manipalglobal
Academy of Data Science

Statistical Techniques for Data Science

Probability

Probability Distribution and Central Limit Theorem

Dr. Subhabaha Pal
Manipal Global Academy of Data Science

Objective

manipalglobal
Academy of Data Science

After attending this session, you will be able to –

- **Explain what is Probability Distribution**
- **Describe Continuous and Discrete Probability Distribution**
- **Describe Binomial and Poisson Distribution**
- **Describe Normal Distribution**
- **Describe Central Limit Theorem**

Probability Distribution

- A probability distribution is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range
- This range will be between the minimum and maximum statistical possible values
- When the values of the random variable are discrete (i.e., it takes only specific values within a range), it is termed as discrete probability distribution
- When the values of the random variable are continuous (i.e., it can take all possible values within a range), it is termed as continuous probability distribution

Discrete Probability Distribution - Example

- A mechanic is examining 3 RAM chips and is detecting whether these chips are non-defective (N) or defective (D)
- The sample space for all possible outcomes are as follows –

$$\Omega = \{NNN,\ NND,\ NDN,\ DNN,\ NDD,\ DDN,\ DND,\ DDD\}$$

- The random variable is defined as X = No. of defectives found out of the 3 RAM chips

Based on the Definition X = No. of defective RAM chips

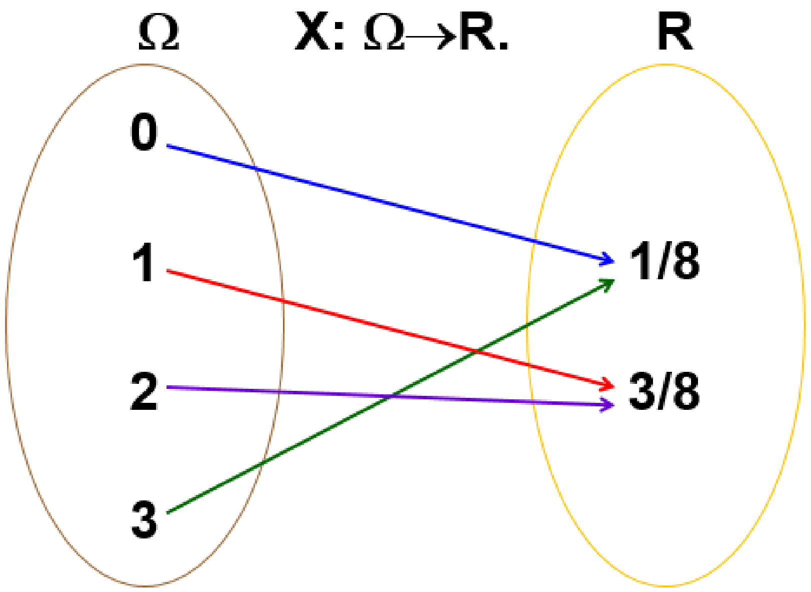
<table>
<tr><th>Outcome</th><th>X</th><th>Frequency (f)</th><th>Probability</th></tr>
<tr><td>NNN</td><td>0</td><td>1</td><td>1/8</td></tr>
<tr><td>NND</td><td>1</td><td rowspan="3">3</td><td rowspan="3">3/8</td></tr>
<tr><td>NDN</td><td>1</td></tr>
<tr><td>DNN</td><td>1</td></tr>
<tr><td>NDD</td><td>2</td><td rowspan="3">3</td><td rowspan="3">3/8</td></tr>
<tr><td>DND</td><td>2</td></tr>
<tr><td>DDN</td><td>2</td></tr>
<tr><td>DDD</td><td>3</td><td>1</td><td>1/8</td></tr>
</table>

Discrete Probability Distribution - Example

- We can write the probability distribution of the random variable X (No. of defectives found out of the 3 RAM chips) as follows –

No. of RAM chips defective	Probability
0	⅛
1	⅜
2	⅜
3	⅛
Total	1

Probability distribution



- It can be written as P(X=0) = 1/8, P(X=1)=3/8, P(X=2)=3/8, P(X=3)=1/8, and we get that p(0)+p(1)+p(2)+p(3)=1, and each p(x)≥0 where x = 0, 1, 2, 3. Here, p(x) = P(X=x)
- The values assumed by the random variable X above are discrete and it is of the form p(x) = P(X=x), which is also called probability mass function

Discrete Probability Distribution – Example

manipalglobal
Academy of Data Science

- Discrete probability distributions can be estimated from relative frequencies
- Number of television sets per household (X) from a sample data of around 1 Lakh house-holds are present below in the following table -

No. of televisions	No. of households	X	P(x)
0	1,218	0	0.012
1	32,379	1	0.319
2	37,961	2	0.374
3	19,387	3	0.191
4	7,714	4	0.076
5	2,842	5	0.028
Total	101,501		1.000

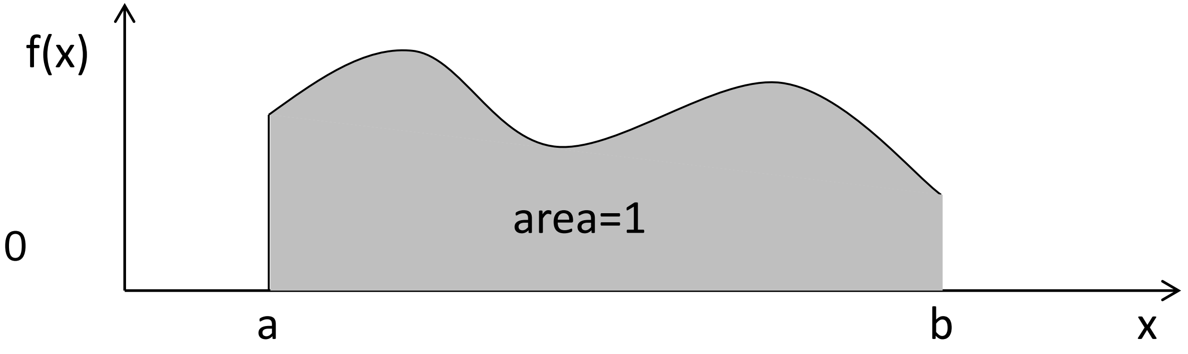
What is the probability that there is at least one television and no more than three in any given household

e.g. $P(X=4) = p(4) = 0.076 = 7.6\%$

"at least one television but no more than three"
$P(1 \leq X \leq 3) = p(1) + p(2) + p(3) = 0.319 + 0.374 + 0.191 = 0.884$

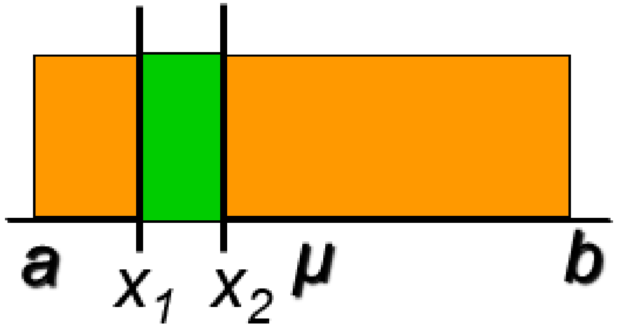
Continuous Probability Distribution

- **The probability distribution in which the random variable X can take on any value within a certain range is termed as Continuous Probability distribution**
- **Since X can take infinite values within a certain range, the probability of X for taking on certain specific value is zero**
- **In case of continuous probability distribution,** we talk about the probability of the random variable assuming a value within a given interval
- The probability of the random variable assuming a value within some given interval from x_1 to x_2 is defined to be the area under the graph of the probability density function between x_1 and x_2
- A function f(x) is called a ***probability density function*** (over the range **a ≤ x ≤ b** if it meets the following requirements –
  1) f(x) ≥ 0 for all **x** between **a** and **b**, and
  2) The total area under the curve between **a** and **b** is 1.0

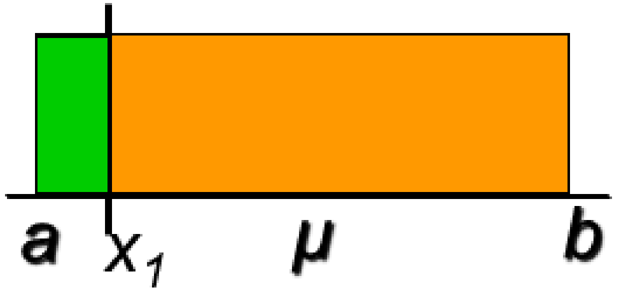


Continuous Probability Distribution – Example

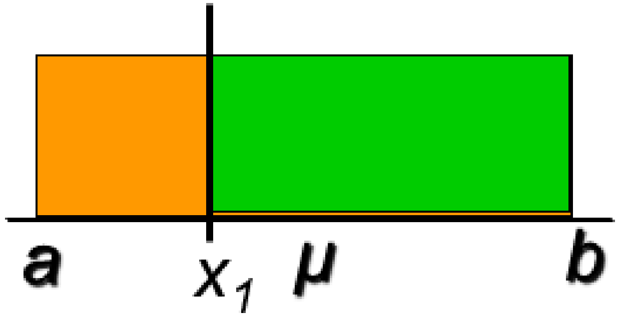
- In case of continuous probability distribution, we talk about the probability of the random variable assuming a value within a given interval



$P(x_1 \leq x \leq x_2)$



$P(x \leq x_1)$



$P(x \geq x_1)$

$P(x \geq x_1) = 1 - P(x \leq x_1)$

Probability Density Function

- A probability density function (PDF) or density of a continuous random variable is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal to the sample
- The PDF is used to specify the probability of the random variable falling within a particular range of values, as opposed to taking on any one value
- The probability is given by the integral of the variable's PDF over the range, i.e., it is given by the area under the density function but above the horizontal axis and between the lowest and greatest values of the range
- The PDF is non-negative everywhere and its integral over the entire space is equal to one
- Suppose X is a continuous variable taking possible values in a certain range and its PDF to be f(x), then the probability that the variable X will fall within a and b (two points within the range of continuous variable X) can be written as –

  F(x) = P(a≤X≤b) = $\int_a^b f(x)\,\mathrm{dx}$
- If the range of X is a to b, then $\int_a^b f(x)\,\mathrm{dx} = 1$

Discrete Probability Distributions

- **We have learnt about Discrete and Continuous Probability distributions in the previous section**
- **Now we will be learning about the following discrete probability distributions in the upcoming section –**
  - **Bernoulli Distribution**
  - **Binomial Distribution**
  - **Poisson Distribution**

Bernoulli Distribution

manipalglobal
Academy of Data Science

- The Bernoulli distribution is the probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability q = 1-p, i.e., the probability distribution of any single experiment that asks a yes-no question (yes makes the value of the variable 1 whereas no makes the value of the variable as 0)
- If X is a random variable with Bernoulli distribution. Then –

  $Pr(X=1) = p = 1-Pr(X=0) = 1-q$
- If X follows a Bernoulli distribution, probability mass function of X can be written as –

$P(X=x) = p^x q^{(1-x)}$, $x = 0, 1$; When $x = 0$, $P(X=0) = p^0 q^{(1-0)} = 1 \times q = q$ and when $x = 1$, $P(X=1) = p^1 q^{(1-1)} = p \times 1 = p$

- If X follows a Bernoulli distribution, the number of successes from a single trial is written as -

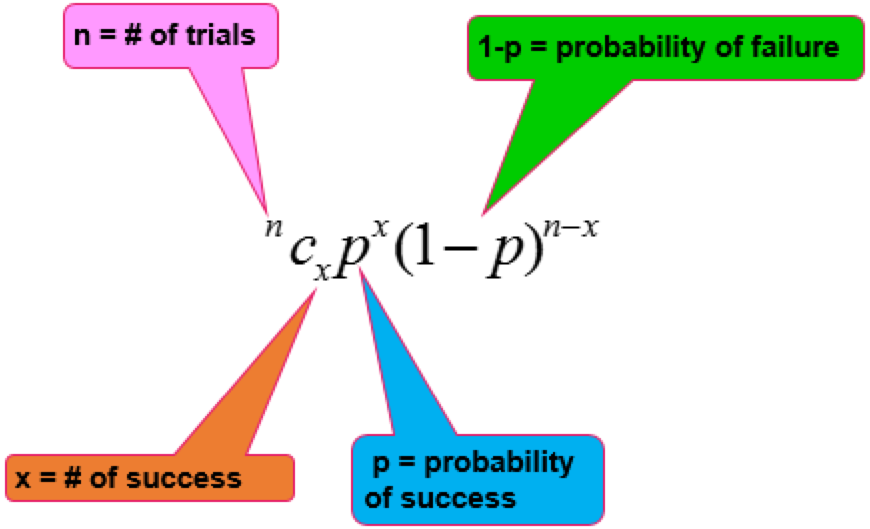
$$p(x) = P(X = x) = p^x q^{1-x}, x = 0, 1$$

- Mean of Bernoulli variable X is p and Variance of Bernoulli variable X is given by pq

Binomial Distribution

- Suppose n independent trials are performed where n is a fixed number and that each experiment results in a success with probability p and failure with probability q = 1-p, then the total number of successes X is a binomial random variable with parameters n and p
- It is generally written as : *X ~ Bin (n, p) {reads: "X is distributed binomially with parameters n and p}*
- The probability that *X* = *x* (i.e., that there are exactly *x* successes) is given by –

$$P(X = x) = \binom{n}{x} p^{x} (1 - p)^{n-x}$$



- The above is the expression for probability mass function of X
- Mean of Binomial Variable X is np and variance of variable X is np(1-p)
- Bernoulli distribution is a special case of Binomial Distribution when n is equal to 1

Binomial Distribution – Example

manipalglobal
Academy of Data Science

- If a coin is tossed 6 times, what is the probability of getting 2 or fewer heads?

Solution – Here, X is the number of successes, i.e., the number of times the head is obtained.

Hence, it is needed to consider the probabilities when X will be 0, 1 or 2.

$$P(X=0)=\binom{6}{0}(0.5)^0(0.5)^6=\frac{6!}{6!\,0!}(0.5)^6=0.015625$$

$$P(X=1)=\binom{6}{1}(0.5)^1(0.5)^5=\frac{6!}{5!\,1!}(0.5)^6=0.09375$$

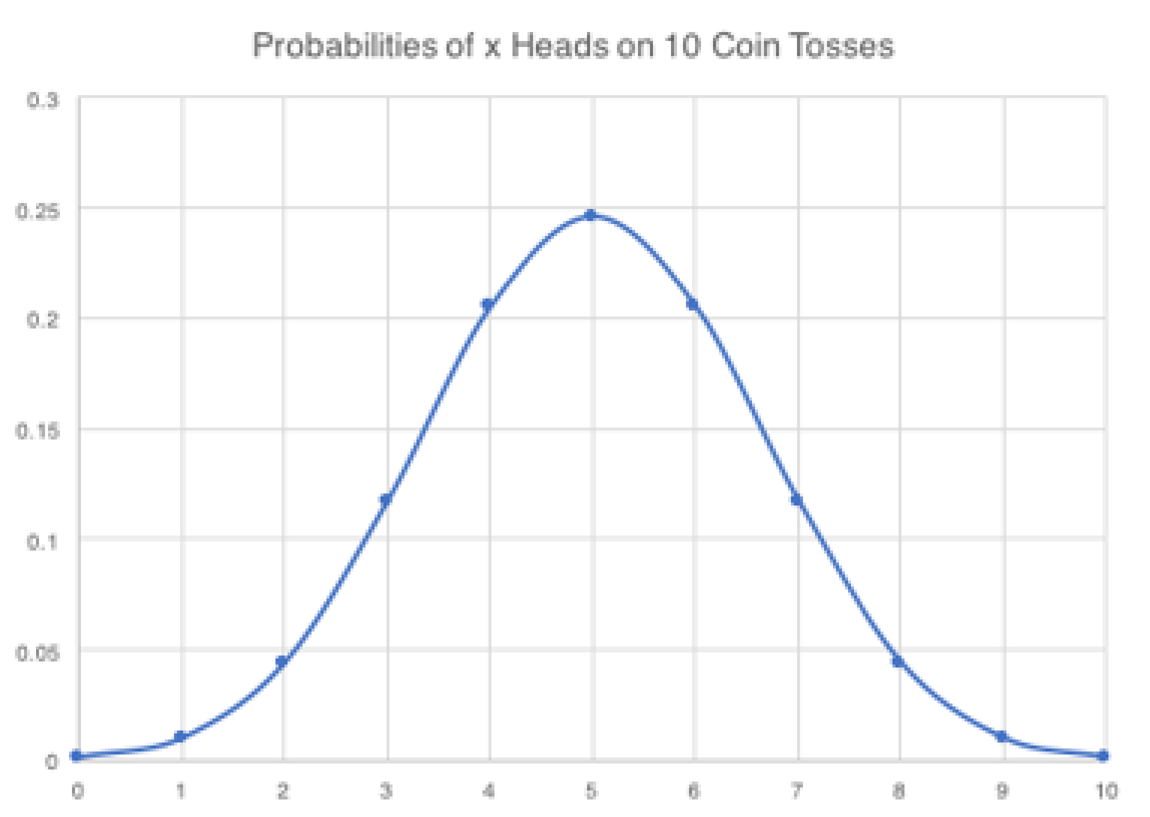
$$P(X=2)=\binom{6}{2}(0.5)^1(0.5)^4=\frac{6!}{4!\,2!}(0.5)^6=0.078125$$

Hence, $P(x \le 2) = \sum p(x) = 0.015625 + 0.09375 + 0.078125 = 0.1875$

Binomial Distribution – Example

A coin is tossed 10 times…	
Getting x heads	Probability of getting
0	0.000976563
1	0.009765625
2	0.043945313
3	0.1171875
4	0.205078125
5	0.24609375
6	0.205078125
7	0.1171875
8	0.043945313
9	0.009765625
10	0.000976563

```
> dbinom (0:10, 10, 0.5)
 [1] 0.0009765625 0.0097656250 0.0439453125 0.1171875000 0.2050781250
 [6] 0.2460937500 0.2050781250 0.1171875000 0.0439453125 0.0097656250
[11] 0.0009765625
>
```



Binomial Distribution – Example

A group of 20 people at a party were asked if they preferred red wine or white wine.

What is the probability that at the Saturday evening party, 8 people preferred red wine?

12.013%

```
Console ~/
> dbinom (8, size = 20, prob = 0.5)
[1] 0.1201344
>
```

What is the probability that at the party, upto 10 people preferred white wine?

58.809%

```
Console ~/
> pbinom (10, size = 20, prob = 0.5)
[1] 0.5880985
> pbinom (10, size = 20, prob = 0.5, lower.tail = TRUE)
[1] 0.5880985
>
```

Poisson Distribution

- If the number of trials are very large, i.e. n tends to ∞ and probability of success p is very small, i.e., p tends to 0, then the distribution of number of successes is said to follow Poisson distribution
- Probability that number of successes X occurs in an infinite sequence of trials is

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

- If X follows Poisson distribution with parameter λ, i.e. X ~ P(λ), its probability mass function is given by –

$$p(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda} \lambda^x}{x\,!}, & x = 0,\ 1,\ 2,\ \ldots \\ 0, & \text{elsewhere} \end{cases}$$

- Mean and variance of X (when X follows Poisson distribution) are both equal to λ

Poisson Distribution - Example

- The manager of a industrial plant is planning to buy a machine of either type A or type B. For each day's operation the number of repairs X, that the machine A needs is a poisson random variable with mean 0.96. The daily cost of operating A is $C_A = 160 + 40 * X^2$. For machine B, let Y be the random variable indicating the number of daily repairs, which has mean 1.12, and the daily cost of operating B is $C_B = 128 + 40 * Y^2$. Assume that the repairs take negligible time and each night the machine are cleaned so that they operate like new machine at the start of each day. Which machine minimizes the expected daily cost ?

Poisson Distribution - Solution

Solution – We know, $V(X) = E(X-E(X))^2 = E(X^2) – 2.E(X).E(X) + (E(X))^2 = E(X^2) – (E(X))^2$

Or, $E(X^2) = V(X) + E(X^2)$.

The expected cost for machine A is, $E(C_A) = E(160 + 40 * X^2) = 160 + 40 * E(X^2) = 160 + 40 * (V(X) + [E(X)]^2) = 160 + 40 * (0.96 + 0.96^2) = 235.264$

The expected cost for machine B is, $E(C_B) = E(128 + 40 * Y^2) = 128 + 40 * E(Y^2) = 128 + 40 * (V(Y) + [E(Y)]^2) = 128 + 40 * (1.12 + 1.12^2) = 222.976$

Hence, the machine type B minimizes expected daily cost.

Poisson Distribution - Example

- The number of calls coming per minute into a hotels reservation center is Poisson random variable with mean 3.

(a) Find the probability that no calls come in a given 1 minute period.

(b) Assume that the number of calls arriving in two different minutes are independent. Find the probability that at least two calls will arrive in a given two minute period.

Poisson Distribution - Solution

Solution. (a) Let X denotes the number of calls coming in that given 1 minute period. $X \sim Poisson(3)$
$P(X = 0) = e^{-3} (3^0/0!) = e^{-3}$

(b) Let X_1 and X_2 be the number of calls coming in the first and second minutes respectively.

We want $P(X_1 + X_2 \geq 2)$. $P(X_1 + X_2 \geq 2) = 1 - P(X_1 + X_2 < 2) = 1 - [P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 0) + P(X_1 = 0, X_2 = 1)] = 1 - [P(X_1 = 0) * P(X_2 = 0) + P(X_1 = 1) * P(X_2 = 0) + P(X_1 = 0) * P(X_2 = 1)] = 1 - e^{-3} e^{-3} - e^{-3}e^{-3}(3^1/1!) - e^{-3}(3^1/1!) e^{-3} = 1 - e^{-6} - 3e^{-6} - 3e^{-6} = 1 - 7e^{-6}$

Poisson Distribution

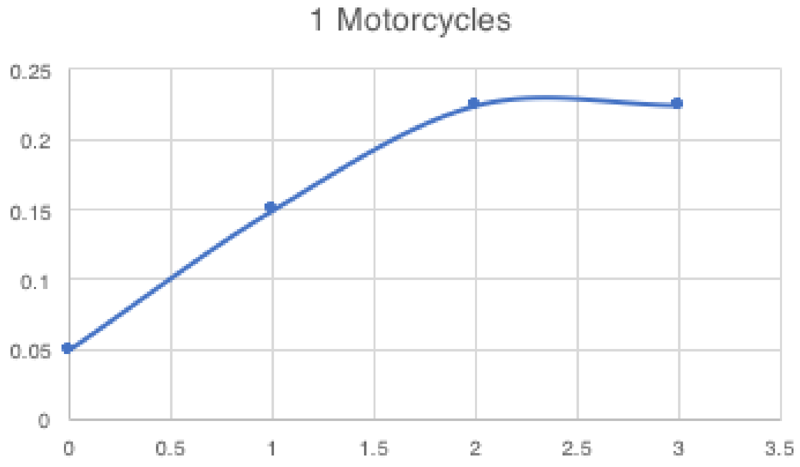
- R > dpois (0:x, lambda = l)
- R > ppois (0:x, lambda = l)

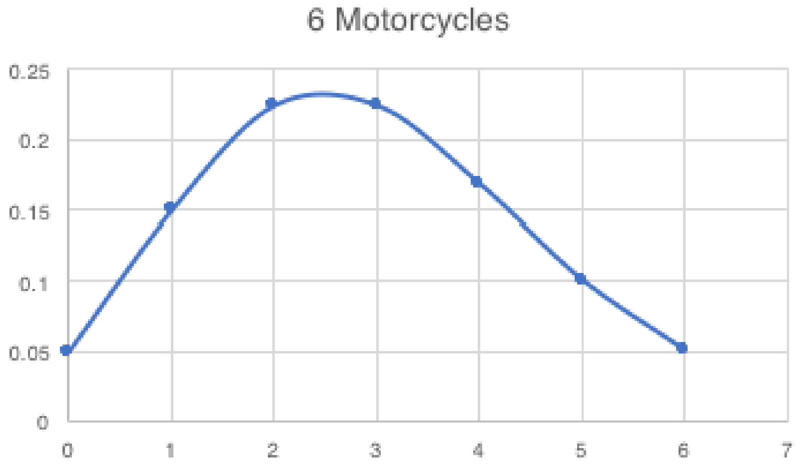
A dealer sells 3 motorcycles on a given day		
x	Probability	C. Probability
=POISSON.DIST(C4,3,FALSE)		
1	0.14936121	0.19914827

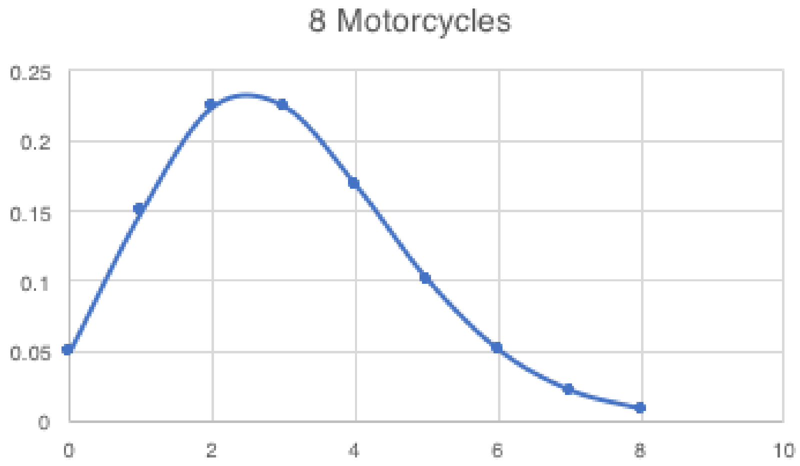
```
Console ~/
> dpois (0:10, lambda = 3)
 [1] 0.0497870684 0.1493612051 0.2240418077 0.2240418077 0.1680313557
 [6] 0.1008188134 0.0504094067 0.0216040315 0.0081015118 0.0027005039
[11] 0.0008101512
> ppois (0:10, lambda = 3)
 [1] 0.04978707 0.19914827 0.42319008 0.64723189 0.81526324 0.91608206
 [7] 0.96649146 0.98809550 0.99619701 0.99889751 0.99970766
>
```

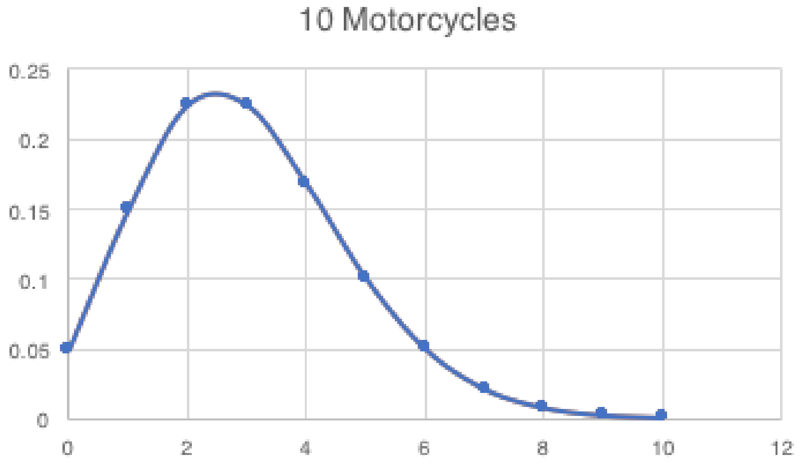
Poisson Distribution

A dealer sells 3 motorcycles on a given day	
x	**Probability**
0	0.04978707
1	0.14936121
2	0.22404181
3	0.22404181
4	0.16803136
5	0.10081881
6	0.05040941
7	0.02160403
8	0.00810151
9	0.00270050
10	0.00081015









Poisson Distribution - Examples

Kumar makes an average of 4 fountain pens per day on his lathe.

What is the probability that on December 5, 2016, he will make 3 pens?

19.537%

```
Console ~/
> dpois (3, lambda = 4)
[1] 0.1953668
>
```

What is the probability that on December 5, 2016, he will make less than 3 pens?

43.347%

```
Console ~/
> ppois (3, 4)
[1] 0.4334701
> ppois (3, lambda = 4, lower.tail = TRUE)
[1] 0.4334701
>
```

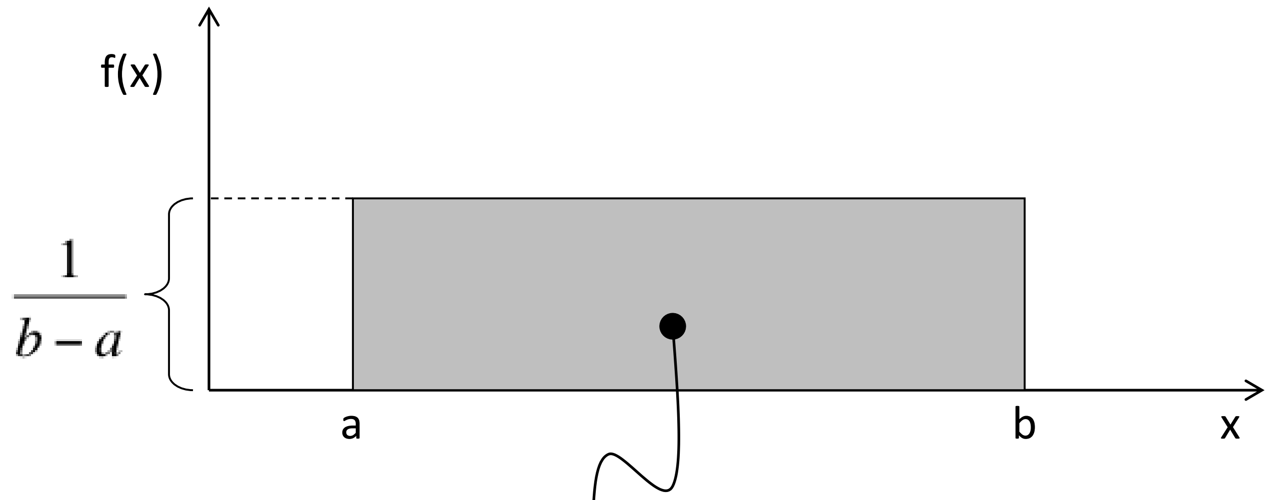
Continuous Probability Distributions

- We have learnt about few Discrete Probability distributions in the previous section
- Now we will be learning about few continuous probability distributions in the upcoming section –
  - Uniform Distribution
  - Normal Distribution
  - Exponential Distribution

Uniform Distribution

- A continuous random variable X is said to follow Uniform distribution, if its PDF f(x) is equal to a constant at all points over the range of X
- The random variable X is said to follow uniform distribution, if –

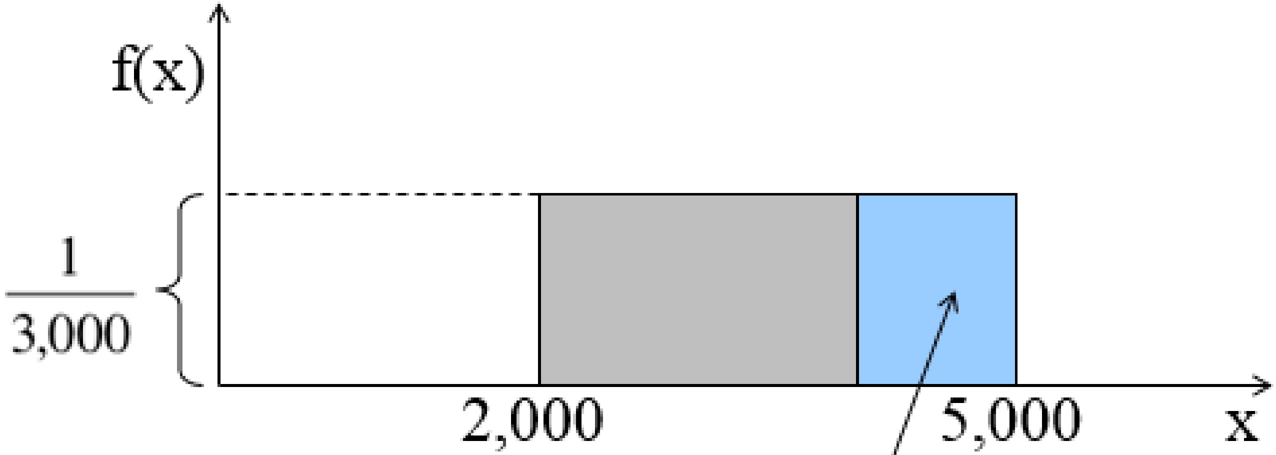
$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b.$$



area = width x height = (b – a) x $\frac{1}{b-a}$ = 1

Uniform Distribution - Example

The amount of gasoline sold daily at a service station is uniformly distributed with a minimum of 2,000 gallons and a maximum of 5,000 gallons.



What is the probability that the service station will sell <u>at least</u> 4,000 gallons?

Algebraically: what is $\mathbf{P(X \geq 4,000)}$?

$\mathbf{P(X \geq 4,000)} = (5,000 - 4,000) \times (1/3000) = .3333$

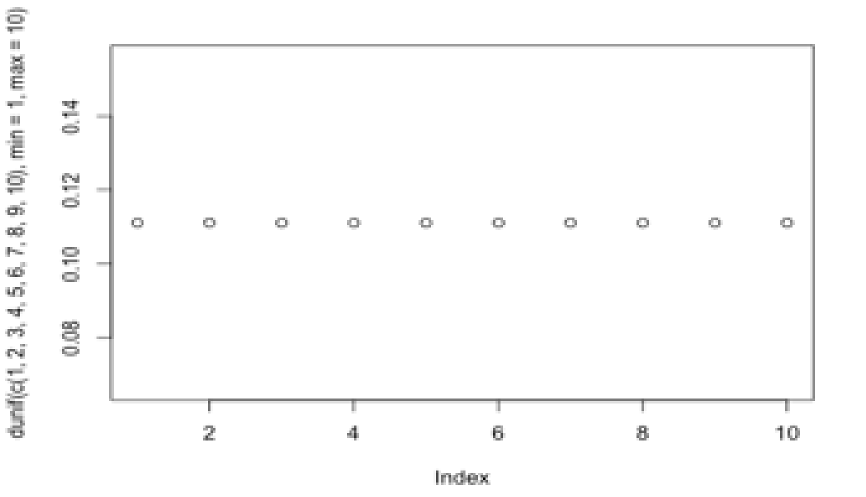
Uniform Distribution - Example

manipalglobal
Academy of Data Science

Uniform Distribution

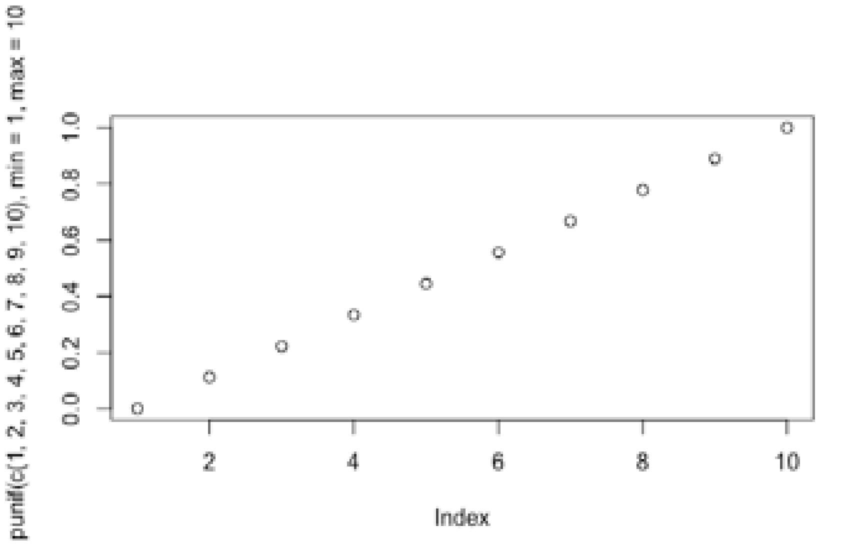
Console ~/

```
> dunif ( c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10), min = 1, max = 10)
 [1] 0.1111111 0.1111111 0.1111111 0.1111111 0.1111111 0.1111111 0.1111111
 [8] 0.1111111 0.1111111 0.1111111
>
```



Console ~/

```
> punif ( c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10), min = 1, max = 10)
 [1] 0.0000000 0.1111111 0.2222222 0.3333333 0.4444444 0.5555556 0.6666667
 [8] 0.7777778 0.8888889 1.0000000
> |
```

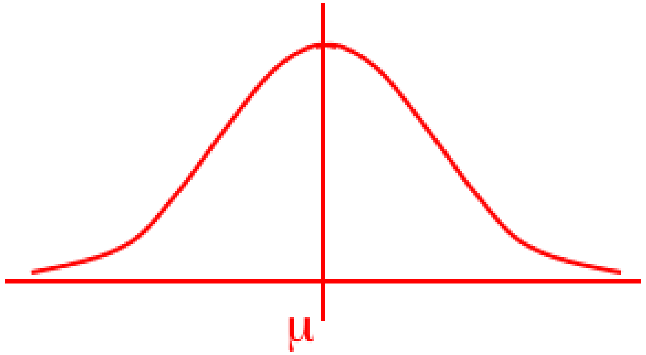


Normal Distribution

- **The Normal distribution is the most important to all probability distributions**
- **The probability density function of a normal random variable is given by –**

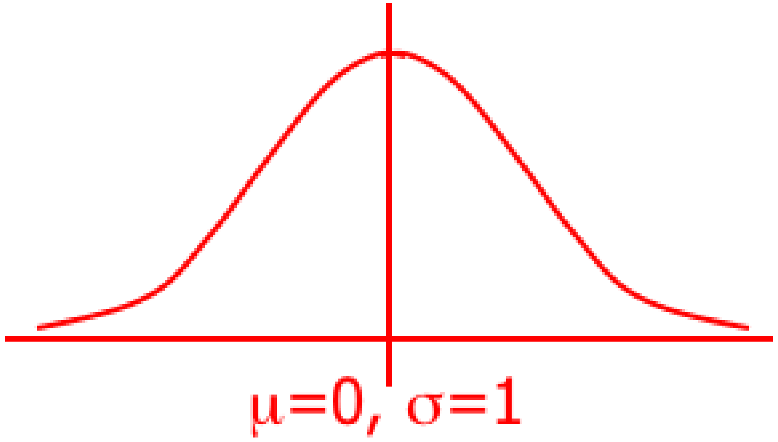
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

- **The probability curve if normal distribution when plotted in a graph looks like bell-shaped curve and it is symmetrical around the mean μ**



- **A normal distribution whose mean is zero and standard deviation is one is called the standard normal distribution**

$$f(x) = \frac{1}{1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-0}{1}\right)^2} \qquad -\infty < x < \infty$$



- **Any normal distribution can be converted to standard Normal distribution**
- **In case of Normal distribution, Mean = Median = Mode**

Normal Distribution

Important things to note:

The normal distribution is fully defined by two parameters:
its standard deviation and mean

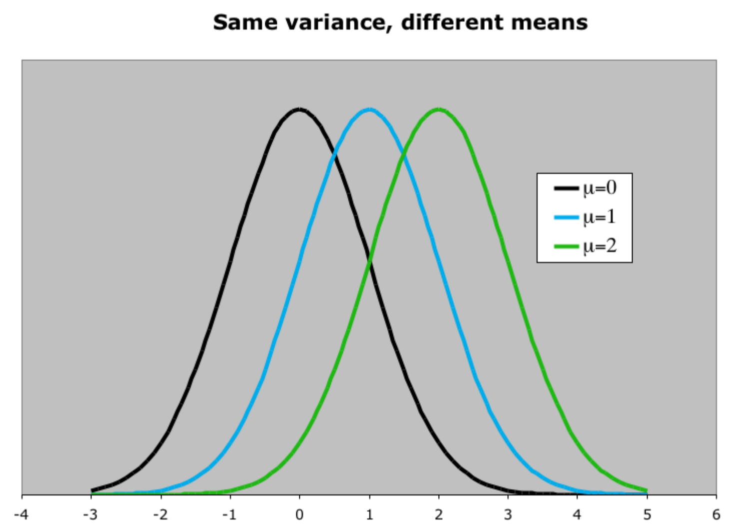
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

The normal distribution is bell shaped and
symmetrical about the mean

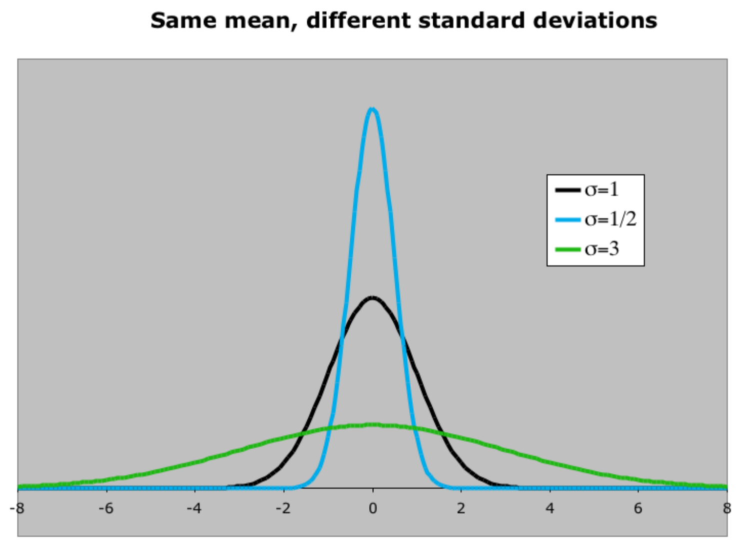
Unlike the range of the uniform distribution ($a \leq x \leq b$)
Normal distributions ***range from minus infinity to plus infinity***

Normal Distribution

➢ **Increasing the mean of the normal distribution shifts the normal curve to the right**

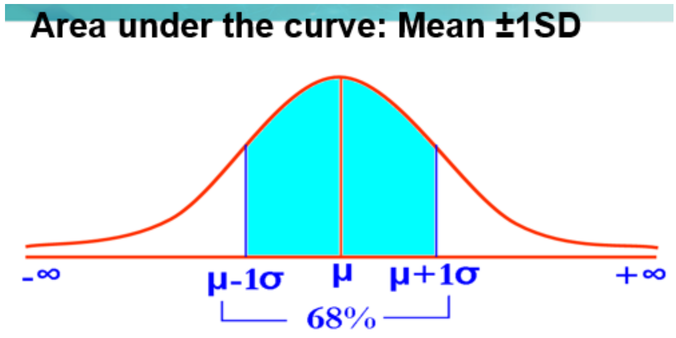


➢ **Increasing the standard deviation of normal distribution Flattens the normal curve**

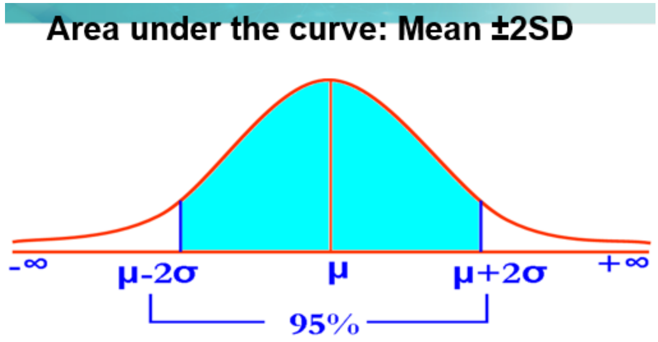


Normal Distribution

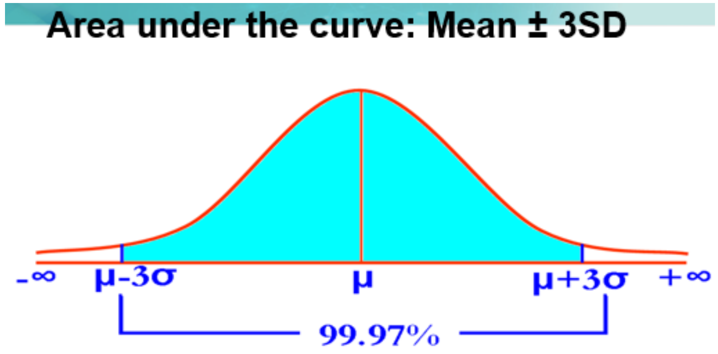
➢ In Normal distribution, 68% of the area under the normal curve lies within the range μ±1σ



➢ In Normal distribution, 95% of the area under the normal curve lies within the range μ±2σ

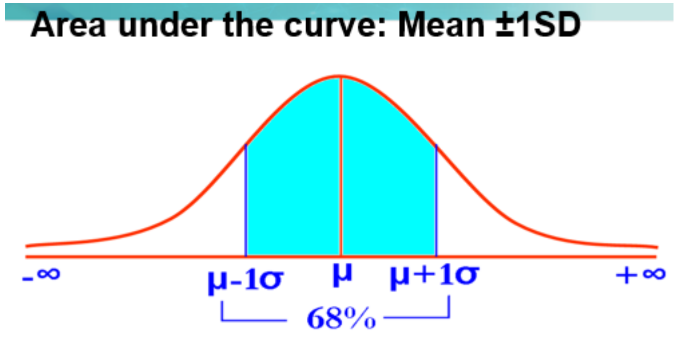


➢ In Normal distribution, 99.97% of the area under the normal curve lies within the range μ±3σ

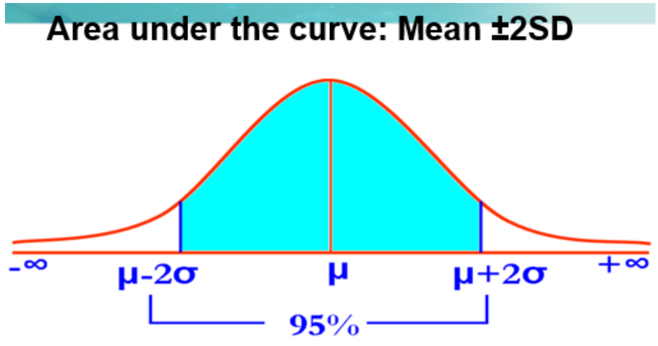


Normal Distribution

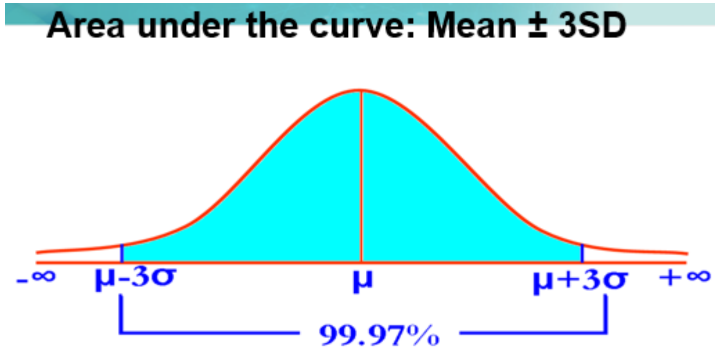
- In Normal distribution, 68% of the area under the normal curve lies within the range $\mu \pm 1\sigma$



- In Normal distribution, 95% of the area under the normal curve lies within the range $\mu \pm 2\sigma$

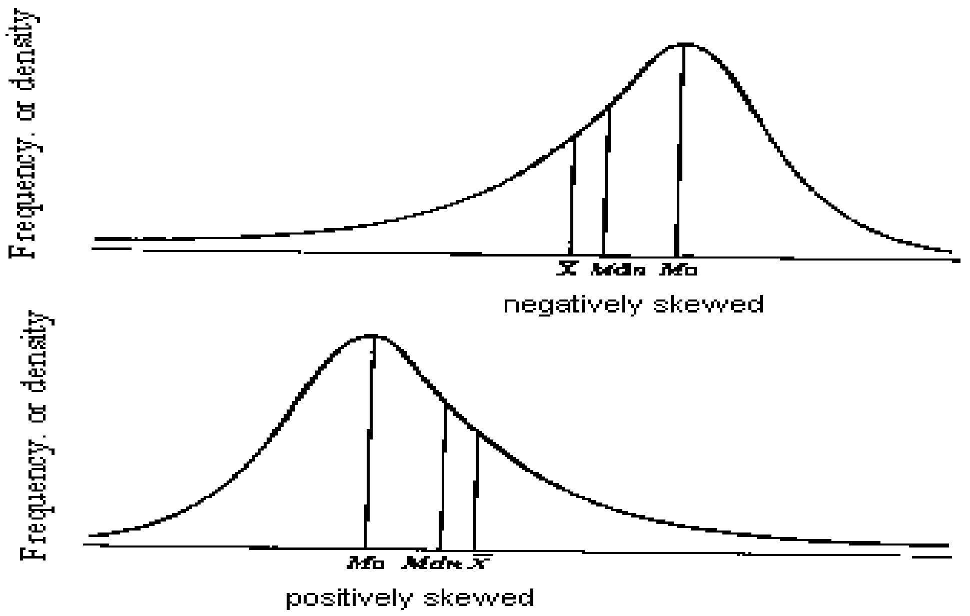


- In Normal distribution, 99.97% of the area under the normal curve lies within the range $\mu \pm 3\sigma$



Skewness

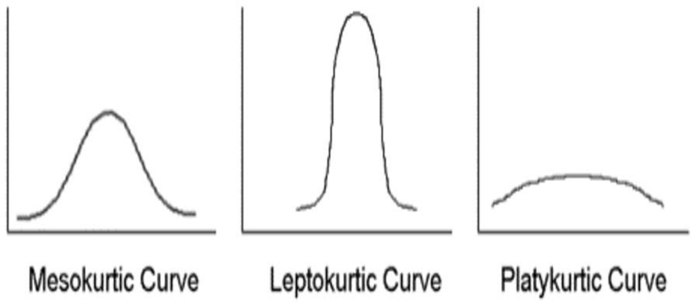
- In probability theory, skewness is a measure of the asymmetry of the probability distribution of a real-valued continuous random variable about its mean
- The skewness value can be positive or negative
- Skewness is measured by the following statistical expression –
- $Skewness = E\left[\left(\frac{x-\mu}{\sigma}\right)^3\right]$
- Positive skew - The right tail is longer; the mass of the distribution is concentrated on the left of the figure - The distribution is said to be right-skewed
- Negative skew - The left tail is longer; the mass of the distribution is concentrated on the right of the figure - The distribution is said to be left-skewed
- If skewness is equal to 0, then distribution is normal
- {(Mean – Mode)/S.D.} also gives some idea about skewness



Kurtosis

- In probability theory, kurtosis measures the peakedness of a curve
- The curves may be leptokurtic, mesokurtic and platykurtic
- The ideal curve is mesokurtic
- Kurtosis is expressed as follows –

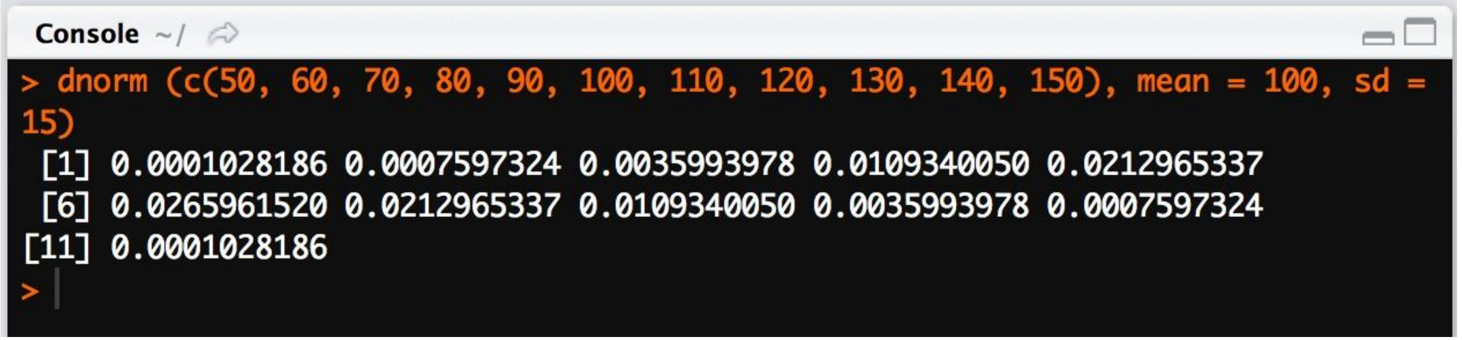
$$Kurtosis\ Coefficient = E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right]$$



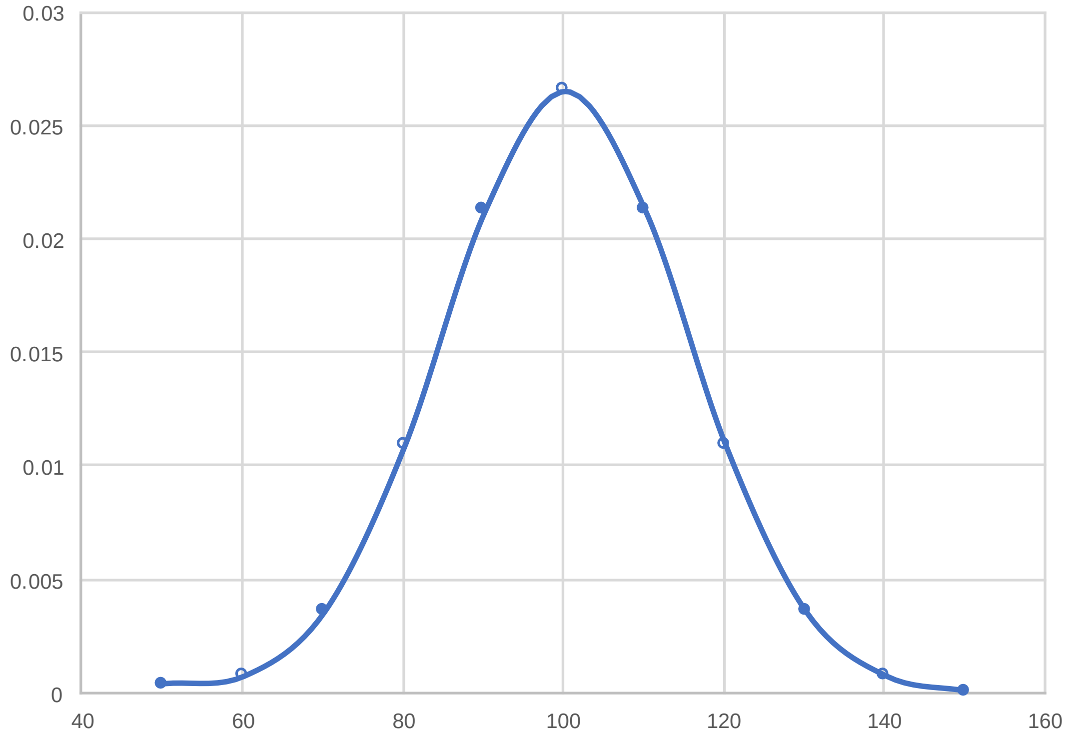
- Kurtosis is a measure of a distribution's peak, which means how much of the distribution is centred on the distributions mean
- The kurtosis coefficient of a normal distribution is 3
- Kurtosis is a measure of how extreme observations are in a data set
- The greater the kurtosis co-efficient, the more peaked the distribution around mean is
- Also, this distribution has fatter tails, which means there is an increase tail risk (extreme results)
- When a distribution's kurtosis is greater than 3, the distribution is leptokurtic and when it is less than 3, it is platykurtic

Normal Distribution - Example

x	Probability
50	0.00010282
60	0.00075973
70	0.00359940
80	0.01093400
90	0.02129653
100	0.02659615
110	0.02129653
120	0.01093400
130	0.00359940
140	0.00075973
150	0.00010282




```
> dnorm (c(50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150), mean = 100, sd =
15)
 [1] 0.0001028186 0.0007597324 0.0035993978 0.0109340050 0.0212965337
 [6] 0.0265961520 0.0212965337 0.0109340050 0.0035993978 0.0007597324
[11] 0.0001028186
>
```

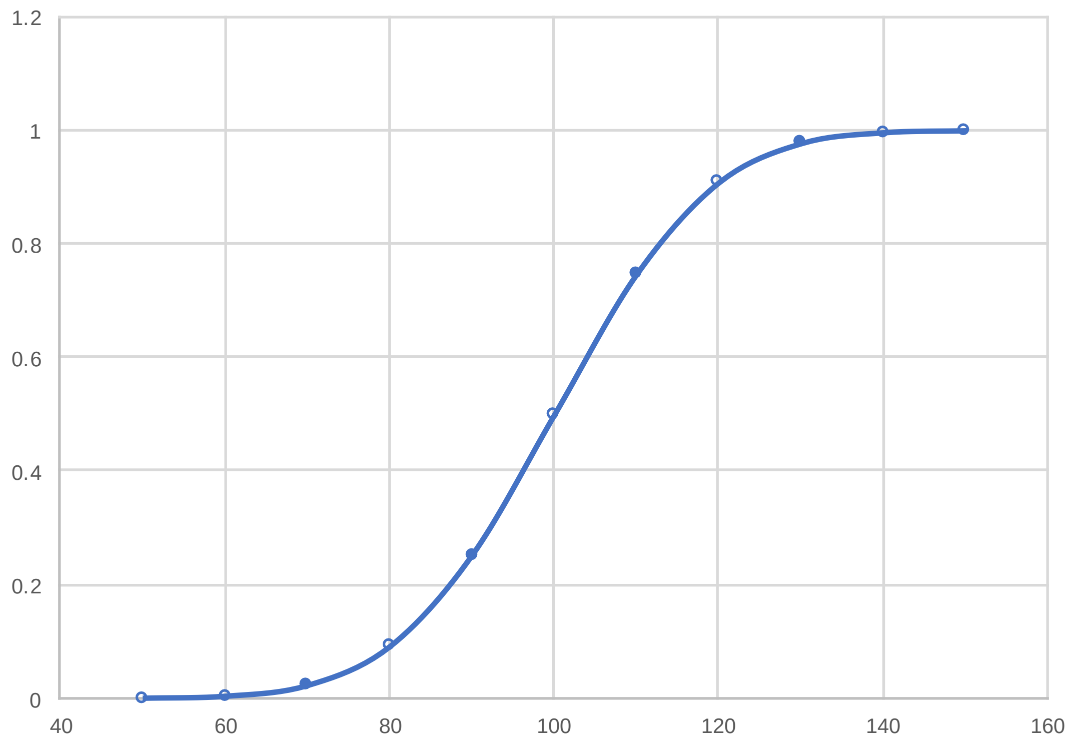


Normal Distribution - Example

x	C. Probability
50	0.00042906
60	0.00383038
70	0.02275013
80	0.09121122
90	0.25249254
100	0.50000000
110	0.74750746
120	0.90878878
130	0.97724987
140	0.99616962
150	0.99957094

Console ~/

```
> pnorm (c(50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150), mean = 100, sd =
15)
 [1] 0.0004290603 0.0038303806 0.0227501319 0.0912112197 0.2524925375
 [6] 0.5000000000 0.7475074625 0.9087887803 0.9772498681 0.9961696194
[11] 0.9995709397
> 
```



Normal Distribution - Example

Buffaloes in India yield an average of 4.71 kg of milk per day. Depending on the region, the milk yield varies with the buffaloes in Meghalaya yielding the least, and those in Punjab yielding the most. The standard deviation of the daily yield is 0.95 kg.

What is the probability that a buffalo yields 3kg milk per day?

8.31%

Console ~/

```
> dnorm (3, mean = 4.71, sd = 0.95)
[1] 0.08310543
>
```

What is the milk yield limit which can be suggested with a probability of 85%?

5.69 kg

Console ~/

```
> qnorm (0.85, mean = 4.71, sd = 0.95)
[1] 5.694612
>
```

Exponential Distribution

- **Exponential distribution is another distribution**
- **An exponentially distributed variable X has the PDF of the following form –**

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

- **Mean and standard deviation of X are equal in case of exponential distribution and it is as follows -**

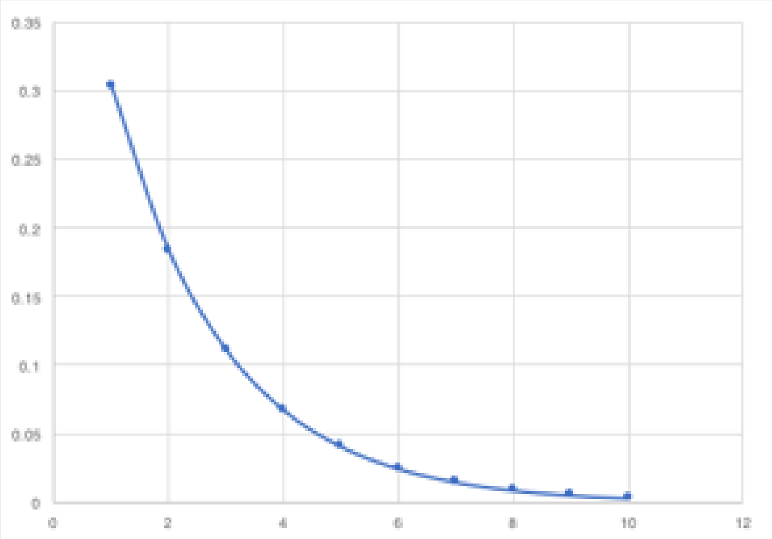
$$\mu = \sigma = \frac{1}{\lambda}$$

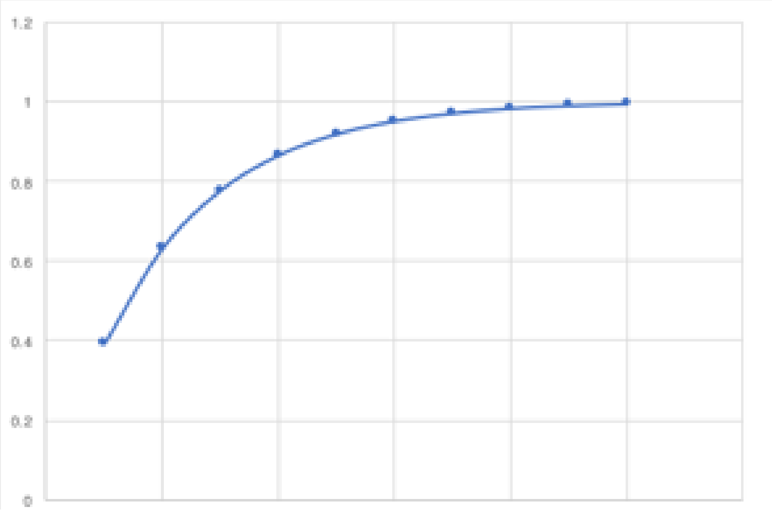
Exponential Distribution - Example

x	Probability	C. Probability
1	0.3032653299	0.3934693403
2	0.1839397206	0.6321205588
3	0.1115650801	0.7768698399
4	0.0676676416	0.8646647168
5	0.0410424993	0.9179150014
6	0.0248935342	0.9502129316
7	0.0150986917	0.9698026166
8	0.0091578194	0.9816843611
9	0.0055544983	0.9888910035
10	0.0033689735	0.9932620530

Console ~/

```
> pexp (1:10, 1/2)
 [1] 0.3934693 0.6321206 0.7768698 0.8646647 0.9179150 0.9502129 0.9698026
 [8] 0.9816844 0.9888910 0.9932621
>
```





Central Limit Theorem

- **A random sample of size n from a given distribution is a set of n independent random variables $X_1, X_2, \ldots X_n$ each having the given distribution, with expectation $E(X_i) = \mu$ and variance $Var(X_i) = \sigma^2$. Such a set of random variables is called independently and identically distributed (iid).**
- **Sample sum: $S = \sum_{i=1}^{n} x_i$; $E(S) = n\mu$; $Var(S) = n\,\sigma^2$**
- **Sample Mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $E(\bar{X}) = n\mu$; $Var(\bar{X}) = \sigma^2/n$**
- **Central Limit Theorem states that the mean and the sum of a random sample of a large enough size (generally, n ≥ 30) from an arbitrary distribution have approximately normal distribution**
- **Given a random sample $X_1, X_2, \ldots X_n$ with $\mu = E(X_i)$ and $\sigma^2 = Var(X_i)$, we have –**
  - **The sample sum $S = \sum_{i=1}^{n} X_i$ is approximately normal $N(n\mu, n\sigma^2)$**
  - **The sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is approximately normal $N(\mu, \sigma^2/n)$**

manipalglobal
Academy of Data Science

manipalglobal
Academy of Data Science