Probability Distributions

Discrete distributions (Finite)

- *) Binomial distribution
- *) Poisson distribution (Infinite)

Continuous distributions

- *) Exponential distribution
- *) Gamma distribution
- *) Weibull distribution
- *) Normal distribution

1) Binomial Distribution

$$\boxed{x \sim B(n, p)}$$

Binomial Distribution :

$$P(x) \Rightarrow nC_x \, p^x q^{n-x}$$

where
$n \Rightarrow$ number of trials
$p \Rightarrow$ probability of success
$q \Rightarrow$ probability of failure
$x \Rightarrow$ number of times getting success

Mean $\Rightarrow np$

Variance $\Rightarrow npq$

To calculate mean

$$M_x(t) = E(e^{tx}) \Rightarrow \sum_{x=0}^{n} (e^{tx} p(x))$$

*) Mean $\Rightarrow \frac{d}{dt}(M_x(t)) \Rightarrow np \quad (\mu_1')$

*) Variance $\Rightarrow \frac{d^2}{dt^2}(M_x(t)) \Rightarrow npq$

$\mu_2' - \mu_1'$

The question which they are asking only is the **p (positive)**

No matter whether the question is positive

Problem: If the chance of running a bus service according to schedule is 0.8, Calculate the probability on a day schedule with 10 services:

- exactly one is late
- at least one is late.

Solution: Let X be a binomial random variable with parameters n and p. Then

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \dots, n.$$

Probability of a bus running according to schedule=0.8=q

therefore, the probability that a bus is late is 0.2=p

Here $n = 10, p = 0.2$.

- $P(X = 1) = \binom{10}{1}(0.2)^1(0.8)^{10-1} = 0.2684$
- $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) =$
  $1 - \binom{10}{0}(0.2)^0(0.8)^{10-0} = 0.8926$

Problem: For a binomial distribution with parameters n=5, p=0.3. Find the probabilities of getting

- at least 3 successes
- at most 3 Sucesses
- exactly 3 failures.

Solution: Let X be a binomial random variable with parameters n and p. Then

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \dots, n.$$

Here $n = 5, p = 0.3$.

- $P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) = 0.1631,$
- $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0.9692$
- $P(X = 2 \text{ sucesses}) = P(X = 3 \text{ failures}) = 0.3087.$

Problem: If the chance of running a bus service according to schedule is 0.8, Calculate the probability on a day schedule with 10 services:

- exactly one is late
- at least one is late.

Solution: Let X be a binomial random variable with parameters n and p. Then

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \ldots, n.$$

Probability of a bus running according to schedule=0.8=q

therefore, the probability that a bus is late is 0.2=p

Here $n = 10, p = 0.2$.

- $P(X = 1) = \binom{10}{1}(0.2)^1(0.8)^{10-1} = 0.2684$
- $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \binom{10}{0}(0.2)^0(0.8)^{10-0} = 0.8926$

Problem: The 10% of the screws produced by an automatic machine are defective, find the probability that os 20 screws selected at random, there are

- exactly two defectives
- at the most three defectives
- at least two defectives and
- between one and three defectives(inclusive).

Problem: 4 coins are tossed and number of heads noted. The experiment is repeated 200 times and the following distribution is obtained.

X:Number of heads	0	1	2	3	4
f:frequencies	62	85	40	11	2

Find the binomial distribution and expected frequency distribution.

Binomial distribution $\rightarrow {}^nC_x p^x q^{n-x}$

Expected frequency distribution $\} \rightarrow N * ({}^nC_x p^x q^{n-x})$

$n \Rightarrow 4$

$N \Rightarrow 200$

$$\text{Mean} \Rightarrow \frac{\Sigma fx}{\Sigma f} \Rightarrow \frac{85 + 80 + 33 + 8}{200} \Rightarrow \frac{206}{200} \Rightarrow 1.03$$

$\lambda \Rightarrow np \Rightarrow 1.03$

$$p = \frac{1.03}{n} \Rightarrow \frac{1.03}{4} \Rightarrow 0.2575$$

$\boxed{p \Rightarrow 0.2575}$

$q \Rightarrow 1 - p$

$\boxed{q \Rightarrow 0.7425}$

X	freq	Binomial distribution	Expected frequency distribution
0	62	60.95	0.30475
1	85	84.33	0.42165
2	40	43.75	0.21875
3	11	10.8	0.054
4	2	0.872	0.0436

$x \Rightarrow 0;\ 4C_0 (0.2575)^0 (0.7425)^4 \Rightarrow 0.30475$

$x \Rightarrow 1;\ 4C_1 (0.2575)^1 (0.7425)^3 \Rightarrow 0.42165$

$x \Rightarrow 2;\ 4C_2 (0.2575)^2 (0.7425)^2 \Rightarrow 0.21875$

$x \Rightarrow 3;\ 4C_3 (0.2575)^3 (0.7425)^1 \Rightarrow 0.054$

$x \Rightarrow 4;\ 4C_4 (0.2575)^4 (0.7425)^0 \Rightarrow 0.0436$

Problem:

Fit a binomial distribution for the following data:

x	0	1	2	3	4	5	6	Total
f	5	18	28	12	7	6	4	80

$\boxed{n \Rightarrow 6}$

$$\text{Mean} \Rightarrow \frac{\Sigma fx}{\Sigma f} \Rightarrow \frac{(0+18+56+36+28+30+24)}{80}$$

$$\lambda \Rightarrow np \Rightarrow \frac{192}{80} \Rightarrow 2.4$$

$$6p \Rightarrow 2.4$$

$\boxed{p \Rightarrow 0.4}$ and $\boxed{q \Rightarrow 0.6}$

$x \Rightarrow 0;\quad 6C_0 (0.4)^0 (0.6)^6 \Rightarrow (0.6)^6 \Rightarrow 0.046656$

$x \Rightarrow 1;\quad 6C_1 (0.4)^1 (0.6)^5 \Rightarrow 0.186624$

$x \Rightarrow 2;\quad 6C_2 (0.4)^2 (0.6)^4 \Rightarrow 15 * 0.020736 \Rightarrow 0.31104$

$x \Rightarrow 3;\quad 6C_3 (0.4)^3 (0.6)^3 \Rightarrow 20 * 0.013824 \Rightarrow 0.27648$

$x \Rightarrow 4;\quad 6C_4 (0.4)^4 (0.6)^2 \Rightarrow 15 * (9.216 * 10^{-3}) \Rightarrow 0.13824$

$x \Rightarrow 5;\quad 6C_5 (0.4)^5 (0.6)^1 \Rightarrow 6 * (6.144 * 10^{-3}) \Rightarrow 0.03686$

$x \Rightarrow 6;\quad 6C_6 (0.4)^6 (0.6)^0 \Rightarrow 1 * (0.4)^6 \Rightarrow 4.096 * 10^{-3}$

x	f	Binomial distribution	Expected frequency for binomial distribution
0	5	0.046656	0.046656 * 80 → 3.73
1	18	0.186624	0.186624 * 80 → 14.93
2	28	0.31104	0.31104 * 80 → 24.88
3	12	0.27648	0.27648 * 80 → 22.12
4	7	0.13824	0.13824 * 80 → 11.06
5	6	0.03686	0.3686 * 80 ⇒ 2.95
6	4	$4.096 * 10^{-3}$	$4.096 * 10^{-3} * 80 \Rightarrow 0.33$

2) Poisson Distribution

Poisson distribution

Probability $P(x) \Rightarrow \frac{e^{-\lambda}\lambda^{x}}{x!}$ $(x \Rightarrow 0, 1, 2 \ldots \infty)$

$$\boxed{x \sim P(\lambda)}$$

*) Number of trials $\boxed{n \Rightarrow \infty}$

*) probability of success in each trial is very small $\}$ $\boxed{p \rightarrow 0}$

*) Mean = variance = $\lambda = np$

To calculate mean

$$M_x(t) \Rightarrow E[e^{tx}]$$

$$M_x(t) \Rightarrow \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$\mu_1' \Rightarrow \frac{d}{dt}\left(M_x(t)\right) \Rightarrow \lambda$$

$$\mu_2' \Rightarrow \lambda^2 + \lambda$$

$$\text{Variance} \Rightarrow \mu_2' - (\mu_1')^2$$

$$\Rightarrow \lambda^2 + \lambda - (\lambda^2)$$

$$\boxed{\text{Variance} \Rightarrow \lambda}$$

Poisson Distribution

QUICK RUNDOWN

* Discrete distribution
* Describes the number of events occuring in a fixed time interval or region of opportunity
* Requires only one parameter, λ
* Bounded by 0 and ∞

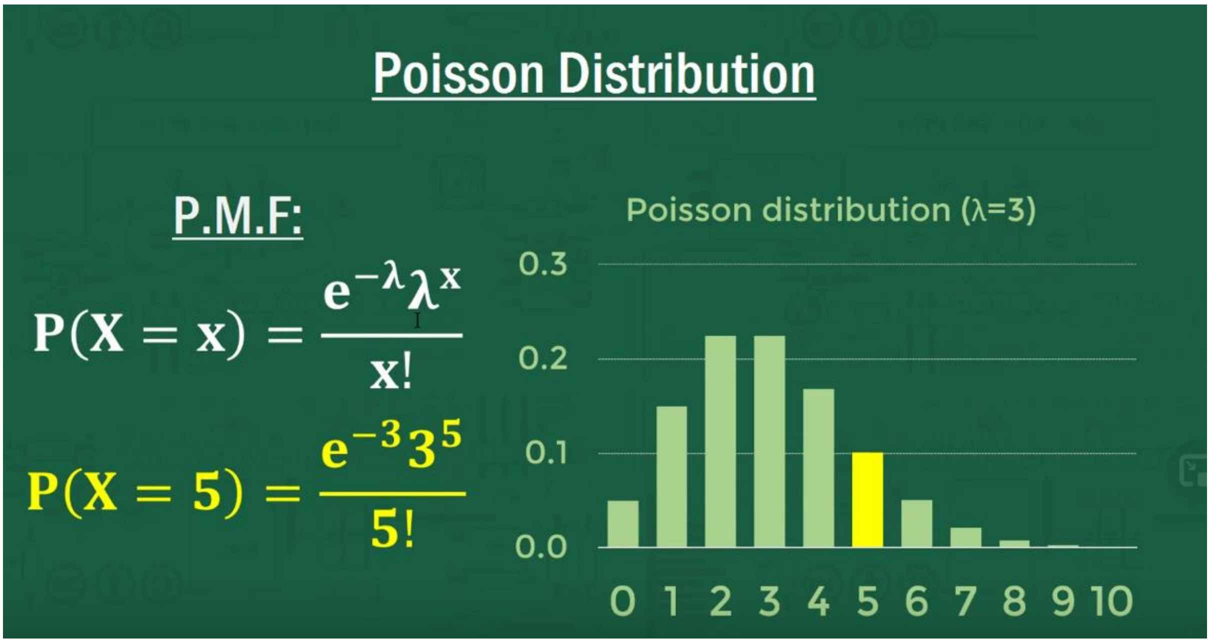
Poisson Distribution

ASSUMPTIONS

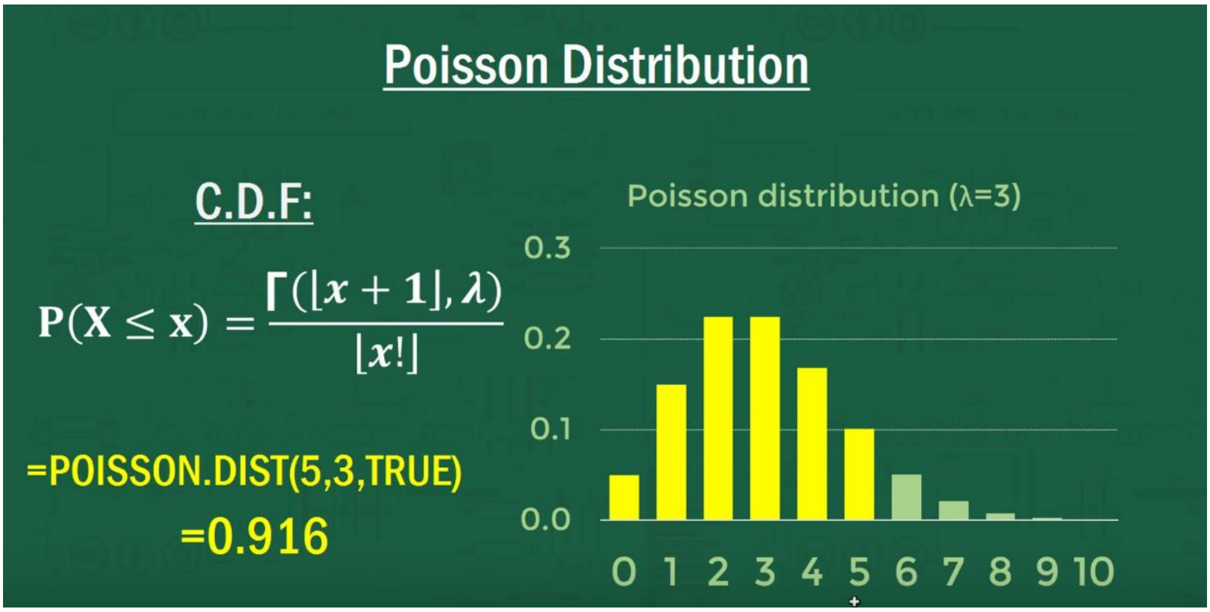
* The rate at which events occur is constant
* The occurrence of one event does not affect the occurrence of a subsequent event (ie. events are independent)

The probability of an event occurring in a certain time interval should be exactly the same for every other time interval of that same length

Probability Mass function



Cumulative Distribution function



Question

Exclusive Vines import Argentinian wine into Australia. They've begun advertising on Facebook to direct traffic to their website where customers can order wine online. The number of click-through sales from the ad is Poisson distributed with a mean of 12 click-through sales per day.

Poisson Distribution

Find the probability of getting:

(a) Exactly 10 click-through sales in the first day

(b) At least 10 click-through sales in the first day

(c) More than one sale in the first hour

(a) P(10 click-through sales in the first day) =

$$P(X = 10) = \frac{e^{-12}12^{10}}{10!}$$

$$= 0.105$$

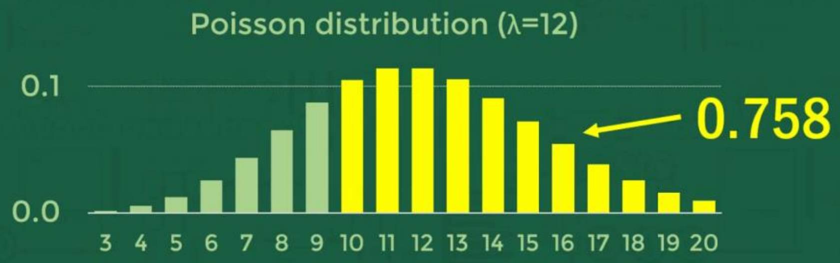
=POISSON.DIST(10,12,FALSE)

(a) P(10 click-through sales in the first day) =

Poisson Distribution

(b) P(At least 10 click-through sales on the first day) =

$$\mathbf{P(X \geq 10) = ???}$$



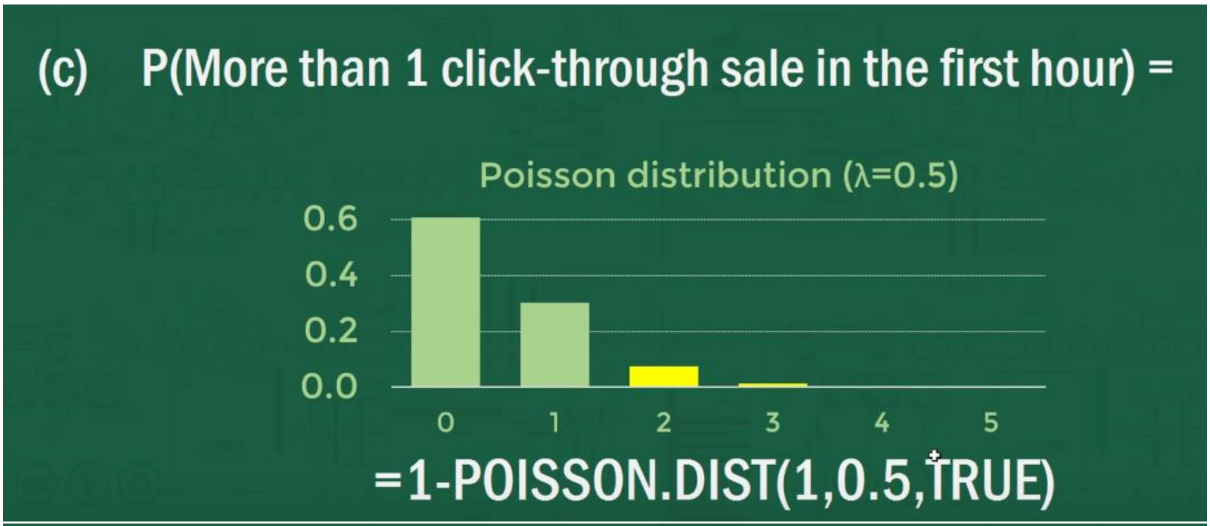
$$\mathbf{P(X \geq 10) = 0.758}$$

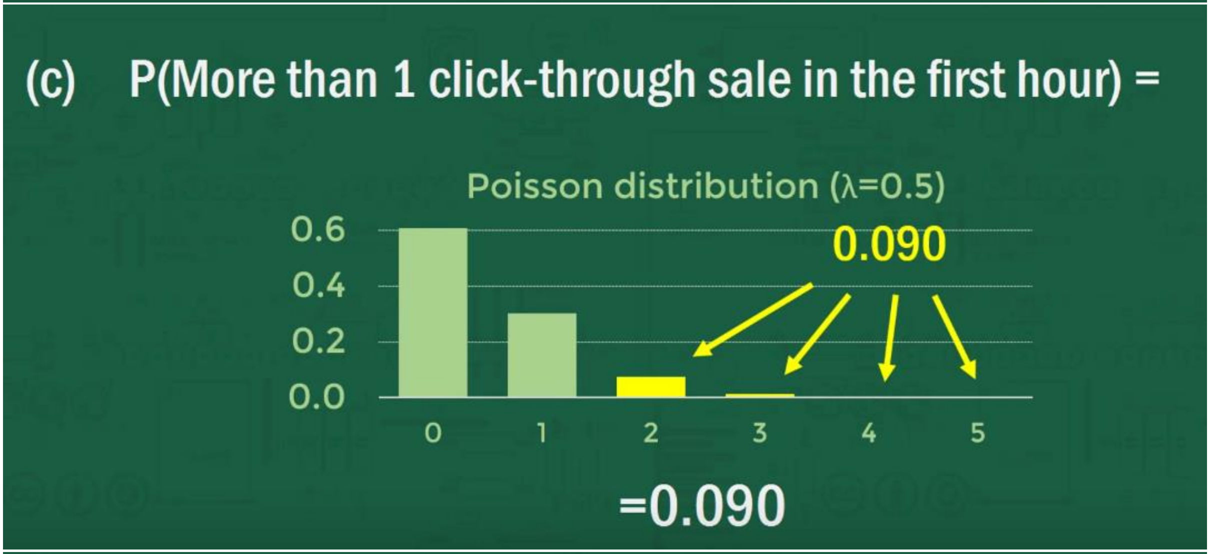
(c) P(More than 1 click-through sale in the first hour) =

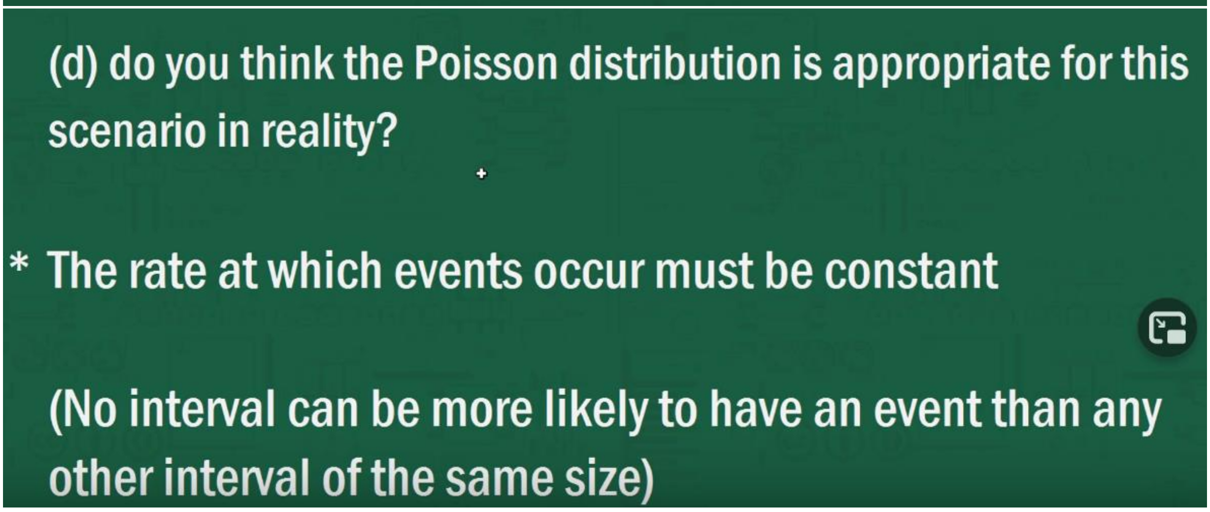
$$\boldsymbol{\lambda = \frac{12}{24} = 0.5\ \mathit{sales\ per\ hour}}$$

1day (24 hrs) → mean = 12

1 hr → mean = ?







In reality, the clicks on facebook ads about wine is not constant, it will not follow poisson distribution. But its better to use these distributions to get a clear picture.

2. If the probability of a defective fuse from a manufacturing unit is 2%, in a box of 200 fuses, find the probability that a) exactly 4 fuses are defective b) more than 3 fuses are defective.

P=0.02, n=200

Mean = λ =np=200*0.02=4

$$a)\ P(X=4)=\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$=\frac{e^{-4}(4)^{4}}{4!}=0.1952$$

$$b)\ P(x>3)=1-P(x\le 3)$$

$$=1-[p(x=3)+p(x=2)+p(x=1)+p(x=0)]$$

$$=1-e^{-4}[\frac{4^3}{3!}+\frac{4^2}{2!}+\frac{4^1}{1!}+\frac{4^0}{0!}]=0.5669$$

Problems:

1. The number of monthly breakdowns of a computer is a random variable having a Poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month.

a) Without a breakdown

b) With only one breakdown and

c) With atleast one breakdown

Let X denotes the number of breakdowns of the computer in a month.

X follows a Poisson distribution with mean λ =1.8

$$P(X=x)=\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$=\frac{e^{-1.8}(1.8)^{x}}{x!}$$

$a)\ p(x=0)=e^{-1.8}=0.1653$

$b)\ p(x=1)=e^{-1.8}(1.8)=0.2975$

$c)\ p(x\geq 1)=1-p(x=0)=0.8347$

Fitting of Poisson Distribution

1. Fit a Poisson distribution to the following data and calculate the theoretical frequencies.

Deaths	0	1	2	3	4
Frequency	122	60	15	2	1

x	f	Poisson distribution	Expected frequency
0	122	0.6065	121
1	60	0.3032	61
2	15	0.07581	15
3	2	0.01263	3
4	1	$1.5795 * 10^{-3}$	0

$$\text{Mean} \Rightarrow \frac{\sum fx}{\sum f} \Rightarrow \frac{100}{200} \Rightarrow 0.5$$

$$x \Rightarrow 0 ; \quad \frac{e^{-0.5}(0.5)^{0}}{0!} \Rightarrow 0.6065$$

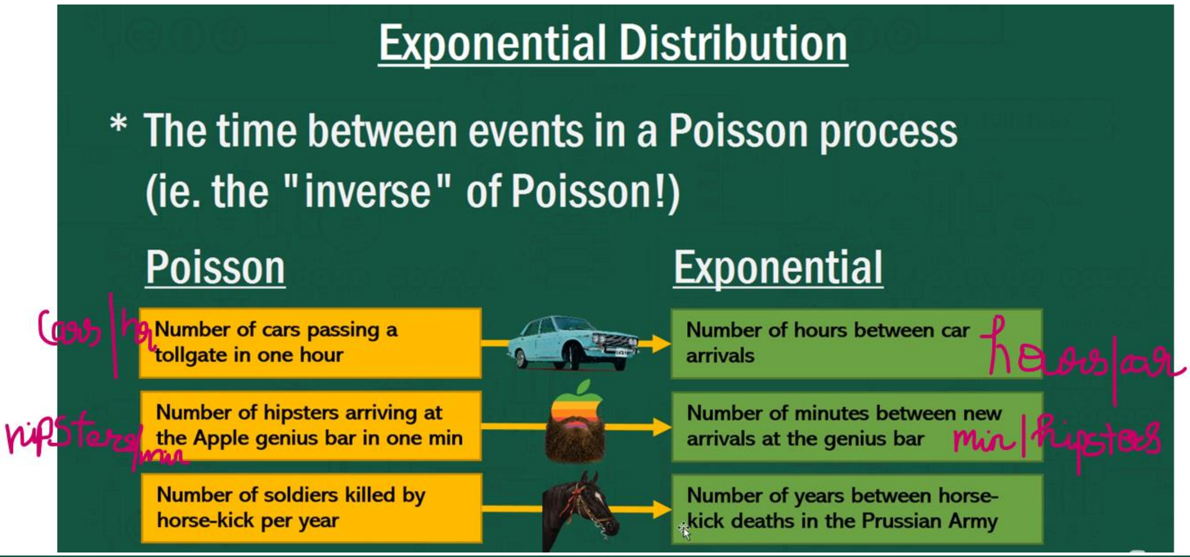
$$x \Rightarrow 1 ; \quad \frac{e^{-0.5}(0.5)^{1}}{1!} \Rightarrow 0.3032$$

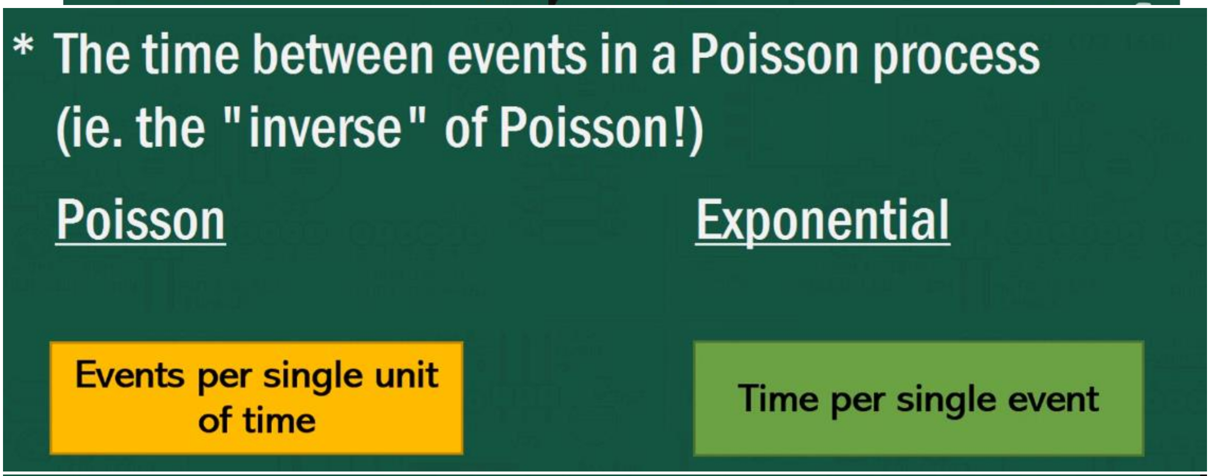
$$x \Rightarrow 2 ; \quad \frac{e^{-0.5}(0.5)^{2}}{2!} \Rightarrow 0.07581$$

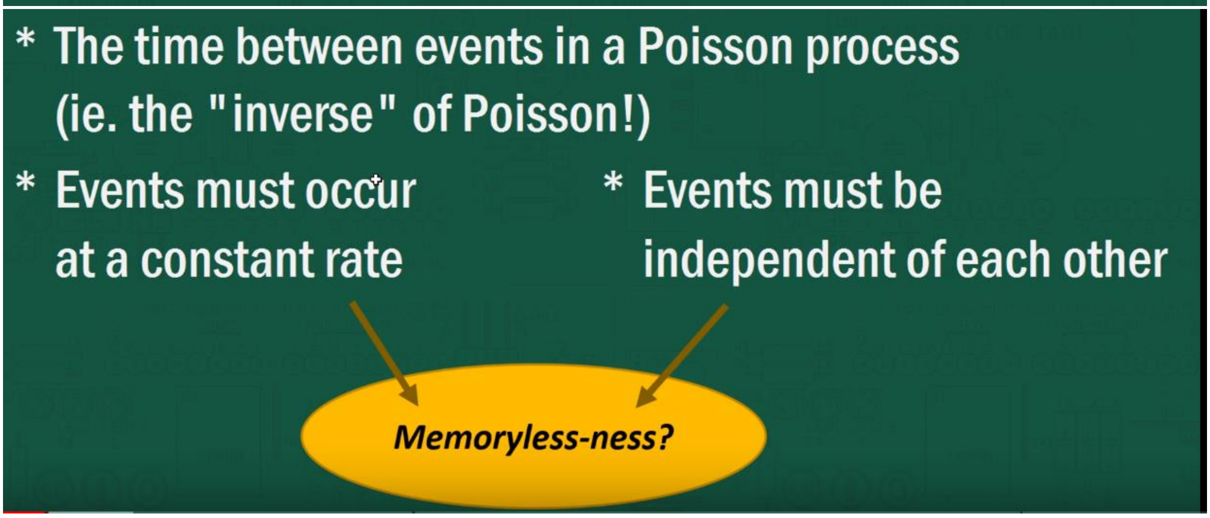
$$x \Rightarrow 3 ; \quad \frac{e^{-0.5}(0.5)^{3}}{3!} \Rightarrow 0.01263$$

$$x \Rightarrow 4 ; \quad \frac{e^{-0.5}(0.5)^{4}}{4!} \Rightarrow 1.57950 * 10^{-3}$$

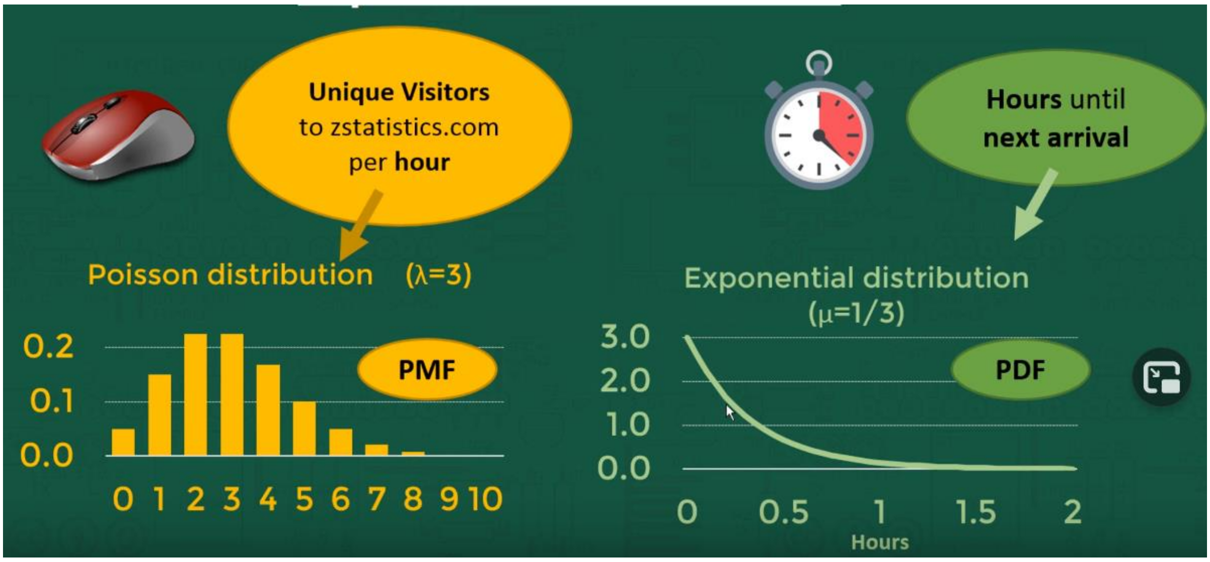
3) Exponential Distribution





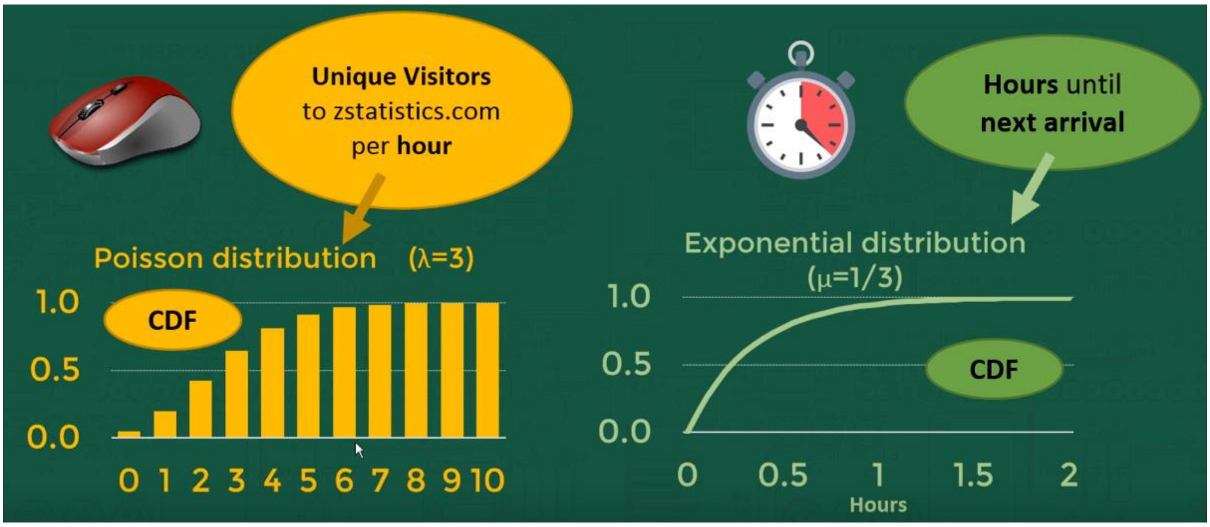


Probability Mass function and Probability Distribution function



Number of people

Cumulative Distribution function

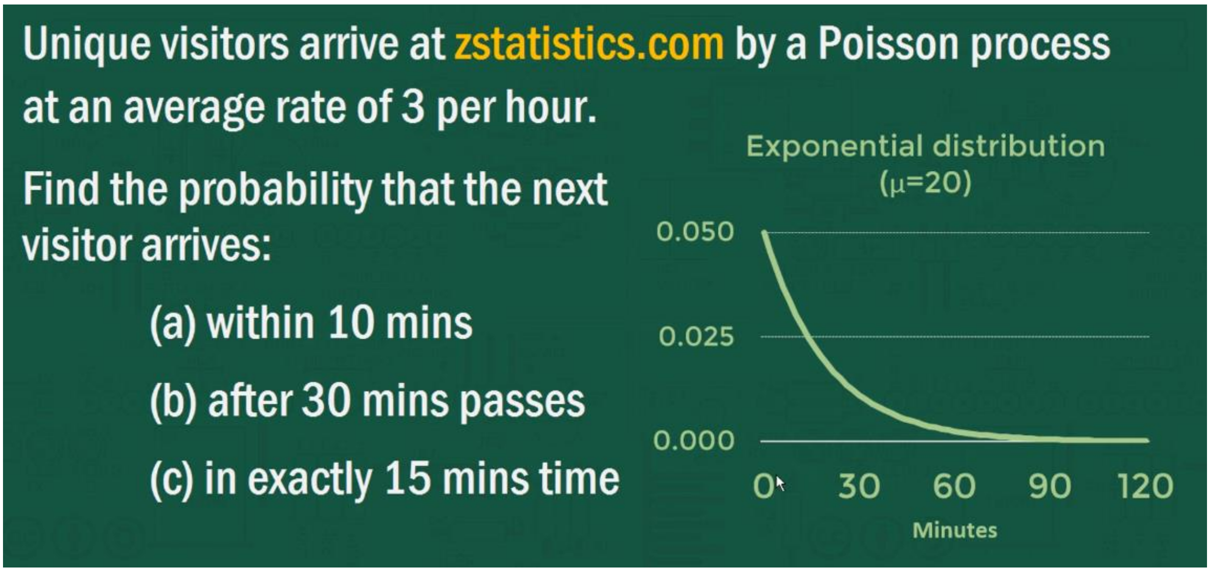


Number of people

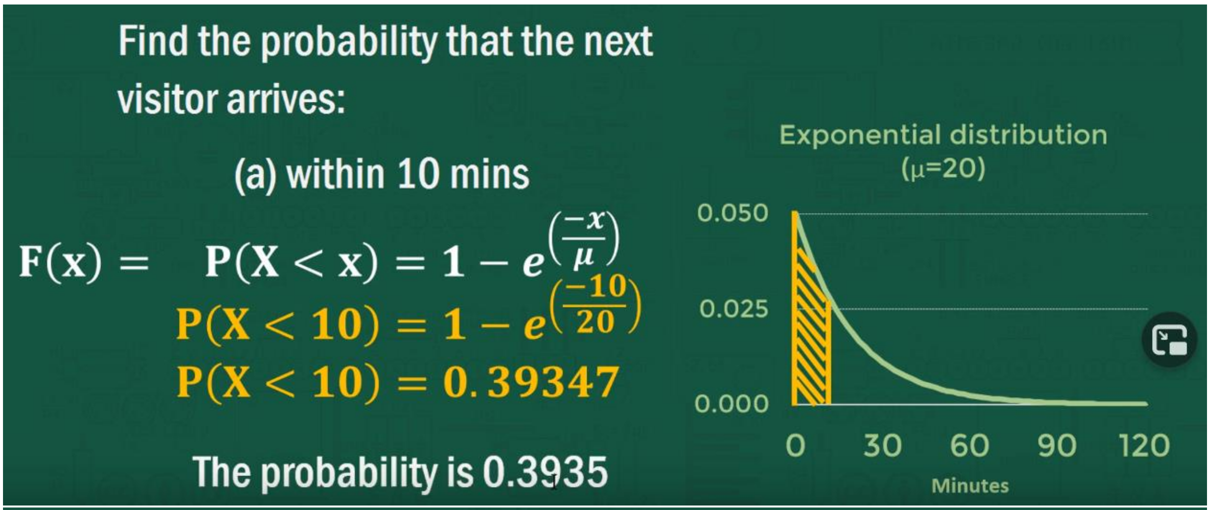
Why poison is a discrete distribution and exponential is continuous distribution??

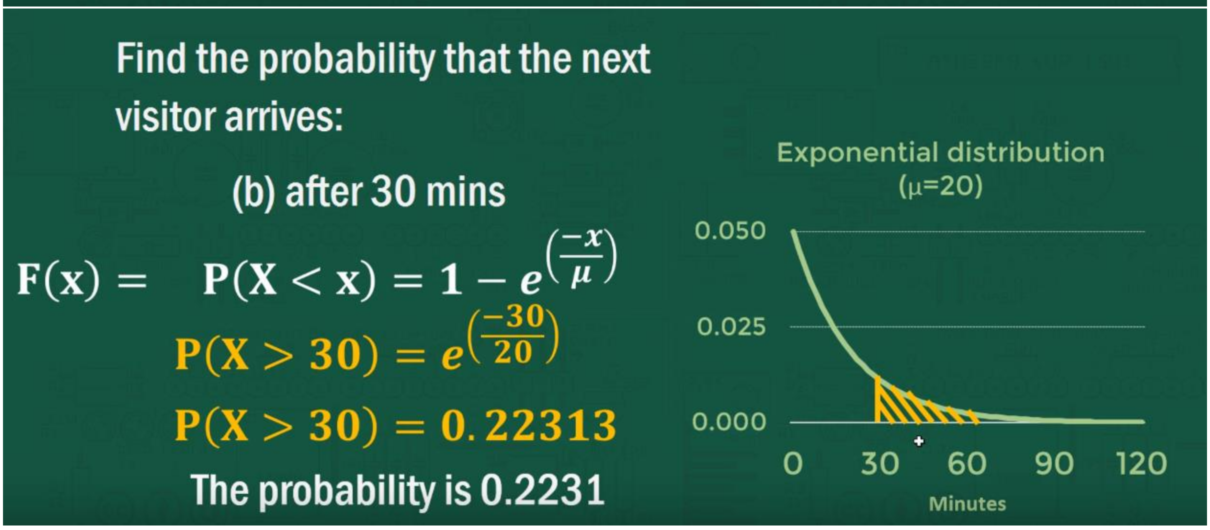
Number of people is fixed → 1/5/200 but number of hour → 3.59099 hrs

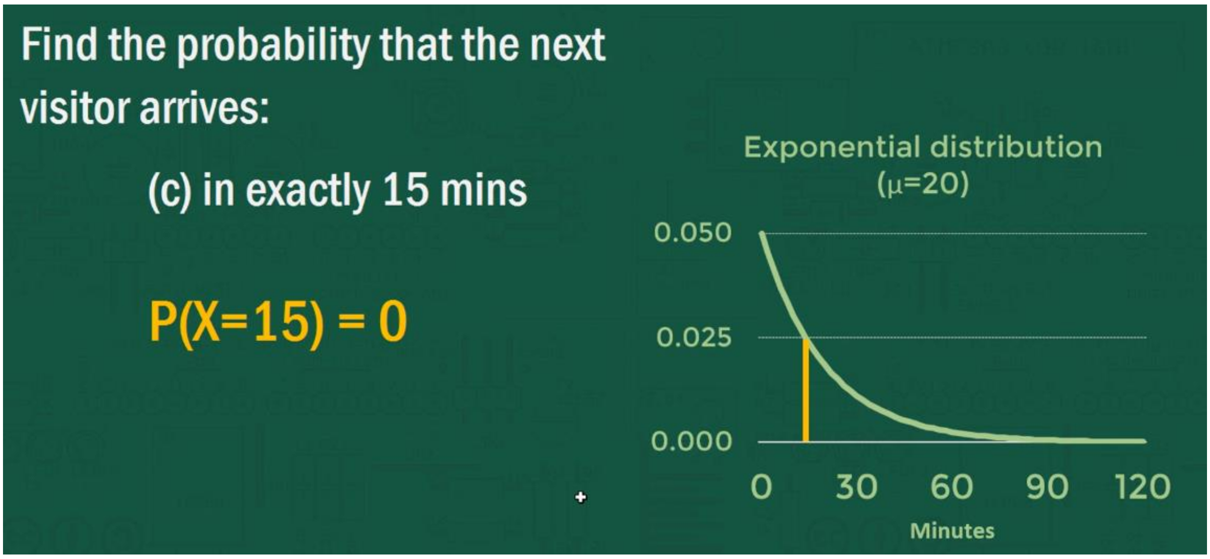
Question



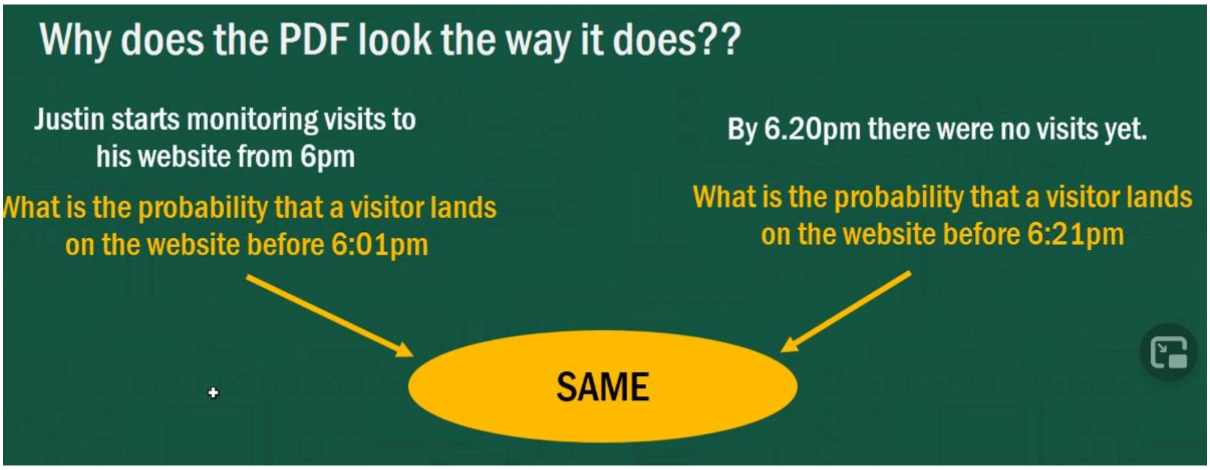
20 minutes is the average time between visits to the web site Zstatistics.com





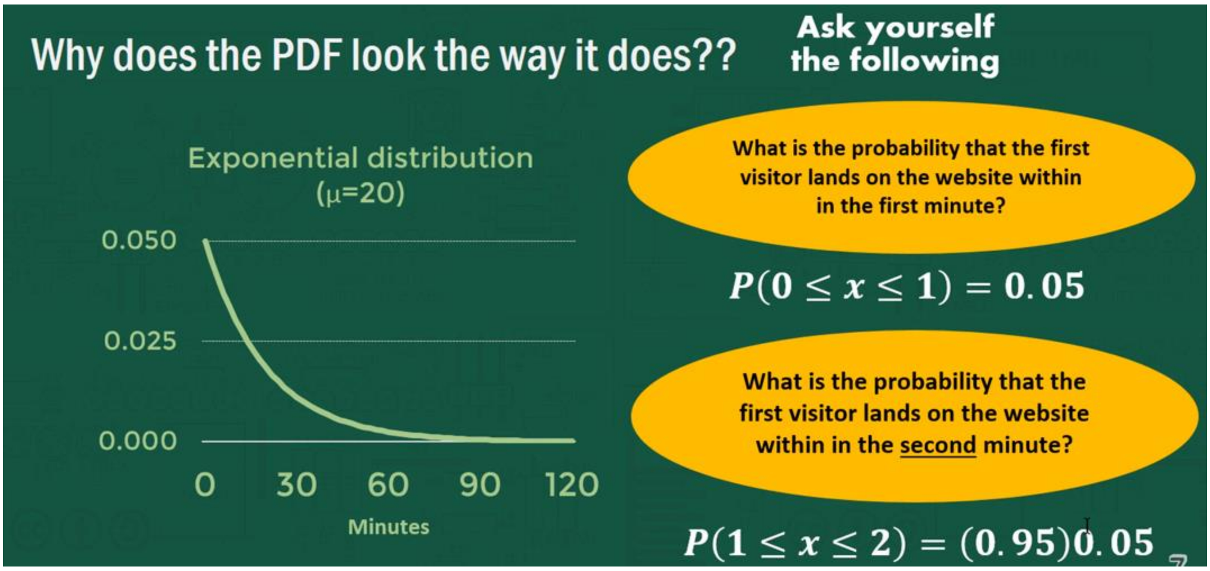


We can't able to finish a task in exactly in 15mins (i.e 15.000000000000minute) We can complete before it / after it but not exact. That is the biggest curse fo continuous distribution.

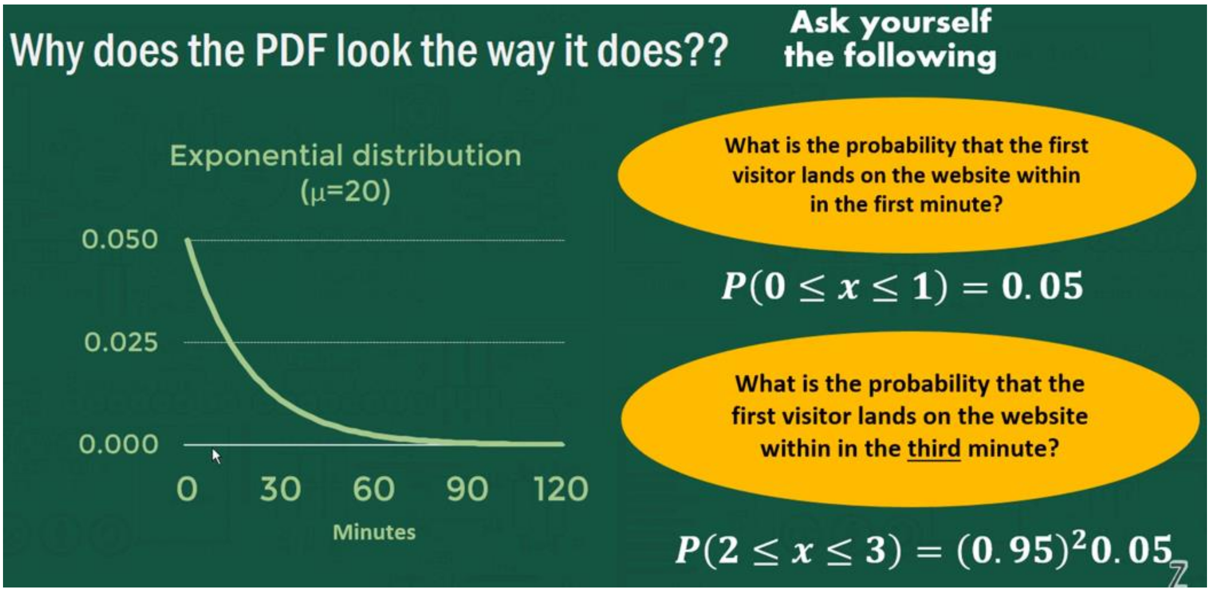


Principle of Memory-lessness in Exponential distribution comes into picture.

Exponential distribution doesn't take care if u have been waiting for 20minutes already with no visits. But the still the same probability in the next minute.



First visitor should not land on the website within the 1st minute and the second visitor must land on the website within the 2nd minute.



For each successive minute the height of the curve will be dropped by 95% and will trail off to almost 0 but never reaching 0.

Exponential distribution

The continuous random variable X has an exponential distribution, with parameter β, if its density function is given by

$$f(x;\beta) = \begin{cases} \frac{1}{\beta} e^{\frac{-x}{\beta}}, & x > 0 \\ 0, & \text{elsewhere}, \end{cases} \tag{1}$$

where $\beta > 0$. The mean and variance of the exponential distribution are $\mu = \beta$ and $\sigma^2 = \beta^2$. Put $\lambda = \frac{1}{\beta}$in the equation (1). Then

$$f(x;\lambda) = \begin{cases} \lambda e^{-x\lambda}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Exponential Distribution:

Definition: A continuous random variable X is said to follow an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$f(x|\lambda) = \lambda e^{-\lambda x}$, $x > 0$, $\lambda > 0$.

It is also known as negative exponential distribution.

- Mean of the exponential distribution is $1/_{\lambda}$.
- Variance of the exponential distribution is $1/_{\lambda^2}$.

Problem : The life of an electric bulb is exponentially distributed with failure rate λ = 1/3 (one failure in every 3000 hours on the average)

Find **(a)** the probability that the lamp will last between 2000 and 3000 hours

(b) P[1.5 ≤ x ≤ 3.0] **(c)** P[x > 3.5 / x > 2.5]

(a) **Let X denote the life time of an electric bulb**

$$P[X > 3] = \int_3^\infty \frac{1}{3} e^{-(x/3)} dx = \frac{1}{3}\left[-\frac{e^{-(x/3)}}{1/3}\right]_3^\infty = e^{-1} = 0.3679$$

(b) $$P[1.5 \le X \le 3.0] = \int_{1.5}^{3.0} \frac{1}{3} e^{-(x/3)} dx = \left[-e^{-x/3}\right]_{1.5}^{3.0} = \left[e^{-1/2} - e^{-1}\right] = 0.2386$$

(c) $$P[X > 3.5 / X > 2.5] = P[X > 1] = \int_1^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = \left| e^{-\frac{1}{3}} \right| = 0.717$$

Example:

The mileage which car owners get with a certain kind of radial tire is a RV having an exponential distribution with mean 40,000 km. Find the probabilities that one of these tires will last (i) at least 20,000 km and (ii) at most 30,000 km.

Solution:

Let X denote the mileage obtained with the tire

$$f(x) = \frac{1}{40,000} e^{-x/40,000} \quad x > 0$$

$$\text{(i) } P(X \geq 20,000) = \int_{20,000}^{\infty} \frac{1}{40,000} e^{-x/40,000} dx$$

$$= \left[-e^{-x/40,000}\right]_{20,000}^{\infty}$$

$$= e^{-0.5} = 0.6065$$

$$\text{(ii) } P(X \leq 30,000) = \int_{0}^{30,000} \frac{1}{40,000} e^{-x/40,000} dx$$

$$= \left[-e^{-x/40,000}\right]_{0}^{30,000}$$

$$= 1 - e^{-0.75} = 0.5270$$

Exercise:

The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = 1/2$.

(a) What is the probability that the repair time exceeds 2 h?

(b) What is the conditional probability that a repair takes at least 10 h given that its duration exceeds 9 h?

HINT:

$X \sim ED(\lambda)$, with $\lambda = 1/2$

(a) $P(X > 2)$

(b) $P(X \geq 10 / X \geq 9) = \dfrac{P(X \geq 10, X \geq 9)}{P(X \geq 9)}$

Given: $\lambda \Rightarrow 1/2$

Solution:

a) $P(X>2) \Rightarrow 1 - P(X \le 2)$

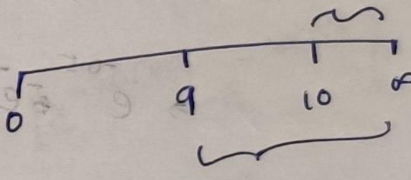
$$\Rightarrow 1 - \int_0^2 \frac{1}{2} e^{-\frac{1}{2}x}\, dx$$

$$\Rightarrow 1 - \frac{1}{2}\left[\frac{e^{-\frac{1}{2}x}}{-\frac{1}{2}}\right]_0^2$$

$$\Rightarrow 1 - \frac{1}{2}\left(\frac{-2}{1}\right)\left[e^{-1} - e^{-0}\right]$$

$$\Rightarrow 1 + \left[e^{-1} - 1\right]$$

$$\boxed{P(X>2) \Rightarrow 0.36787}$$



b) $P(X \geq 10 \mid (X \geq 9))$

$$\hookrightarrow \frac{P(X \geq 10) \cap P(X \geq 9)}{P(X \geq 9)} \Rightarrow \frac{P(X \geq 10)}{P(X \geq 9)}$$

$$P(X \geq 10) \Rightarrow \int_{10}^{\infty} \frac{1}{2} e^{-\frac{1}{2}x}\, dx$$

$$\Rightarrow \frac{1}{2}\left[\frac{e^{-\frac{1}{2}x}}{-\frac{1}{2}}\right]_{10}^{\infty}$$

$$\Rightarrow -\left[e^{-\frac{1}{2}x}\right]_{10}^{\infty} \Rightarrow -\left[e^{-\infty} - e^{-5}\right]$$

$P(X \geq 10) \Rightarrow 0.0067379$

$$P(X \geq 9) \Rightarrow \int_{9}^{\infty} \frac{1}{2} e^{-\frac{1}{2}x}\, dx$$

$$\Rightarrow \frac{1}{2} \left[\frac{e^{-\frac{1}{2}x}}{-\frac{1}{2}} \right]_{9}^{\infty}$$

$$\Rightarrow - \left[e^{-\frac{1}{2}x} \right]_{9}^{\infty}$$

$$\Rightarrow - \left[e^{-\infty} - e^{-4.5} \right] \Rightarrow e^{-4.5}$$

$$P(X \geq 9) \Rightarrow 0.0111089$$

$$P(X \geq 10 \mid X \geq 9) \Rightarrow \frac{P(X \geq 10)}{P(X \geq 9)} \Rightarrow \frac{e^{-5}}{e^{-4.5}} \Rightarrow e^{-5+4.5} \Rightarrow e^{-0.5}$$

$$\boxed{P(X \geq 10 \mid X \geq 9) \Rightarrow 0.60653}$$

4) Gamma Distribution

A CRV X is said to follow gamma distribution if its p.d.f. is given by

$$f(x) = \begin{cases} \dfrac{a^m x^{m-1} e^{-ax}}{\Gamma m} & ; x \geq 0 \\ 0 & ; \text{otherwise} \end{cases}$$

Here X is a gamma variate with parameter a and m.

$\Gamma m, a$ or $G(m, a)$

As f(x) is a probability density function

1) $f(x) \geq 0$

2) $\displaystyle\int_{-\infty}^{+\infty} f(x)\, dx = 1$

$\hookrightarrow \displaystyle\int_{-\infty}^{0} \Rightarrow 0 + \int_{0}^{\infty} \Rightarrow 1$

$$f(x) \Rightarrow \int_0^\infty \frac{a^m x^{m-1} e^{-ax}}{\Gamma m}\, dx$$

$x \Rightarrow 0$ to ∞ y also 0 to ∞

Let $y = ax$; $x = y/a$

$dy = a\,dx$

$dx = \frac{dy}{a}$

$$f(x) \Rightarrow \int_0^\infty \frac{a^m x^{m-1} e^{-ax}}{\Gamma m}\, dy/a$$

$$f(x) \Rightarrow \int_0^\infty \frac{a^m (y/a)^{m-1} e^{-y}}{\Gamma m}\, dy/a$$

$$\boxed{\Gamma(m) \Rightarrow \int_0^\infty (x^{m-1} e^{-x})\,dx}$$

Gamma function

$$\Rightarrow \frac{a^m}{\Gamma m} \int_0^\infty \frac{\frac{y^{m-1}}{a^{m-1}}}{1} e^{-y}\, dy/a$$

$$f(x) \Rightarrow \frac{\cancel{a^m}}{\Gamma m} \int_0^\infty \frac{y^{m-1} e^{-y}}{\cancel{a^m} a^{-1}}\, dy/a \Rightarrow \frac{1}{\Gamma m} \int_0^\infty y^{m-1} e^{-y} \cancel{a} \left(\frac{dy}{\cancel{a}}\right)$$

$$f(x) \Rightarrow \frac{1}{\Gamma m} \int_0^\infty (e^{-y} y^{m-1})\, dy \Rightarrow \frac{\Gamma m}{\Gamma m} \Rightarrow 1$$

$\therefore$ $f(x)$ is a pdf