

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) $1 - \text{stats.norm.cdf}(x = 50, \text{loc} = 45, \text{scale} = 8)$
0.2676

2. The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean $\mu = 38$ and Standard deviation $\sigma = 6$. 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.

We have a normal distribution with $\mu = 38$ and $\sigma = 6$. Let X be the number of employees. So according to question

a) Probability of employees greater than age of 44 = $\Pr(X > 44)$

$$\Pr(X > 44) = 1 - \Pr(X \leq 44).$$

$$Z = (X - \mu) / \sigma = (X - 38) / 6$$

Thus the question can be answered by using the normal table to find

$$\Pr(X \leq 44) = \Pr(Z \leq (44 - 38) / 6) = \Pr(Z \leq 1) = 84.1345\%$$

$$\text{Probability that the employee will be greater than age of 44} = 100 - 84.1345 = 15.86\%$$

So the probability of number of employees between 38-44 years of age = $\Pr(X < 44) - 0.5 = 84.1345 - 0.5 = 34.1345\%$

Therefore the statement that "More employees at the processing center are older than 44 than between 38 and 44" is TRUE.

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

Probability of employees less than age of 30 = $\Pr(X < 30)$.

$$Z = (X - \mu) / \sigma = (30 - 38) / 6$$

Thus the question can be answered by using the normal table to find

$$\Pr(X \leq 30) = \Pr(Z \leq (30 - 38) / 6) = \Pr(Z \leq -1.333) = 9.12\%$$

So the number of employees with probability 0.912 of them being under age 30 = $0.0912 \times 400 = 36.48$ (or 36 employees).
Therefore the statement B of the question is also TRUE.

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 $2X_1$ and $X_1 + X_2$? Discuss both their distributions and parameters.

As we know that if $X \sim N(\mu_1, \sigma_1^2)$, and $Y \sim N(\mu_2, \sigma_2^2)$ are two independent random variables then $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$, and $X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.
Similarly if $Z = aX + bY$, where X and Y are as defined above, i.e Z is linear combination of X and Y , then $Z \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$.

Therefore in the question

$$2X_1 \sim N(2\mu, 4\sigma^2) \text{ and}$$

$$X_1 + X_2 \sim N(\mu + \mu, \sigma^2 + \sigma^2) \sim N(2\mu, 2\sigma^2)$$

$$2X_1 - (X_1 + X_2) \sim N(4\mu, 6\sigma^2)$$

4. Let $X \sim N(100, 20^2)$. 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. 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).

$$Z * \sigma + \mu = X$$

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

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

So, option D is correct.

5. Consider a company that has two different divisions. The annual profits from the two divisions are independent and have distributions $\text{Profit}_1 \sim N(5, 3^2)$ and $\text{Profit}_2 \sim N(7, 4^2)$ 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

```
import pandas as pd
import numpy as np
from scipy import stats
```

```
mean = 5+7
print('Mean profit in rs',mean*45,'millions')
      Mean profit in rs 540 millions
```

```
sd = np.sqrt((9)+(16))
print('sd profit in rs', sd*45,'millions')
```

```
sd profit in rs 225.0 millions
```

```
print('Range profit is',(stats.norm.interval(.95,540,225)), 'millions')
```

```
      Range profit is (99.00810347848784, 980.9918965215122) m
illions
```

- B. Specify the 5th percentile of profit (in Rupees) for the company

```
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
```

- C. Which of the two divisions has a larger probability of making a loss in a given year?

```
stats.norm.cdf(0,5,3)
0.0477903522728147
```

```
stats.norm.cdf(0,7,4)
0.040059156863817086
```

Division 1 with 4.77% making more loss compare to Division 2 with 4%