**Topics: Normal distribution, Functions of Random Variables**

1. The time required for servicing transmissions is normally distributed with *μ* = 45 minutes and *σ* = 8 minutes. The service manager plans to have work begin on the transmission of a customer’s car 10 minutes after the car is dropped off and the customer is told that the car will be ready within 1 hour from drop-off. What is the probability that the service manager cannot meet his commitment?
2. 0.3875
3. 0.2676
4. 0.5
5. 0.6987

**Ans**🡪 We have a normal distribution with = 45 and = 8.0.

Let X be the amount of time it takes to complete the repair on a customer's car. To finish in one hour, you must have X <= 50 so the question is to find P(X > 50).

P(X > 50) = 1 - P(X <= 50).

Z = (X - µ)/ = (50 - 45)/8.0 = **0.625**

stats.norm.cdf(z) = **0.734**

Probability that the service manager will not meet his demand will be = 100 - 73.4 = **26.6% or 0.2676**.

**ANS** - **B. 0.2676**

1. The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean *μ* = 38 and Standard deviation *σ* =6. For each statement below, please specify True/False. If false, briefly explain why.
2. More employees at the processing center are older than 44 than between 38 and 44.

**Ans**🡪

Mean(*μ)* = 38(loc) and Standard deviation (*σ)* =6(scale)

**-** Probability of employees greater than age of 44= P(X>44)

P(X > 44) = 1 - P(X <= 44).

Z = (X -µ)/ = (44 - 38)/6

Z = **1.0**

Thus, the question can be answered by using the normal table to find

stats.norm.cdf(z) = **84.1345%**

Probability that the employee will be greater than age of 44 = 100-84.1345=15.86%

So, the probability of number of employees between 38-44 years of age

= P(X<44)-0.5=84.1345 – 50.00

= **34.1345%**

Therefore, the statement that more employees at the processing center are older than 44 than between 38 and 44 is **TRUE**.

1. A training program for employees under the age of 30 at the center would be expected to attract about 36 employees.

**Ans**🡪

Mean(*μ)* = 38(loc) and Standard deviation (*σ)* =6(scale)

Probability of employees less than age of 30 = P(X<30).

Z = (X -µ)**/** = (30 - 38)/6

Z **= -1.3333**

Thus, the question can be answered by using the normal table to find

stats.norm.cdf(z) = **9.12%**

So, the number of employees with probability 0.912 of them being under age 30

= 0.0912\*400

**= 36.48** (or **36** employees).

Therefore, the statement B of the question is also **TRUE**.

**Code**🡪

import numpy as np

import pandas as pd

from scipy import stats

from scipy.stats import norm

# p(X>44); Employees older than 44 yrs of age

1-stats.norm.cdf(44,loc=38,scale=6)

# p(38<X<44); Employees between 38 to 44 yrs of age

stats.norm.cdf(44,38,6)-stats.norm.cdf(38,38,6)

# P(X<30); Employees under 30 yrs of age

stats.norm.cdf(30,38,6)

# No. of employees attending training program from 400 nos. is N\*P(X<30)

400\*stats.norm.cdf(30,38,6)

1. If *X1* ~ *N*(μ, σ2) and *X*2 ~ *N*(μ, σ2) are *iid* normal random variables, then what is the difference between 2 *X*1 and *X*1 + *X*2? Discuss both their distributions and parameters.

**ANS**🡪

As we know that if X ~ N (µ1, σ 1^2) and Y ~ N (µ2, σ 2^2) are two independent random variables then X + Y ~ N (µ1 + µ2, σ 1^2 + σ 2^2) and X - Y ~ N (µ1 - µ2, σ 1^2 + σ 2^2 ) .

Similarly if Z = aX + bY , where X and Y are as defined above, i.e. Z is linear combination of X and Y , then Z ~ N(aµ1 + bµ2, a^2 σ 1^2 + b^2 σ 2^2 ).

Thus, following the property of multiplication, we get

2X1~ N(2 µ,2^2\*σ ^2)

2X1~ N(2 µ,,4\*σ ^2)

and following the property of addition,

X1+X2 ~ N(µ + µ, σ ^2 + σ ^2 )

X1+X2 ~ N(2u, 2\*σ ^2 )

And the difference between the two is given by

**2X1-(X1+X2) = N(2 µ -2 µ,2\*σ ^2+,,4\*σ ^2)=N(0,6\*σ ^2)**

The mean of 2X1 and X1+X2 is same but the var(**σ**) of is 2X1 is 2 times more than the variance of X1+X2.

**The difference between the two says that the two given variables are identically and independently distributed**.

1. Let X ~ N(100, 202). Find two values, *a* and *b*, symmetric about the mean, such that the probability of the random variable taking a value between them is 0.99.
2. 90.5, 105.9
3. 80.2, 119.8
4. 22, 78
5. 48.5, 151.5
6. 90.1, 109.9

ANS🡪

Mean(loc)=100 , SD(scale)=1, C=99%=0.99.

**For finding 2 values** , **interval (confidence, loc=0, scale=1)**

**Output= (48.48341392902199, 151.516586070978)**

**So, option D is correct.**

**Code🡪**

import numpy as np

import pandas as pd

from scipy import stats

from scipy.stats import norm

stats.norm.interval(0.99,100,20)

output=(48.48341392902199, 151.516586070978)

**or**

**By stats 🡪**

Since we need to find out the values of a and b, which are symmetric about the mean, such that the probability of random variable taking a value between them is 0.99, we have to work out in reverse order.

The Probability of getting value between a and b should be 0.99.

So, the Probability of going wrong, or the Probability outside the a and b area is 0.01(i.e. 1-0.99).

The Probability towards left from a = -0.005 (i.e. 0.01/2).

The Probability towards right from b = +0.005 (i.e. 0.01/2).

So since we have the probabilities of a and b, we need to calculate X, the random variable at a and b which has got these probabilities.

By finding the Standard Normal Variable Z (Z Value), we can calculate the X values.

Z=(X- ¼ ) / σ

For Probability 0.005 the Z Value is -2.57 (from Z Table).

Z \* σ + ¼ = X

Z(-0.005)\*20+100 = -(-2.57)\*20+100 = **151.4**

Z(+0.005)\*20+100 = (-2.57)\*20+100 = **48.6**

**So, option D is correct.**

1. Consider a company that has two different divisions. The annual profits from the two divisions are independent and have distributions Profit1 ~ N(5, 32) and Profit2 ~ N(7, 42) respectively. Both the profits are in $ Million. Answer the following questions about the total profit of the company in Rupees. Assume that $1 = Rs. 45
2. Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.

ANS🡪

**Mean Profit = Rs 540 Million**

**Standard Deviation = Rs 225.0 Million**

**Range = Rs (99.00810347848784, 980.9918965215122) in Millions**

Code🡪

import numpy as np

from scipy import stats

from scipy.stats import norm

# Mean profits = Mean1 + Mean2

Mean = 5+7

print('Mean Profit = Rs', Mean\*45,'Million') #because 1$=45rs

# Variance = SD1^2 + SD2^2

SD = np.sqrt((9)+(16))

print('Standard Deviation = Rs', SD\*45, 'Million')

# Rupee range for 95% probability for the annual profit of the company.

print('Range = Rs',(stats.norm.interval(0.95,540,225)),'in Millions')

1. Specify the 5th percentile of profit (in Rupees) for the company.

ANS🡪

**5th percentile of profit = 170.0 Million**

Code🡪

import numpy as np

from scipy import stats

from scipy.stats import norm

#for 5th Percentile, we use the formula X=μ + Zσ; wherein from z table, 5 percentile = -1.645

X= 540+(-1.645)\*(225)

print('5th percentile of profit = ',np.round(X,), 'Million')

1. Which of the two divisions has a larger probability of making a loss in a given year?

ANS🡪

**P(Division 1 ) = 0.0477903522728147**

**P(Division 2) = 0.040059156863817086**

Code🡪

import numpy as np

from scipy import stats

from scipy.stats import norm

# Probability of Division 1 making loss P(X<0)

stats.norm.cdf(0,5,3)

# Probability of Division 2 making loss P(X<0)

stats.norm.cdf(0,7,4)