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

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

**Que.1.The time required for servicing transmissions is normally distributed with  = 45 minutes and  = 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?**

**A. 0.3875**

**B. 0.2676**

**C. 0.5**

**D. 0.6987**

**Answer-** 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 d" 50 so the question is to find Pr(X > 50).

Pr(X > 50) = 1 - Pr(X d" 50).

Z = (X - )/ = (X - 45)/8.0 ##Thus the question can be answered by using the normal table to find

Pr(X d" 50) = Pr(Z d" (50 - 45)/8.0) = Pr(Z d" 0.625)=73.4%

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

**Que.2. The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean meu = 38 and Standard deviation  =6. For each statement below, please specify True/False. If false, briefly explain why.**

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

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

Answer- The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean = 38 and Standard deviation = 6.

For each statement below, please specify True/False. If false, briefly explain why.

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

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

**Answer** - We have a normal distribution with = 38 and = 6. Let X be the number of employees. So according to question

a)Probabilty of employees greater than age of 44= Pr(X>44)

Pr(X > 44) = 1 - Pr(X d" 44).

Z = (X - )/ = (X - 38)/6

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

Pr(X d" 44) = Pr(Z d" (44 - 38)/6) = Pr(Z d" 1)=84.1345%

Probabilty 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 = Pr(X<44)-0.5=84.1345-0.5= 34.1345%

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

b) Probabilty of employees less than age of 30 = Pr(X<30).

Z = (X - )/ = (30 - 38)/6

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

Pr(X d" 30) = Pr(Z d" (30 - 38)/6) = Pr(Z d" -1.333)=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

**3. If X1 ~ N(μ, σ2) and X2 ~ N(μ, σ2) are iid normal random variables,**

**then what is the difference between 2 X1 and X1 + X2? Discuss both their distributions and parameters.**

**Answer--**

As we know that if X <" N(µ1, Ã 1^2 ), and Y <" N(µ2, Ã 2^2 ) are two independent random variables then X + Y <" N(µ1 + µ2, Ã 1^2 + Ã 2^2 ) , and X " Y <" N(µ1 " µ2, Ã 1^2 + Ã 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^2Ã 1^2 + b^2Ã 2^2 ).

Therefore in the question

2X1~ N(2 u,4 Ã ^2) and

X1+X2 ~ N(µ + µ, Ã ^2 + Ã ^2 ) ~ N(2 u, 2Ã ^2 )

2X1-(X1+X2) = N( 4µ,6 Ã ^2)

**4. 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.**

**A. 90.5, 105.9**

**B. 80.2, 119.8**

**C. 22, 78**

**D. 48.5, 151.5**

**E. 90.1, 109.9**

Answer--

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 (ie. 1-0.99).

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

The Probability towards right from b = +0.005 (ie. 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 ) / Ã

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

Z \* Ã + 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.

**5. 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**

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

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

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

import numpy as np

from scipy import stats

from scipy.stats import norm

Mean = 5+7

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

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

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

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

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

print('5th percentile of profit (in Million Rupees) is',np.round(X,))

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

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

print('5th percentile of profit (in Million Rupees) is',np.round(X,))

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

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