

Gaussian/Normal Distribution - Lab

Introduction

In this lab we shall learn how to generate random normal distributions in python. We shall look into visualising a histogram and building a density function using the formula as well as seaborn's built in functions.

Objectives

You will be able to:

- Generate random normal distributions in python with given parameters
- Calculate the density function for normal distributions
- Use seaborn to visualize distributions with histograms and density functions

A quick refresher!

Here's the formula for calculating normal distribution density function.

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

μ = Mean
 σ = Standard Deviation
 $\pi \approx 3.14159\dots$
 $e \approx 2.71828\dots$

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

```
In [13]: # Generate a random normal variable with given parameters , n=5000
import numpy as np

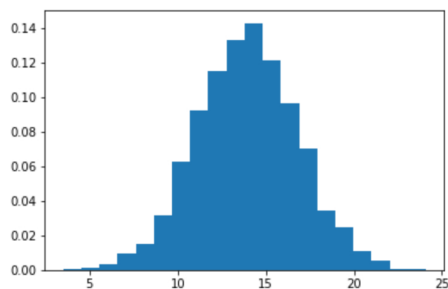
mu, sigma = 14, 2.8
N = 5000
s = np.random.normal(mu, sigma, N)
```

Calculate a normalized histogram for this distribution in matplotlib - use bin size = 20.

Get the bin positions and count for each bin

Refer to [official documentation](#) to view input and output options for `plt.hist()`

```
In [10]: import matplotlib.pyplot as plt
# Create the bins and histogram
count, bins, ignored = plt.hist(s, 20, normed=True)
```



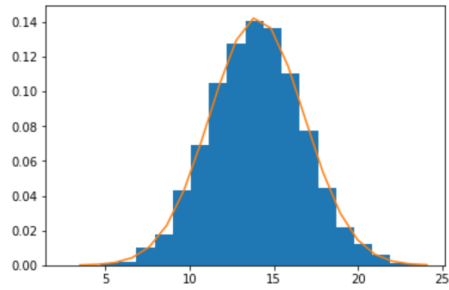
Calculate the density function (using above formula) with mu, sigma and bin information calculated above .

```
In [7]: # Calculate the normal Density function
density = 1/(sigma * np.sqrt(2 * np.pi)) * np.exp( - (bins - mu)**2 / (2 * sigma**2))
```

Plot the histogram and density function

```
In [14]: # Plot histogram along with the density function
plt.hist(s, 20, normed=True)
```

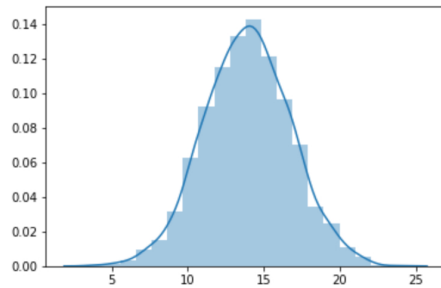
```
plt.plot(bins, density)
plt.show()
```



Visualize the distribution using seaborn and plot the KDE

```
In [12]: import seaborn as sns
sns.distplot(s, bins=20, kde=True)
```

```
Out[12]: <matplotlib.axes._subplots.AxesSubplot at 0x1a19d4cda0>
```



Summary

In this lab we saw how to generate random normal distributions in python using numpy. We also looked into calculating the density for gaussian distributions using the general formula as well as seaborn's kde. We shall now move on to see how we can analyze such variables for answering analytical questions.