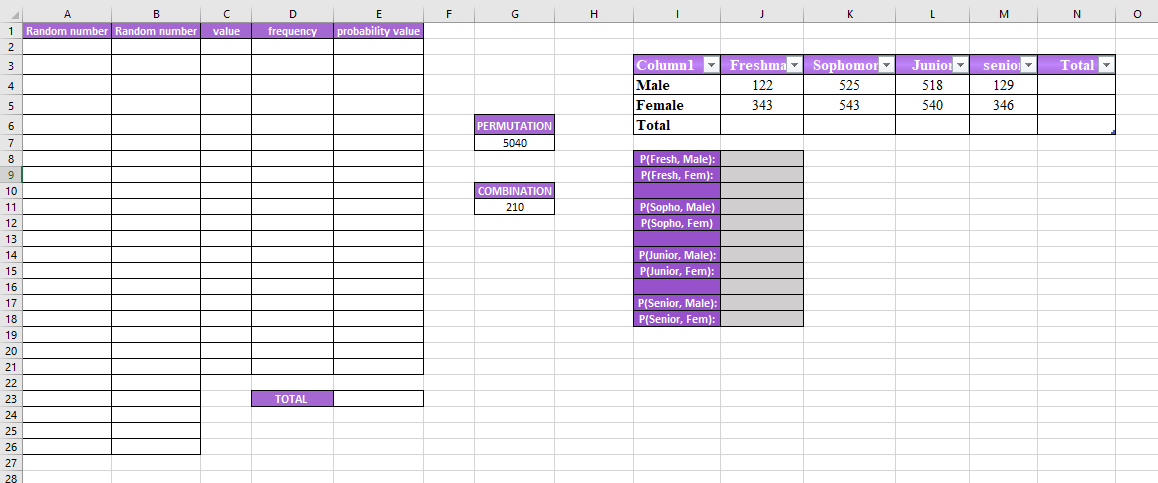
|  |  |
| --- | --- |
| **Practical No. 1** | Introduction to Probability in Excel. |
| **Practical No. 2** | Discover Probability using Formulas. |
| A | Design a spreadsheet to demonstrate the association Between Two Variables by Computing the Covariance and Correlation Coefficient. |
| B | Design and spreadsheet experiment to compute the probability using the geometric distribution formula. |
| **Practical No. 3** | Random Variables and Distribution Functions |
| A | Expected Value, Mean, and Variance Using Excel |
| B | Create a spreadsheet application to Compute Binomial Probabilities. [Hint: Use BINOM DIST] |
| **Practical No. 4** | Probability Distribution and Law |
| A | Create a spreadsheet application to Poisson Probability Distribution. [Hint: Use POISSON] |
| B | Create a spreadsheet application to implement joint probability law. |
| **Practical No. 5** | Mathematical Expectation and Chebychev’s Theorem |
| A | Create a spreadsheet application to compute the expectation of a Function of a  Random Variable |
| B | Create a spreadsheet application to apply Chebychev’s Theorem |
| **Practical No. 6** | Conditional Expectation and Generating Functions |
| A | Create a spreadsheet application to compute Conditional Expectation and Conditional Variance. |

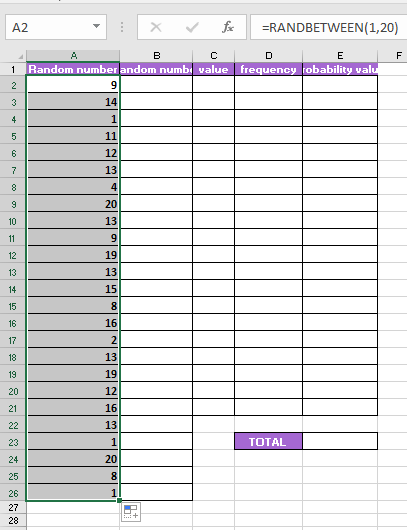
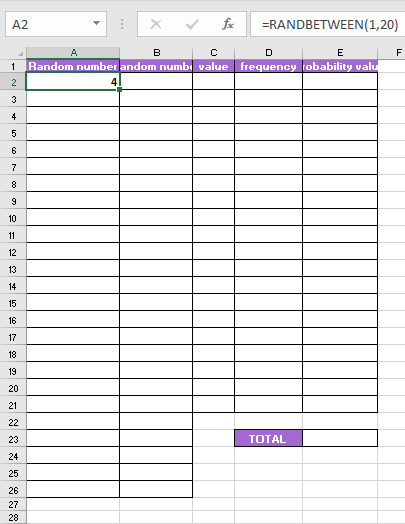
**Practical No. 1**

**Aim:** Introduction to Probability.

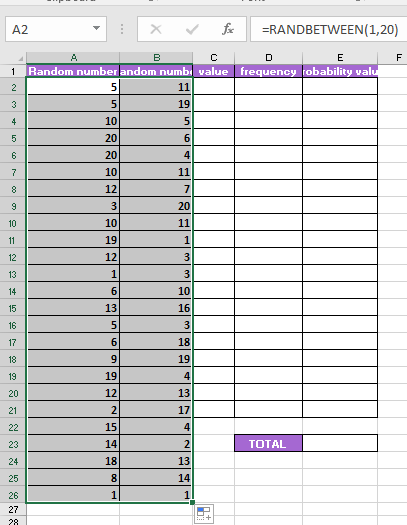
**Step1:** Create an Excel Sheet with below data.



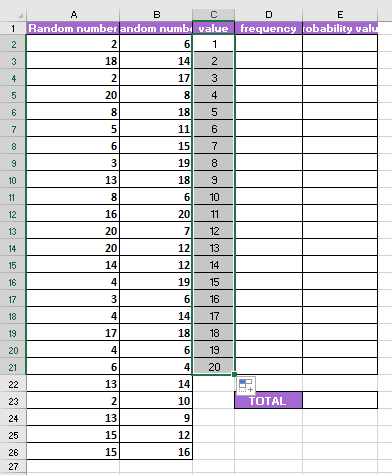
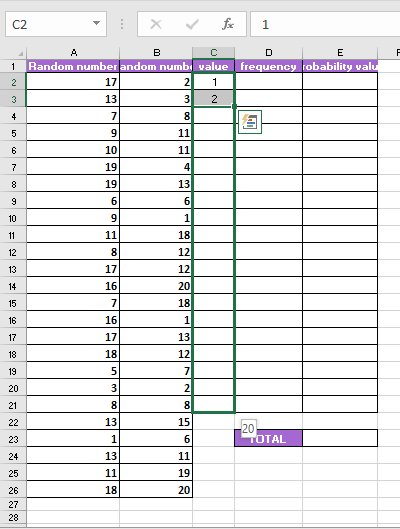
**Step2:** In cell A2, type **“=RANDBETWEEN (1,20)”**, and drag this formula from A2 to A26. You have now created 25 random numbers between 1 and 20.



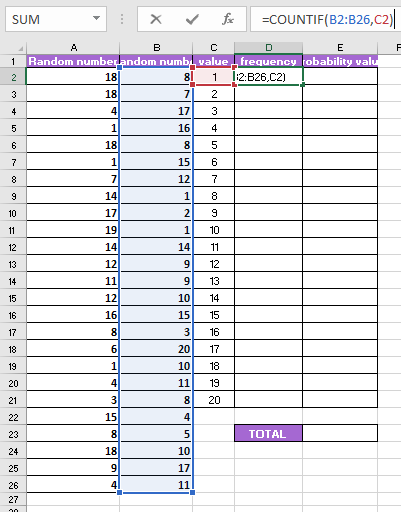
Select column 🡪 copy them in B2 cell



**Step3:** In C2 cell type 1 to 20 numbers in order.



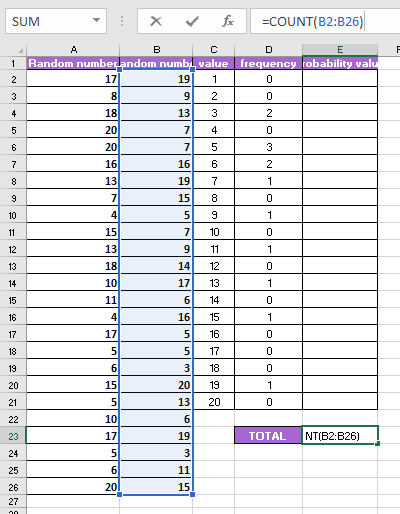
**Step4:** In D2 cell type the formula: **“=COUNTIF($B$2:$B$26,C2)”,** and drag this formula down till D21.



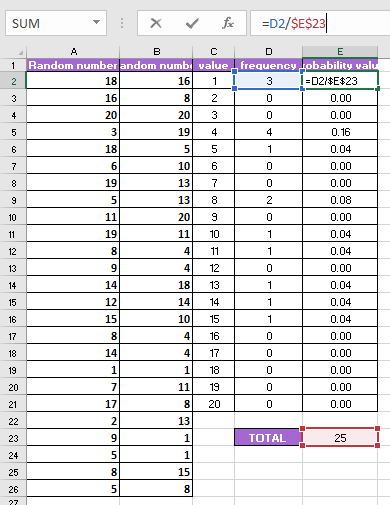
**Step5:** In cell D22, we wish to calculate the total number of values in our data set. We can do this in two ways.

**Total count “=COUNT(B2:B20)”**

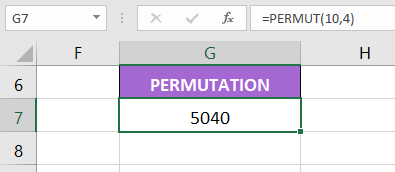
**Total “=SUM(B2:B20)”**



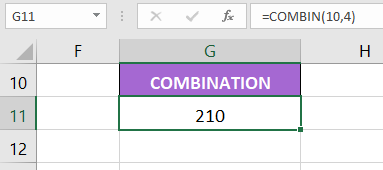
**Step6:** In cell E2, calculate the probability of selecting the value 1 in our data set. Type the formula: **“=D2/$D$20”**, and drag this formula down till D21.



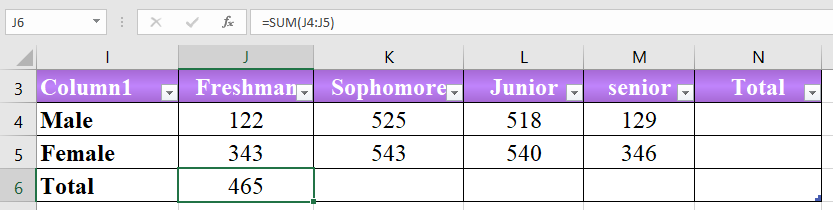
**Step7:** In cell G2, type the formula: **“=PERMUT(10,4)”** calculate the number of permutations choosing 4 out of 10 objects.



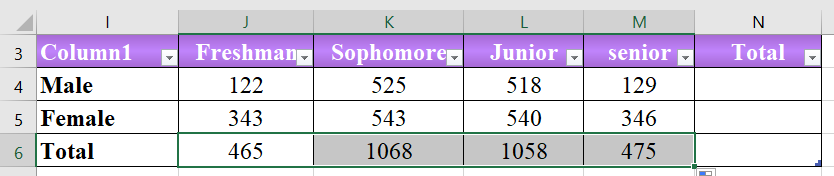
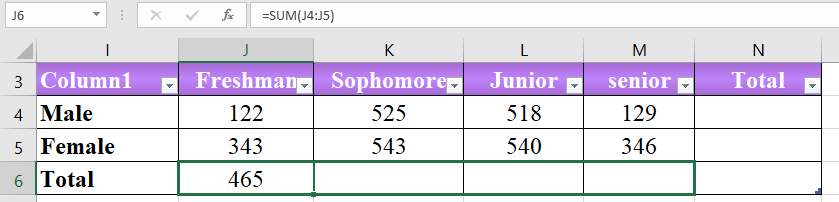
**Step8:**  In cell G6, type the formula: **“=COMBIN (10,4)”** to calculate the number of combinations choosing 4 out of 10 objects.



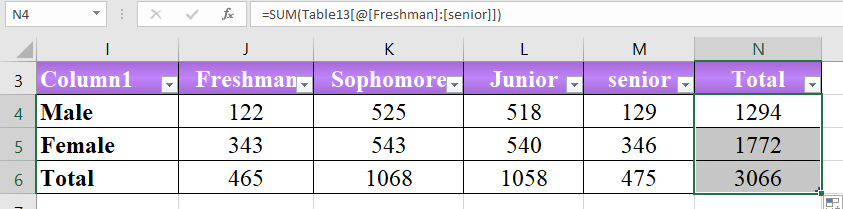
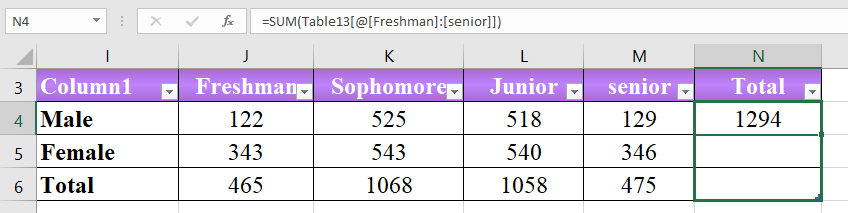
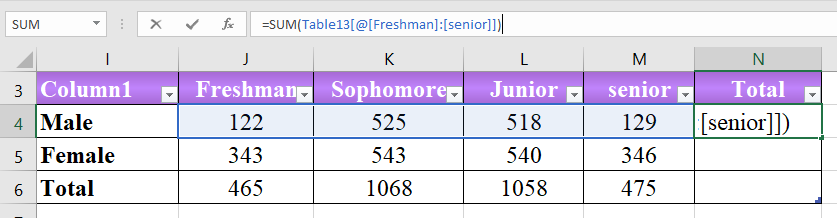
**Step9:** In cell N2, calculate the total sum of the male, type the formula:” **=SUM(J4:J5)”**



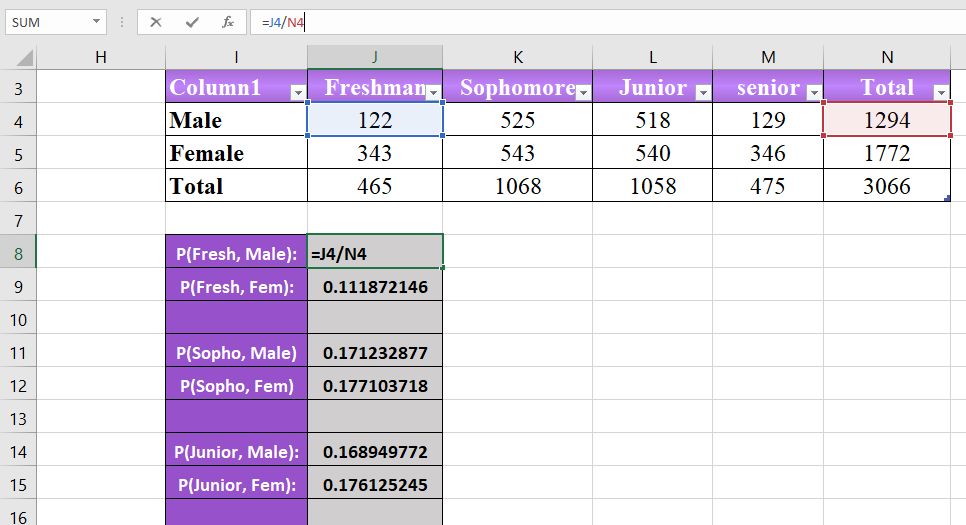
**Drag till M6**

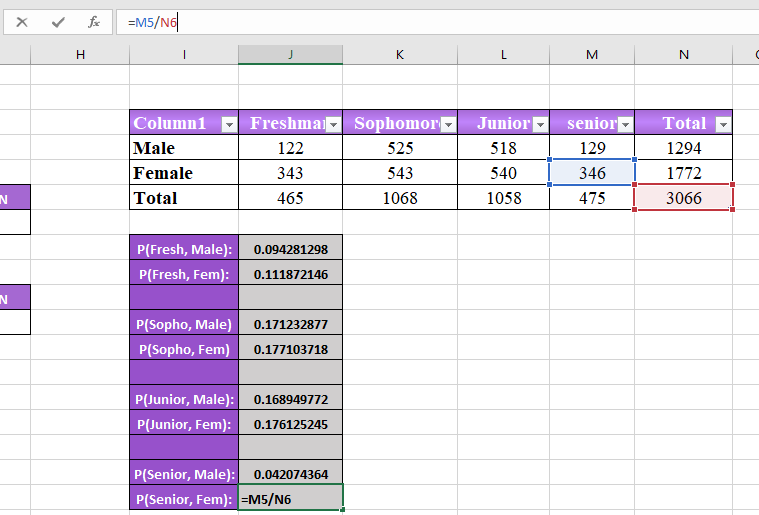
****

**Step 10 :**

****

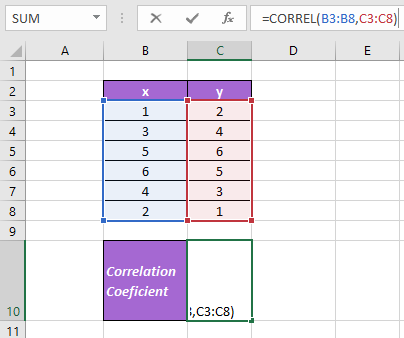
**Step12:** same as step 11, calculate the probability of male and female for each category.





**Practical No. 2**

**Aim:** Discover Probability using Formulas. (Calculating correlation with spreadsheets.)



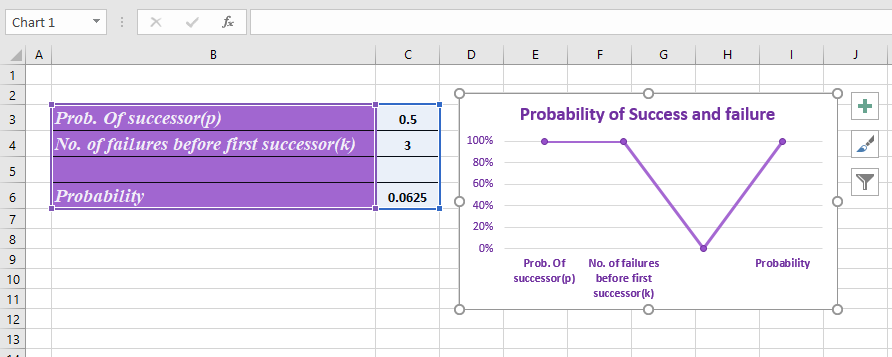
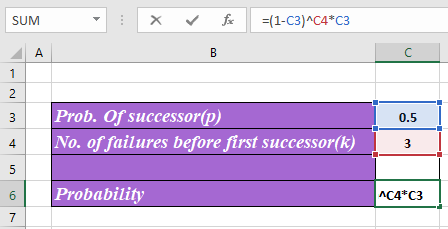
1. Design a spreadsheet to demonstrate the association Between Two Variables by Computing the Covariance and Correlation Coefficient.

If a random variable *X* follows a geometric distribution, then the probability of experiencing *k* failures before experiencing the first success can be found by the following formula:

**P(X=k) = (1-p)kp**

* **k:** number of failures before first success
* **p:** probability of success on each trial

The following examples show how to calculate probabilities related to the geometric distribution in Excel.



1. Design and spreadsheet experiment to compute the probability using the geometric distribution formula.

**Eaxmple 1: Flipping a coin**

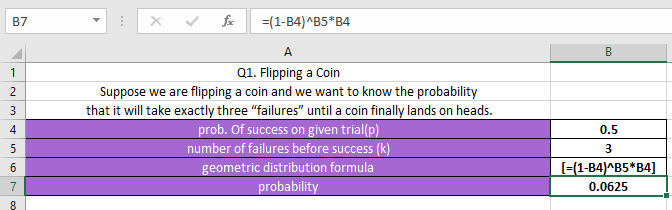
Suppose we are flipping a coin and we want to know the probability that it will take exactly three “failures” until a coin finally lands on heads.

We would use the following formula to calculate this probability:

**Step1:** In excel sheet in B1 cell type the probability success of given trial is 0.5

In cell B2 type number of failures before first success i.e.3

**Step2:** To find out Probability use the formula: **“=(1-B1)^B2\*B1”**.

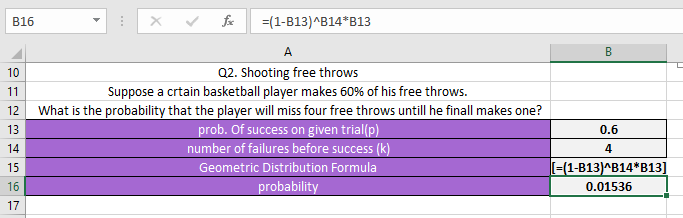


The probability that we will experience three “failures” until a coin finally lands on heads is **0.0625**.

**Example 2: Shooting Free Throws**

Suppose a certain asketball player makes 60% of his free throws. What is the probability that the player will miss four free throws until he finally makes one?

We would use the following formula to calculate this probability:

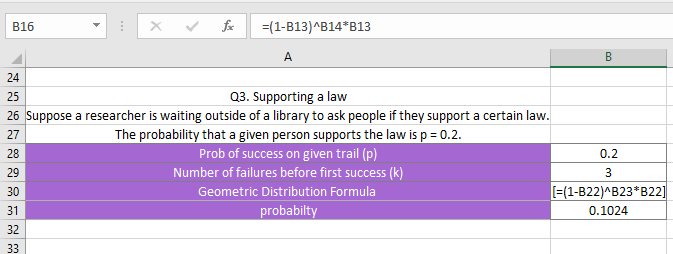


The probability that the player will miss four free throws until he finally makes one is **0.01536**.

**Example 3: Supporting a Law**

Suppose a researcher is waiting outside of a library to ask people if they support a certain law. The probability that a given person supports the is p=0.2. What is the probability that the fourth person researcher talks to is the first person to support the law?

We would use the following formula to calculate this probability:



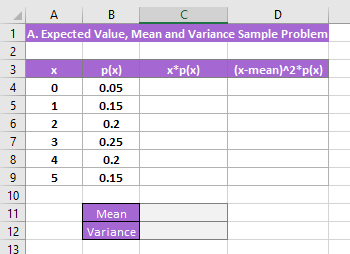
The probability that the fourth person researcher talks to is the first person to support the law is **0.1024**.

**Practical No. 3**

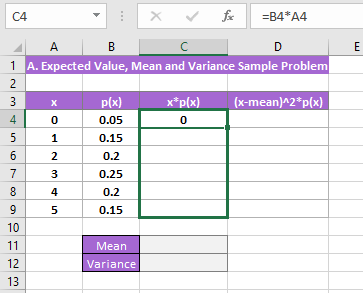
**Aim:** Random Variables and Distribution Functions.

1. Expected Value, Mean, and Variance Using Excel.

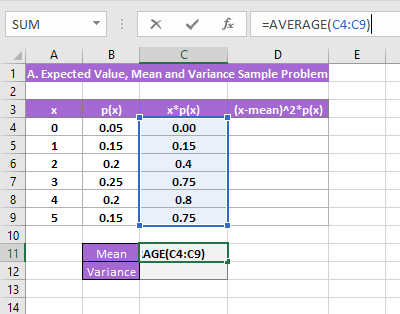
* Harrington Health Food stocks 5 loaves of Neutro-Bread. The probability distribution has been entered into the Excel spreadsheet, as shown below.

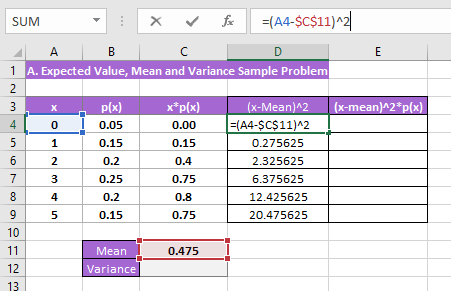
****

* You want to know how many loaves Harrington will sell on average and the variance of the distribution.
* You are going to be calculating the mean and the variances using expected value.
* To calculate expected value, you want to sum up the products of the X’s (Column A) times their probabilities (Column B).
* Start in cell C4 and type the formula: **“B4\*A4”**.
* Then drag that cell doen to cell C9 and do the auto fill; this gives us each of the individual expected values, as shown below.

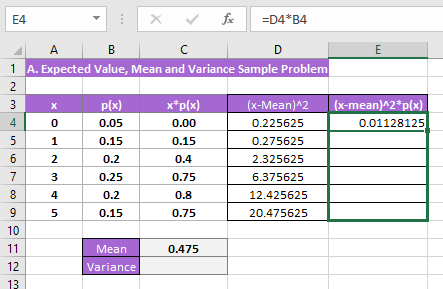


* Now to calculate the mean, you are going to just sum up that column by typing the formula: **“=SUM(C4:C9)”** into cell C10, hit the enter key or click the checkmark icon, and our mean is 2.85, as shown below.

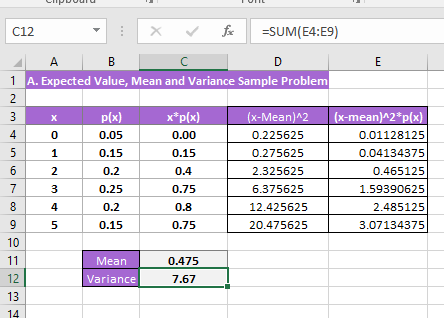




* Now that you have the mean, we can calculate the variance.
* The variance is calculates by taking each X value and subtracting the eman (cell C10), and then squaring that and then multiplying that times the probability of that X or (x-mean)2\*p(x).
* Start off in cell D4 and type the formula: **“=A4-$C$10)^2\*B4”** and hit the Enter Key or click the checkmark icon. Be sure to put absolute cell references for the mean so the cell does not change (A4-$C$10).
* Now just that down to cell D9 and auto fill like you did before; this gives you the variances, as shown below.



* Now to calculate the variance of the distribution, sum up that column by typing **“=SUM(D4:D9)”** into cell D11, hit the Enter Key or click the checkmark icon, and our variance is 2.0275, as shown below.

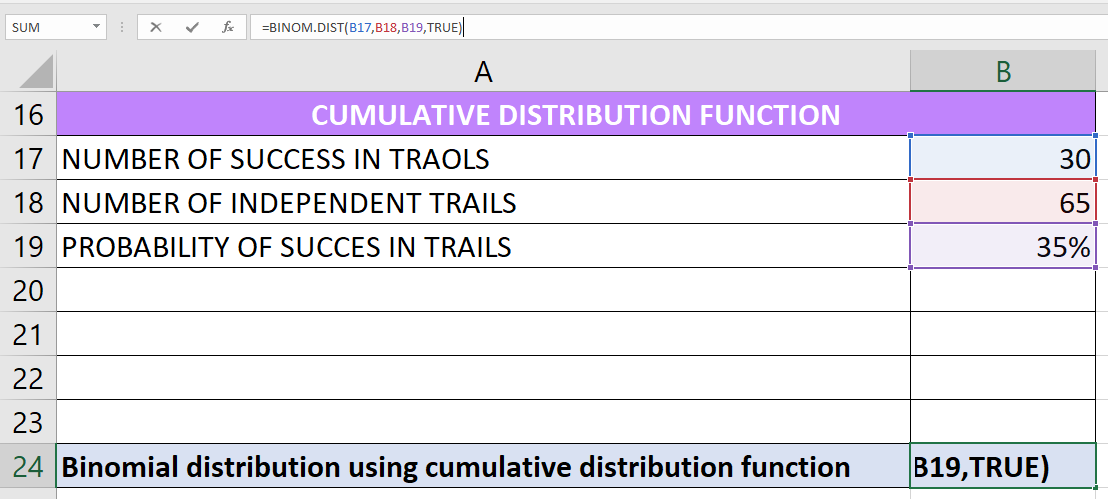


1. Create a spreadsheet application to Compute Binomial Probabilities. [Hint: Use BINOM DIST]

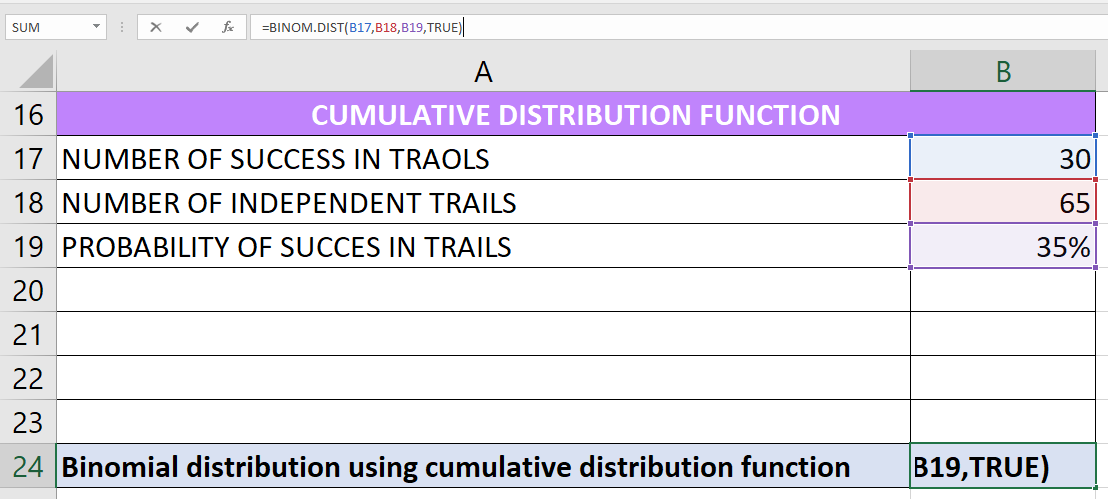
* The binomial distribution is a statistical measure that is frequently used to indicate the probability of a specific number of successes occurring from a specific number of independent trials.
* The two forms used are:
* The Probability Mass Function – Calculates the probability of there being exactly x successes from n independent trails.
* The Cumulative Distribution Function – Calculates the probability of there being at most x successes from n independent trials.
* Formula for Binomial Disstribution: **“BINOM.DIST(number\_s,trials,probability\_s,cumulative)”**
* The BINOM.DIST uses the following arguments:
* **Number\_s** (required argument) – This is the number of successes in trials.
* **Trials** (required argument) – This is the number of independent trials. It must be greater than or equal to 0.
* **Probability\_s** (required argument) – This is the number of successes in each trial.
* **Cumulative**(required argument) – This is a logical value that determins the form of the function. It can either be:
* **TRUE** – Uses the Cumulative Distribution Function.
* **FALSE** – Uses the Probability Mass Function.

Example:

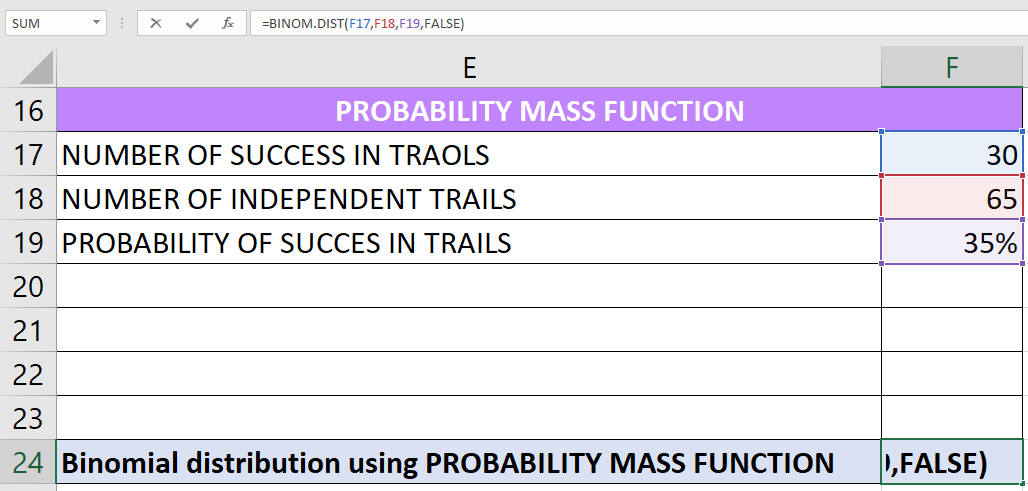
* Suppose we are given the following data:



* The formula for calculating binomial distribution using the cumulative distribution function is shown below: **“=BINOM.DIST(H3,H4,H5,TRUE)”**.



* The formula for calculating binomial distribution using the probability mass function is shown below: **“=BINOM.DIST(H3,H4,H5,FALSE)”**.



**Practical No. 4**

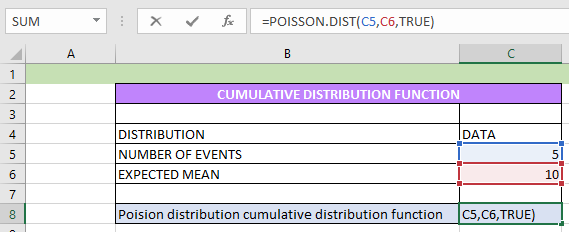
**Aim:** Probability Distribution and Law.

1. Create a spreadsheet application to Poisson Probability Distribution. [Hint: Use BINOM DIST]

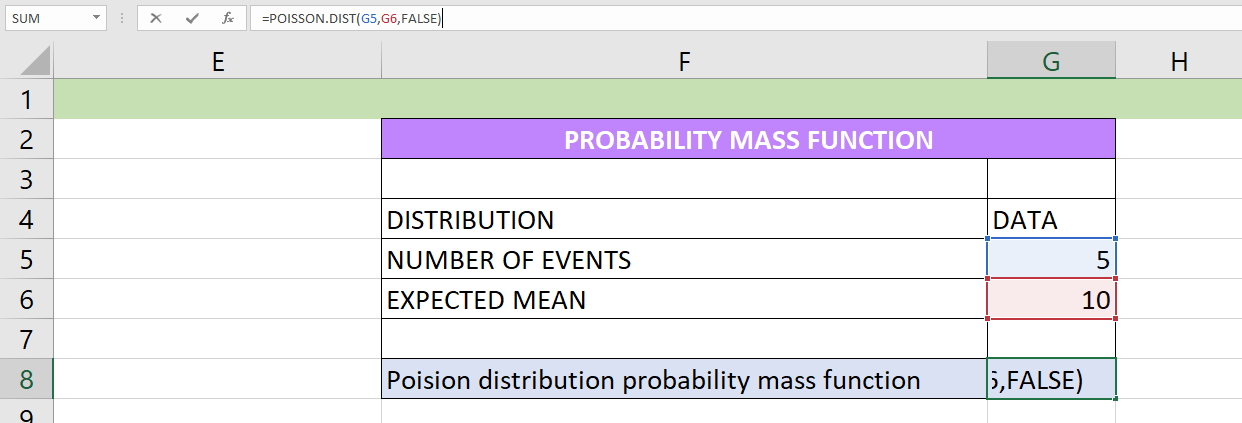
* Formula for Poisson Disstribution: **“POISSON.DIST(x,mean,cumulative)”**
* The POISSON.DIST functionuses the following arguments:
* **X** (required argument) – This is the number of events for which we want to calculate the probability. The value must be greater than or equal to 0.
* **Mean** (required argument) – This is the number of expected events. The argument must be greater than or equal to zero.
* **Cumulative** (required argument) – This is the logical argument that specifies the type of distribution to be calculated. It can either be:
* **TRUE** – Returns the cumulative Poisson probability that the number of random events occurring will be between zero and x inclusive.
* **FALSE** – Returns the probability mass function that the number of events occurring will be exactly x.

Example:

* Suppose we are given the following data:
* Number of events: 5
* Expected mean: 10
* To find out the Cumulative Poisson probability, we will use the following formula: **“=POISSON.DIST(B3,B4,TRUE)”**.



* To find out the Cumulative Poisson probability, we will use the following formula: **“=POISSON.DIST(B3,B4,FALSE)”**.



1. Create a spreadsheet application to implement just probability law.

* Probabilities and inverse probabilities in Excel.
* T-Distribution probabilities and inverse probabilities
* Normal Distribution probabilities and inverse probabilities
* Other distributions
* Random draws from distributions such as the normal.
* These are useful in introductory statistics classes.
* Their main use in analysis of economics data is to obtain p-values (use TINV) or critical values (use TDIST).
* **PROBABILITIES AND INVERSE PROBABILITIES.**

We consider the standard normal distribution as an example.

Let X be random variale, x be a value of the random variable, and p be a probability. Then:

* A probability such Pr(X<=x) is given by the cumulative distribution function.

So the excel command includes “DIST”

e.g. TDIST for the T distribution

e.g. NORMSDIST for the standard normal distribution

e.g. NORMDIST for the normal distribution

* A value of x such that Pr(X<=x) = p for some specified value of p is called the inverse of the cumulative distribution function.

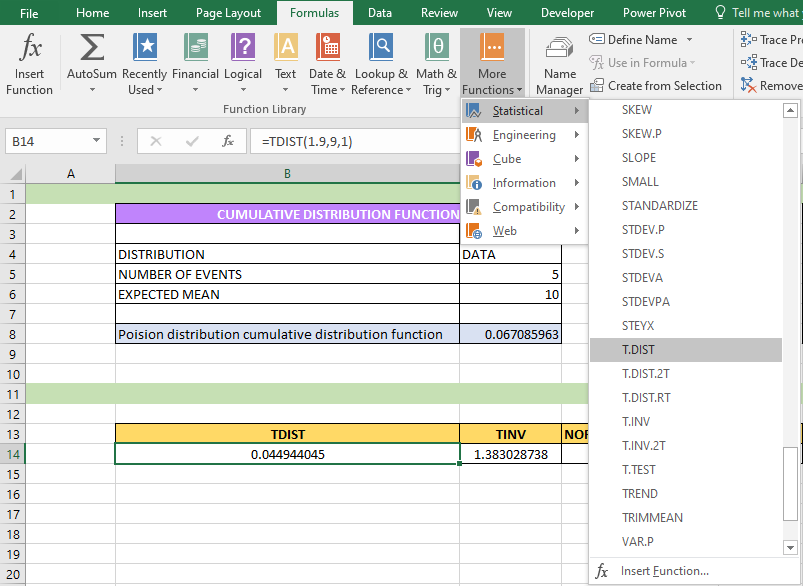
So the excel command includes “DIST”

e.g. TINV for the T distribution

e.g. NORMSINV for the standard normal distribution

e.g. NORMINV for the normal distribution

* These functions are given in Formulas Tab | Function Library Group | More Functions | Statistical.



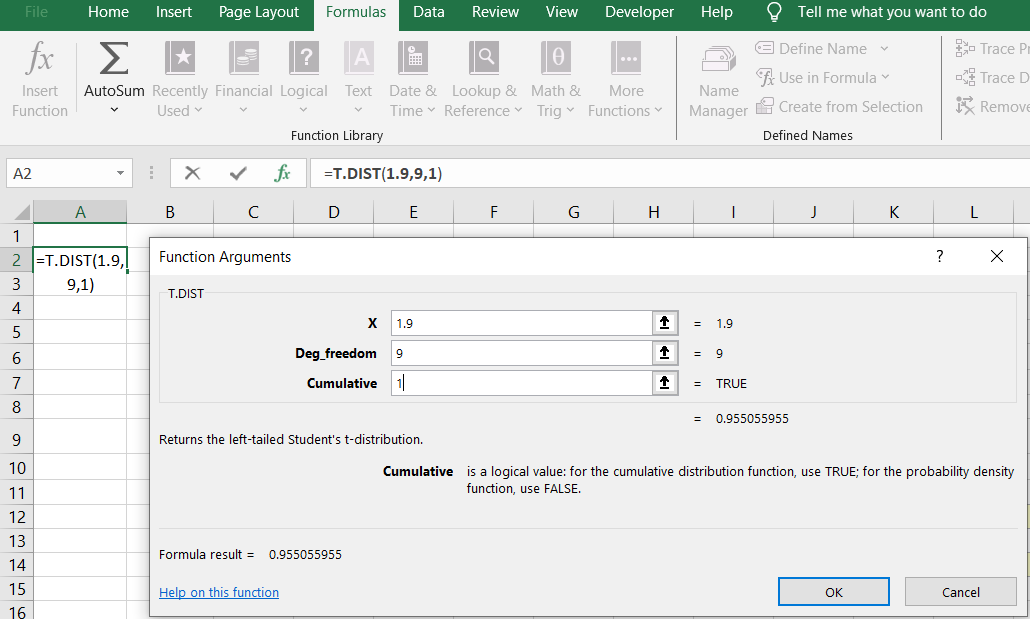
* **T-DISTRIBUTION PROBABILITIES AND INVERSE PROBABILITIES.**
* These are the most commonly-used probabilities in statistical analysis of economics data.
* These use the TDIST and TINV functions.
* TDIST gives the probability of being in the right tail i.e. Pr(X>x), or of being in both tails i.e. Pr(|X|>x).
* TINV considers the inverse of the probability of being in both tails.

1. Find Pr(X<=1.9) when x is t-distributed with 9 degrees of freedom.

This is 1-Pr(X>1.9) where Excel function TDIST gives Pr(X>1.9).

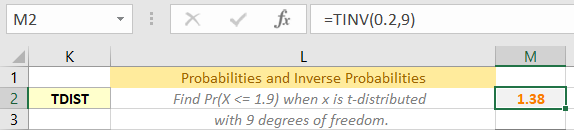
Choose Formulas Tab | Function Library Group | More Functions | Statistical | TDIST.

Fill in the Functions Arguments Tab:



This gives result that Pr(X>1.9)=0.0449.

Much simpler is to directly type in the cell = 1 – TDIST(1.9,9,1) and hit <enter>.



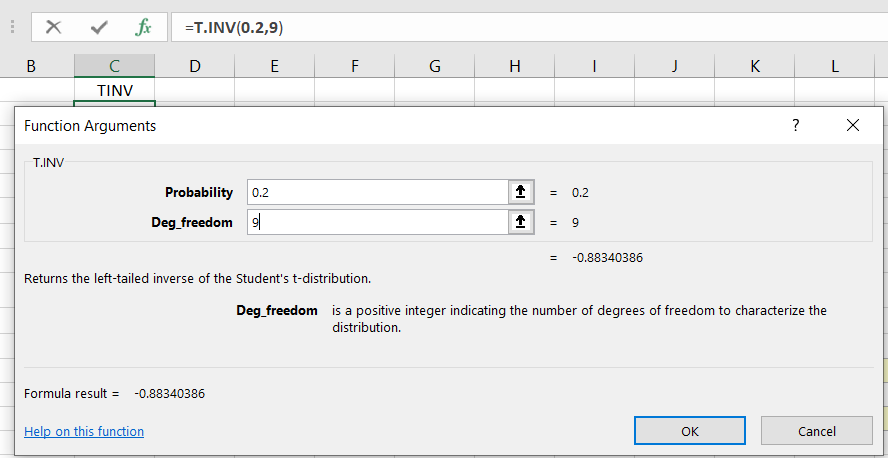
1. Find the value x\* such that Pr(X<=x\*)=0.9 when x is t-distributed with 9 degrees of freedom.

This is the same value as that for which Pr(|X|>=x\*)=0.2.

(Since there is probability 0.1 in the tail right tail and probability 0.1 in the left tail).

Choose Formulas Tab | Function Library Group | More Functions | Statistical | TINV.

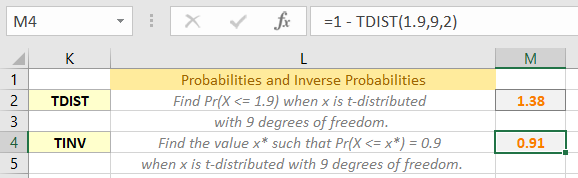
Fill in the Functions Arguments Tab:



This gives result that Pr(|X|>1.383)=0.2.

So Pr(X<=1.383)=0.9.

Much simpler is to directly type in the cell = TINV(0.2,9) and hit <enter>to get Pr(|X|>1.383)=0.2.

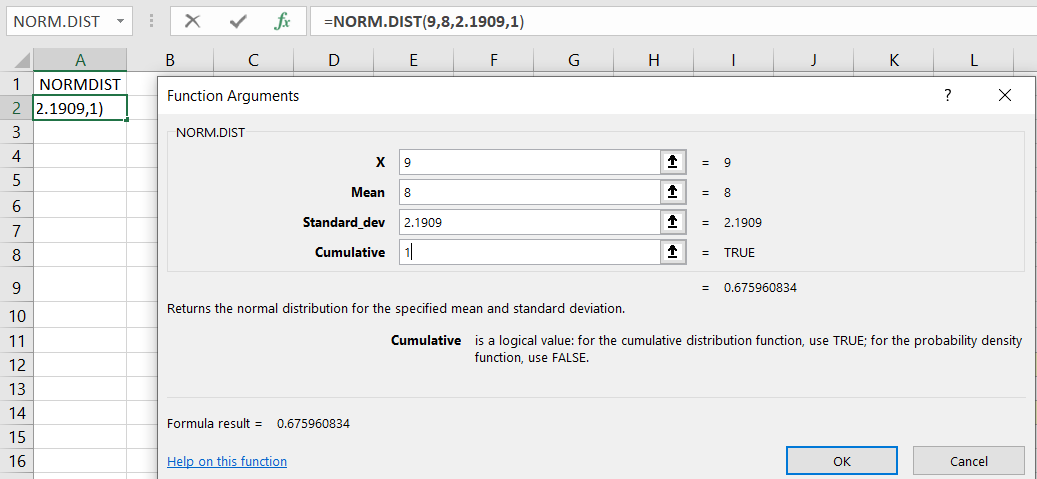


* **STANDARD NORMAL DISTRIBUTION PROBABILITIES AND INVERSE PROBABILITIES.**
* **NORMAL DISTRIBUTION PROBABILITIES AND INVERSE PROBABILITIES.**
* The standard normal sets the mean to 0 and standard deviation to 1.
* Here we consider the normal distribution with other values for the mean µ and standard deviationσ.
* The functions used are NORMDIST and NORMINV.

1. Find Pr(X <= 9) when x is normal with mean µ =8 and variance 4.8. Here standard deviation = **=** σ = sqrt(4.8) = 2.1909.

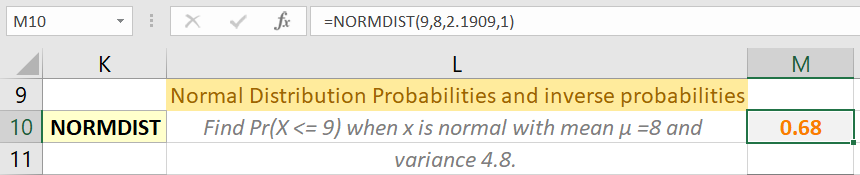
Choose Formulas Tab | Function Library Group | More Functions | Statistical | TINV.

Fill in the Functions Arguments Tab:



This gives result that Pr(X>9)=0.67596 for X normally distributed with mean 8 and variance 4.8.

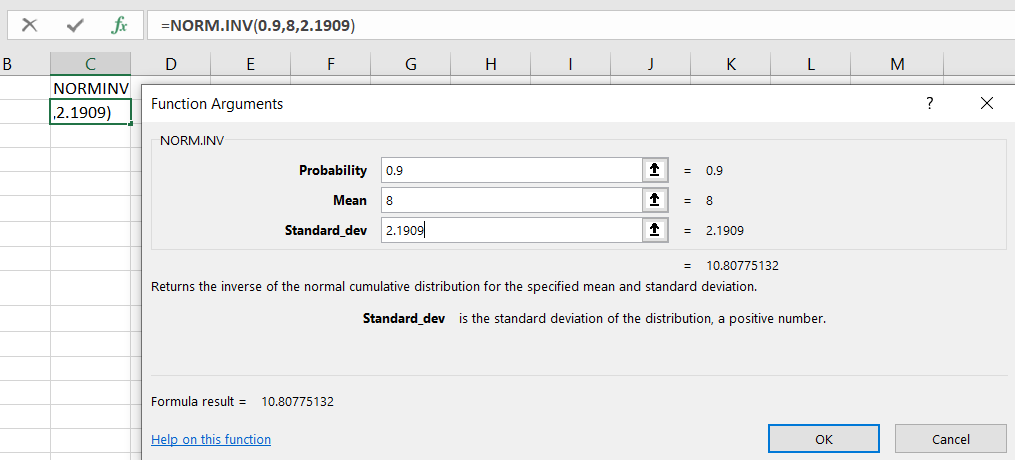
Much simpler is to directly type in the cell = NORMDIST(9,8,2.1909,1) and hit <enter>.



1. Find the value x\* such that Pr(X <= x\*) = 0.9 when x is normal with mean µ =8 and variance 4.8, so standard deviation **=** σ = sqrt(4.8) = 2.1909.

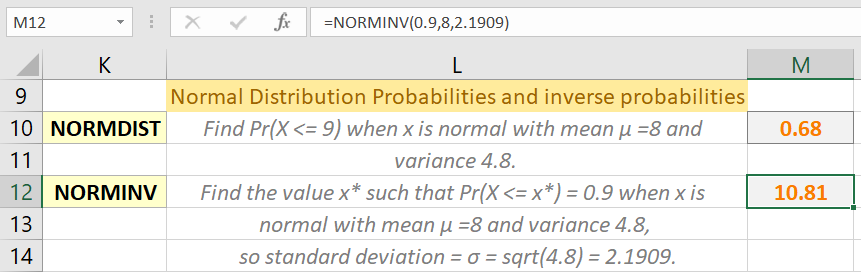
Choose Formulas Tab | Function Library Group | More Functions | Statistical | TINV.

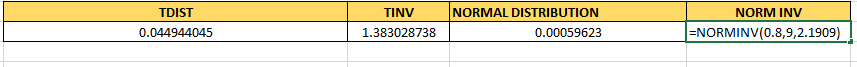
Fill in the Functions Arguments Tab:



This gives result x\*=10.8077. i.e. Pr(X>10.8077)=0.9 when x is normal with mean µ =8 and variance 4.8.

Much simpler is to directly type in the cell = NORMINV(0.9,8,2.1909) and hit <enter>.



****

* **OTHER DISTRIBUTIONS.**

Excel provides probabilities for the following distributions (in Formulas Tab | Function Library Group | More Functions | Statistical), presented in approximate order of most commonly used in the analysis of economics data:

* Normal: NORMDIST, NORMINV
* Standard normal: NORMSDIST, NORMSINV
* t-distribution: TDIST, TINV
* F-distribution: FDIST, FINV
* Chi-square: CHIDIST, CHIINV
* Lognormal: LOGNORMDIST, LOGINV
* Binomial: BINOMDIST, CRITBINOM
* Hypergeometric: HYPGEOMDIST
* Beta: BETADIST, BETAINV
* Gamma: GAMMADIST, GAMMAINV
* Exponential: EXPONDIST
* Weibull: WEIBULL
* Poisson: POISSON
* Negative binomial: NEGBINOMDIST
* **RANDOM** **NUMBER GENERATION**.

It can be useful to generate a random sample of observations from a specified distribution, such as the standard normal.

Use Data Tab | Analysis Group | Data Analysis.

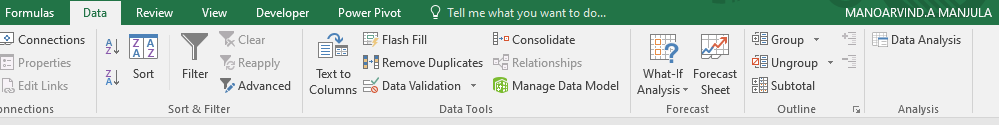
This permits generation from:

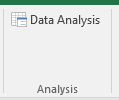
* Normal
* Uniform
* Bernoulli (0 or 1)
* Binomial
* Poisson
* Discrete (you provide the values and probabilities for a discrete distribution with finite number of possible values)
* Patterned

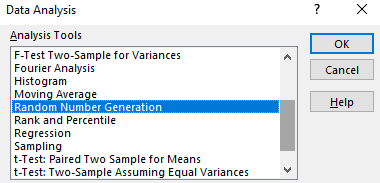
Example: Normal

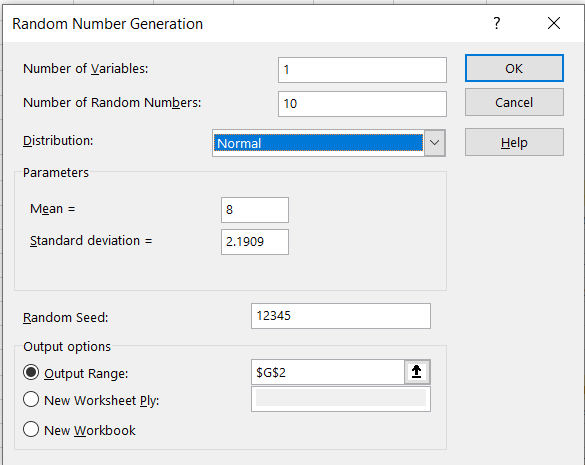
Generate 1000 vlues of x where x is normal with mean mu = 8 and variance 4.8, so standard deviation = sigma = sqrt(4.8) = 2.1909.

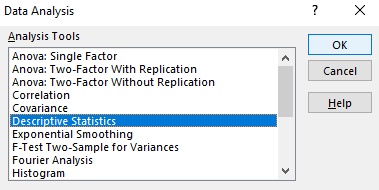
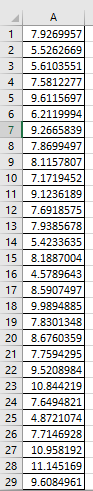
Choose Data | Analysis | Data Analysis | Random Number Generation.  
Then in the Random Number Generation dialog box fill in:

****

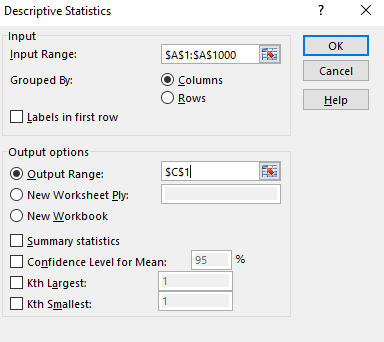
****

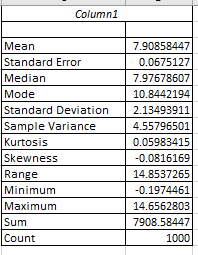
****

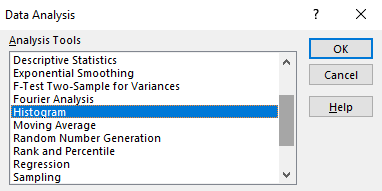


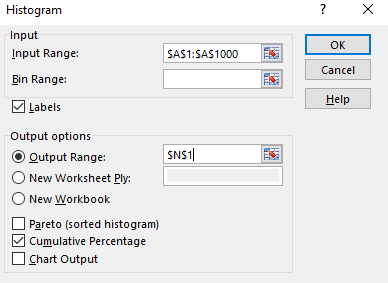
****

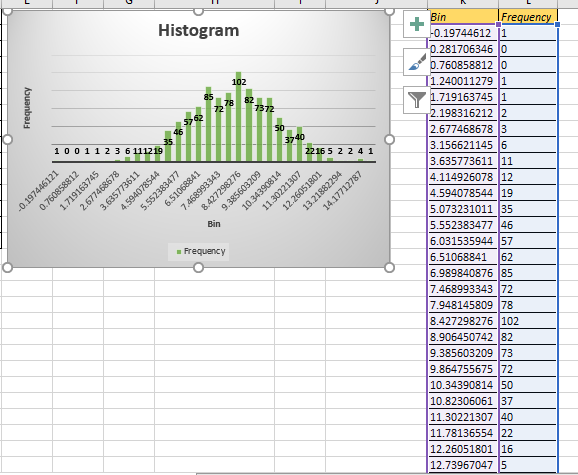
The 1,000 random draws have sample mean close to 8, sample variance close to 4.8, and histogram that is close to bell-shaped curve.

****

****

****

****

****

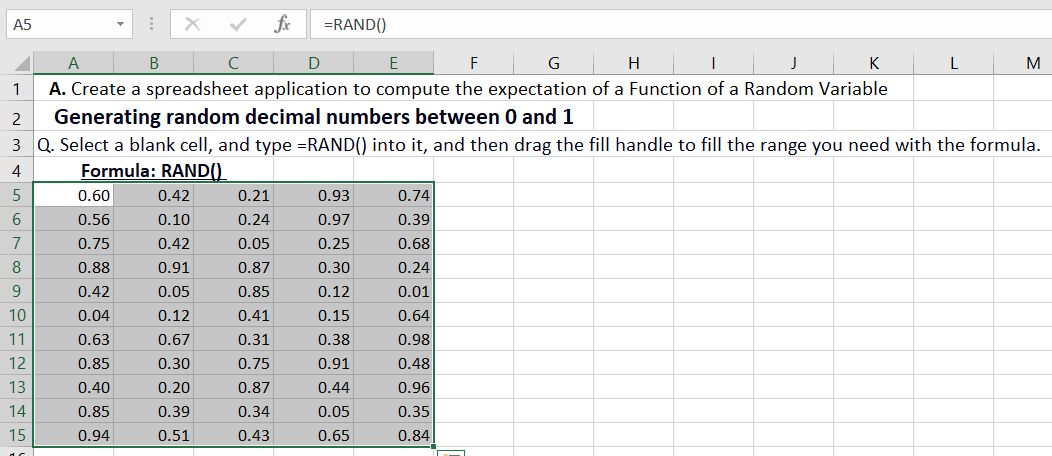
**Practical No. 5**

**Aim:** Mathematical Expectation and Chebyshev’s Theorem.

1. Create a spreadsheet application to compute the expectation of a Function of a Random Variable.

* Generating random decimal numbers between 0 and 1.

1. Select a blank cell, and type the formula: **“=RAND()”** into it, and then drag the fill the range you need with the formula.

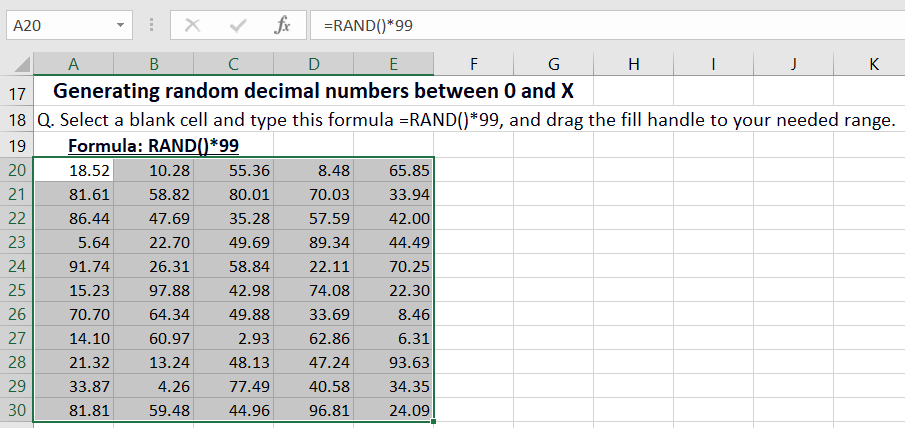


1. Then select the range you have applied the formula, and click the Increase Decimal button or Decrease Decimal button under Home Tab to specify the decimal numbers.

* Generating random numbers between 0 and X.

To randomize decimal number between 0 and X (X indicates any maximum number you need), you can use this formula: **“=RAND()\*X”**. For instance, I want to generate random decimal numbers between 0 and 99, now do as below:

1. Select a blank cell and type this formula: **“=RAND()\*99”**, and drag the fill handle to your needed range. See screenshot:

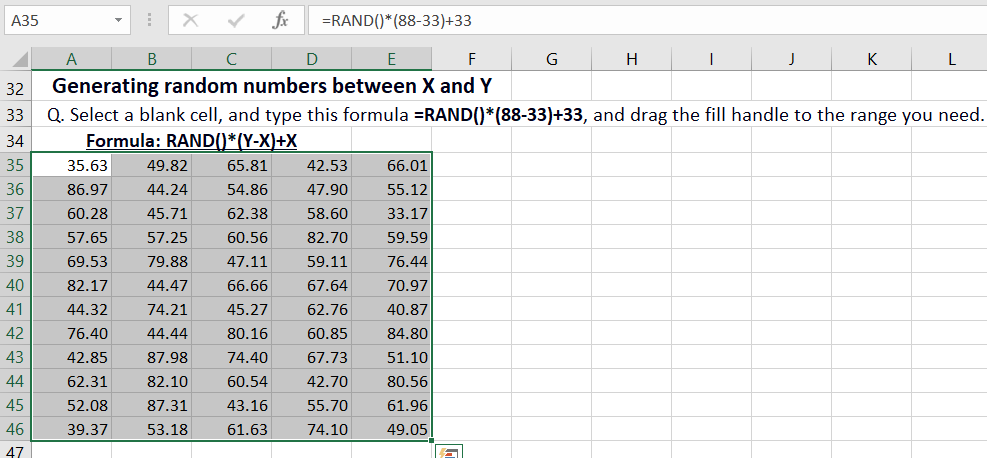


1. Then click Home Tab, and go to click Increase Decimal button or Decrease Decimal button to specify the decimal numbers.

* Generating random numbers between X and Y.

If you want to insert random numbers between X and Y, you can use this formula: **“=RAND()\*(Y-X)+X”** (X and Y indicate any number, and X<Y). For instance, I will insert random numbers between 33 and 88, now you can do this:

1. Select a blank cell, and type this formula: **“=RAND()\*(88-33)+33”**, and drag the fill handle to the range you need. See Screenshot:



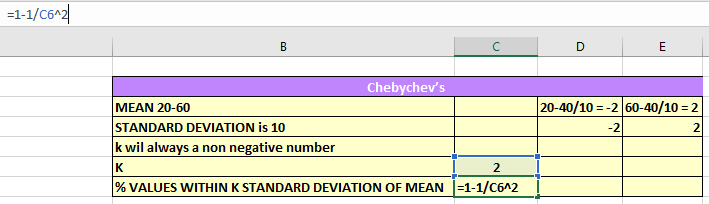
1. Then click Home Tab, and go to click Increase Decimal button or Decrease Decimal button to specify the decimal numbers.
2. Create a spreadsheet application to apply Chebyshev’s Theorem

**Example 1:** Use Chebyshev’s Theorem to find what percentage of values will fall between 20 and 60 for a dataset with a mean of 40 and a standard deviation of 10. To begin with, decide the incentive for k. We can do this by figuring out the number of standard deviations away 20 and 60 that are from the mean:

(20 – mean) / standard deviation = (20 – 40) / 10 = -20 / 10 = -2

(60 – mean) / standard deviation = (60 – 40) / 10 = 20 / 10 = 2

The qualities 20 and 60 are 2 standard deviations underneath or more than the mean, individually. In this way, k = 2. We can then utilize the accompanying equation in Excel to find the base level of values that fall inside 2 standard deviations of the mean for this dataset:



The level of values that fall between 20 and 60 for this dataset will be something like 75%.

**Example 2:** Use Chebyshev’s Theorem to find what percentage of values will fall between 10 and 40 for a dataset with a mean of 25 and a standard deviation of 5. To begin with, decide the incentive for k. We can do this by figuring out the number of standard deviations away 10 and 40 that are from the mean:

(10 – mean) / standard deviation = (10 – 25) / 5 = -15 / 5 = -3

(40 – mean) / standard deviation = (40 – 25) / 5 = 15 / 5 = 3

The qualities 10 and 40 are 3 standard deviations underneath or more than the mean, individually. Hence, k = 3. We can then utilize the accompanying recipe in Excel to find the base level of values that fall inside 3 standard deviations of the mean for this dataset: