Central Limit Theorem - Part III

December 1, 2017

0.0.1 Central Limit Theorem - Part III

You saw how the **Central Limit Theorem** worked for the sample mean in the earlier concept. However, let's consider another example to see a case where the **Central Limit Theorem** doesn't work...

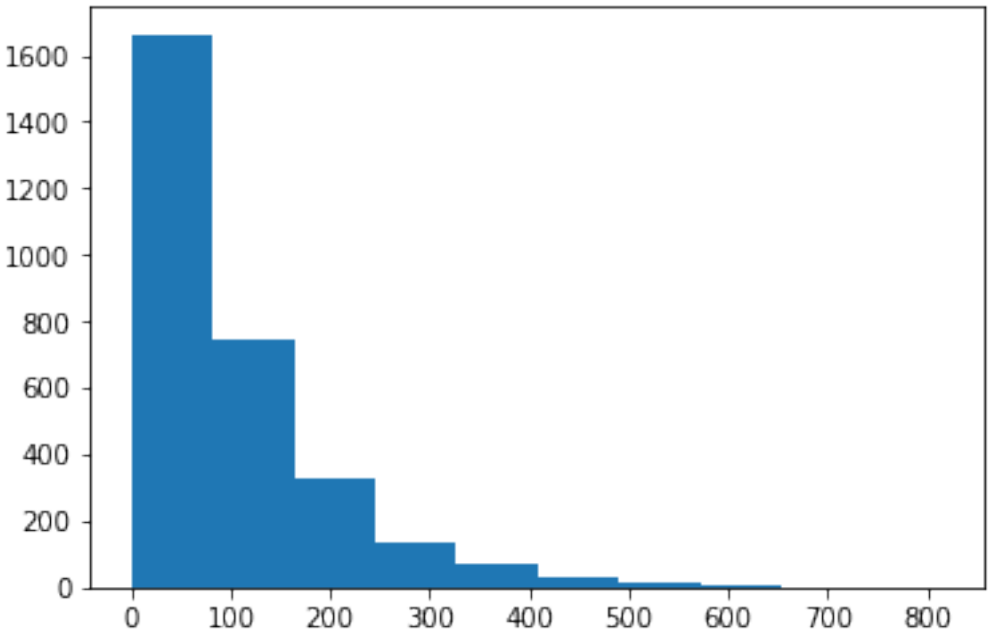
Work through the questions and use the created variables to answer the questions that follow below the notebook.

Run the below cell to get started.

```
In [1]: import numpy as np
        import matplotlib.pyplot as plt

        %matplotlib inline
        np.random.seed(42)

        pop_data = np.random.gamma(1,100,3000)
        plt.hist(pop_data);
```



```
In [6]: pop_data.var()

Out[6]: 9955.7693930654896

In [7]: np.var(pop_data)

Out[7]: 9955.7693930654896
```

1. In order to create the sampling distribution for the variance of 100 draws of this distribution, follow these steps:

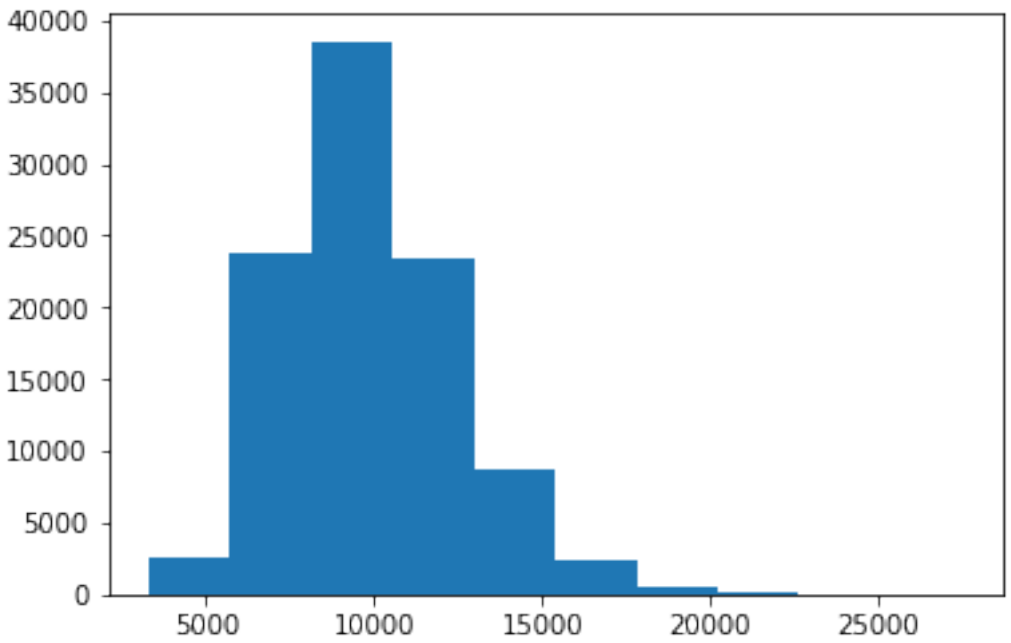
a. Use numpy's **random.choice** to simulate 100 draws from the `pop_data` array. b. Compute the variance of these 100 draws. c. Write a loop to simulate this process 10,000 times, and store each variance into an array called **var_size_100**. d. Plot a histogram of your sample variances. e. Use **var_size_100** and **pop_data** to answer the quiz questions below.

```
In [2]: np.random.choice(pop_data, size=100).var()

Out[2]: 12098.749344943082

In [3]: x = []
        for _ in range(100000):
            x.append(np.random.choice(pop_data, size=100).var())

In [5]: plt.hist(x);
```



```
In [9]: np.mean(x)
```

```
Out[9]: 9865.7382823026128
```

```
In [10]: np.var(x)
```

```
Out[10]: 6485388.9806262478
```