STATISTICS - 2 - Assignment

By Prakash Ghosh

In [1]:
```
# import the libraries
import pandas as pd
import numpy as np
import scipy as sci
import matplotlib.pyplot as plt
import math
from scipy.stats import binom
```

Problem Statement 1:

A test is conducted which is consisting of 20 MCQs (multiple choices questions) with every MCQ having its four options out of which only one is correct. Determine the probability that a person undertaking that test has answered exactly 5 questions wrong.

In [26]:
```
# Define the variable
total_num_of_question = 20

# probality of wrong answer of a single question
p_wrong_answer = 3/4  # as one of the 4 question is right
print('Probality of wrong answer of a single question:\t', p_wrong_answer)

print('\nProbality of 5 wrong answers')
print('----------------------------')
print('\nSolution -1 - Mathemetical Model')
print('\t Using Formula: 20C5 * power((1-p_wrong_answer),15) * power(p_wrong_answer, 5)')

p_five_wrong_answer = (math.factorial(20)/(math.factorial(15)*math.factorial(5))) \
                      * math.pow((1-p_wrong_answer),15) * math.pow(p_wrong_answer,5)
print('\nProbality of 5 wrong answers',p_five_wrong_answer)

print('\nSolution -2 - Using scipy - probability mass function')
print('\t Using Formula: binom.pmf(5,total_num_of_question,p_wrong_answer)')

p_five_wrong_answer=binom.pmf(5,total_num_of_question,p_wrong_answer)
print('\nProbality of 5 wrong answers',p_five_wrong_answer)
```

```
Probality of wrong answer of a single question:  0.75

Probality of 5 wrong answers
----------------------------

Solution -1 - Mathemetical Model
         Using Formula: 20C5 * power((1-p_wrong_answer),15) * power(p_wrong_answer, 5)

Probality of 5 wrong answers 3.4264958230778575e-06

Solution -2 - Using scipy - probability mass function
         Using Formula: binom.pmf(5,total_num_of_question,p_wrong_answer)

Probality of 5 wrong answers 3.4264958230778435e-06
```

Problem Statement 2:

A die marked A to E is rolled 50 times. Find the probability of getting a “D” exactly 5 times.

In [29]:
```
# Define the variable
total_num_of_roll = 50

# probality of getting D in single roll
p_D = 1/5  # as in a roll, D has equal chance among A to E
print('Probality of getting D in single roll:\t', p_D)

print('\nProbality of getting D exactly 5 times')
print('---------------------------------------')
print('\nSolution -1 - Mathemetical Model')
print('\t Using Formula: 50C5 * power((1-p_D),45) * power(p_D, 5)')

p_five_D = (math.factorial(50)/(math.factorial(45)*math.factorial(5))) \
                        * math.pow((1-p_D),45) * math.pow(p_D,5)
print('\nProbality of getting D exactly 5 times',p_five_D)

print('\nSolution -2 - Using scipy - probability mass function')
print('\t Using Formula: binom.pmf(5,total_num_of_roll,p_D)')

p_five_D=binom.pmf(5,total_num_of_roll,p_D)
print('\nProbality of 5 wrong answers',p_five_D)
```

```
Probality of getting D in single roll:   0.2

Probality of getting D exactly 5 times
---------------------------------------

Solution -1 - Mathemetical Model
         Using Formula: 50C5 * power((1-p_D),45) * power(p_D, 5)

Probality of getting D exactly 5 times 0.029531204310524292

Solution -2 - Using scipy - probability mass function
         Using Formula: binom.pmf(5,total_num_of_roll,p_D)

Probality of 5 wrong answers 0.029531204310523224
```

Problem Statement 3:

Two balls are drawn at random in succession without replacement from an urn containing 4 red balls and 6 black balls. Find the probabilities of all the possible outcomes.

In [49]:
```
# Define the variable
total_num_of_balls = 10
num_of_red_balls = 4
num_of_black_balls = 6

# Two balls are drawn at random in succession without replacement

# If Red ball is denoted by R and Balck ball is denoted by B then
# Possible outcome - RR, RB, BR, BB

# probability of first ball red = 4/10
# probability of second ball red = 3/9 [when first ball is red]
# probability of second ball red = 4/9 [when first ball is black]

# probability of first ball black = 6/10
# probability of second ball black = 5/9 [when first ball is black]
# probability of second ball black = 6/9 [when first ball is red]

probablity_RR = (4/10) * (3/9)
probablity_RB = (4/10) * (6/9)
probablity_BR = (6/10) * (4/9)
probablity_BB = (6/10) * (5/9)

# Create a Dataframe with the Probability distribution and random variable

lst_color=['RR','RB','BR','BB']

df_probablity=pd.DataFrame({'Color':lst_color,
                'Probablity':[probablity_RR,probablity_RB,probablity_BR,probablity_BB]})
print(df_probablity)

# Plot the Probabalitis distributions
plt.bar(df_probablity.Color,df_probablity.Probablity,width=.3)
plt.xlabel('Color of the Balls')
plt.xticks(lst_color)
plt.ylabel('Probabality')
plt.title('\nProbabalitis distribution Plot\n')
plt.show()
```

	Color	Probablity
0	RR	0.133333
1	RB	0.266667
2	BR	0.266667
3	BB	0.333333