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?

   A. 0.3875
   B. 0.2676
   C. 0.5
   D. 0.6987

Soln:

```
In [4]: #Soln:
"""the servicing work will began after 10 min of drop off so 45+10
which will now take more than the usual time so new mew is 55 minutes
and the porbability that it will take more than 1 hour to complete"""
mean = 55
std = 8
q1 = stats.norm.sf(60, loc = mean, scale = std)
print("""The probability that the service manager cannot meet his commitment is""",np.round(q1,5))

The probability that the service manager cannot meet his commitment is 0.26599
```

2. 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.
   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

Soln:

```
In [5]: #Soln:
mean = 38
std1 = 6
q2_lessthan_38 = stats.norm.cdf(38, loc = mean, scale = std1)


q2_less_than_44 = stats.norm.cdf(44, loc = mean,  scale = std1)


q2_betweeen_38_and_44 = (q2_less_than_44 - q2_lessthan_38)
print('The probability of employee age betweeen 38 and 44 is',np.round(q2_betweeen_38_and_44*100,2),'%')

q2_morethan_44 = 1-stats.norm.cdf(44, loc = mean, scale = std1)
print('The probability of employee age more than 44 is',np.round(q2_morethan_44*100,2),'%')

true_or_false = (q2_morethan_44 > q2_betweeen_38_and_44)
print('Answer:',true_or_false)

q2b = stats.norm.cdf(30, loc = mean, scale = std1)
print("""A training program for employees under the age of 30 at the center would be expected to attract about"""
      ,np.round((q2b*400),0),'employees')

The probability of employee age betweeen 38 and 44 is 34.13 %
The probability of employee age more than 44 is 15.87 %
Answer: False
A training program for employees under the age of 30 at the center would be expected to attract about 36.0 employees
```

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

Soln:

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

Soln:

```
In [8]: # Given
mean = 100
std = 20
# p(a<x<b)
#To 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"""
# Solution
""" From the above details,we have to exclude .005% area from each
left and right tails. Hence, we want to find the .005th and the
.995th percentiles Z score values"""

# Z value for .005 percentiles
z_005_ = np.round(stats.norm.ppf(0.005),4)
z_005_

# Z value for .99 percentiles
z_99_ = np.round(stats.norm.ppf(0.995),4)
z_99_

#z = (x_bar - mew) / std
#x_bar = (z*std) + mew
a = np.round((z_005_*std) + mean,1)
b = np.round((z_99_*std) + mean,1)
print("""The two values of a and b, symmetric about the mean, are such that the probability of the random variable taking a value
between them is 0.99:""",a,b)
```

```
The two values of a and b, symmetric about the mean, are such that the probability of the random variable taking a value
between them is 0.99: 48.5 151.5
```

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 5^{th} 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?

Soln:

Given that:

$1 = Rs. 45

$$Profit_1 \sim N(5, 3^2)$$
$$Profit_2 \sim N(7, 4^2)$$

Thus,

Company's profit:

$$P \sim N(5+7, 3^2+4^2) = N(12, 5^2)$$

A):

95% of the probability lies between 1.96 standard deviations of the mean.

Thus range is:

$$= (12 - 1.96 \times 5, 12 + 1.96 \times 5)$$

$$= (\$2.2M, \$22.8M)$$

$$= (\text{Rs. } 99\text{M}, \text{Rs. } 1026\text{M})$$

B): Fifth percentile is calculated as:

$$P(Z \leq \frac{p-12}{5}) = 0.05$$

From p values of z score table, we get:

$$\frac{p-12}{5} = -1.644$$
$$p = 12 - 8.22 = 3.78$$

Thus at $3.78M dollars, or Rs. 170.1M amount, 5th percentile of profit lies.

Or 5th percentile of profit is Rs. 170.1 Million.

C): Loss is when profit < 0

Thus: $p < 0$

The first division of company, thus have larger probability of making a loss in a given year