Samit Dhawal

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

Ans: Option B is correct. The calculation is shown in the attached ipynb file.

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.

    Ans: False, as we know in normal distribution, the majority of values lie around the mean, and the distribution is symmetrical. Given that mean age is 38, so there will be more employees between 38 and 44(within one standard deviation) than older than 44 (beyond one standard deviation).

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

    Ans: False, to determine the number of employees expected to be under the age of 30, we need additional information such as proportion or percentage of employees in that age range. The mean and standard deviation are not sufficient to provide information about the number of employees in a specific age category.

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.

Ans: From the properties of normal random variables,

If $X_1 \sim N(\mu, \sigma^2)$ and $X_2 \sim N(\mu, \sigma^2)$ are two independent identically distributed random variables then,

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

  $Z \sim N(a\mu, a^2\sigma^2)$

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

  $Z \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + a^2\sigma_2^2)$

Distribution:

- $2X_1$: The random variable $2X_1$ follows a normal distribution. Multiplying a random variable by a constant scales the distribution but does not change its shape.
- $X_1 + X_2$: The random variable $X_1 + X_2$ follows a normal distribution as well. The sum of two independent normal random variables is also a normal random variable.

Parameters:

- For both $2X_1$ and $X_1 + X_2$, the mean (μ) remains the same as the original distribution of X_1 and X_2, which is denoted by μ.
- The variance of $2X_1$ is four times the variance of X_1, from $Z \sim N(a\mu, a^2\sigma^2)$, a=2 So, $2^2 \Rightarrow 4\sigma^2$.

  The variance of $X_1 + X_2$ is the sum of the variances of X_1 and X_2, from $Z \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + a^2\sigma_2^2)$, a=1, b=1 So, $1^2 + 1^2 \Rightarrow 2\sigma^2$

Conclusion:

- $2X_1$ follows a normal distribution with mean 2μ and variance $4\sigma^2$.
- $X_1 + X_2$ follows a normal distribution with mean 2μ and variance $2\sigma^2$.
- Overall, mean of $2X_1$ and $X_1 + X_2$ are same, but variance of $2X_1$ is 2 times more than the variance of $X_1 + X_2$.

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

Ans: Explained in the attached ipynb with calculation.

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

Ans: Given,

Profit1 follows a normal distribution with mean1 = 5 million dollars and variance1 = 3^2 million dollars.
Profit2 follows a normal distribution with mean2 = 7 million dollars and variance2 = 4^2 million dollars.
$1 = Rs. 45

stdev1 = sqrt(variance1)
stdev2 = sqrt(variance2)

A. Solution
   To specify a rupee range centered on the mean with 95% probability for the annual profit of the company.

   Converting mean profits from dollars to rupees.
   Mean1 = mean1*45 = 5*45 = 225 million rupees
   Mean2 = mean2*45 = 7*45 = 315 million rupees.

Converting Standard deviation from dollars to rupees.
Stdev1 = stdev1*45 = 3*45 = 135 million rupees.
Stdev2 = stdev2*45 = 4*45 = 180 million rupees.

Range1: (Mean1 - 1.96*Stdev1, Mean1 + 1.96*Stdev1)
(225 - 1.96*135, 225 + 1.96*135)
(-39.60, 489.6)

Range2: (Mean2 – 1.96*Stdev2, Mean2 + 1.96*Stdev2)
(315 – 1.96*180, 315 + 1.96*180)
(-37.80, 667.8)

1.96 is value of z-score for two tailed normal distribution, with 95% confidence level.

B. To specify the 5th percentile of profit (in rupees) for the company:

Division1:
Percentile1: Mean1 + z * Stdev1, where z is the z-score for 5th percentile (~ -1.645)
= 225 + (-1.645 * 135)
= 2.925 million rupees

Divison2:
Percentile2: Mean2 + z * Stdev2,
= 315 + (-1.645 * 180)
= 18.899 million rupees

At the 5th percentile, Division1 tends to have lower profits than Division2. Means Division2 has higher probability of generating profits above around Rs. 18.9 million compared to Division1.

C. To determine which division has a larger probability of making a loss in a given year:

Compare the probabilities of negative profits(losses) for each division using the mean and standard deviation.

Division 1(Profit1):
Calculating the z-score corresponding to zero for division 1:
z1 = (0-Mean1) / Stdev1 => -1.6666666666666667

Using the cumulative distribution function (CDF) of the standard normal distribution to calculate the probability of Profit1 being less than zero.

Probability1 = stats.norm.cdf(z1)

```
# C.
Mean1 = 225
Stdev1 = 135
z1 = (0-Mean1)/Stdev1
prob1 = stats.norm.cdf(z1)
prob1
```

✓ 0.0s

```
0.0477903522728147
```

Division 2(Profit2)

Calculating the z-score corresponding to zero for division 2:

z2 = (0-Mean2) / Stdev2 => -1.6666666666666667

Using the cumulative distribution function (CDF) of the standard normal distribution to calculate the probability of Profit1 being less than zero.

Probability2 = stats.norm.cdf(z2)

```
Mean2 = 315
Stdev2 = 180
z2 = (0-Mean2)/Stdev2
prob2 = stats.norm.cdf(z2)
prob2
```

✓ 0.0s

```
0.040059156863817086
```

Probabilities of loss of Probability1 is higher than the Probability2.