Topics: Normal distribution, Functions of Random Variables

1. The time required for servicing transmissions is normally distributed with $\mu = 45$ minutes and $\sigma = 8$ 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

➔ Ans : (B) 0.2676

```
In [1]: from scipy import stats

In [3]: 1-stats.norm.cdf(x=50,loc=45,scale=8)
Out[3]: 0.26598552904870054

In [ ]: # 26.59% The service manager cannot meet his commitment
```

2. 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.

➔ True

```
In [3]: from scipy import stats

In [ ]: # Mean=38 & SD=6

In [4]: stats.norm.cdf(44,38,6) # Probability less than 44 is 84.13%
Out[4]: 0.8413447460685429

In [5]: 1-stats.norm.cdf(44,38,6)  # Probability greater than 44 is 15.86%
Out[5]: 0.15865525393145707

In [9]: stats.norm.cdf(38,38,6)  # Probability less than 38 is 50%
Out[9]: 0.5

In [6]: stats.norm.cdf(44,38,6) - stats.norm.cdf(38,38,6) # Probability between 38 and 44 is 34.13%
Out[6]: 0.3413447460685429

In [ ]: # More employees at the processing center are older than 44 than between 38 and 44. Therefore given statement is true
```

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

- ➔ True

```
In [8]: stats.norm.cdf(30,38,6) # The probability at the age 30 is 09%
Out[8]: 0.09121121972586788

In [9]: # If the probablity at the age 30 is 9% means 9% of total employee is 36 because 400*.09= 36 . Therefore given statement is true

In [ ]:
```

3. If $X_1 \sim N(\mu, \sigma^2)$ and $X_2 \sim N(\mu, \sigma^2)$ 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.

- ➔ 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) , and X − Y ~ N(μ1 − μ2, σ1^2 + σ2^2).
- ➔ 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^2σ1^2 + b^2σ2^2).
- ➔ Therefore in the question

  2X1~ N(2 u,4 σ^2) and

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

  2X1-(X1+X2) = N(4μ,6 σ^2)

4. Let X ~ N(100, 20^2). 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

- ➔ D – 48.5, 151.5
- ➔ 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.

$Z=(X- \mu) / \sigma$

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

$Z * \sigma + \mu = X$

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

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

5. Consider a company that has two different divisions. The annual profits from the two divisions are independent and have distributions $Profit_1 \sim N(5, 3^2)$ and $Profit_2 \sim N(7, 4^2)$ 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?

In [1]:
```
import pandas as pd
import numpy as np
from scipy import stats
```

In [2]:
```
# Mean profit from 2 different divisions of the company = mean1 + mean2
Mean = 5+7
print('Mean profit in Rs',Mean*45,'Million')
```

```
Mean profit in Rs 540 Million
```

In [3]:
```
# Variance profit from 2 different divisions of the company = SD^2 = SD1^2 + SD2^2
SD = np.sqrt((9)+(16))
print('SD profit in Rs',SD*45,'Million')
```

```
SD profit in Rs 225.0 Million
```

In [6]:
```
# A. Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.
print('Range profit is',(stats.norm.interval(.95,540,225)),'Million')
```

```
Range profit is (99.00810347848784, 980.9918965215122) Million
```

In [10]:
```
# 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 Rs',np.round(X),'Millions')
```

```
5th percentile of profit in Rs 170.0 Millions
```

In [11]:
```
# C. Which of the two divisions has a larger probability of making a loss in a given year?
```

In [12]:
```
# Probability of Division 1 making a loss P(X<0)
stats.norm.cdf(0,5,3)
```

Out[12]:
```
0.0477903522728147
```

In [13]:
```
# Probability of Division 2 making a loss P(X<0)
stats.norm.cdf(0,7,4)
```

Out[13]:
```
0.040059156863817086
```

In [ ]:
```
# Division 1  with 4.77% making more loss compare to Division2 with 04%
```