STAT 515 Lec 8

The Poisson and Exponential Distributions

Karl Gregory

The Poisson distribution

Suppose X is the number of occurrences per unit of time or space of an event, where the occurrences

1. are independent
2. occur randomly but at a constant rate over the entire time/space.

We often treat the mechanism which generates the counts X over a span of time or region in space under a mathematical model called a *Poisson process*. We have in mind random variables like those in the following examples:

Example. Let X be the number of customers entering a store in an hour.

Example. Let X be the number of earthquakes per decade in a region.

Example. Let X be the number of weeds growing per square meter of a field.

Example. Let X be the number of bird nests per acre in a habitat.

For such random variables as in these examples, which are counts that could take the values $0, 1, 2, \ldots,$ we often posit a probability distribution called the Poisson distribution, which was first suggested by the subject of the portrait in Figure 1: Siméon Denis Poisson.

Definition: Poisson distribution

The *Poisson distribution with mean* λ is the probability distribution with probability mass function given by

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!},$$

where $\lambda > 0$. If X is a random variable with this distribution, we write $X \sim \text{Poisson}(\lambda)$.

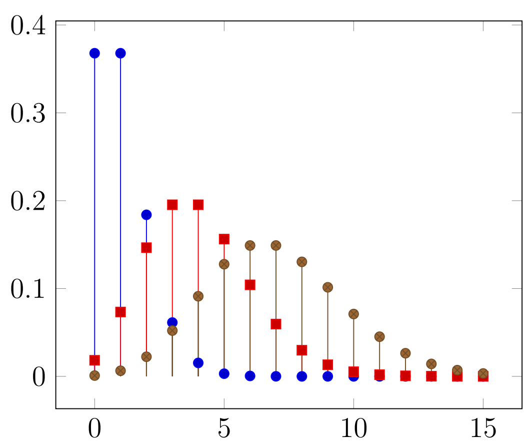
Figure 1: Siméon Denis Poisson (1781 - 1840)

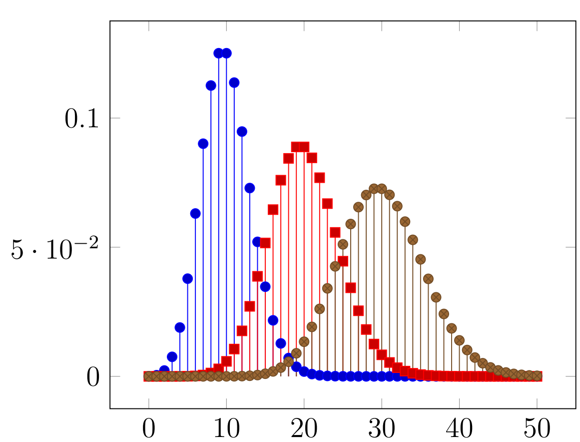
Result: Mean and variance of the Poisson distribution

If $X \sim \text{Poisson}(\lambda)$, then

$$\mathbb{E}X = \lambda \quad \text{and} \quad \operatorname{Var} X = \lambda.$$

The values of the probability mass function of the Poisson(λ) distribution are plotted below for λ taking the values 1, 4, and 7 in the first plot and the values 10, 20, and 30. in the second plot. Note that the distributions are "centered" at these λ values in the sense that if they were sitting on a see-saw with a fulcrum positioned at λ, the see-saw would not tip in either direction.





Exercise. Suppose the number of car accidents on any given day in a mid-sized city follows a Poisson distribution with a mean of 20 accidents per day.

1. Find the probability that there are exactly 10 accidents on a given day.

Answer: Letting X be the number of accidents on a given day, we have $X \sim \text{Poisson}(\lambda)$ with $\lambda = 20$. So the probability that there are exactly 10 accidents on a given day is

$$P(X = 10) = \frac{(20)^{10} e^{-20}}{10!} = \texttt{dpois(x = 10,lambda = 20)} = 0.005816307.$$

2. Find the probability that there are 12 or more accidents on a given day.

   Answer: We have

$$\begin{aligned} P(X \geq 12) &= 1 - P(X \leq 11) \\ &= 1 - [P(X = 0) + P(X = 1) + \cdots + P(X = 11)] \\ &= 1 - \sum_{x=0}^{11} \frac{(20)^x e^{-20}}{x!} \\ &= \texttt{1 - ppois(q=11,lambda=20)} \\ &= 0.9786132. \end{aligned}$$

Keeping in mind that our random variable X represents a count per unit time or space, we may suppose that if we make our counts over larger units of time or space, the counts will tend to be greater. A hallmark of Poisson processes is that the distribution of the counts within a unit of time or space scales with the size of the unit. That is, we have the following:

Result: Scaling the Poisson time/space interval

Suppose $X \sim \text{Poisson}(\lambda)$ comes from a Poisson process such that X is the number of occurrences per unit of time or space of an event. Then if Y is the number of occurrences of the event per t units of time or space, we have $Y \sim \text{Poisson}(t\lambda)$.

Exercise. Suppose the number of car accidents on any given day in a mid-sized city follows a Poisson distribution with a mean of 20 accidents per day. What is the probability that there are no more than 130 car accidents in a given week?

Answer: Let $X \sim \text{Poisson}(\lambda)$, with $\lambda = 20$, be the number of car accidents on a given day and let Y be the number of car accidents in a given week. Then $Y \sim \text{Poisson}(140)$, since $7(20) = 140$. We have

$$\begin{aligned} P(Y \leq 130) &= P(Y = 0) + P(Y = 1) + \cdots + P(Y = 130) \\ &= \sum_{y=0}^{130} \frac{(140)^y e^{-140}}{y!} \\ &= \texttt{ppois(q=130,lambda=140)} \\ &= 0.2124409. \end{aligned}$$

The Exponential distribution

Suppose X comes from a Poisson process where the expected number of occurrences per unit of time or space is λ. Let Y be the time between any two occurrences of the event or the time until the first occurrence. How might we get the probability density function of the random variable Y? Begin by writing, for some time y into the future.

$$P(Y > y) = P(\text{"no occurrences before time } y\text{"}).$$

Now,

- In 1 unit of time or space, we "expect" λ occurrences.
- In y units of time or space, we "expect" $y\lambda$ occurrences.

Here "expect" is in the sense of expected value. Following this logic, we can see that the number of occurrences in y units of time or space follows a Poisson distribution with mean $y\lambda$. Thus we may write

$$P(Y > y) = P(\text{"no occurrences before time } y\text{"}) = \frac{(y\lambda)^0 e^{-y\lambda}}{0!} = e^{-y\lambda}.$$

From this we get

$$P(Y \leq y) = P(\text{"first occurrence is before time } y\text{"}) = 1 - P(Y > y) = 1 - e^{-y\lambda}.$$

Recalling the definition of the cumulative distribution function (cdf), we see that we have derived the cdf F of the random variable Y as

$$F(y) = 1 - e^{-y\lambda}. \tag{1}$$

By its definition, the probability density function f of the random variable Y must satisfy

$$F(y) = \int_{-\infty}^{y} f(t)dt \quad \text{for all } y.$$

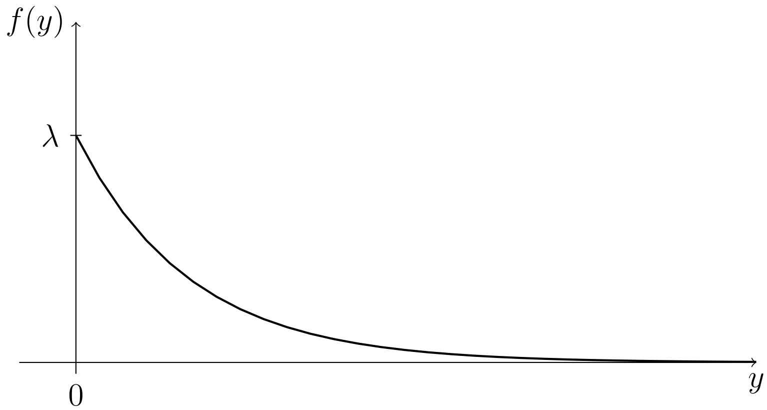
Using calculus, we find that f must be the function

$$f(y) = \lambda e^{-y\lambda},$$

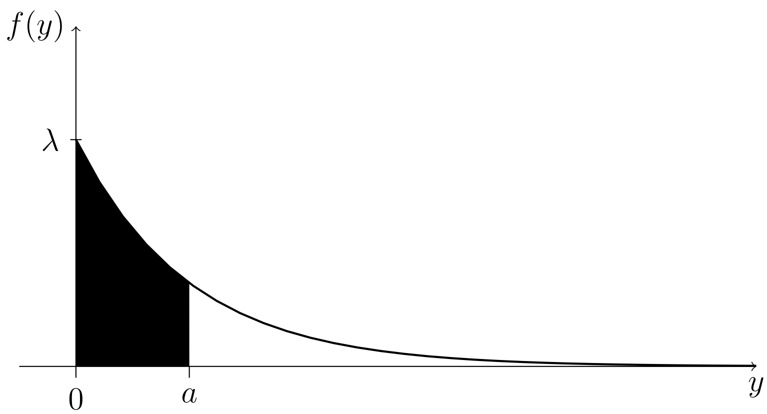
which we find by taking the derivative of $F(y)$ with respect to y:

$$\frac{d}{dy}F(y) = \frac{d}{dy}[1 - e^{-y\lambda}] = -e^{-y\lambda}(-\lambda) = \lambda e^{-y\lambda}.$$

The pdf of Y has a shape like the curve below:



For any a the the probability $P(Y \leq a)$ is given by $F(a)$, which is the area under the pdf to the left of a, as depicted below.



Definition: Exponential distribution

The continuous probability distribution with pdf given by

$$f(y) = \lambda e^{-y\lambda}$$

is called the *exponential distribution with mean* $1/\lambda$. If a random variable Y has this distribution, we write $Y \sim \text{Exponential}(\lambda)$.

If $Y \sim \text{Exponential}(\lambda)$ then

$$\mathbb{E}Y = \frac{1}{\lambda} \quad \text{and} \quad \text{Var}(Y) = \frac{1}{\lambda^2}.$$

Exercise. Suppose the occurrence of blown-out tires lying along the freeway can be regarded as a Poisson process such that for every mile, the expected number of blown-out tires is 1/3.

1. Let X be the number of tires we see in the next mile.

   (a) Find $P(X = 2)$.

   Answer: We have $X \sim \text{Poisson}(1/3)$, so

$$P(X = 2) = \frac{(1/3)^2 e^{-(1/3)}}{2!} = \texttt{dpois(2,1/3)} = 0.0398.$$

   (b) Find $P(X \geq 1)$.

   Answer: We have

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{(1/3)^0 e^{-(1/3)}}{0!} = \texttt{1 - dpois(0,1/3)} = 0.283.$$

2. Let W be the number of tires we see in the next 12 miles.

   (a) Find $P(W = 2)$.

   Answer: We have $W \sim \text{Poisson}(4)$, so

$$P(W = 2) = \frac{4^2 e^{-4}}{2!} = 0.1465.$$

   (b) Find $P(W \geq 1)$.

   Answer: We have

$$P(W \geq 1) = 1 - P(W = 0) = 1 - \frac{4^0 e^{-4}}{0!} = 1 - e^{-4} = 0.982.$$

3. Let Y be the distance traveled before the first blown-out tire is seen.

   (a) What is the distribution of Y?

   Answer: Since the occurrences of blown-out tires lying on the freeway is a Poisson process such that the expected number of blown-out tires in any one-mile segment is 1/3, the distances between blown-out tires would follow an exponential distribution with mean 3. That is $Y \sim \text{Exponential}(1/3)$.

   (b) Find $P(Y = 5)$.

   Answer: Since Y is a continuous random variable, $P(Y = 5) = 0$.

(c) Find $P(Y \leq 5)$.

Answer: We can use the cdf of the Exponential(1/3) distribution from equation (1). we have

$$P(Y \leq 5) = 1 - e^{-(1/3)(5)} = \texttt{pexp(5,1/3)} = 0.811.$$

(d) Find $P(Y > 10)$.

Answer: We can again use the cdf of the Exponential(1/3) distribution.

$$P(Y > 10) = 1 - P(Y \leq 10) = 1 - [1 - e^{-(1/3)(10)}] = \texttt{1-pexp(10,1/3)} = 0.036.$$