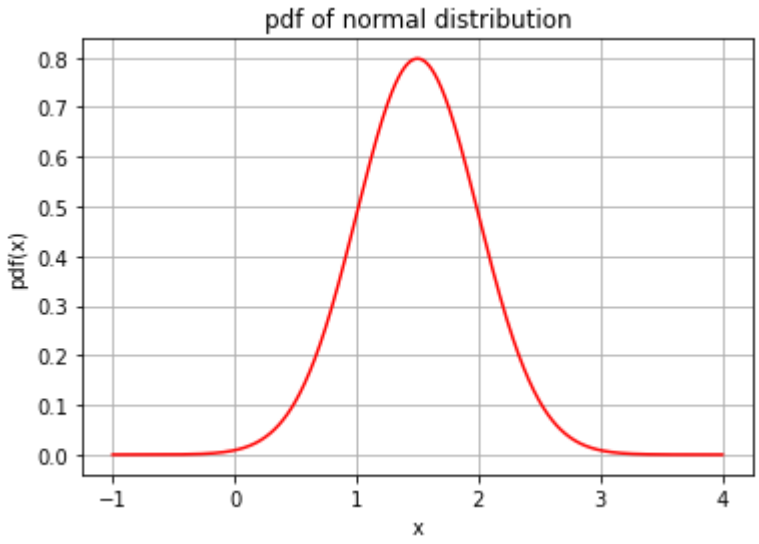
In [1]:

```
# Importing necessary libraries
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
import csv
import pandas as pd
```

Question 1

In [2]:

```
# declaring input variables
mean = 1.5
sigma = 0.5
x = np.linspace(-1, 4, 1000)

# creating gaussian distribution
gaussian = stats.norm(mean, sigma)

# plots pdf of normal distribution with mean = 1.5 and standard deviation = 0.5
plt.plot(x, gaussian.pdf(x), color = 'red')
plt.xlabel('x')
plt.ylabel('pdf(x)')
plt.title('pdf of normal distribution')
plt.grid()
plt.show()
```



In [3]:

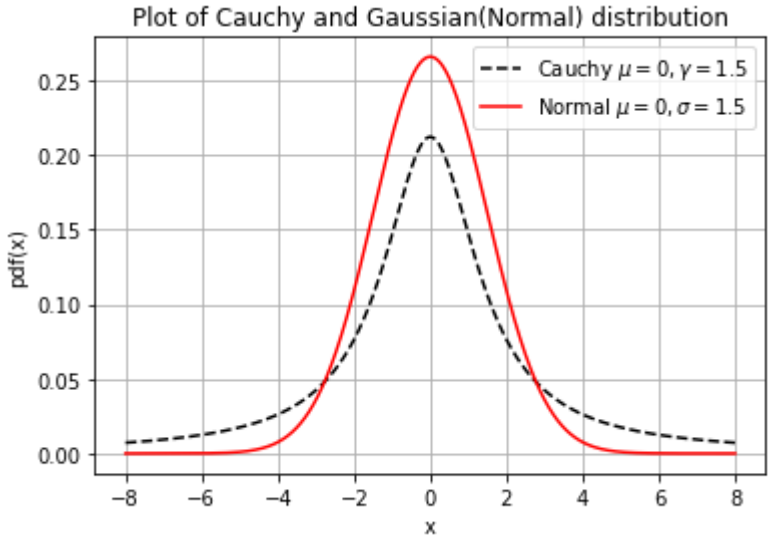
```
sample = gaussian.rvs(1000)

# finding sample mean
sample_mean = np.sum(sample)/np.size(sample)
# finding sample variance
sample_var = np.sum(np.square(sample - sample_mean))/(np.size(sample) - 1)

# finding sample skew
sample_skew = stats.skew(sample)
# finding sample kurtosis
sample_kurtosis = stats.kurtosis(sample)

# printing all the values
print("Mean =", sample_mean)
print("Variance =", sample_var)
print("skewness =", sample_skew)
print("Kurtosis =", sample_kurtosis)

# standard deviation using MAD
MAD = np.median(np.abs(sample - np.median(sample)))
std_dev = 1.482 * MAD

# printing MAD and standard deviation
print("\nMAD", MAD)
print("standard deviation using MAD =", std_dev)

# using sigma_G
q25 = np.percentile(sample, 25)
q75 = np.percentile(sample, 75)
sigma_g = 0.7413 * (q75 - q25)
# print sigma_G
print("sigma_G =", sigma_g)
```

```
Mean = 1.486743356467604
Variance = 0.2370822381951789
skewness = 0.020172971873505058
Kurtosis = -0.09673897988446845

MAD 0.319511697703584
standard deviation using MAD = 0.4735163359967115
sigma_G = 0.4730858366597332
```

Question 2

In [4]:

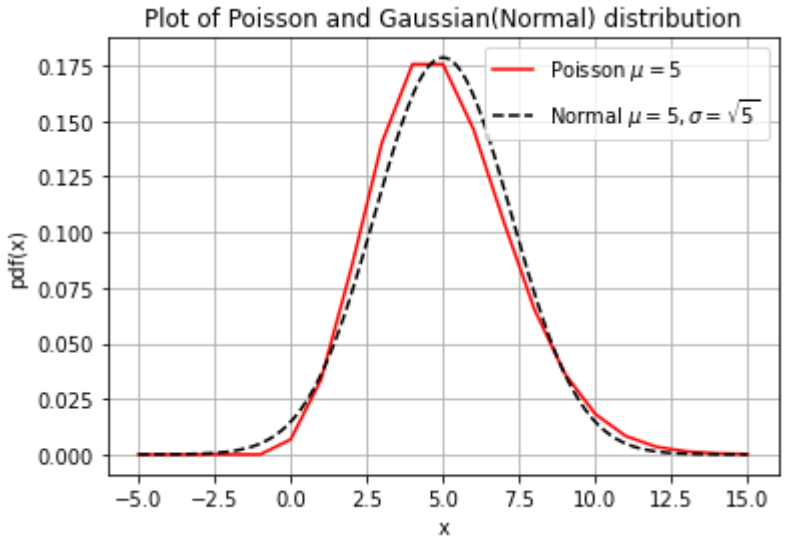
```
x = np.linspace(-8, 8, 2000)
# creating cauchy and gaussian distributions
prob1 = stats.cauchy(0, 1.5)
prob2 = stats.norm(0, 1.5)

# plotting both distributions
plt.plot(x, prob1.pdf(x), label='Cauchy $\mu=0,\gamma=1.5$', ls='--', color='black')
plt.plot(x, prob2.pdf(x), label='Normal $\mu=0,\sigma=1.5$', color='red')
plt.xlabel('x')
plt.ylabel('pdf(x)')
plt.legend(loc='best')
plt.title("Plot of Cauchy and Gaussian(Normal) distribution")
plt.grid()
plt.show()
```



Question 3

In [5]:

```
# creating 2 different x-coordinates for poisson and normal distribution
x1 = np.arange(-5, 16)
x2 = np.linspace(-5, 15, 200)
# creating distribution curves
prob1 = stats.poisson(5)
prob2 = stats.norm(5, np.sqrt(5))

# plot
plt.plot(x1, prob1.pmf(x1), label='Poisson $\mu=5$', color='red')
plt.plot(x2, prob2.pdf(x2), label='Normal $\mu=5,\sigma= \sqrt{5}$', ls='--', color='bl
ack')
plt.xlabel('x')
plt.ylabel('pdf(x)')
plt.legend(loc='best')
plt.title("Plot of Poisson and Gaussian(Normal) distribution")
plt.grid()
plt.show()
```



Question 4

In [6]:

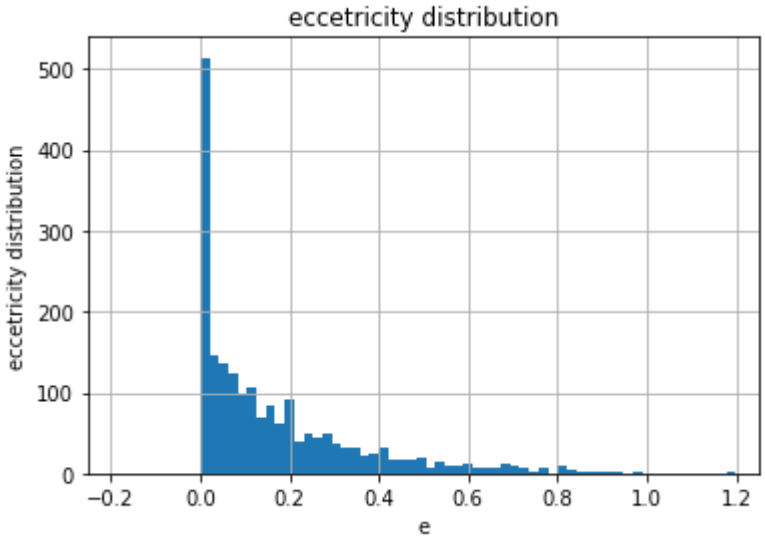
```
# declaring input variables
data = np.array([0.8920, 0.881, 0.8913, 0.9837, 0.8958])
change = np.array([0.00044, 0.009, 0.00032, 0.00048, 0.00045])

# computing required values
weighted_mean = np.sum(data/(change**2))/np.sum(1/(change**2))
uncertainity_mean = 1/np.sum(1/(change**2))

# printing required values
print("Weighted mean lifetime=", weighted_mean)
print("Uncertainity of the mean =", uncertainity_mean)
```

```
Weighted mean lifetime= 0.9089185199574897
Uncertainity of the mean = 4.128510743662294e-08
```

Question 5

In [7]:

```
# reading csv file
dataframe = pd.read_csv(r'exoplanet.eu_catalog.csv')
data = pd.DataFrame(dataframe, columns=['eccentricity'])
data = data.dropna(axis=0)
data = data.to_numpy()

# plot
plt.hist(data, bins=150)
plt.xlim(-0.25, 1.25)
plt.title('eccetricity distribution')
plt.ylabel('eccetricity distribution')
plt.xlabel('e')
plt.grid()
plt.show()
```



In [8]:

```
# removing rows with value 0 since boxcox function accepts non-zero values
data = data[data != 0.0]
# Gaussianizing eccentricity distribution
e1,_ = stats.boxcox(data)

plt.figure()
plt.hist(e1, bins=150)
plt.xlabel('e')
plt.ylabel('eccetricity distribution')
plt.title('Gaussianized eccetricity distribution')
plt.grid()
plt.show()
```