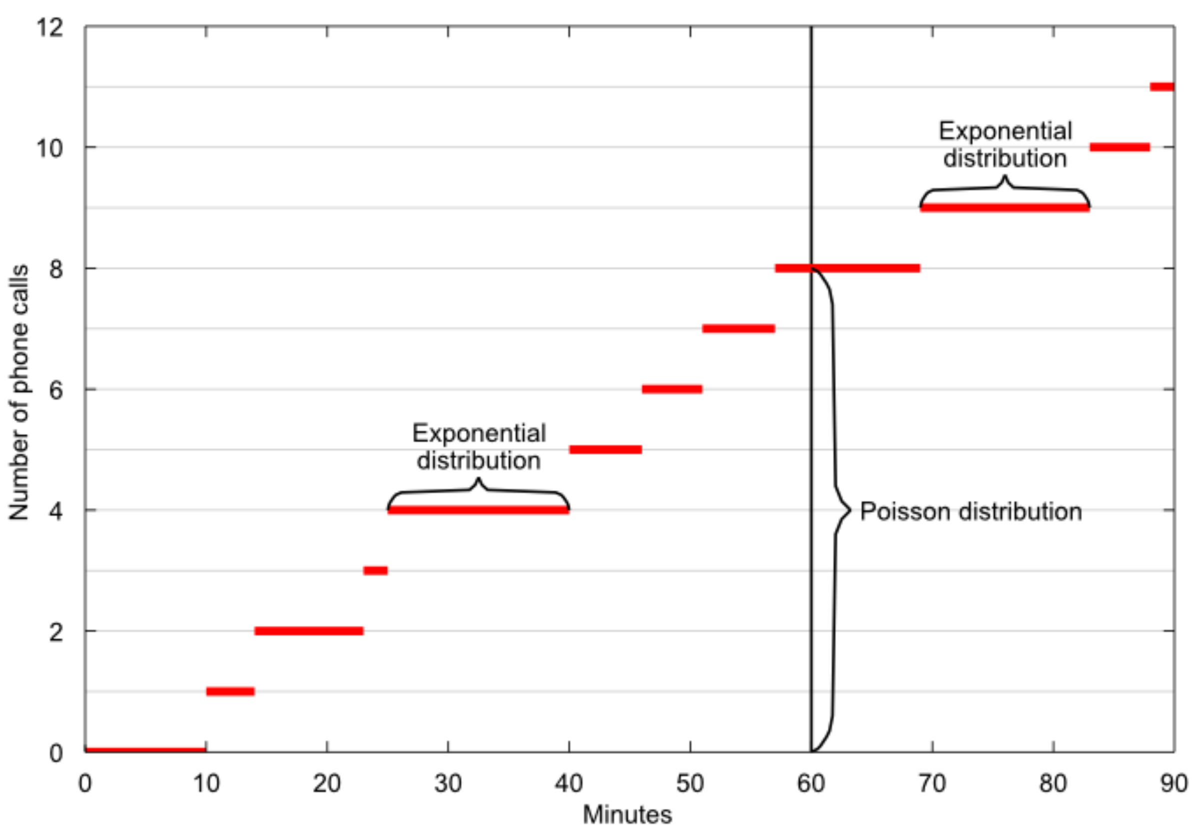
Poisson distribution

The Poisson distribution is related to the exponential distribution. Suppose an event can occur several times within a given unit of time. When the total number of occurrences of the event is unknown, we can think of it as a random variable. This random variable has a Poisson distribution if the time elapsed between two successive occurrences of the event has an exponential distribution and it is independent of previous occurrences.

A classical example of a random variable having a Poisson distribution is the number of phone calls received by a call center. If the time elapsed between two successive phone calls has an exponential distribution and it is independent of the time of arrival of the previous calls, then the total number of calls received in one hour has a Poisson distribution.



The concept is illustrated by the plot above, where the number of phone calls received is plotted as a function of time. The graph of the function makes an upward jump each time a phone call arrives. The time elapsed between two successive phone calls is equal to the length of each horizontal segment and it has an exponential distribution. The number of calls received in 60 minutes is equal to the length of the segment highlighted by the vertical curly brace and it has a Poisson distribution.

The following sections provide a more formal treatment of the main characteristics of the Poisson distribution.

Definition

A Poisson random variable is characterized as follows.

Definition Let be a discrete random variable. Let its support be the set of non-negative integer numbers:

Let . We say that has a **Poisson distribution** with parameter if its probability mass function is

where is the factorial of .

Relation to the exponential distribution

The relation between the Poisson distribution and the exponential distribution is summarized by the following proposition.

Proposition The number of occurrences of an event within a unit of time has a Poisson distribution with parameter if the time elapsed between two successive occurrences of the event has an exponential distribution with parameter and it is independent of previous occurrences.

Proof

Denote by the number of occurrences of the event and by

τ_1 the time elapsed before the first event occurs
τ_2 the time elapsed between the first and the second occurrence of the event
$\vdots$
τ_n the time elapsed between the $(n-1)$-th and the n-th occurrence of the event
$\vdots$

Note that there are at least occurrences of the event (i.e.,) within a unit of time if and only if the sum of the times elapsed between the occurrences is less than one unit of time. In other words, the events

and

coincide. Therefore,

for any . Thus, the distribution of can be derived from the distribution of the waiting times . We are going to prove that the assumption that the waiting times are exponential implies that has a Poisson distribution. Denote by the sum of waiting times:

Since the sum of independent exponential random variables with common parameter is a Gamma random variable with parameters and , then is a Gamma random variable with parameters and , i.e., its probability density function is

where

and the last equality stems from the fact that we are considering only integer values of . We need to integrate the density function to compute the probability that is less than :

The last integral can be computed integrating by parts times:

$$
\begin{aligned}
&\int_0^1 z^{x-1}\exp(-\lambda z)dz \\
&= \left[-\frac{1}{\lambda}z^{x-1}\exp(-\lambda z)\right]_0^1 + \int_0^1 (x-1)z^{x-2}\frac{1}{\lambda}\exp(-\lambda z)dz \\
&= -\frac{1}{\lambda}\exp(-\lambda) + (x-1)\frac{1}{\lambda}\int_0^1 z^{x-2}\exp(-\lambda z)dz \\
&= -\frac{1}{\lambda}\exp(-\lambda) + (x-1)\frac{1}{\lambda}\left\{\left[-\frac{1}{\lambda}z^{x-2}\exp(-\lambda z)\right]_0^1 + \int_0^1 (x-2)z^{x-3}\frac{1}{\lambda}\exp(-\lambda z)dz\right\} \\
&= -\frac{1}{\lambda}\exp(-\lambda) - (x-1)\frac{1}{\lambda^2}\exp(-\lambda) + (x-1)(x-2)\frac{1}{\lambda^2}\int_0^1 z^{x-3}\exp(-\lambda z)dz \\
&= \ldots \\
&= -\sum_{i=1}^{x-1}\frac{(x-1)!}{(x-i)!}\frac{1}{\lambda^i}\exp(-\lambda) + \frac{(x-1)!}{1}\frac{1}{\lambda^{x-1}}\int_0^1 \exp(-\lambda z)dz \\
&= -\sum_{i=1}^{x-1}\frac{(x-1)!}{(x-i)!}\frac{1}{\lambda^i}\exp(-\lambda) + \frac{(x-1)!}{\lambda^{x-1}}\left[-\frac{1}{\lambda}\exp(-\lambda z)\right]_0^1 \\
&= -\sum_{i=1}^{x-1}\frac{(x-1)!}{(x-i)!}\frac{1}{\lambda^i}\exp(-\lambda) - \frac{(x-1)!}{\lambda^x}\exp(-\lambda) + \frac{(x-1)!}{\lambda^x}
\end{aligned}
$$

Multiplying by , we obtain

Thus, we have obtained

$$
P(X \geq x) = P(\tau_1 + \ldots + \tau_x \leq 1) = 1 - \sum_{j=0}^{x-1}\frac{\lambda^j}{j!}\exp(-\lambda)
$$

But this is exactly what we get when has a Poisson distribution:

Expected value

The expected value of a Poisson random variable is

Proof

It can be derived as follows:

$$
\begin{aligned}
& \mathrm{E}[X] \\
&= \sum_{x \in R_X} x p_X(x) \\
&= \sum_{x=0}^{\infty} x \exp(-\lambda) \frac{1}{x!} \lambda^x \\
&= 0 + \sum_{x=1}^{\infty} x \exp(-\lambda) \frac{1}{x!} \lambda^x && \text{(the first term of the sum is zero since } x = 0\text{)} \\
&= \sum_{y=0}^{\infty} (y+1) \exp(-\lambda) \frac{1}{(y+1)!} \lambda^{y+1} && \text{(by changing variable: } y = x - 1\text{)} \\
&= \sum_{y=0}^{\infty} (y+1) \exp(-\lambda) \frac{1}{(y+1)y!} \lambda \lambda^y && \text{(since } (y+1)! = (y+1)y!\text{)} \\
&= \lambda \sum_{y=0}^{\infty} \exp(-\lambda) \frac{1}{y!} \lambda^y \\
&= \lambda \sum_{y=0}^{\infty} p_Y(y) && (p_Y \text{ is the pmf of a Poisson r.v. with parameter } \lambda) \\
&= \lambda && \text{(the sum of a pmf over its support is 1)}
\end{aligned}
$$

Variance

The variance of a Poisson random variable is

Proof

Moment generating function

The moment generating function of a Poisson random variable is defined for any :

Proof

Characteristic function

The characteristic function of a Poisson random variable is

Proof

Distribution function

The distribution function of a Poisson random variable is

where is the floor of , i.e. the largest integer not greater than .

Proof

Values of are usually computed by computer algorithms. For example, the MATLAB command:

```
poisscdf(x,lambda)
```

returns the value of the distribution function at the point `x` when the parameter of the distribution is equal to `lambda`.

Solved exercises

Below you can find some exercises with explained solutions.

Exercise 1

The time elapsed between the arrival of a customer at a shop and the arrival of the next customer has an exponential distribution with expected value equal to 15 minutes. Furthermore, it is independent of previous arrivals. What is the probability that more than 6 customers arrive at the shop during the next hour?

Solution

If a random variable has an exponential distribution with parameter , then its expected value is equal to . Here

Therefore, . If inter-arrival times are independent exponential random variables with parameter , then the number of arrivals during a unit of time has a Poisson distribution with parameter . Thus, the number of customers that will arrive at the shop during the next hour (denote it by) is a Poisson random variable with parameter . The probability that more than 6 customers arrive at the shop during the next hour is

$$\begin{aligned} P(X > 6) &= 1 - P(X \leq 6) \\ &= 1 - F_X(6) \qquad \text{(where } F_X \text{ is the distribution function of } X\text{)} \\ &= 1 - \exp(-4) \sum_{s=0}^{6} \frac{4^s}{s!} \simeq 0.1107 \end{aligned}$$

and the value of can be calculated with a computer algorithm, for example with the MATLAB command

Exercise 2

At a call center, the time elapsed between the arrival of a phone call and the arrival of the next phone call has an exponential distribution with expected value equal to 15 seconds. Furthermore, it is independent of previous arrivals. What is the probability that less than 50 phone calls arrive during the next 15 minutes?

Solution

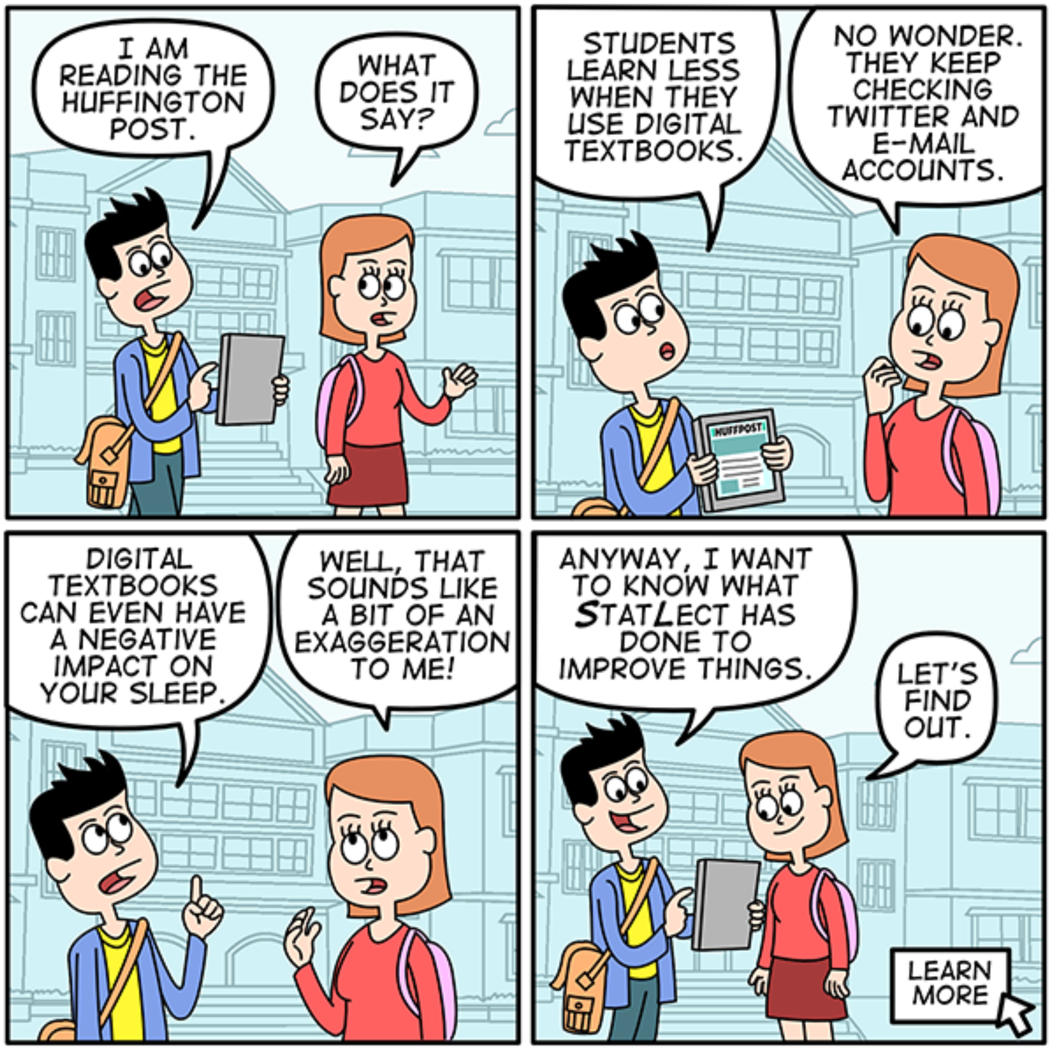
If a random variable has an exponential distribution with parameter , then its expected value is equal to . Here

where, in the last equality, we have taken 15 minutes as the unit of time. Therefore, . If inter-arrival times are independent exponential random variables with parameter , then the number of arrivals during a unit of time has a Poisson distribution with parameter . Thus, the number of phone calls that will arrive during the next 15 minutes (denote it by) is a Poisson random variable with parameter . The probability that less than 50 phone calls arrive during the next 15 minutes is

$$
\begin{aligned}
\mathrm{P}(X < 50) &= \mathrm{P}(X \leq 49) \\
&= F_X(49) \qquad \text{(where } F_X \text{ is the distribution function of } X\text{)} \\
&= \exp(-60) \sum_{s=0}^{49} \frac{60^s}{s!} \simeq 0.0844
\end{aligned}
$$

and the value of can be calculated with a computer algorithm, for example with the MATLAB command

```
poisscdf(49,60)
```



The book

Most of the learning materials found on this website are now available in a traditional textbook format.

LEARN MORE