Ivana Zdravevska

MATH333-017 Probability and Statistics

Prof. Matthew Michal

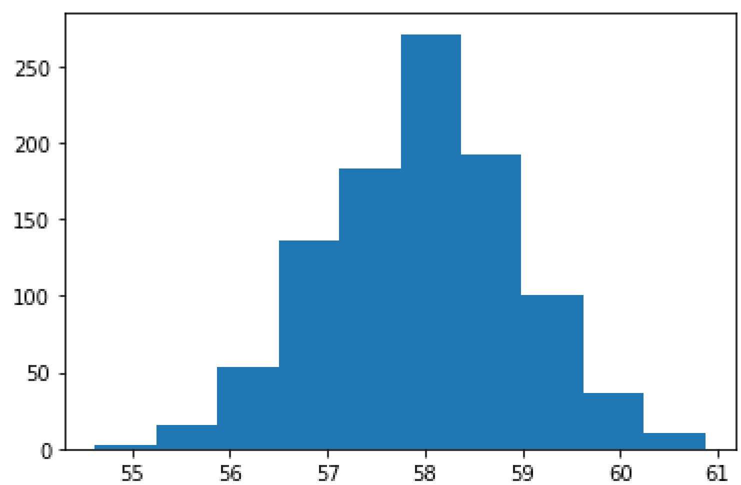
Project 5

In [47]:
```
from numpy import mean
from scipy.stats import norm
import matplotlib.pyplot as plt
import statistics

array1_norm = norm.rvs(loc=58,scale=10, size=100)
mean_norm = array1_norm.mean()
print("The mean of a normal distribution array with a mean of 58, standa
rd deviation of 10, and a size of 100 is " + str(mean_norm))

means = [mean(norm.rvs(loc=58,scale=10, size=100)) for _i in range(1000
)]
plt.hist(means)
plt.show()
print('The mean of the sample means is {}'.format(mean(means)))
print('The standard deviation of the sample means is {}'.format(statisti
cs.stdev(means)))
```

```
The mean of a normal distribution array with a mean of 58, standard dev
iation of 10, and a size of 100 is 57.693801920609765
```



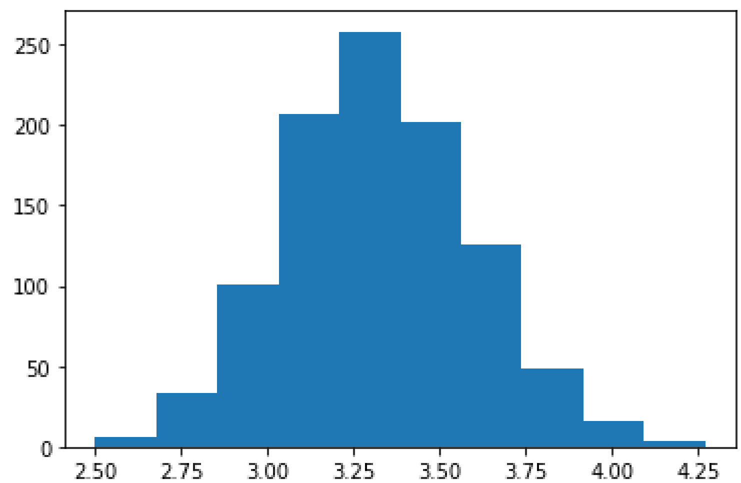

```
The mean of the sample means is 57.953916060320765
The standard deviation of the sample means is 0.9803550195810288
```

In [10]:
```
from numpy import mean
from scipy.stats import geom
import matplotlib.pyplot as plt
import statistics

array1_geom = geom.rvs(p=0.3, size=100)
mean_geom = array1_geom.mean()
print("The mean of a geometric distribution array with a probability of
 0.3 and a size of 100 is " + str(mean_geom))

means2 = [mean(geom.rvs(p=0.3, size=100)) for _i in range(1000)]
plt.hist(means2)
plt.show()
print('The mean of the sample means is {}'.format(mean(means2)))
print('The standard deviation of the sample means is {}'.format(statisti
cs.stdev(means2)))
```

```
The mean of a geometric distribution array with a probability of 0.3 an
d a size of 100 is 3.92
```




```
The mean of the sample means is 3.32452
The standard deviation of the sample means is 0.2767134617974701
```

When calculating the theoretical standard deviation of a sample, we should use 1000 for the value of n. Since the variance of a sample mean is σ^2/n, the standard deviation is σ/squareroot(n). This means that as n increases, the variance shrinks. Therefore, as the size increases we get more precise estimates of the group means. The standard error decreases leading to an increase in higher accuracy. The smaller the standard error, the more powerful the statistical test is since it brings us closer to the true value and away from an error. With n being bigger, the variance shrinks which means that the measure of the spread in the sample also decreases giving us data that is less spread out and more accurate.