**BATCH:-DATA SCIENCE**

Abstract

**Normal Distribution+ function SET+2**

Assignment2

Basics stats

**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:-** The work begin after 10 min, so the average time increase from 45min to 55min.

for normal distribution :-

z = (X-μ)/б

= (60-55)/8

= 0.625

In R software for probability finding we use function called pnorm

As we want to find the probability of service manager cannot meet his commitment, So we should write below command.

1-pnorm(0.625)

**=0.2676**

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:-**

1. **More employees at the processing center are older than 44 than between 38 and 44:-**

***:-FALSE***

**Explanation:-** Mean = 38 ,SD = 6

Z score = (Value - Mean)/SD

Z score for 44  = (44 - 38)/6  = 1  =>  84.13 %

=> People above 44 age = 100 - 84.13 =  15.87%  ≈  63    out of 400

Z score for 38  = (38 - 38)/6 = 0 => 50%

Hence People between 38 & 44  age = 84.13 - 50 = 34.13 % ≈  137 out of 400

Hence More employees at the processing center are older than 44 than between 38 and 44. is F**ALSE**

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

**:-** Z score for 30  = (30 - 38)/6 =  -1.33  =  9.15  %   ≈ 36 out of 400

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

***ANS:-***

1. Var (k\*X1) = k ^2 \* Var (X)
2. 2) If X and Y are independent,  
   V(X + Y) = V(X) + V(Y)

Given that X1 ~ N(µ, σ2) and X2 ~ N(µ, σ2)  
Assume that X1 and X2 have same µ =1, σ =2

Var(2\*X1) = 2^2\*Var(1,2^2)  
= 4\*Var(1,4)  
= Var(4,16)  
  
Var(X1 + X2) = Var(X1) + Var (X2)  
= Var(1, 2^2) + Var(1, 2^2)  
= var (1,4) + var (1,4)  
= var (2,8)

***It means that Var (2\*X1) has 2 times more variance value than Var(X1+X2).***

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:-** Two values symmetric about mean for the given standard normal distribution are[48.5,151.5]

**Step-by-step explanation:**

Given:  p(a<x<b) = 0.99 ,m ean =100,standardDeviation = 20

To Find:

Identify symmetric values for the standard normal distribution such that the area enclosed is .99

From the above details,we have to excluded area of .005 in each of the left and right tails. Hence, we want to find the 0.5th and the 99.5th percentiles Z score values

Using Python

Z value is given as stats.norm.ppf(pvalue)

Z value at 0.5th percentile is given as

                                         Z(0.5) = stats.norm.ppf(0.005)= -2.576

Z value at 99.5 percentile is given as

                         Z(99.5) = stats.norm.ppf(0.995) = 2.576

Z = (x - 100)/20 = > x = 20z+100

      a = -(20\*2.576) + 100= 48.5

      b = (20\*2.576)+100= 151.5

***Two values symmetric about mean for the given standard normal distribution are[48.5,151.5]***

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:-**

**Explanation:**

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**