no trees on a given $10m^2$ area?
- (b) We fire a bullet in a random direction from the middle of the forest. What is the probability that the bullet will fly at least 50 meters before it hits a tree?

Problem 8

In a forest, there are on average 10 trees per $100\mathrm{m}^2$. Let us assume that each tree has diameter 20 cm on the ground level. (Ignore the possibility that they may overlap.)

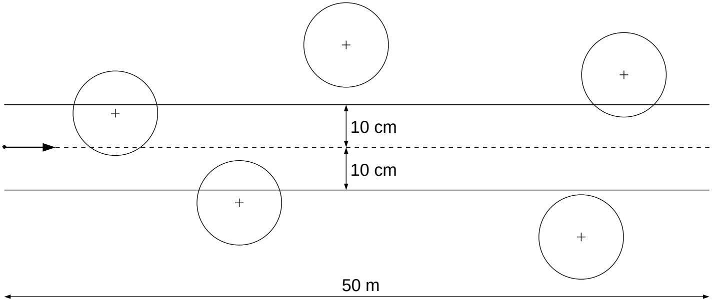
a. What is the probability that there are no trees on a given $10m^2$ area?
b. We fire a bullet in a random direction from the middle of the forest. What is the probability that the bullet will fly at least 50 meters before it hits a tree?

Solution.

(a) The center points of the trees form a 2-dimensional Poisson process with rate $10/100 = 0.1$ trees per m^2, so the number of trees in a given $10\mathrm{m}^2$ area is $X \sim \mathrm{POI}(10\mathrm{m}^2 \times 0.1\frac{1}{\mathrm{m}^2} = 1)$, and

$$\mathbb{P}(X = 0) = \frac{1^0}{0!}e^{-1} \approx 0.368.$$

(b) The bullet will hit a tree if the center of the tree is closer to the path of the bullet than 10 cm, because the trees have radius 10 cm.



Problem 8

(b) This forms a strip of area $50\mathrm{m} \times 20\mathrm{cm}$ around the path of the bullet. The bullet will fly at least 50 meters exactly when this strip contains 0 tree center points.

(b) This forms a strip of area $50\mathrm{m} \times 20\mathrm{cm}$ around the path of the bullet. The bullet will fly at least 50 meters exactly when this strip contains 0 tree center points.

This strip has area $50\mathrm{m} \times 20\mathrm{cm} = 10\mathrm{m}^2$, so the number of trees in this areas is $X \sim \mathrm{POI}(10\mathrm{m}^2 \times 0.1\frac{1}{\mathrm{m}^2} = 1)$, and

$$\mathbb{P}(X = 0) = \frac{1^0}{0!}e^{-1} \approx 0.368$$

(as already computed in part (a)).

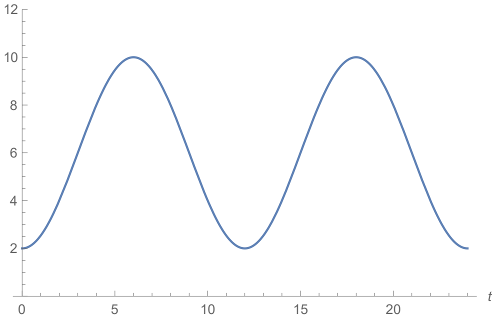
We count trucks on a road. Truck traffic is inhomogeneous, the density of trucks during the day has rate function (number of trucks per hour):

$$\lambda(t) = 6 - 4\cos\left(\frac{\pi}{6}t\right) \qquad t \in [0, 24]$$

1. Plot the rate function. At what time are the maximum points?
2. What is the average number of trucks passing through the road during one day?
3. What is the probability that during 12:00 and 13:00, exactly 3 trucks pass?

Solution.

(a) This is an inhomogeneous Poisson process. The rate function changes between 2 and 10, with maximums at $t = 6$ and $t = 18$.



Problem 9

(b) The total number of trucks passing in one day is

$$X \sim \mathrm{POI}\left(\int_0^{24} \lambda(t)\mathrm{d}t = 144\right),$$

which has mean 144, so that's the average number of trucks during one day.

Problem 9

(b) The total number of trucks passing in one day is

$$X \sim \mathrm{POI}\left(\int_0^{24} \lambda(t)\mathrm{d}t = 144\right),$$

which has mean 144, so that's the average number of trucks during one day.

(b) The number of trucks passing between 12:00 and 13:00 is

$$Y \sim \mathrm{POI}\left(\int_{12}^{13} \lambda(t)\mathrm{d}t = 4 - \frac{3}{\pi} \approx 3.045\right),$$

and

$$\mathbb{P}(Y = 3) = \frac{\left(4 - \frac{3}{\pi}\right)^3}{3!} e^{4 - \frac{3}{\pi}} \approx 0.224.$$