**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:

The work begin after 10 min, so the average time increase from 45min to 55min.

for normal distribution :-

z = (X-μ)/б

= (60-55)/8

= 0.625

In R software for probability finding we use function called pnorm

As we want to find the probability of service manager cannot meet his commitment, So we should write below command.

1-pnorm(0.625)

=0.2659

So answer is option D

Python code:

**from** scipy **import** stats

**from** scipy.stats **import** norm

*# Find Z-Scores at X=50; Z = (X - µ) / σ*

Z**=**(50**-**45)**/**8

Z

0.625

*# Find probability P(X>50) = 1-stats.norm.cdf(abs(z\_score))*

1**-**stats**.**norm**.**cdf(abs(0.625))

0.26598552904870054

*# OR Find probability P(X<=50); p\_value=stats.norm.cdf(abs(z\_score))*

p\_value**=**stats**.**norm**.**cdf(abs(0.625))

p\_value

0.7340144709512995

*# P(X>50) = 1 - P(X<=50)*

1**-**0.734

0.266

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

Ans:

Mean = 38

SD = 6

Z score = (Value - Mean)/SD

Z score for 44  = (44 - 38)/6  = 1  =>  84.13 %

=> People above 44 age = 100 - 84.13 =  15.87%  ≈  63    out of 400

Z score for 38  = (38 - 38)/6 = 0 => 50%

Hence People between 38 & 44  age = 84.13 - 50 = 34.13 % ≈  137 out of 400

Hence More employees at the processing center are older than 44 than between 38 and 44. is FALSE

Z score for 30  = (30 - 38)/6 =  -1.33  =  9.15  %   ≈ 36 out of 400

Hence A training program for employees under the age of 30 at the center would be expected to attract about 36 employees – TRUE

Python code:

**from** scipy **import** stats

**from** scipy.stats **import** norm

*# A. More employees at the processing center are older than 44 than between 38 and 44.*

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

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

0.15865525393145707

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

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

0.3413447460685429

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

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

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

0.09121121972586788

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

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

36.484487890347154

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:

The Normal distribution has two parameters, the mean, µ, and the variance .

µ and satisfy −∞ < µ < ∞, > 0.

We write X ∼ Normal(µ, ), or X ∼ N(µ, )

**Probability density function, fx(x):**

|  |
| --- |
| fx(x) = for - |

**Distribution function, FX(x):**

There is no closed form for the distribution function of the Normal distribution.

If X ~ Normal(, then FX(x) can only be calculated by computer.

R command : FX(x) = pnorm (x, mean = , sd=sqrt()).

**Mean and Variance:**

For X ~ Normal() ,

|  |
| --- |
| E(X) = |

**Linear transformations:**

If X ∼ Normal(µ, ), then for any constants a and b,

|  |
| --- |
| aX + b Normal (a + b, ) |

In particular, put a = and b = -, then

|  |
| --- |
| X Normal() => () |

Z Normal(0,1) is called the standard normal random variable.

**Sums of Normal random variables:**

If X and Y are independent, and X Normal(, ), Y Normal ( , ),

Then

X + Y Normal ( + , + )

More generally ,if ,,,…………, are independent, and Normal( , ) for i=1,2,3,…..,n then

+ +…………+ Normal ( ( +…………+), ( +…………+ ) )

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:

p(a<x<b) = 0.99 ,m ean =100,standardDeviation = 20

Identify symmetric values for the standard normal distribution such that the area enclosed is .99

From the above details,we have to excluded area of .005 in each of the left and right tails. Hence, we want to find the 0.5th and the 99.5th percentiles Z score values

Using Python

Z value is given as stats.norm.ppf(pvalue)

Z value at 0.5th percentile is given as

                          Z(0.5) = stats.norm.ppf(0.005)= -2.576

Z value at 99.5 percentile is given as

                         Z(99.5) = stats.norm.ppf(0.995) = 2.576

Z = (x - 100)/20 = > x = 20z+100

      a = -(20\*2.576) + 100= 48.5

      b = (20\*2.576)+100= 151.5

Two values symmetric about mean for the given standard normal distribution are[48.5,151.5]

Python code:

from scipy import stats

from scipy.stats import norm

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

(48.48341392902199, 151.516586070978)

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.
3. Specify the 5th percentile of profit (in Rupees) for the company
4. Which of the two divisions has a larger probability of making a loss in a given year?

Ans:

Python code:

import numpy as np

from scipy import stats

from scipy.stats import norm

*# Mean profits from two different divisions of a company = Mean1 + Mean2*

Mean **=** 5**+**7

print('Mean Profit is Rs', Mean**\***45,'Million')

Mean Profit is Rs 540 Million

*# Variance of profits from two different divisions of a company = SD^2 = SD1^2 + SD2^2*

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

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

Standard Deviation is Rs 225.0 Million

*# A. Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.*

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

Range is Rs (99.00810347848784, 980.9918965215122) in Millions

*# B. Specify the 5th percentile of profit (in Rupees) for the company*

*# To compute 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 (in Million Rupees) is',np**.**round(X,))

5th percentile of profit (in Million Rupees) is 170.0

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

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

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

0.0477903522728147

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

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

0.040059156863817086