**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:

Since work begins 10 mins after the car dropped, the time left to complete work is 50 min.

Probability of the Service manager cannot meet his the commitment = P( X>50) =1 – P(X<50)

First of find out the z score value

z = (50-45)/8 = 0.625

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

p\_value = 1- stats.norm.cdf(z) = 0.26598552904870054

That is option B

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:

Find probability P(X>44)

p\_value = 1 -stats.norm.cdf(44,38,6)

p\_value = 0.15865525393145707

Probability of employees at the processing center is older than 44 = 15.86%

#Find probability P(38<X<44)

p1 = stats.norm.cdf(44,38,6)

p2 = stats.norm.cdf(38,38,6)

p3 = p1 – p2

p3 = 0.3413447460685429

Probability of employees at the processing center is between 38 and 44 = 34.13%

The statement is false. Because Probability of employees at the processing center is older than 44 is only 15.86% but in between 38 and 44 is more than 15.86 that is 34.13.

1. The statement is true.

#Find probability P(X<30)

p\_value = stats.norm.cdf(30,38,6)

p\_value = 0.09121121972586788

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

400\*p\_value = 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:

According to the **Central Limit Theorem**, any **large sum** of **independent**, **identically distributed(iid)** random variables is approximately **Normal**. The **Normal distribution** is defined by two parameters, the **mean** μ, and the **variance** σ2, and written as

X ∼N(μ, σ2).

As we know that if X ∼N(μ1, σ1^2 ), and Y ∼N(μ2, σ2^2 ) are two independent random variables, then

* X + Y ∼N(μ1 + μ2, σ1^2 + σ2^2 ) ,
* X − Y ∼N(μ1 − μ2, σ1^2 + σ2^2 ), and
* Z = aX = Z ∼N(aμ1,a^2\* σ1^2)

Given to find, 2 *X*1,Thus, following the property of multiplication, we get

2 *X*1 ~ N(2 μ,22 σ2) =>2 *X*1 ~ N(2 μ,4 σ2)

and following the property of addition,

*X*1 + *X*2 ~( μ+ μ, σ2 + σ2 )=> *X*1 + *X*2 ~(2 μ,2σ2)

And the difference between the two is given by

2 *X*1 -( *X*1 + *X*2) ~ N(2 μ - 2 μ , 2σ2+4 σ2)~ N(0,6 σ2)

The Normal distribution has two parameters, the mean, µ, and the variance, σ2. µ and σ2 satisfy −∞ < µ < ∞, σ2> 0. We write X ∼ Normal (µ, σ2) or X ∼ N(µ, σ2 ).

The mean of 2 *X*1  and *X*1 + *X*2 is same but the var(σ2) of  2 *X*1  is 2 times more than the variance of *X*1 + *X*2 . 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: D

Mean =100

Standard deviation = 20

Confidence level = 99%

#Find Confidence interval of mean 100, std 20 and confidence level 99%

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:

Mean= (5+7)\*45 = 540

Standard deviation = sqrt(9+16) = 5\*45 = 225

stats.norm.interval(0.95,540,225)

(99.00810347848784, 980.9918965215122)

Rupee ranges in between [9.9 to 98.1] Crore Rupees.

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

# To compute 5th Percentile, we use the formula X=μ + Zσ;

Z\_score = stats.norm.ppf(0.05)

X= 540+(Z\_score)\*(225)

X

169.9079339359186

5th percentile of profit (in Rupees) for the company is approximately 17 crore rupees

# 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

Division 1 has a larger probability of making a loss in a given year