Poisson distribution

Typically a time or space bound activity

Football game

Average goals every 90 mins is 2.5

We may be interested in probability of 1 goal in the last 30 mins

Rate = 2.5 goals / 90 mins

Rate = 1.25 goals / 45 mins

Rate = 0.833 goals / 30 mins

Customers entering a store

Some 100 customers arrive every day

We may be interested in probability of 10 customers in the next hour

Rate = 100/day

Support centre phone calls

Some 100 calls every hour.

We may be interested in knowing optimal number of staff

Rate = 100/hour

Rate = 1.66/minute

Poisson distribution   Typically a time or space bound activity

Farmers delight

Suppose there are 100 trees every acre of land

Can there be more than 60 trees in half an acre?

Rate = 100 trees / acre

Hospital emergency

Suppose, on average, 5 patients come every hour

What is the probability of more than 10 people next hour?

Rate = 5 patients / hour

Typos

A book might have an average of 3 typos per page

What is the probability of a page having no typos?

Rate = 3/page

Poisson distribution   Typically a time or space bound activity

Rate is the average or expected number of events per interval

This interval is typically time, but can be space, or even “number of pages” etc.

We typically denote Rate = λ

Because it is also the average or expected number, some literature may also use Rate = μ

The value of the rate depends on the way we define our interval

Poisson distribution   Rules deciding Poisson

Counting

The experiment counts the number of occurrences of an event over an interval

Independence

The occurrence of one event does not affect the probability that a second event will occur

Rate

The average rate at which events occur is independent of any occurrences

No Simultaneous events

No two event occur simultaneously

Poisson distribution

A city sees 3 accidents per day on average.

Find the probability that there will be 5 accidents tomorrow

Rate $\lambda = 3$ per day

Let "X" denote the number of accidents tomorrow

We say "X" is Poisson distributed with rate = 3

$E[X] = 3$ $\mu = 3$ same as λ

```python
from scipy.stats import poisson
```

$P[X = 5] =$ `poisson.pmf(k=5, mu=3)`

$= 0.1008$

$$P[X = 5] = \frac{3^5 e^{-3}}{5!} = 0.1008 \qquad P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

Poisson distribution

`poisson.pmf(k, mu)` $P[X=k] = \frac{\lambda^k e^{-\lambda}}{k!}$

Let "X" be the number of typos in a page in a printed book, with mean 3 typos per page. What is the probability that a randomly selected page has at most 1 typo?

$\lambda = 3$

$P[X=0] =$ `poisson.pmf(k=0, mu=3)` $= 0.049$

$$P[X=0] = \frac{3^0 e^{-3}}{0!} = e^{-3} = 0.049$$

$P[X=1] =$ `poisson.pmf(k=1, mu=3)` $= 0.149$

$$P[X=1] = \frac{3^1 e^{-3}}{1!} = 3e^{-3} = 0.149$$

$P[X \leq 1] =$ `poisson.cdf(k=1, mu=3)` $= 0.199$

$$P[X \leq 1] = P[X=0] + P[X=1] = 4e^{-3} = 0.199$$

Poisson distribution

`poisson.pmf(k, mu)` $P[X=k] = \frac{\lambda^k e^{-\lambda}}{k!}$

A shop is open for 8 hours. The average number of customers is 74 - assume Poisson distributed.

Q1) What is the average or expected number of customers in 2 hours?

8 hours — 74 customers

2 hours — ?

$$\frac{2 * 74}{8} = 18.5$$

Q2) What is the probability that in 2 hours, there will be at most 15 customers?

$P[X \leq 15] =$ `poisson.cdf(k=15, mu=18.5)` $= 0.249$

Q3) What is the probability that in 2 hours, there will be at least 7 customers?

$P[X \geq 7] = 1 - P[X \leq 6] =$ `1 - poisson.cdf(k=6, mu=18.5)` $= 0.99$

Poisson distribution

`poisson.pmf(k, mu)` $P[X = k] = \dfrac{\lambda^k e^{-\lambda}}{k!}$

You receive 240 messages per hour on average - assume Poisson distributed.

Q1) What is the average or expected number of messages in 30 seconds?

1 hour (3600 seconds) — 240 messages

30 seconds — ?

$$\frac{30 * 240}{3600} = 2$$

Q2) What is the probability of one message arriving over a 30 second time interval?

If we consider 30 seconds as one unit interval, then $\lambda = 2$

$P[X = 1] =$ `poisson.pmf(k=1, mu=2)` $= 0.27$

$$P[X = 1] = \frac{(2)^1 e^{(-2)}}{1!} = 0.27$$

Q3) What is the probability that there are no messages in 15 seconds?

$$\lambda = \frac{15 * 240}{3600} = 1$$

$P[X = 0] =$ `poisson.pmf(k=0, mu=1)` $= 0.367$

$$P[X = 0] = \frac{(1)^1 e^{(-1)}}{1!} = 0.367$$

Q4) What is the probability that there are 3 messages in 20 seconds?

$$\lambda = \frac{20 * 240}{3600} = 1.33$$

$P[X = 3] =$ `poisson.pmf(k=3, mu=1.33)` $= 0.104$

$$P[X = 3] = \frac{(1.33)^3 e^{(-1.33)}}{3!} = 0.104$$

Poisson distribution

`poisson.pmf(k, mu)` $P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$

Suppose we receive 3 support tickets every 20 days.

Q1) What is the average or expected number of tickets in 1 day?

20 days — 3 tickets

1 day — ?

$$\frac{3}{20} = 0.15$$

Q2) What is the probability that there will not be more than 1 ticket in a day?

$P[X \leq 1] =$ `poisson.cdf(k=1, mu=0.15)` $= 0.989$

Poisson distribution

`poisson.pmf(k, mu)` $P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$

There are 80 students in a kinder garden class.

Each one of them has 0.015 probability of forgetting their lunch on any given day.

Q1) What is the average or expected number of students who forgot lunch in the class?

1 student — 0.015

80 students — ?

$$\frac{80 * 0.015}{1} = 1.2$$

Q2) What is the probability that exactly 3 of them will forget their lunch today?

$P[X = 3] =$ `poisson.pmf(k=3, mu=1.2)` $= 0.086$

$$P[X = 3] = \frac{(1.2)^3 e^{-1.2}}{3!} = 0.086$$

Binomial distribution

Here, $n = 80, p = 0.015$, and $k = 3$ $\lambda = np$

$P[X = 3] =$ `binom.pmf(k=3, n=80, p=0.015)` $= 0.086$

$E[X] =$ `binom.expect(args=(80, 0.015))` $= 1.2$

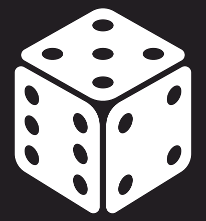
Poisson distribution

Poisson approximation to Binomial

Binomial trials “n” is at least 30
Probability of success “p” is at most 0.05

Binomial $E[X] = np$

$$\lambda = np$$



How many times can we expect to get a 6 if we throw 600 times?

1 throw	$\frac{1}{6}$
600 throws	?

$$(600)\frac{1}{6} = 100$$

How many times can are we expected to throw to get the first 6?

1 throw	$\frac{1}{6}$
?	1

$$\frac{1}{1/6} = 6$$

Geometric distribution

What is the probability that the first heads comes in the 7^{th} toss?

T, T, T, T, T, T, H

$$P[X = 7] = (1 - p)^6 p$$

What is the probability of first heads comes in the k^{th} toss?

$$P[X = k] = (1 - p)^{k-1} p$$

Geometric distribution

$$P[X = k] = (1-p)^{k-1}p$$

Suppose we throw a dice till the first time we get 6

Q1) What is the probability that we have to throw 4 times?

$$P[X=4] = \left(\frac{5}{6}\right)^3 \frac{1}{6} = 0.0964$$

```
from scipy.stats import geom
```

$P[X=4] =$ `geom.pmf(k=4, p=1/6)` $= 0.0964$

Q2) What is the expected number of throws to get the first 6?

1 throw $\quad \frac{1}{6}$

? $\quad 1$

$$\frac{1 * 1}{\frac{1}{6}} = 6$$

`geom.expect(args=(1/6,))` $= 6$

$$E[X] = \frac{1}{p}$$

Geometric distribution

`geom.pmf(k, p)` $P[X=k] = (1-p)^{k-1}p$

$$E[X] = \frac{1}{p}$$

I am playing a game where the prob of winning a prize is 0.7

What is the probability that I win the prize on the 4th attempt?

$$P[X=4] = (0.3)^3(0.7) = 0.0189$$

```python
from scipy.stats import geom
```

$P[X=4] =$ `geom.pmf(k=4, p=0.7)` $= 0.0189$

What is the probability that I don't win in the first two attempts

$$P[X>2] = 1 - P[X \le 2]$$

$P[X>2] =$ `1 - geom.cdf(k=2, p=0.7)` $= 0.09$

What is the expected number of attempts to win the prize

$$E[X] = \frac{1}{0.7} = 1.42$$

$E[X] =$ `geom.expect(args=(0.7,))` $= 1.42$