index

January 11, 2022

1 The Normal Distribution - Lab

1.1 Introduction

In this lab, you'll learn how to generate random normal distributions in Python. You'll learn how to visualize a histogram and build a density function using the formula.

1.2 Objectives

You will be able to:

- Use `numpy` to generate a random normal distribution
- Calculate the density function for normal distributions with a Python function
- Plot and interpret density plots and comment on the shape of the plot

1.3 A quick refresher!

Here's the formula for the normal distribution density function once more:

$$N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Here, - μ is the mean - σ is the standard deviation - $ 3.14159 $ - $ e 2.71828 $

1.4 First generate a normal distribution containing 5000 values with $\mu = 14$ and $\sigma = 2.8$

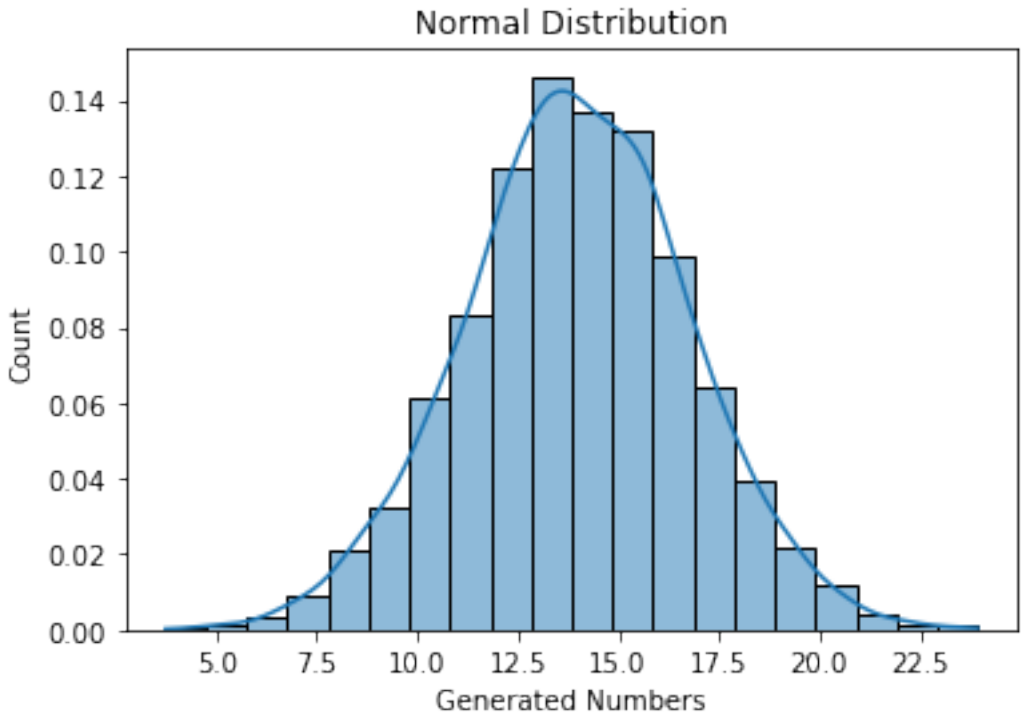
[44]:
```
# Generate a random normal variable with given parameters , n=5000
import numpy as np
mu = 14
sigma = 2.8
n = 5000
s = np.random.normal(mu, sigma, size = n)
```

1.5 Calculate a normalized histogram for this distribution in matplotlib, with bin size = 20

Make sure to get the bin positions and counts for each of the obtained bins. You can use official documentation to view input and output options for `plt.hist()`

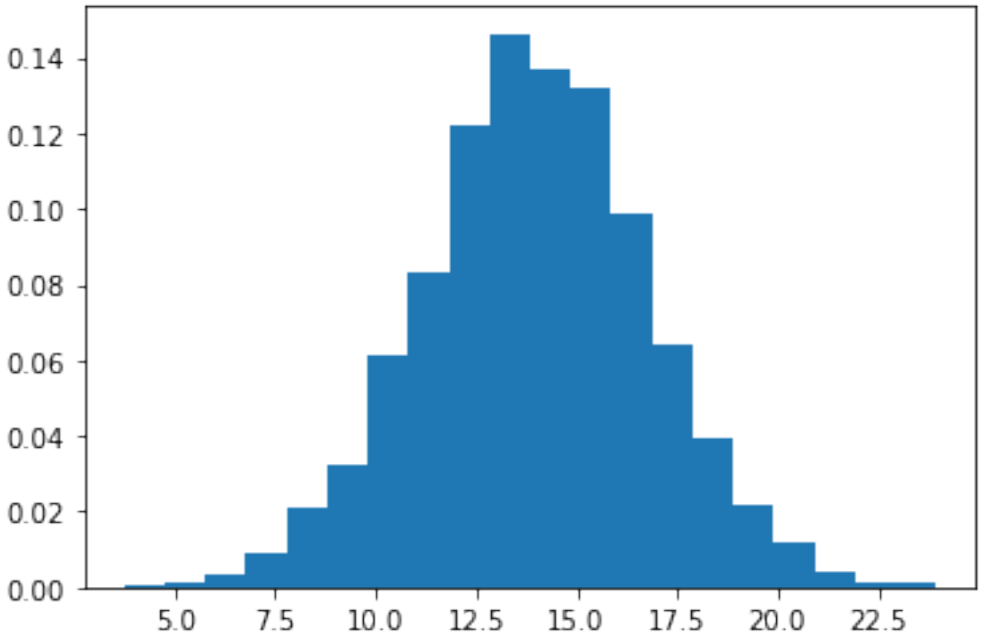
[45]:
```
# Calculate a histogram for above data distribution
import seaborn as sns


hist = sns.histplot(s,kde=True,stat="density", bins = 20)
hist.set(xlabel ="Generated Numbers", ylabel = "Count", title ='Normal␣
 ↪Distribution');
```



[46]:
```
# From GitHub Solution

import matplotlib.pyplot as plt
%matplotlib inline
# Create the bins and histogram
count, bins, ignored = plt.hist(s, 20, density=True)
```



1.6 Use the formula to calculate the density function with μ, σ and bin information obtained before

$$N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

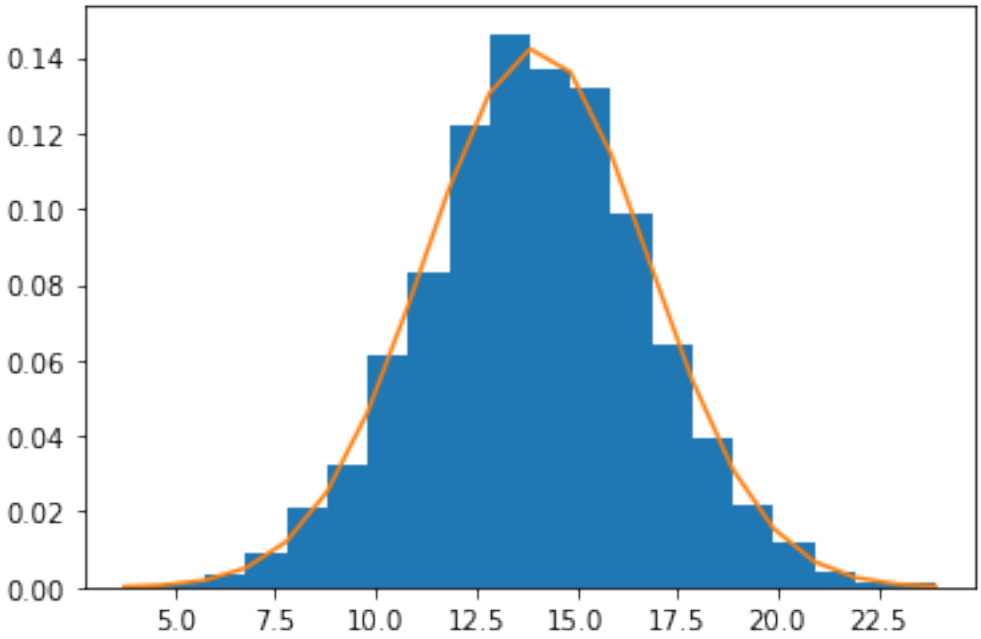
[55]:
```
# Calculate the normal Density function
density =  1 / (sigma * np.sqrt(2*np.pi)) * np.exp(-(bins-mu)**2 / (2 *␣
 ↪sigma**2))
```

1.7 Plot the histogram and density function

[58]:
```
# Plot histogram along with the density function

# From GitHub
import matplotlib.pyplot as plt
%matplotlib inline

plt.hist(s, bins = 20, density = True)
plt.plot(bins, density)
plt.show();
```

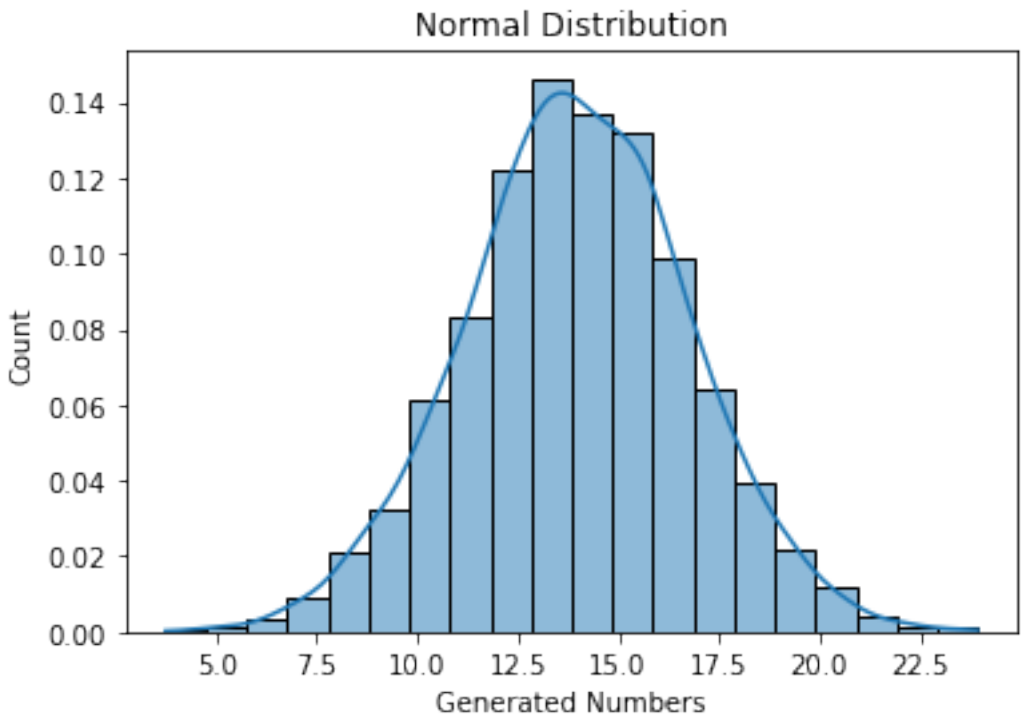


1.8 Visualize the distribution using seaborn and plot the KDE

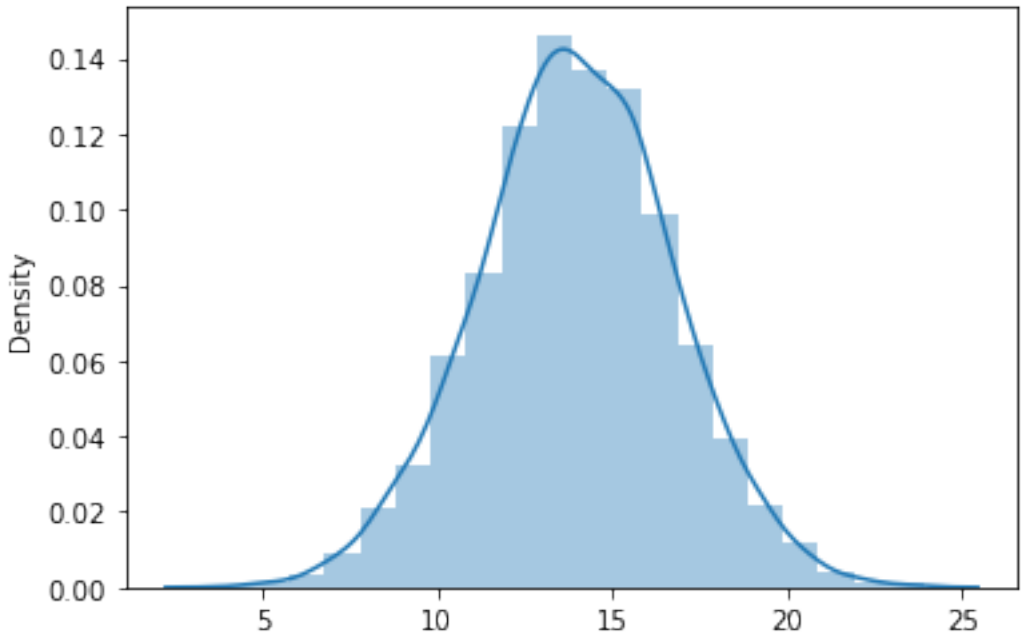
[71]:
```
# Use seaborn to plot the histogram with KDE

import seaborn as sns


hist = sns.histplot(s, bins = 20, kde = True, stat="density")
hist.set(xlabel ="Generated Numbers", ylabel = "Count",
         title ='Normal Distribution');
```



[69]:
```
import seaborn as sns
# import warnings
# warnings.filterwarnings(action='ignore', category=FutureWarning)
sns.distplot(s, bins=20, kde=True);
```



Note: Pay attention to the results of `sns.distplot` and `sns.histplot`. As we can see, the results of `sns.distplot` is already normalized but for `sns.histplot` we need to put the argument `stat="density` to normalize the data

1.9 Summary

In this lab, you learned how to generate random normal distributions in Python using Numpy. You also calculated the density for normal distributions using the general formula as well as seaborn's KDE. Next, you'll move on to learn about the standard normal distribution and how normal distributions are used to answer analytical questions.