

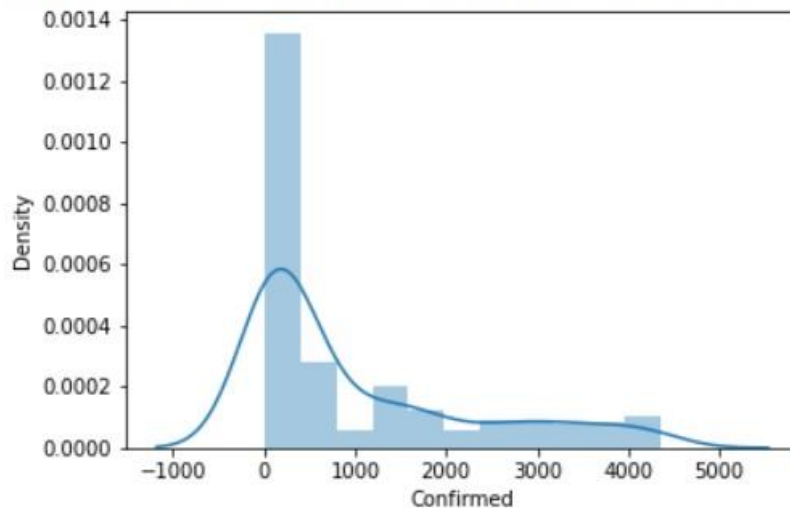
Computer Assignment 2

Odisha

(a) For each of the two variables, draw a histogram and overlay the normal density to observe the departure if any from normality to the two variables in earlier assignment.

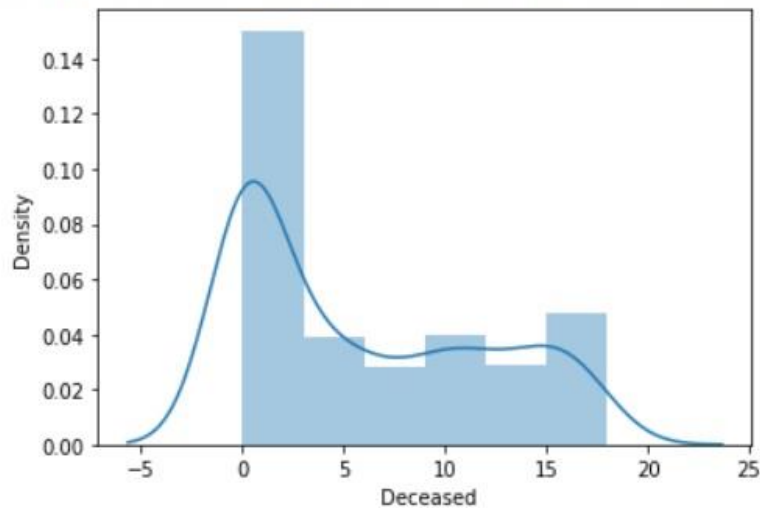
```
▶ sns.distplot(data['Confirmed'])
```

```
↳ /usr/local/lib/python3.7/dist-packages/seaborn/distribution
warnings.warn(msg, FutureWarning)
<matplotlib.axes._subplots.AxesSubplot at 0x7f9adb4022d0>
```



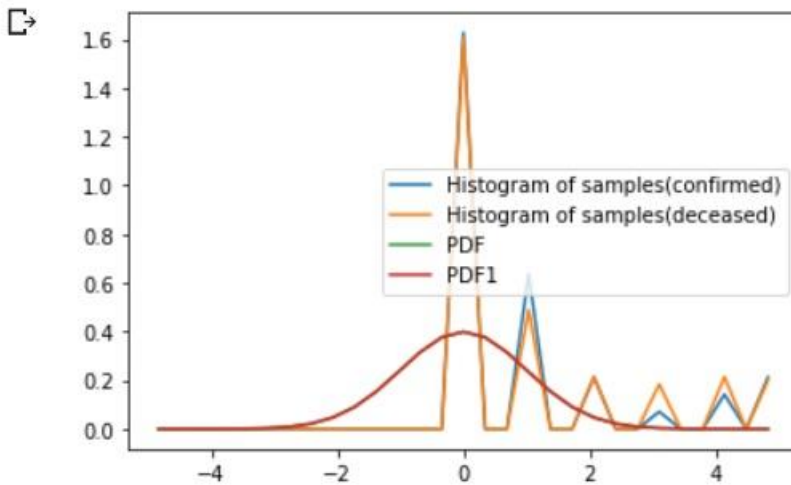
```
▶ sns.distplot(data['Deceased'])
```

```
↳ /usr/local/lib/python3.7/dist-packages/seaborn/distribution
warnings.warn(msg, FutureWarning)
<matplotlib.axes._subplots.AxesSubplot at 0x7f9ada191290>
```



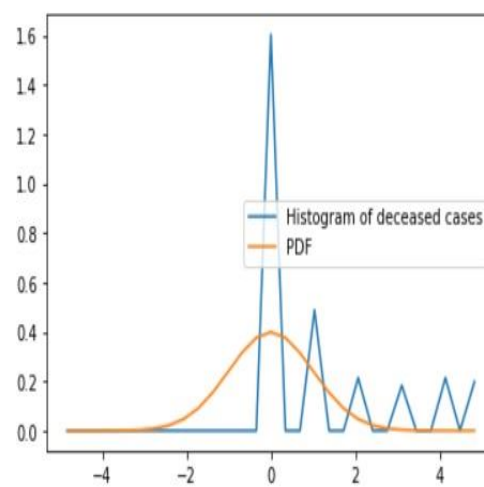
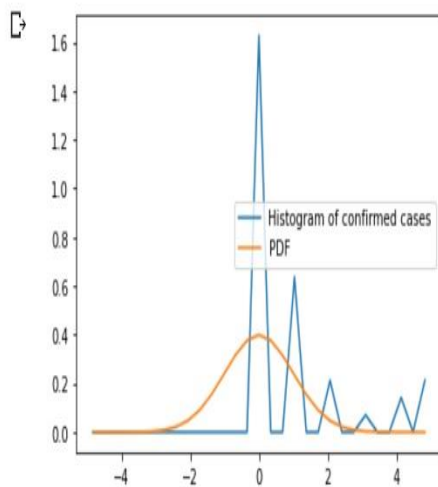
Looking at the above graphs we can see the curve is not bell-shaped, hence these variables do not follow normal distribution.

(b) Draw the bivariate normality pdf and also assess the shape of the contour for normality and comment on the correlation between the two variables.



This plot clearly gives the difference between the normal curve and original distribution. It is clear that these two variables do not follow normal distribution.

To see separately, following graphs can be analyzed:



Again, normality is not followed. These plots indicate a positive correlation between confirmed and Deceased cases.

(c) Use approach (2) above to check if the two variables are normally distributed.

As per approach 2 we need to check:

1. If 68.3% of sample lies within: (sample mean-stddev, sample mean+stddev)
2. If 95.4% of sample lies within: (sample mean - 2*stddev, sample mean+ 2*stddev)

When checked using python, it was found that both failed the first test, i.e, whether 68.3% of data lies within the above-mentioned range. Hence, they do not follow normality.

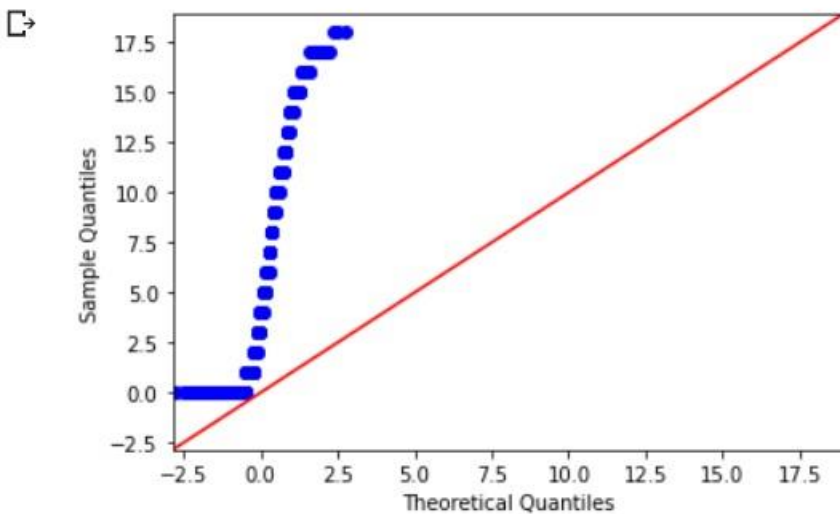
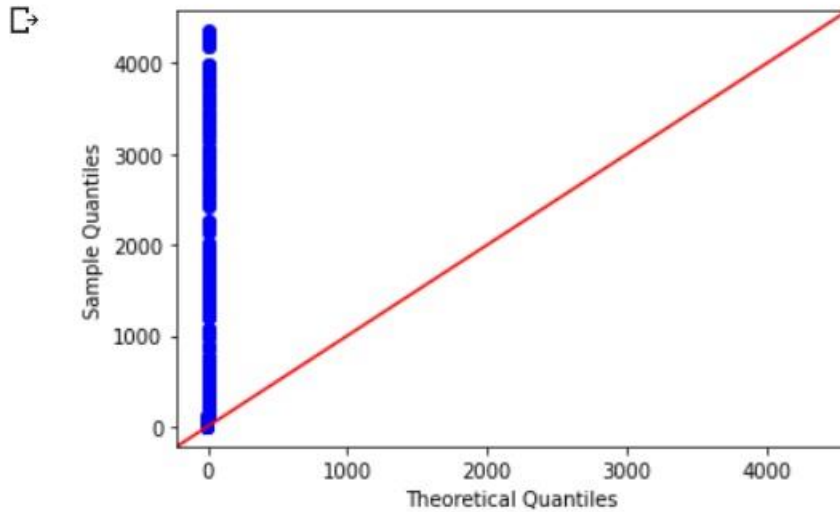
```
▶ count=0
for i in data['Confirmed']:
    if i in range(int(confirmed_avg-confirmed_std),int(confirmed_avg+confirmed_std)):
        count=count+1
print("Confirmed cases: ")
normalitytest1(count)
```

```
☞ Confirmed cases:
Normality test1 failed
```

```
[26] count=0
for i in data[' Deceased']:
    if i in range(int(deceased_avg-deceased_std),int(deceased_avg+deceased_std)):
        count+=1
print("Deceased Cases: ")
normalitytest1(count)
```

```
Deceased Cases:
Normality test1 failed
```

(d) Estimate a Q-Q plot for each of the variables and verify further the departure if any from normality.



For both the plots, the sample lies above and very far away from 45-degree line, hence these variables are not following normality.

Tools used: Python

Python Libraries Used: pandas, numpy, seaborn, statsmodel, pylab