$P(K > 2)$

$P(K \le 2) \longrightarrow$ Poisson.cdf( )

pmf

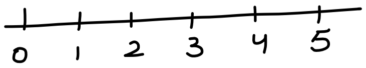
cdf

Advance Distribution

Exponential

poisson dist

Discrete prob. dist



$P(K=1)$
$P(K=2)$
$P(K=3)$
$P(K=4)$

point prob [Poisson.pmf]

poisson dist : prob. of # of events in the given time interval

(Rate) $\longrightarrow$ Avg event observed in the given time interval

Quiz: (Read the following questions and be ready with your answers. I will launch the poll at 09:08 PM)

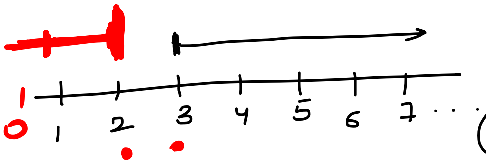
Poisson Dist $\longrightarrow$ Discrete prob dist < pmf, cdf

Sumit is head over heels for Ankita, and he can't resist talking to her on the phone. On average, Sumit calls Ankita 5 times per hour. Assume that the number of times Sumit calls Ankita in a given hour follows a Poisson distribution.

What is the probability that during a particular hour-long boring lecture at IIT Kanpur, Sumit will call Ankita:

a) Exactly 3 times? (Poor Ankita, she might get annoyed!)
b) At most 2 times? (Sumit is trying to control himself, bless his heart!)

A) 0.11, 0.43
B) 0.91, 0.41
C) 0.14, 0.12 ✓
D) 0.5, 0.5



$\lambda = 5$ calls / hr }

① k=3 times

poisson.pmf(k=3, mu=5)

P(at least 3 calls)

= 1 − poisson.cdf(k=2, mu=5)

② at most ≤

P(k ≤ 2) = possion.cdf(k=2, mu=5)

Quiz: You receive 240 messages per hour on average - assume Poisson distributed. What is the average number of messages <u>per second</u>?

Rate

$\lambda \longrightarrow$ 240 messages / hr

$\lambda \longrightarrow$ 240 messages / 60 min

$\lambda \longrightarrow$ 240 messages / 3600 secs

3600 s $\longrightarrow$ 240 m

1 s $\longrightarrow \frac{240}{3600}$

1 s $\longrightarrow$ 0.067 messages

$\left[\lambda = \underline{0.067 \text{ messages/sec}}\right]$

Rate expressed in time interval of seconds

=

You receive 240 messages per hour on average - assume Poisson distributed. What is the probability of having no message in 10 seconds?

$k = 0$

poisson.pmf ($k=0$, mu$=0.67$)

$= 0.51$

$\lambda = 240$ messages/hr

$\lambda = 240$ messages/3600 sec

Rate for the time interval of 10 secs

3600 sec $\longrightarrow$ 240 messages

1 sec $\longrightarrow$ $\frac{240}{3600}$

10 sec $\longrightarrow$ $\frac{240}{3600} \times 10$ messages

$= 0.67$

$\lambda = 0.67$ messages/10 sec

Rate $= 0.67$ m/10sec

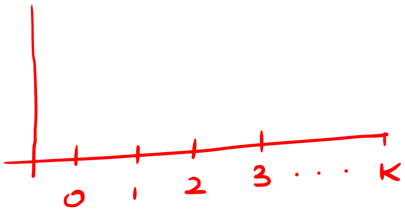
You receive 240 messages per hour on average - assume Poisson distributed. What's the probability of waiting more than 10 sec for the next message?

Poisson Dist (Discrete prob. dist)

prob of k events in the given time interval

Rate (λ) = avg events observed in the given interval

K - Discrete
$\{0, 1, 2 \ldots k\}$



For ex: I receive 10 calls/min, what is the prob that I will receive 2 calls in the next minute.

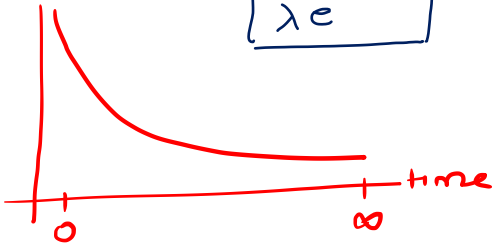
Exponential Dist (continuous prob. dist)

prob. of the time interval

Scale (s) = time b/w two events

time - continuous RV

$$\boxed{\lambda e^{-\lambda x}}$$



For ex: I receive 10 calls/min. what's the prob. that I need to wait for more than 10 mins to receive the next call

You receive 240 messages per hour on average - assume Poisson distributed. What is the average number of messages per second?

Scale

$\lambda = 240$ messages / hr

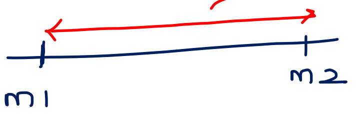
$$\boxed{Scale = \frac{1}{\lambda}}$$

Rule : If the question is asking for the time in unit of (t) then find scale in t

$\lambda = 0.067$ messages / sec

$$Scale = \frac{1}{\lambda} = \frac{1}{0.067} = 15 \text{ secs}$$

avg wait time b/w two messages is 15 sec → scale



Exponential prob. → what's the prob that I need to wait for more than 10 sec to receive the next messages

You receive 240 messages per hour on average - assume Poisson distributed. What's the probability of waiting more than 10 sec for the next message?

Let's solve it using exponential

$\lambda = 240$ messages / hr

prob of waiting more than 10 sec for next message → ?

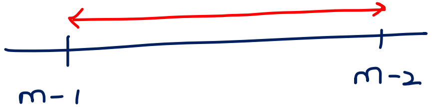
Q: I receive 240 m / hr, what's the prob that I receive 300 message in next one hr.

given interval in the question is sec

$\lambda = 0.067 / \text{sec}$

$\text{Scale} = \frac{1}{\lambda} = \frac{1}{0.067} = 15 \text{ sec}$

on an avg I need to wait for 15 secs to receive next message



Question →

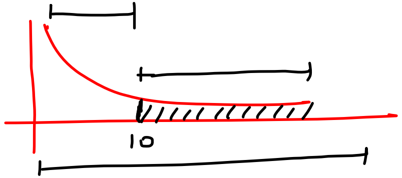
Given → scale — Avg wait time b/w two events

Asked → what's the prob. of waiting for t time

exponential dist

$P(t > 10 \text{ sec}) = 1 - \text{expon.cdf}(k = 10, \text{scale} = 15)$

$= 0.513$



If a call center receives 4 calls per minute, what is the probability of waiting less than or equal to 30 seconds, i.e. 0.5 min for the next call?

$\lambda = 4$ calls / min

$\lambda = 4$ calls / 60 sec

60 sec $\longrightarrow$ 4

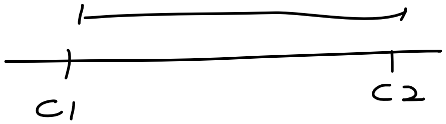
1 sec $\longrightarrow \frac{4}{60}$

$= 0.06$

$\lambda = 0.066$ calls / sec

Scale $= \frac{1}{\lambda} = \frac{1}{4} = 0.25$ mins

on on avg there is a wait time of 0.25 min b/w two calls



Scale $= \frac{1}{\lambda} = 15.1$ sec

expon.cdf (x = 30 ↘sec, scale = 15.1 ↘sec)

$p(t \leq 0.5) =$ expon.cdf (x = 0.5 ↘min, Scale = 0.25 ↘min)

$= 0.86$

$=$

Applications of the Exponential Distribution:

1. Time between telephone calls
2. Time between machine breakdowns
3. Time between successive job arrivals at a computing centre

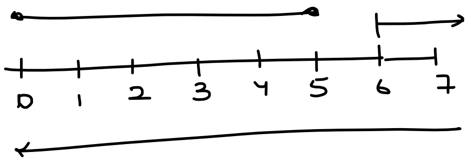
queue   Scale = 5 mins   on on avg I need to wait for 5 mins ~~to be~~ for the next person to be served

what the prob that I wait for mor than 20 min to be served

Accidents occur with a Poisson distribution at an average of 4 per week. i.e. λ = 4

1. Calculate the probability of more than 5 accidents in any one week
2. What is the probability that at least two weeks will elapse between accident?

$\lambda = 4$ accident / week

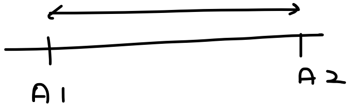


① $P(K > 5) = 1 - P(K \leq 5)$

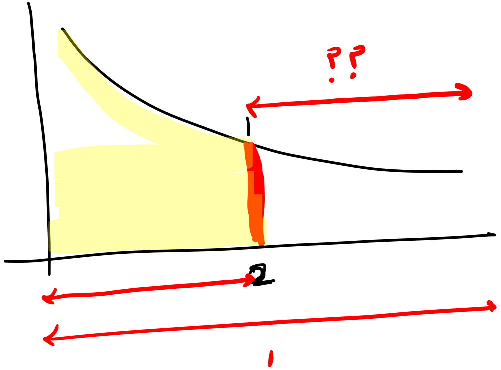
$= 1 - \text{poisson.cdf}(k = 5, mu = 4)$

$= 0.21$

B/w two accident there is a wait time of 0.25 weeks



② $\text{Scale} = \frac{1}{\lambda} = \frac{1}{4} = 0.25$ weeks



$P(t \geq 2) = 1 - P(t \leq 2)$

$= 1 - \text{expon.cdf}(x = 2, \text{Scale} = 0.25)$

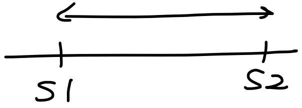
$= 0.00033$

$\boxed{\leq 2 = < 2}$

Quiz:

Suppose you have a system that fails, on average, every 50 hours.
What is the probability that the system will fail within the first 20 hours?

there is a wait time of 50 hrs b/w two consec. system failure



$\lambda = 1/50$ hrs

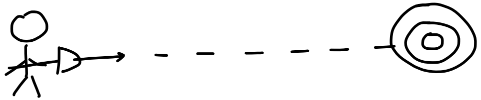
$$\text{Scale} = \frac{1}{\lambda} = \frac{1}{1/50} = 50$$

$$P(t \leq 20) = \text{expon.cdf}(x = 20, \text{Scale} = 50)$$

$$= 0.329$$

Geometric Distribution

Discrete prob. Dist $\longrightarrow$ prob. of first success



$P(\text{hit}) = 0.25$

what's the prob thate he will get the first hit in the 3rd trail

$$= \underbrace{(1-0.25)}_{f}\;\underbrace{(1-0.25)}_{f}\;\underbrace{0.25}_{S}$$

$$= (1-0.25)^2(0.25)$$

$$\left[\,P(X=n) = (1-p)^{n-1}\,p\,\right]$$

what's the prob. that first success will be in n'th trail that means in all $(n-1)$ trails, the exp. results in a failure

Quiz:

You are flipping a biased coin with a 30% chance of getting heads until you succeed. What is the probability of getting heads on the 2nd flip?

$p(\text{head}) = 0.3$

$p(\text{fail}) = 0.7$

$\underline{0.7} \quad \underline{0.3} = 0.21$

Quiz:

In a factory that produces light bulbs, there's a 5% chance that any given bulb is defective. What is the probability of needing 1 or 2 bulbs to find the first defective bulb?

$P(\text{defective}) = 0.05$

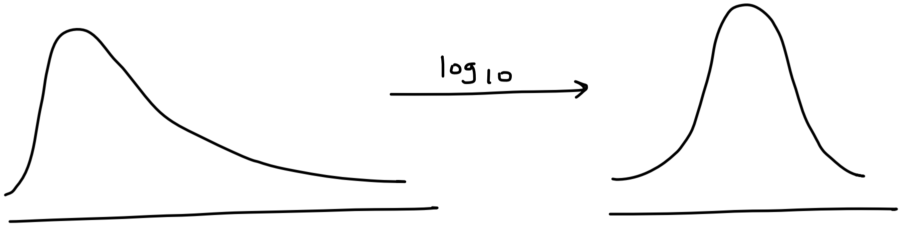
$$\left.\begin{array}{l} P(x=1) = 0.05 \\ P(x=2) = (1-0.05) \times 0.05 \end{array}\right] \longrightarrow \begin{array}{l} 0.05 + (1-0.05) \times 0.05 \\ = 0.0975 \end{array}$$

$\text{geom.cdf}(k=2, P=0.05)$

Detailed discussion to be happen during machine learning

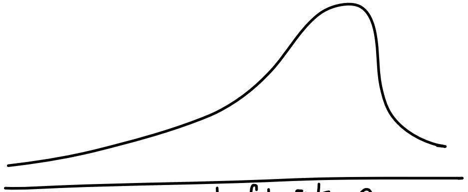
Box Cox Transformation

$$y_i^{(\lambda)} = \begin{cases} \dfrac{y_i^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln(y_i) & \text{if } \lambda = 0, \end{cases}$$



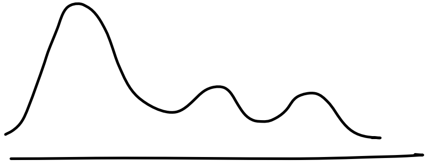
Right skew

Log-normal

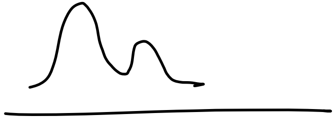


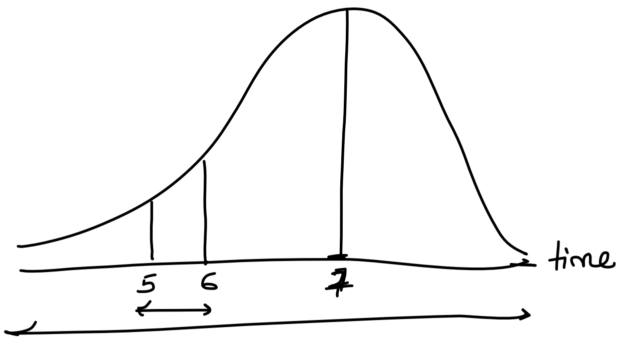
Left skew

$$\text{Non-normal Dist} \xrightarrow[\text{transfo-}]{\text{Box-cox}} \left[\text{Normal Dist}\right]$$



multi-model





$$\frac{1}{10} = \frac{1}{10} = 0.1$$

$$\frac{0}{10} = 0$$

$$P = \frac{\#\text{ of events}}{\text{total no of events}}$$

$$\left\{ P = \frac{\text{length of seg.}}{\text{total length}} \right\}$$