**Probability Basics - Session 1**

# Practice Exercise Solutions

|  |  |
| --- | --- |
| Blood Group | People |
| A | 50 |
| B | 65 |
| O | 70 |
| AB | 15 |

P(O) = Freq of O / Total Frequencies = 70/200 =0.35

1. Sample space: {HH, HT, TH, TT}

Random Variable, X : Number of heads observed.

Rx = {0,1,2}

X is a discrete R.V.

Px (X=0) = P(TT) = 1/4

Px (X=1) = P(HT, TH) = 2/4

Px (X=2) = P(HH) = 1/4

1. import matplotlib.pyplot as plt

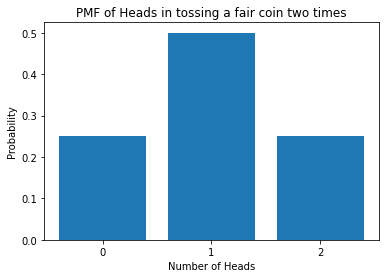
pmf = {'0':0.25, '1': 0.5, '2': 0.25}

plt.bar(pmf.keys(),pmf.values())

plt.xlabel("Number of Heads")

plt.ylabel("Probability")

plt.title("PMF of Heads in tossing a fair coin two times")



1. **Gender** is categorical so we can use PMF to understand its distribution

**Age and Annual Income:** are continuous so PDF

**Spending Score:** has countable values in range 1-100 so PMF can be used.

But this dataset has 200 records so not many values will be repeated.

It is better to observe the value counts in such a case.

Since, not many values are repeated this can be treated as continuous R.V. and PDF can be used.

min = df['Spending Score (1-100)'].min()

max = df['Spending Score (1-100)'].max()

df['SpendPower'] = pd.cut(x=df['Spending Score (1-100)'],

                      bins=[min,20,60,max],

                      labels=['Low', 'Medium', 'High'])

1. **NO**

gender\_spend = pd.crosstab(index=[df["Gender"],df['SpendPower']], columns="count")

print((100\*(gender\_spend.loc['Male','High'])/gender\_spend['count'].sum())['count'])

print((100\*(gender\_spend.loc['Female','High'])/gender\_spend['count'].sum())['count'])

13.636363636363637

17.67676767676768

1. df['SpendPower'].value\_counts()

x=df['SpendPower'].value\_counts()/df.shape[0]#.plot.hist()

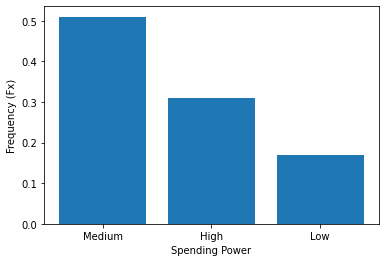
plt.bar(x.index,x.values)

plt.xlabel("Spending Power")

plt.ylabel("Frequency (Fx)")

#find probability of P(X=Child,Baby)

x['Low']+x['High']



**0.48**

1. From frequency table it is evident that medium group are dominant so we can send out offers like Coupon for next purchases since they are more likely to come back

Since low spending customers form around 18% to encourage them to spend more offers like cashbacks and discounts can be promoted.

For high spending customers, luxury products and exclusive experiences can be promoted.