**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**

Answer : The given problem involves a scenario where the average time for a certain task is 45 minutes (represented by μ) with a standard deviation of 8 minutes (represented by σ). The work begins after 10 minutes, resulting in an average time increase from 45 minutes to 55 minutes.

To calculate the z-score, we use the formula z = (X - μ) / σ, where X represents the value we want to find the probability for. In this case, we want to find the probability associated with a time of 60 minutes.

Substituting the values into the formula, we have z = (60 - 55) / 8 = 0.625.

To find the corresponding p-value, we use the cumulative distribution function (CDF) of the standard normal distribution. In Python, the CDF can be calculated using the stats.norm.cdf() function. So, p\_value = stats.norm.cdf(abs(0.625)).

After performing the calculations, the obtained p-value is approximately 0.7340144.

Since the p-value represents the probability of observing a value as extreme as or more extreme than the one obtained, we subtract the p-value from 1 to obtain the probability of observing a value less extreme. Therefore, 1 - 0.7340144 = 0.2659856, which can be rounded to 0.266.

Ans ~ 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.**

Answer :

we have 400 clerical employees,

μ = 38, Standard deviation

σ =6

A) To determine if there are more employees at the processing center older than 44 than between 38 and 44, we can use z-scores.

Z-score for 44:

Z score = (Value - Mean)/SD

Z = (44 - 38) / 6 = 1

The z-score of 1 corresponds to a cumulative probability of approximately 0.8413 or 84.13%. This means that approximately 84.13% of the employees are younger than 44. Therefore, the percentage of employees older than 44 is 100% - 84.13% = 15.87%.

Number of employees older than 44: 15.87% \* 400 = 63 employees.

Z-score for 38:

Z = (38 - 38) / 6 = 0

The z-score of 0 corresponds to a cumulative probability of 0.5 or 50%. This means that 50% of the employees fall between the ages of 38 and 44.

Number of employees between 38 and 44: (84.13% - 50%) \* 400 = 34.13% \* 400 = 137 employees.

Based on these calculations, we can conclude that there are more employees between the ages of 38 and 44 (137 employees) compared to those older than 44 (63 employees). Hence, the statement that there are more employees at the processing center older than 44 than between 38 and 44 is false.

B) To find the number of employees a training program for employees under the age of 30 would attract, we can use the z-score formula.

Z-score for 30:

Z score = (Value - Mean)/SD

Z = (30 - 38) / 6 = -1.33

The z-score of -1.33 corresponds to a cumulative probability of approximately 0.0918 or 9.18%. This means that about 9.18% of the employees are younger than 30.

Number of employees under the age of 30: 9.18% \* 400 = 36 employees.

Therefore, the statement that a training program for employees under the age of 30 would attract about 36 employees is correct.

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.**

Answer :

We are given two independent and identically distributed random variables: X1 ~ N(μ, σ^2) and X2 ~ N(μ, σ^2). According to the Central Limit Theorem, when we have a large sum of independent and identically distributed random variables, the distribution tends to be approximately normal.

To find the distribution of 2X1, we can apply the properties of multiplication. Using these properties, we can determine that 2X1 follows a normal distribution with mean 2μ and variance 4σ^2, denoted as N(2μ, 4σ^2).

Similarly, using the properties of addition, we can determine the distribution of X1+X2. It follows a normal distribution with mean (μ + μ) = 2μ and variance (σ^2 + σ^2) = 2σ^2, denoted as N(2μ, 2σ^2).

Now, if we consider the difference between 2X1 and (X1+X2), we have 2X1 - (X1+X2). By applying the properties of subtraction, we can simplify it as 2X1 - (X1+X2) = X1 - X2.

The resulting distribution of X1 - X2 is a normal distribution with mean (2μ - 2μ) = 0 and variance (2σ^2 + 4σ^2) = 6σ^2, denoted as N(0, 6σ^2).

In summary, both 2X1 and X1+X2 have the same mean, which is 2μ. However, the variance of 2X1 is twice as large as the variance of X1+X2. Specifically, the variance of 2X1 is 4σ^2, whereas the variance of X1+X2 is 2σ^2.

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**

Answer :

Given that the probability of the random variable X falling within the range (a, b) is 0.99, where the mean (μ) is 100 and the standard deviation (σ) is 20, we can use the properties of the standard normal distribution to determine the values of a and b.

First, we find the z-values corresponding to the percentiles. The z-value at the 0.5th percentile (lower tail) is -2.576, and the z-value at the 99.5th percentile (upper tail) is 2.576. These z-values can be calculated using the stats.norm.ppf() function in Python.

Next, we use the z-value formula to find the corresponding values of X:

x = 20z + 100

Substituting the z-values, we get:

For the lower bound (a):

a = -(20 \* 2.576) + 100 = 48.5

For the upper bound (b):

b = (20 \* 2.576) + 100 = 151.5

Therefore, the two values symmetric about the mean for the given standard normal distribution are 48.5 and 151.5. These represent the lower and upper bounds, respectively, of the range (a, b) for which the probability is 0.99.

D. 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?**

Answer :

we have $1=Rs. 45

Profit1 ~ N(5, 3^2)

Profit2 ~ N(7, 4^2)

Thus,

Company’s profit:

P~N(5+7,3^2+4^2) = N(12, 5^2)

import numpy as np

from scipy import stats

from scipy.stats import norm

# Mean profits from two different divisions of a company = Mean1 + Mean2

Mean = 5+7

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

Mean Profit is Rs 540 Million

# Variance of profits from two different divisions of a company = SD^2 = SD1^2 + SD2^2

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

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

a).

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

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

Range is Rs (99.00810347848784, 980.9918965215122) in Millions

b).

#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 Million Rupees) is',np.round(X,))

5th percentile of profit (in Million Rupees) is 170.0

c).

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

# Probability of Division 1 making a loss P(X<0)

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

0.0477903522728147

# Probability of Division 2 making a loss P(X<0)

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

0.040059156863817086

In summary:

a) The range within which 95% of the probability lies for the company's profit is approximately 99.008 million Rupees to 980.992 million Rupees.

b) The 5th percentile of profit for the company is approximately 170 million Rupees.

c) Profit1 has a slightly higher probability (approximately 0.0478) of making a loss compared to Profit2 (approximately 0.0401)