**10. Write a program to calculate chi-square value using Python. Report your observation**

**Chi-Square Test**

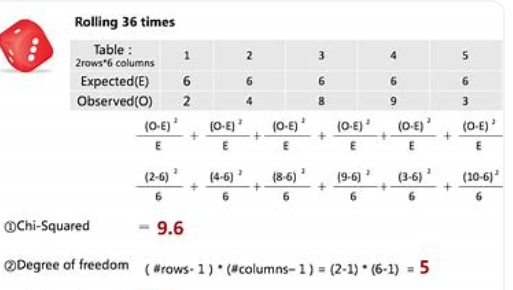
The Chi-Square test is a statistical procedure for determining the difference between observed and expected data. This test can also be used to determine whether it correlates to the categorical variables in our data. It helps to find out whether a difference between two categorical variables is due to chance or a relationship between them.

**chi2:** The test statistic

**p:** The p-value of the test

**dof:** Degrees of freedom

**expected:** The expected frequencies, based on the marginal sums of the table



import pandas as pd

import numpy as np

from scipy.stats import chi2\_contingency

import seaborn as sns

import matplotlib.pyplot as plt

%matplotlib inline

df = pd.DataFrame({'Gender' : ['M', 'M', 'M', 'F', 'F'] \* 10, 'isSmoker' : ['Smoker', 'Smoker', 'Non-Smpoker', 'Non-Smpoker', 'Smoker'] \* 10 })

df.head()

contigency= pd.crosstab(df['Gender'], df['isSmoker'])

contigency

contigency\_pct = pd.crosstab(df['Gender'], df['isSmoker'], normalize='index')

contigency\_pct

plt.figure(figsize=(12,8))

sns.heatmap(contigency, annot=True, cmap="YlGnBu")

# Chi-square test of independence.

c, p, dof, expected = chi2\_contingency(contigency)

# Print the p-value

print(p)

