**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

**Solution: B. 0.2676, which is close to the value that we get.**

# Let's suppose we need to complete in 50 mins as the work starts after 10 minutes and expectation is it should be completed in 1 hour i.e., 60 mins.

# In that case Mean or MU = 45 and Sigma or standard deviation is 8 mins.

1- stats.norm.cdf (50, loc=45, scale=8) # 0.26598552904870054

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.

**Solution: False**

# Probability greater i.e., "Older than 44"

1-stats.norm.cdf(44, loc=38, scale=6) # 0.15865525393145707

# Probability "Between 38 and 44"

stats.norm.cdf(44, loc=38, scale=6) - stats.norm.cdf(38, loc=38, scale=6) # 0.3413447460685429

So, Statement is **False** as there are more employees "Between 38 and 44"

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

**Solution: True**

"Between 38 and 44"

# Probability of lesser than 30

stats.norm.cdf(30, loc=38, scale=6) # 0.09121121972586788

0.09121121972586788 \* 400 # 36.484487890347154

**=** 36.48

So, answer is **True**

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.

**Solution: The difference between 2 *X*1 and *X*1 + *X*2 is *N*(0, 6σ2)**

**Explanation:**

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).

Given *X1* ~ *N*(μ, σ2) and *X*2 ~ *N*(μ, σ2) are two independent identically distributed random variables.

**Properties of normal random variables**

if *X* ~ *N*(μ1, σ12) and *Y* ~ *N*(μ2, σ22) are two independent identically distributed random variables then

* The sum of normal random variables is given by

*X+Y* ~ *N*(μ1 + μ2 , σ12 + σ22),

* The difference of normal random variables is given by

*X-Y* ~ *N*(μ1 – μ2 , σ12 + σ22),

* When Z = aX, the product of X is given by

Z~ *N*(aμ1, a2σ12)

* When Z = aX + bY, the linear combination of X and Y is given by

Z~ *N*(aμ1 – bμ2 , a2σ12 + b2σ22),

Using the property of Multiplication, provided earlier, we calculate *2X1* and get –

*2X1* ~ *N*(2μ, 22σ2) = *2X1* ~ *N*(2μ, 4σ2)

Using the property of Addition, provided earlier, we calculate *X1  + X2*  and get –

*X1  + X2*  ~ *N*(μ + μ, σ2 + σ2) ~ *N*(2μ, 2σ2)

Using the property of Difference, we calculate difference of *2X1* and *X1  + X2* and get –

*2X1 – (X1  + X2* ) ~ *N*(2μ - 2μ, 2σ*1*2 + 4σ*2*2) ~ *N*(0, 6σ2)

**The conclusion is that the mean of *2X1* and *(X1  + X2* ) is same but the var(σ2) of *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

Solution: **The correct solution is D. 48.5, 151.5**

stats.norm.interval (0.99, loc=100, scale=20)

(48.48341392902199, 151.516586070978)

**Checking:**

stats.norm.cdf (151.5, loc=100, scale=20) - stats.norm.cdf (48.5, loc=100, scale=20)

= 0.9899759913364774

and matches 0.99, the expected value

Alternate Method – (Using the reverse logic)

The expected probability of the random variables taken between a and b should be 99% or 0.99.

The probability of the condition not matching is 1 – 0.99 = 0.1

Since it is a normal distribution, this value between points a and b should be in the middle with the 2 tails combining to the value to 0.1 i.e., 0.05 and 0.05 on both sides.

For the first point ‘a’ the percentage is 0.05 and for the second point ‘b’, the percentage is 0.995, so we can use the ppf formula –

# Getting the point a based on the percentage 0.005

stats.norm.ppf(0.005,loc=100,scale=20) # 48.483413929021985

= 48.483413929021985

# Getting the point b based on the percentage 0.995 i.e., 0.1 – 0.005 = 0.995

stats.norm.ppf(0.995,loc=100,scale=20) # 151.516586070978

= 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

Mean\_Profit\_Sum\_D (in millions) = 5 + 7

# Converting to Rupees - $1 = Rs.45

Mean\_Profit\_Sum\_R (in millions) = Mean\_Profit\_Sum\_D \* 45 = 540

**The Sum of the Mean Profit is Rs. 540 million**

Variance\_Profit\_Sum\_D (in millions) = (3\*\*2) + (4\*\*2)

SD\_Profit\_D (in millions) = √((3\*\*2) + (4\*\*2)) = √25 = 5

# Converting to Rupees - $1 = Rs.45

SD\_Profit\_R (in millions) = 5 \* 45 = 225

**The Standard Deviation of Profit is Rs. 225.0 million**

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

**Solution**: **(99.00810347848784, 980.9918965215122)**

# Rupee Range with 95% probability in Millions

Range\_95\_R = stats.norm.interval(0.95,loc=540,scale=225)

**The Rupee range with 95% probability for annual profit of company in Millions is (99.00810347848784, 980.9918965215122)**

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

**Solution: The 5th percentile of profit for the company is 169.88 million.**

X = MU - Z-Score \* Sigma

Mean (MU): 540 million

Standard Deviation (Sigma): 225 million

Z-Score for 5%: -1.645

X (in Millions) = 540 + ((-1.645) \* 225) = 169.88

**The 5th percentile of profit for the company is 169.88 million.**

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

**Solution: Probability of Department 1 making a loss P(Dept1\_loss) is larger**

The Probability of Department 1 making a loss i.e., P(Dept1\_loss)

P\_Dept1\_loss = stats.norm.cdf (0,5,3)

**The probability of Department 1 making a loss is 0.04779**

The Probability of Department 2 making a loss i.e., P(Dept2\_loss)

P\_Dept2\_loss = stats.norm.cdf (0,7,4)

**The probability of Department 2 making a loss is 0.04006**

**Probability of Department 1 making a loss P(Dept1\_loss) is larger**