

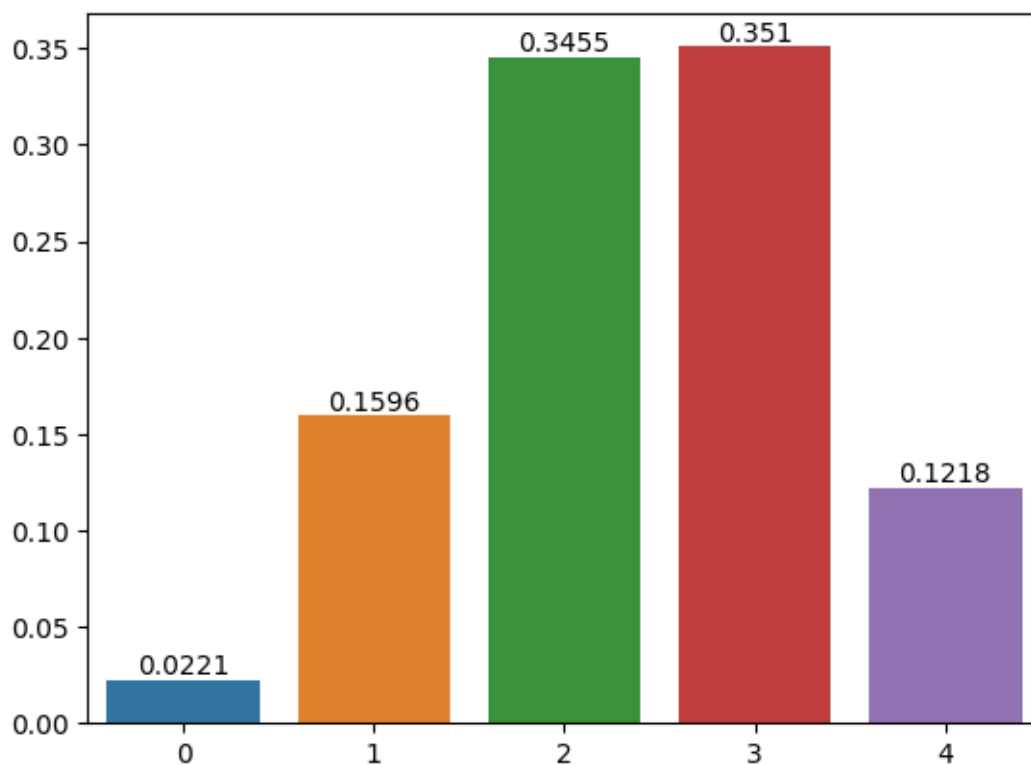
distributions-2-may23-beginner-mon2

November 24, 2023

```
[1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[2]: red_values = []
for person in range(10000):
    rolls = np.random.choice(['R', 'R', 'R', 'B', 'B'], size=4)
    num_red = np.count_nonzero(rolls=='R')
    red_values.append(num_red)
x = pd.value_counts(red_values, normalize=True)
```

```
[7]: ax = sns.barplot(x=x.index, y=x.values)
for i in ax.containers:
    ax.bar_label(i,)
```



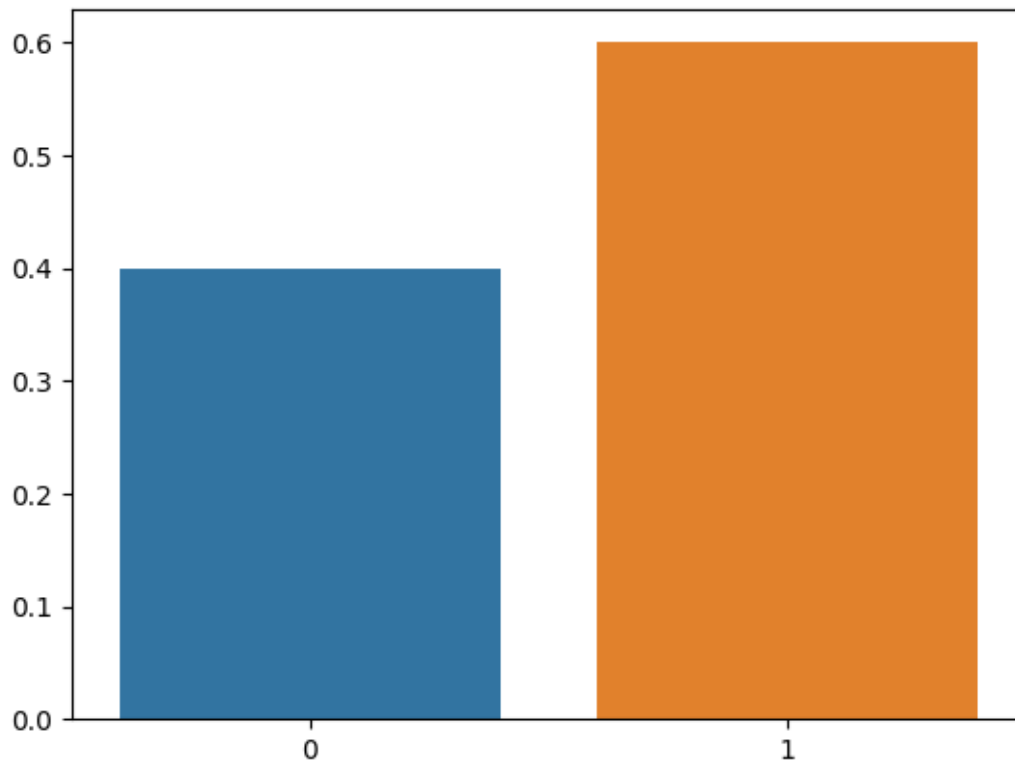
```
[5]: from scipy.stats import binom
```

```
[6]: binom.pmf(n=4, p=3/5, k=0)
```

```
[6]: 0.025599999999999994
```

```
[8]: x = [0, 1]  
y = [2/5, 3/5]  
sns.barplot(x=x, y=y)
```

```
[8]: <Axes: >
```



```
[9]: !wget https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/035/126/  
original/weight-height.csv?1684995383 -O weight-height.csv
```

```
--2023-11-24 16:52:19-- https://d2beiqkhq929f0.cloudfront.net/public_assets/ass  
ets/000/035/126/original/weight-height.csv?1684995383  
Resolving d2beiqkhq929f0.cloudfront.net (d2beiqkhq929f0.cloudfront.net)...  
99.84.178.132, 99.84.178.172, 99.84.178.226, ...  
Connecting to d2beiqkhq929f0.cloudfront.net
```

(d2beiqkhq929f0.cloudfront.net)|99.84.178.132|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 428120 (418K) [text/plain]
Saving to: 'weight-height.csv'

weight-height.csv 100%[=====>] 418.09K 555KB/s in 0.8s

2023-11-24 16:52:20 (555 KB/s) - 'weight-height.csv' saved [428120/428120]

```
[10]: df_hw = pd.read_csv('weight-height.csv')
      df_height = df_hw['Height']
      df_height.head()
```

```
[10]: 0    73.847017
      1    68.781904
      2    74.110105
      3    71.730978
      4    69.881796
      Name: Height, dtype: float64
```

```
[12]: df_hw.shape
```

```
[12]: (10000, 3)
```

```
[13]: df_height.mean()
```

```
[13]: 66.36755975482124
```

```
[19]: np.mean(df_height.sample(5000))
```

```
[19]: 66.36485094233258
```

```
[20]: df_height.sample(5000)
```

```
[20]: 8752    62.890520
      8835    63.178726
      1569    66.141697
      9679    61.593362
      5020    62.286838
      ...
      1956    69.342366
      252    65.929422
      3854    71.017948
      3245    65.035786
      2679    69.806291
      Name: Height, Length: 5000, dtype: float64
```

```
[21]: 1 - binom.cdf(n=10, p=1/4, k=3)
```

```
[21]: 0.22412490844726562
```

```
[22]: binom.pmf(n=10, k=2, p=1/4)
```

```
[22]: 0.28156757354736334
```