

1 Poisson Distribution

Poisson Probability mass function

1.1 Plot of the Poisson PMF

The horizontal axis is the index k , the number of occurrences. The function is only defined at integer values of k . The connecting lines are only guides for the eye. Cumulative distribution function Plot of the Poisson CDF The horizontal axis is the index k , the number of occurrences. The CDF is discontinuous at the integers of k and flat everywhere else because a variable that is Poisson distributed only takes on integer values.

Any French speaker will notice that "Poisson" means "fish", but really there's nothing fishy about this distribution. It's actually pretty straightforward. The name comes from the mathematician Simon-Denis Poisson (1781-1840).

The Poisson Distribution is very similar to the Binomial Distribution. We are examining the number of times an event happens. The difference is subtle. Whereas the Binomial Distribution looks at how many times we register a success over a fixed total number of trials, the Poisson Distribution measures how many times a discrete event occurs, over a period of continuous space or time. There isn't a "total" value n . As with the previous sections, let's examine a couple of experiments or questions that might have an underlying Poisson nature.

- How many pennies will I encounter on my walk home?
- How many children will be delivered at the hospital today?
- How many mosquito bites did you get today after having sprayed with insecticide?
- How many angry phone calls did I get after airing a particularly distasteful political ad?
- How many products will I sell after airing a new television commercial?
- How many people, per hour, will cross a picket line into my store?
- How many alien abduction reports will be filed this year?
- How many defects will there be per 100 metres of rope sold?

What's a little different about this distribution is that the random variable X which counts the number of events can take on any non-negative integer value. In other words, I could walk home and find no pennies on the street. I could also find one penny. It's also possible (although unlikely, short of an armored-car exploding nearby) that I would find 10 or 100 or 10,000 pennies.

Instead of having a parameter p that represents a component probability like in the Bernoulli and Binomial distributions, this time we have the parameter "lambda" or λ which represents the "average or expected" number of events to happen within our experiment. The probability mass function of the Poisson is given by

$$P(N = k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

1.2 An example

We run a restaurant and our signature dish (which is very expensive) gets ordered on average 4 times per day. What is the probability of having this dish ordered exactly 3 times tomorrow? If we only have the ingredients to prepare 3 of these dishes, what is the probability that it will get sold out and we'll have to turn some orders away?

The probability of having the dish ordered 3 times exactly is given if we set $k=3$ in the above equation. Remember that we've already determined that we sell on average 4 dishes per day, so $\lambda=4$.

$$P(N = k) = \frac{e^{-\lambda} \lambda^k}{k!} = \frac{e^{-4} 4^3}{3!} = 0.195$$

Here's a table of the probabilities for all values from $k=0..6$:

Value for k Probability $f(k)$ 0 0.0183 1 0.0733 2 0.1465 3 0.1954 4 0.1954 5 0.1563 6 0.1042 Now for the big question: Will we run out of food by the end of the day tomorrow? In other words, we're asking if the random variable $X \leq 3$. In order to compute this we would have to add the probabilities that $X=4$, $X=5$, $X=6$,... all the way to infinity! But wait, there's a better way!

The probability that we run out of food $P(X \leq 3)$ is the same as 1 minus the probability that we don't run out of food, or $1 - P(X > 3)$. So if we total the probability that we sell zero, one, two and three dishes and subtract that from 1, we'll have our answer. So,

$1 - P(X \leq 3) = 1 - (P(X=0) + P(X=1) + P(X=2) + P(X=3)) = 1 - 0.4335 = 0.5665$ In other words, we have a 56.65%

Poisson Mean

We calculate the mean as follows:

$$\begin{aligned} E[X] &= \sum_{i=0}^{\infty} f(x_i) \cdot x_i = \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \cdot x \\ E[X] &= \frac{e^{-\lambda} \lambda^0}{0!} \cdot 0 + \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \cdot x \\ E[X] &= 0 + e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda \lambda^{x-1}}{(x-1)!} \\ E[X] &= \lambda e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \\ \text{Remember that } e^{\lambda} &= \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \end{aligned}$$

$$E[X] = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Variance[edit] We derive the variance using the following formula:

$$\text{Var}[X] = E[X^2] - (E[X])^2$$

We have already calculated $E[X]$ above, so now we will calculate $E[X^2]$ and then return to this variance formula:

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{\infty} f(x_i) \cdot x_i^2 \\ E[X^2] &= \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \cdot x^2 \\ E[X^2] &= 0 + \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \cdot x^2 \\ E[X^2] &= \lambda \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \cdot (x+1) \\ E[X^2] &= \lambda \left(\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \right) \end{aligned}$$

The first sum is $E[X] = \lambda$ and the second we also calculated above to be 1.

$$E[X^2] = \lambda[\lambda + 1] = \lambda^2 + \lambda$$

Returning to the variance formula we find that

$$\text{Var}[X] = (\lambda^2 + \lambda) - (\lambda)^2 = \lambda$$

External links[edit]

Examples

Notation

The Poisson distribution is characterised by a single parameter " λ ". In many books it is denoted

1.3 Uses (Time Period)

- the number of deaths by horse kicking in the Prussian army (first application)
- birth defects and genetic mutations
- rare diseases (like Leukemia, but not AIDS because it is infectious and so not independent)
 - especially in legal cases
- car accidents in a year.
- traffic flow and ideal gap distance
- spread of an endangered animal in Africa
- failure of a machine in one month

Uses (Durations)

hairs found in McDonald's hamburgers

number of typing errors on a page

1.4 Definitions

The probability distribution of a Poisson random variable X representing the number of successes occurring in a given time interval or a specified region of space is given by the formula:

where

$x = 0, 1, 2, 3, \dots$

$e = 2.71828$ (Use your calculator's "e" button)

λ = mean number of successes in the given time interval (or given duration)

Equation The Probability of k arrivals in a unit period is given as follows.

Approaches We will looking at computing probabilities using two approaches 1) using the formula 2) using the tables.

X is the random variable used to describe the "arrivals" in a "certain period".

Changing time periods

Durations We can extend the idea of "arrivals" beyond time periods.