

Assignment - 2 - Set 2 - (Basic Statistic Level - 2)

Qus.no.2

```
In [10]: 1 from scipy import stats
          2 from scipy.stats import norm
```

```
In [5]: 1 # A. More employees at the center are older than 44 than between 38 and 44
```

```
In [3]: 1 # P(X>44); Employees older than 44 yrs of age.
          2 1-stats.norm.cdf(44,loc = 38,scale = 6)
```

```
Out[3]: 0.15865525393145707
```

```
In [4]: 1 # P(38<X<44); Employees between 38 to 44 yrs of age
          2 stats.norm.cdf(44,38,6)-stats.norm.cdf(38,38,6)
```

```
Out[4]: 0.3413447460685429
```

```
In [6]: 1 # B. A training program for employees under the age of 30 at the center would be expected to attract about 36 employees.
```

```
In [8]: 1 # P(X<30); Employees under 30 yrs of age .
          2 stats.norm.cdf(30,38,6)
```

```
Out[8]: 0.09121121972586788
```

```
In [9]: 1 # No.of employees attending training program from 400 nos.is N*P(X<30)
          2 400*stats.norm.cdf(30,38,6)
```

```
Out[9]: 36.484487890347154
```

Qus.no.4

```
In [1]: 1 import pandas as pd
          2 import numpy as np
          3 import matplotlib.pyplot as plt
          4 import scipy.stats as stats
          5 import seaborn as sns
          6 import statsmodels.api as smf
          7 import warnings
          8 warnings.filterwarnings('ignore')
```

```

In [2]: 1 # Given
2 mew = 100
3 std = 20
4 # P(a<x<b)
5 # To Find =
6 """ two values, a and b, symmetric about the mean,such that the probabilityof the random
7 variable taking a value between them is 0.99."""
8 # Solution
9 """ From the above details, we have to exclude .005% area from each left and right tails.
10 Hence,we want to find the .005th and the .995th percentiles Z score value """
11
12 # Z value for .005 percentiles
13 z_005_ = np.round(stats.norm.ppf(0.005),4)
14 z_005_
15
16 # Z value for .99 percentiles
17 z_99_ = np.round(stats.norm.ppf(0.995),4)
18 z_99_
19
20 #z = (x_bar-new)/std
21 #x_bar = (z*std)+mew
22 a = np.round((z_005_*std) + mew,1)
23 b = np.round((z_99_*std) + mew,1)
24 print("""two values, a and b, symmetric about the mean,such that the probabilityof the random
25 variable taking a value between them is 0.99.""",a,b)

```

two values, a and b, symmetric about the mean,such that the probabilityof the random variable taking a value between them is 0.99. 48.5 151.5

Qus.no.5

```

In [3]: 1 import numpy as np
2 from scipy import stats
3 from scipy.stats import norm

```

```

In [6]: 1 # Mean profits from two different divisions of a company = Mean 1 + Mean 2
2 Mean=5+7
3 print('Mean profits is Rs',Mean*45,'Million')

```

Mean profits is Rs 540 Million

```

In [7]: 1 # Variance of profits from two different divisions of a company = SD^2 = SD1^2 + SD2^2
2 SD = np.sqrt((9)+(16))
3 print('Standard Deviation is Rs',SD*45,'Million')

```

Standard Deviation is Rs 225.0 Million

```

In [8]: 1 # A. Specify a Rupee range (centered on the mean) such that it contains 95% probability
2 #for the annual profit of the company.
3 print('Range is Rs',(stats.norm.interval(0.95,540,225)), 'in Million')

```

Range is Rs (99.00810347848784, 980.9918965215122) in Million

```

In [9]: 1 # B. Specify the 5th percentile of profit (in Rupees) for the company to compute 5th percentile,
2 # we use the formula  $X = \mu + Z\sigma$ ; where in from z table ,5 percentile = -1.645
3 x= 540+(-1.645)*(225)
4 print('5th percentile of profit (in Million Rupees)is',np.round(x,1))

```

5th percentile of profit (in Million Rupees)is 169.9

```

In [10]: 1 # C.Which of the two divisions has a larger probability of making a loss in a given year?

```

```

In [11]: 1 # Probability of divisions 1 making a Loss P(X<0)
2 stats.norm.cdf(0,5,3)

```

Out[11]: 0.0477903522728147

```

In [12]: 1 #Probability of divisions 2 making a Loss P(X<0)
2 stats.norm.cdf(0,7,4)

```

Out[12]: 0.040059156863817086

In []:

1