PROBABILITY FOR COMPUTING

PRACTICAL FILE

RAMANUJAN COLLEGE

UNIVERSITY OF DELHI

SUBMITTED BY :

NAME : Mayank Nautiyal

ROLL NO. : 24570038

EXAMINATION ROLL NO : 24020570045

CLASS : BSC(H) COMPUTER SCIENCE SEM 2

SUBMITTED TO:

DR. AAKASH

ASSISTANT PROFESSOR

(OPERATIONAL RESEARCH),

DEPARTMENT OF COMPUTER SCIENCE,

RAMANUJAN COLLEGE,

UNIVERSITY OF DELHI,

CR PARK MAIN ROAD,BLOCK H, KALKAJI

NEW DELHI 110019

Acknowledgement

I would like to take this opportunity to acknowledge everyone who has helped us in every stage of this project.

I am deeply indebted to my mathematics Professor, Dr Aakash for his guidance and suggestions in completing this project. The completion of this project was possible under his guidance and support.

I am also very thankful to my parents and my friends who have boosted me up morally with their continuous support.

At last but not least, I am very thankful to God almighty for showering his blessings upon me.

Binomial distribution

Binomial Distribution in Probability gives information about only two types of possible outcomes i.e. Success or Failure. Binomial Probability Distribution is a discrete probability distribution used for the events that give results in ***'Yes or No' or 'Success or Failure'.***

the probability of success (usually denoted as "p") and the probability of failure (usually denoted as "q") is constant for each trial.

Binomial Distribution Formula

The Binomial Distribution Formula which is used to calculate the probability, for a random variable X = 0, 1, 2, 3,….,n is given as

$$P(X = k) = \overbrace{\binom{n}{k} p^k (1-p)^{n-k}}^{\text{Probability Mass Function for a Binomial}}$$

Probability that our
variable takes on the
value k

$$\text{Mean} \quad \mu = n \cdot p$$

$$\text{Variance} \quad \sigma^2 = n \cdot p \cdot q$$

$$\text{Std. Dev.} \quad \sigma = \sqrt{n \cdot p \cdot q}$$

, k = 0, 1, 2, 3….

Where ,

- *p is probability of success*
- *q is probability of failure and q = 1 – p*
- *p, q > 0 such that p + q = 1*
- *n is the number of independent trials*
- *k is the number of success*
- *nCk is the number of ways to obtain k success in n trials.*

Understanding Binomial Distribution Symmetry

- *A binomial distribution is symmetric if its success probability, p, is equal to 0.5 (i.e., the probability of failure, q, is also 0.5). When p = 0.5, the distribution is said to be balanced, which means that the chances of success and failure are equal. In this case, the given binomial distribution has p = 0.5, which means that it is* ***symmetric****. The shape of the probability distribution will be* ***symmetric and not skewed***.

Determining the Skewness of the Distribution

- *Since we know that the distribution is symmetric when p = 0.5, it follows that the* ***skewness is zero****. Skewness is a measure of the asymmetry of a distribution. A distribution is* ***positively skewed*** *(skewed right) when there are more smaller values, and* ***negatively skewed*** *(skewed left) when there are more larger values. A symmetric distribution has a skewness of zero because there is an equal balance of larger and smaller values in the distribution. This means that the shape of the probability distribution is not skewed to the left or the right. Therefore, the shape of the probability distribution in this case is symmetric due to p = 0.5.*

how to implement in excel

= BINOM.DIST(number_s , trials, probability_s, cumulative)

= BINOM.DIST(k ,n ,p, FALSE)

The BINOM.DIST utilizes the accompanying contentions:

1. ***Number_s*** *(required argument)* **–** *This is the number of successes in trials.*
2. ***Trials*** *(required argument)* **–** *This is the number of independent trials. It must be greater than or equal to 0.*

3. ***Probability_s*** *(required argument)* **–** *This is the probability of success in each trial.*
4. ***Cumulative*** *(required argument)* **–** *This is a logical value that determines the form of the function. It can either be:*
    1. *TRUE – Uses the cumulative distribution function.*
    2. *FALSE – Uses the probability mass function*

1. Plotting and fitting of Binomial distribution and graphical representation of probabilities.

Binomial distribution

Binomial Distribution in Probability gives information about only two types of possible outcomes i.e. Success or Failure. Binomial Probability Distribution is a discrete probability distribution used for the events that give results in ***'Yes or No' or 'Success or Failure'.***

the probability of success (usually denoted as "p") and the probability of failure (usually denoted as "q") is constant for each trial.

Binomial Distribution Formula

The Binomial Distribution Formula which is used to calculate the probability, for a random variable X = 0, 1, 2, 3,….,n is given as

$$P(X = k) = \overbrace{\binom{n}{k} p^k (1-p)^{n-k}}^{\text{Probability Mass Function for a Binomial}}$$

Probability that our variable takes on the value k

$$\begin{aligned} &\textbf{Mean} && \mu = n \cdot p \\ &\textbf{Variance} && \sigma^2 = n \cdot p \cdot q \\ &\textbf{Std. Dev.} && \sigma = \sqrt{n \cdot p \cdot q} \end{aligned}$$

, k = 0, 1, 2, 3….

Where ,

- *p is probability of success*

- *q is probability of failure and q = 1 – p*
- *p, q > 0 such that p + q = 1*
- *n is the number of independent trials*
- *k is the number of success*
- *nCk is the number of ways to obtain k success in n trials.*

Understanding Binomial Distribution Symmetry

- *A binomial distribution is symmetric if its success probability, p, is equal to 0.5 (i.e., the probability of failure, q, is also 0.5). When p = 0.5, the distribution is said to be balanced, which means that the chances of success and failure are equal. In this case, the given binomial distribution has p = 0.5, which means that it is **symmetric**. The shape of the probability distribution will be **symmetric and not skewed**.*

Determining the Skewness of the Distribution

- *Since we know that the distribution is symmetric when p = 0.5, it follows that the **skewness is zero**. Skewness is a measure of the asymmetry of a distribution. A distribution is **positively skewed** (skewed right) when there are more smaller values, and **negatively skewed** (skewed left) when there are more larger values. A symmetric distribution has a skewness of zero because there is an equal balance of larger and smaller values in the distribution. This means that the shape of the probability distribution is not skewed to the left or the right. Therefore, the shape of the probability distribution in this case is symmetric due to p = 0.5.*

how to implement in excel

= BINOM.DIST(number_s, trials, probability_s, cumulative)

= BINOM.DIST(k ,n, p, FALSE)

The BINOM.DIST utilizes the accompanying contentions:

5. ***Number_s** (required argument) **–** This is the number of successes in trials.*
6. ***Trials** (required argument) **–** This is the number of independent trials. It must be greater than or equal to 0.*
7. ***Probability_s** (required argument) **–** This is the probability of success in each trial.*

8. ***Cumulative*** *(required argument)* ***–*** *This is a logical value that determines the form of the function. It can either be:*

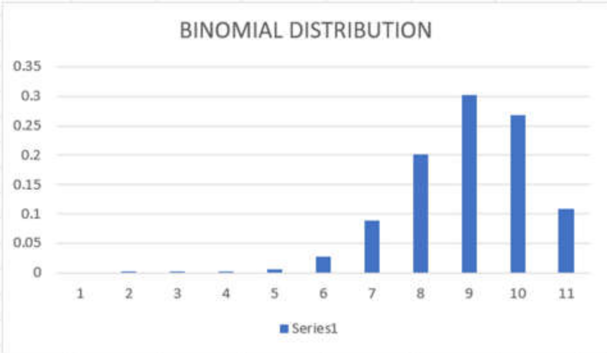
    1. *TRUE – Uses the cumulative distribution function.*

    2. *FALSE – Uses the probability mass function*

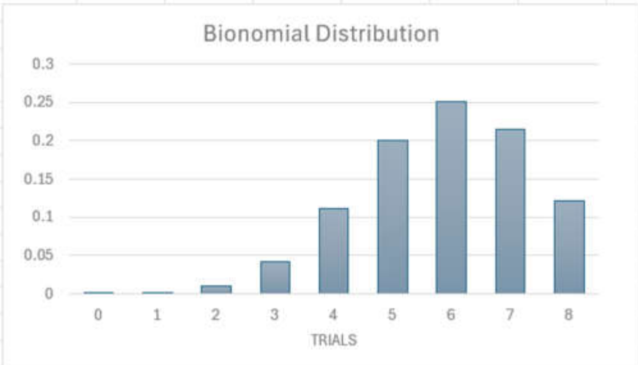
1. Plotting and fitting of Binomial distribution and graphical representation of probabilities.

Suppose there is a quiz competition in school and 50% chances are of getting answer right .If 10 students at random are participating and right answer is defined as a success. The probability distribution of the number of success during these 10 trials with probability p=0.8. then plot the graph of this probability in excel ?

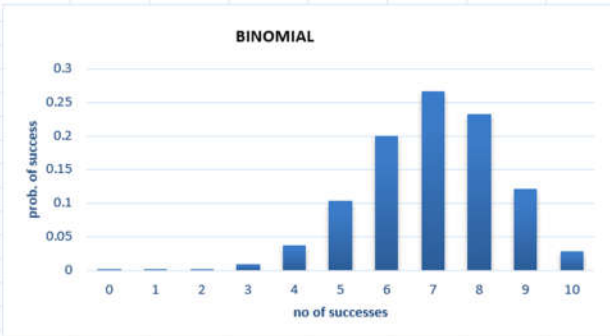
Number of Correct Answers (X=K)	P(X=K) in Student Quiz (n=10, p=0.8)
Binomial Distribution - daily customers	
Probability of Success p	0.8
Number of Trials n	10
0	1.024E-07
1	4.096E-06
2	7.3728E-05
3	0.000786432
4	0.005505024
5	0.026424115
6	0.088080384
7	0.201326592
8	0.301989888
9	0.268435456
10	0.107374182



Number of Success(X=K)	10
Proabability of success P(X=K)	0.6
Trials	Binomial Distribution of trials
0	0.000104858
1	0.001572864
2	0.010616832
3	0.042467328
4	0.111476736
5	0.200658125
6	0.250822656
7	0.214990848
8	0.120932352



Number of Correct Answers (X=K)	P(X=K) in chocolate (n=10, p=0.7)
Binomial Distribution - daily customers	
Probability of Success p	0.7
Number of Trials n	10
0	0.000005905
1	0.000137781
2	0.001446701
3	0.009001692
4	0.036756909
5	0.102919345
6	0.200120949
7	0.266827932
8	0.233474441
9	0.121060821
10	0.028247525



Multinomial distribution

The multinomial distribution is a multivariate generalization of the ***binomial distribution****. Consider a trial that results in exactly one of some fixed* ***finite number*** k *of possible outcomes, with* *probabilities* $p_1, p_2, \ldots, p_k$ *(so that* $p_i \geq 0$ *for* $i = 1, \ldots, k$ *and* $\sum_{i=1}^{k} p_i = 1$*), and there are* n *independent trials. Then let the random variables* X_i *indicate the number of times outcome number* i *was observed over the* n *trials. Then* $X = (X_1, X_2, \ldots, X_k)$ *follows a multinomial distribution with parameters* **n** *and* ***p****, where* ***p*** *=(*$p_1, p_2, \ldots, p_k$*).*

Multinomial Distribution Formula

$$p(x_1, x_2 \ldots \ldots x_k) = \left[\frac{n!}{x_1! \cdot x_2! \ldots\ldots x_k!}\right] \cdot p_1^{x_1} \cdot p_2^{x_2} \ldots\ldots p_k^{x_k}$$

Mean	$E\{X_i\} = np_i$
Variance	$\mathrm{Var}(X_i) = np_i(1 - p_i)$ $\mathrm{Cov}(X_i, X_j) = -np_i p_j \quad (i \neq j)$

When ***X*** = (x_1, x_2, …, x_k) follows a multinomial distribution with the PMF given above, X_i follows a binomial distribution with n trials and success probability p_i.

how to implement in excel

Multinomial = MULTINOMIAL(X1,X2,X3)

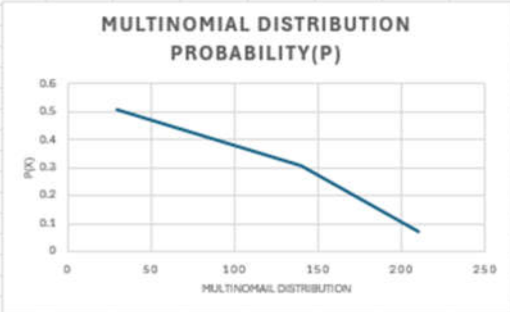
Probability = MULTINOMIAL*PRODUCT(p1^X1,p2^X2,p3^X3)

2. Plotting and fitting of Multinomial distribution and graphical representation of probabilities.

Suppose that a bag contain 8 balls 3 red, 1 green and 4 blue to reach in a bag to pull ball at random and then pull the ball back and pull out another ball .experiment is repeated at total of 10 time . What is the probability outcome will result in 4 red and 6 blues?

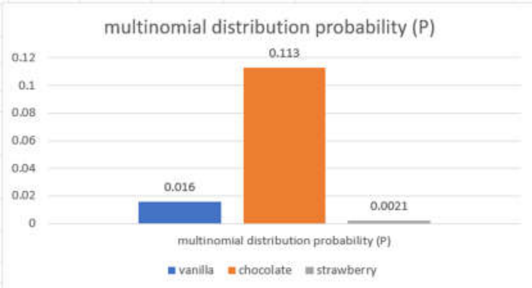
		1st case	2nd case	3rd case
Red	X1	4	3	2
Green	X2	0	1	1
Blue	X3	6	3	2
P_Red	P1	0.325	0.325	0.325
P_greer	P2	0.15	0.15	0.15
P_blue	P3	0.5	0.5	0.5
	Multinomia	210	140	30
	P1^X1	0.0111566	0.034328	0.105625
	P2^X2	0.0005063	0.15	0.15
	P3^X3	0.0625	0.125	0.25
	Probability	0.0741629	0.309328	0.505625

Multinmial	multinomial distribution proabability(P)
210	0.0741629
140	0.3093281
30	0.505625



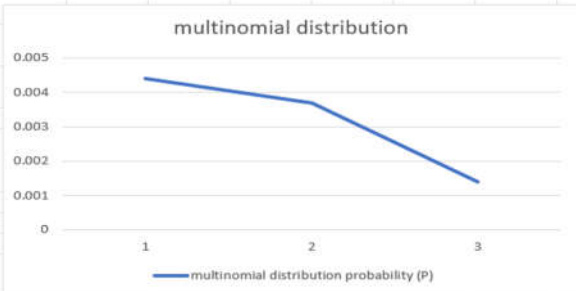
		1st case	2nd case	3rd case
Vanilla	X1	3	4	2
Chocolate	X2	4	3	5
Strawberry	X3	3	3	3
P_Vanilla	P1	0.4	0.4	0.4
P_Chocolate	P2	0.35	0.35	0.35
P_Strawberry	P3	0.25	0.25	0.25
	Multinomial	1680	1680	2520
	P1^X1	0.064	0.1024	0.016
	P2^X2	0.015	0.0429	0.0053
	P3^X3	0.0156	0.0156	0.0156
	Probability	0.016	0.113	0.0021

Multinomial	multinomial distribution probability (P)
1680	0.016
1680	0.113
2520	0.0021



P_Organic (P1)	0.5	0.5	0.5
P_PaidAds (P2)	0.3	0.3	0.3
P_Social (P3)	0.2	0.2	0.2
Multinomial Coeff.	2520	1680	1680
P1^X1	0.0625	0.03125	0.125
P2^X2	0.02187	0.09	0.0081
P3^X3	0.008	0.008	0.008
Probability	**0.0044**	**0.0037**	**0.0014**

Multinomial	**multinomial distribution probability (P)**
1680	0.0044
1680	0.0037
2520	0.0014



Poisson distribution

*The Poisson distribution is a type of discrete probability distribution that determines the likelihood of an event occurring a specific number of times (k) within a designated time or space interval. This distribution is characterized by **a single parameter, λ (lambda),** representing the average number of occurrences of the event.*

Poisson Distribution Formula

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

Mean	$\mu = E(X) = \lambda$
Variance	$\sigma^2 = V(X) = \lambda$
Standard Deviation	$\sigma = \sqrt{\sigma^2} = \sqrt{\lambda}$

Where:

- ***P(X=k)*** *is the probability of observing k events*
- ***e*** *is the base of the natural logarithm (approximately 2.71828)*
- ***λ*** *mean number of success that occur during a specific interval,* ***λ = np***
- ***k*** *is the number of success*

how to implement in excel

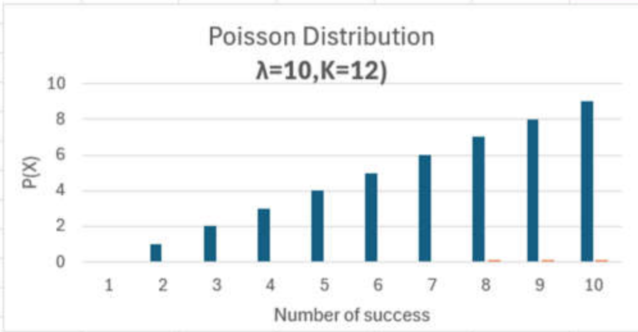
POISSON.DIST(number_s , average, cumulative)

POISSON.DIST(k, λ ,FALSE)

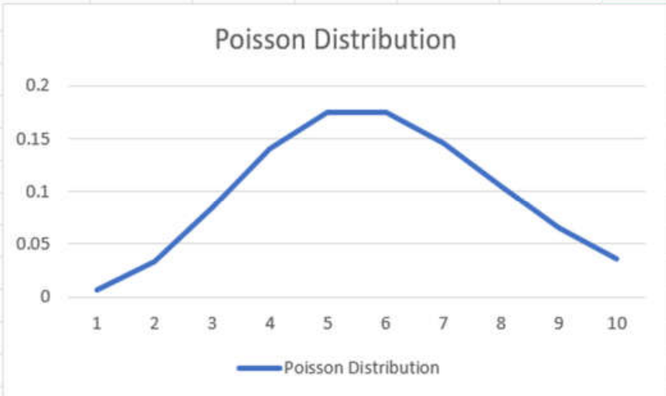
3. Plotting and fitting of Poisson distribution and graphical representation of probabilities.

An electronic store sells on average 8 Desktop in a week. Assuming that purchases are as described above then what is the probability that the store will have turn away potential buyers before the end if the stock 9 computers ? how many computers should the store stock in order to make sure that it has 99% probability of being able to supply a week's demand ?

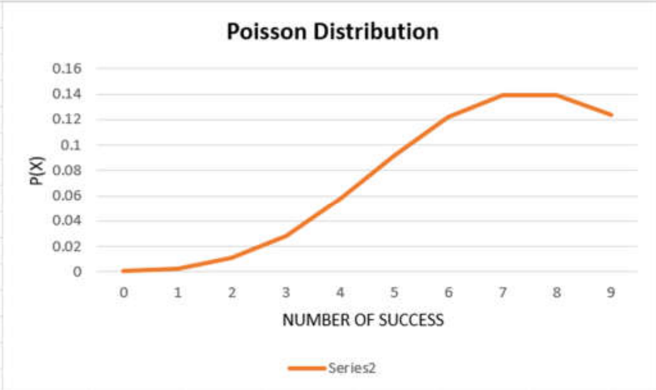
λ(Mean numbet of success)	10
K(Number of Success)	Poisson Distribution
0	4.53999E-05
1	0.000453999
2	0.002269996
3	0.007566655
4	0.018916637
5	0.037833275
6	0.063055458
7	0.090079226
8	0.112599032
9	0.125110036



λ(Mean numbet of success)	5
K(Number of Success)	Poisson Distribution
0	0.006737947
1	0.033689735
2	0.084224337
3	0.140373896
4	0.17546737
5	0.17546737
6	0.146222808
7	0.104444863
8	0.065278039
9	0.036265577



λ(Mean numbet of success)	8
K(Number of Success)	Poisson Distribution
0	0.000335463
1	0.002683701
2	0.010734804
3	0.028626144
4	0.057252288
5	0.091603662
6	0.122138215
7	0.139586532
8	0.139586532
9	0.124076917



Geometric distribution

In a Bernoulli trial, the likelihood of the number of successive failures before success is obtained is represented by a geometric distribution, which is a sort of discrete probability distribution. A Bernoulli trial is a test that can only have one of two outcomes: success or failure. In other words, a Bernoulli trial is repeated until success is obtained and then stopped in geometric distribution.

A geometric distribution is a discrete probability distribution that indicates the likelihood of achieving one's first success after a series of failures. The number of attempts in a geometric distribution can go on indefinitely until the first success is achieved. Geometric distributions are probability distributions that are based on three key assumptions.

- *The trials that are being undertaken are self-contained.*
- *Each trial may only have one of two outcomes: success or failure.*
- *For each trial, the success probability, represented by p, is the same*

Geometric Distribution formula

$$P(X=k) = (1-p)^k p$$

Mean:	$\mu = E(X) = \dfrac{1}{p}$
Variance:	$\sigma^2 = V(X) = \dfrac{(1-p)}{p^2}$

where:

- **k:** number of failures before first success

- **p:** probability of success on each trial

The chance of a trial's success is denoted by p, whereas the likelihood of failure is denoted by q, q = 1 – p in this case. X ~ G (p) represents a discrete random variable, X, with a geometric probability distribution.

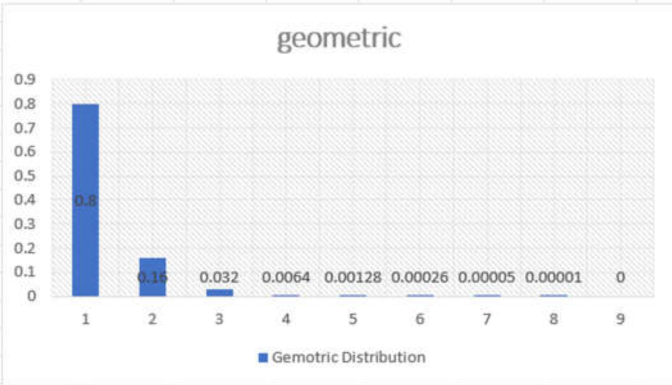
how to implement in excel

Probability = (1-p)^k*p

4. Plotting and fitting of Geometric distribution and graphical representation of probabilities.

Suppose a programmer is waiting outside a PM office to take his views on artificial intelligence like they support AI or not .The probability that a PM supports the AI is p = 0.8 . what is the probability that the fourth PM , the programmer talk to is the first PM to support AI ?

Prob. Of Success on given trial(P)	0.8
Number of failuires before first s	Gemotric Distribution
0	0.8
1	0.16
2	0.032
3	0.0064
4	0.00128
5	0.00026
6	0.00005
7	0.00001
8	0



Uniform Distribution Function

A uniform distribution is a distribution that has constant probability due to equally likely occurring events. It is also known as rectangular distribution (continuous uniform distribution). It has two parameters a and b: a = minimum and b = maximum. The distribution is written as U (a, b).

A uniform distribution is a type of probability distribution where every possible outcome has an equal probability of occurring. This means that all values within a given range are equally likely to be observed.

Uniform Distribution Formula

The probability density function (PDF) of a continuous uniform distribution defines the probability of a random variable falling within a particular interval. For a continuous uniform distribution over the interval [a,b].

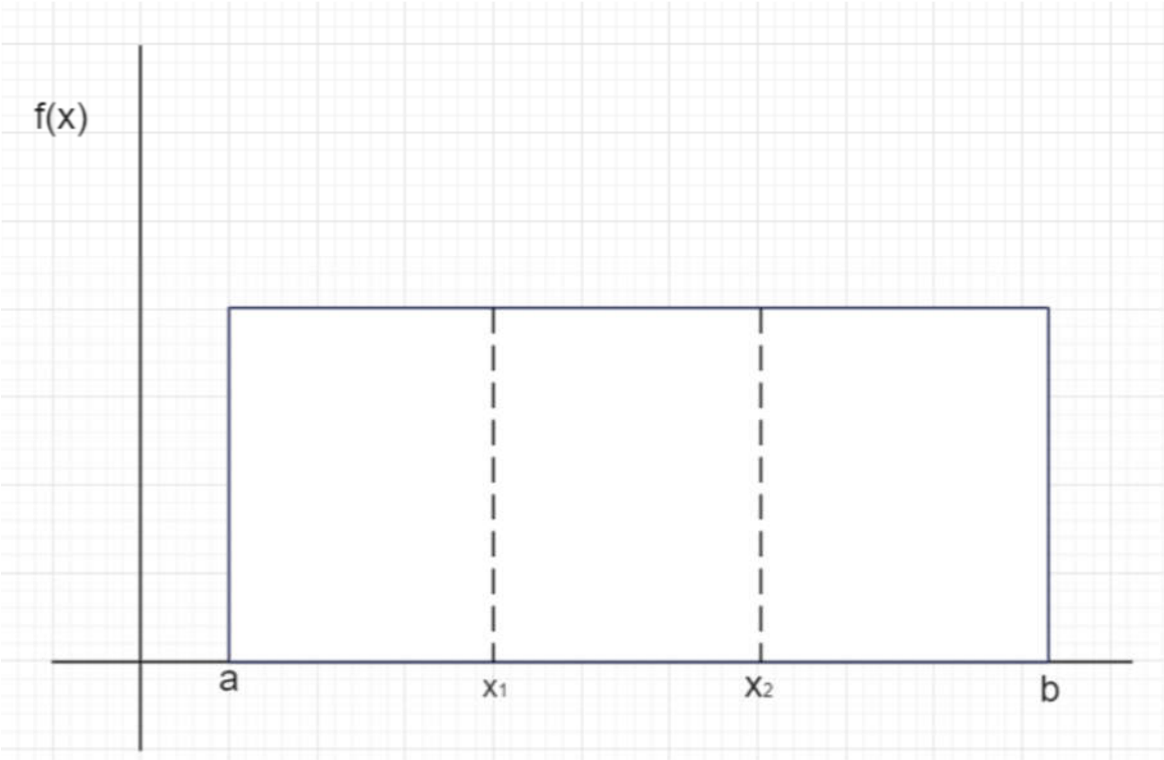
$$f(x) = \frac{1}{b-a} \text{ for } a \le x \le b$$

$$\text{Mean } \mu = \frac{a+b}{2}$$

$$\text{Variance } \sigma^2 = \frac{(b-a)^2}{12}$$

how to implement in excel

$$P = (x_2 - x_1) / (b - a)$$



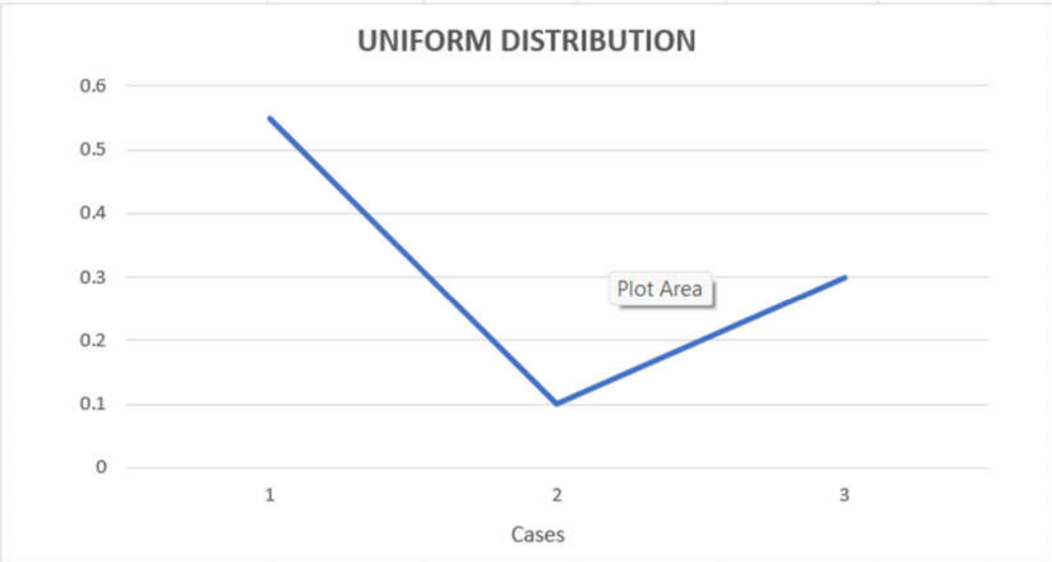
For calculating probability, we need:

1. *a: minimum value in the distribution*
2. *b: maximum value in the distribution*
3. *x_1: the minimum value you're interested in*
4. *x_2: the maximum value you're interested in*

5. Plotting and fitting of Uniform distribution and graphical representation of probabilities.

A sound sensor on x1x2 sensor takes 10 seconds and it will show up the result every time in 30 seconds. If you will use x1x2 sensor then what is the probability that a sensor will show up in 15-26 seconds ?

Machine Temperature Monitoring				
a	10			
b	30			
		case1	case2	case3
Sensor 1 (x1)		15	19	17
Sensor 1 (x2)		26	21	23
uniformal distributed		0.55	0.1	0.3



Exponential random variable

*The support (set of values the Random Variable can take) of an Exponential Random Variable is the set of all positive real numbers. Suppose we are posed with the question- How much time do we need to wait before a given event occurs? The answer to this question can be given in probabilistic terms if we model the given problem using the Exponential Distribution. Since the **time** we need to wait is unknown, we can think of it as a Random Variable. If the probability of the event happening in a given interval is proportional to the length of the interval, then the Random Variable has an exponential distribution. The support (set of values the Random Variable can take) of an Exponential Random Variable is the set of all positive real numbers.*

This distribution can be used to solve following type of real life problems-

- *How long does a shop owner need to wait until a customer enter a shop.*
- *How long will a battery continue to work before it dies.*
- *How long will a computer continue to work before it breakdown.*

$$\mathbf{f(x)} = \begin{cases} \lambda e^{-\lambda * x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$E(X) = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

*Here **λ** is the rate parameter and its effects on the density function .*

e is a constant roughly equal to 2.718

How to Implement in excel

EXPON.DIST(X, lambda, cumulative)

EXPON.DIST(X, lambda, FALSE)