

Question

In an organization, there is a total of 1,00,000 employees. HR Department wants to order XL and L T-shirts for their employees. HR has taken a sample of 500 employees out of which 300 want XL and 200 want L. From this given data HR wants us as a data analyst to tell them the total how many XL and L T-shirts they have to order with a 95% confidence interval?

n= 500

Confidence Interval = 0.95

As per the data, 60% of the sample want XL while 40% of them want L

Let's assume the Standard Deviation of the population is 100 i.e. $\sigma = 100$.

Assume Sample mean is $\bar{X} = 150$

Null Hypothesis: - 60% of them want L T-Shirt

Alternate Hypothesis: - 60% of them don't want L T-Shirt

Calculate Significance level

Significance Value(α) = 1 - Confidence Interval

$$= 1 - 0.95 = 0.05$$

$$z \alpha/2 = z 0.05/2 = z 0.025$$

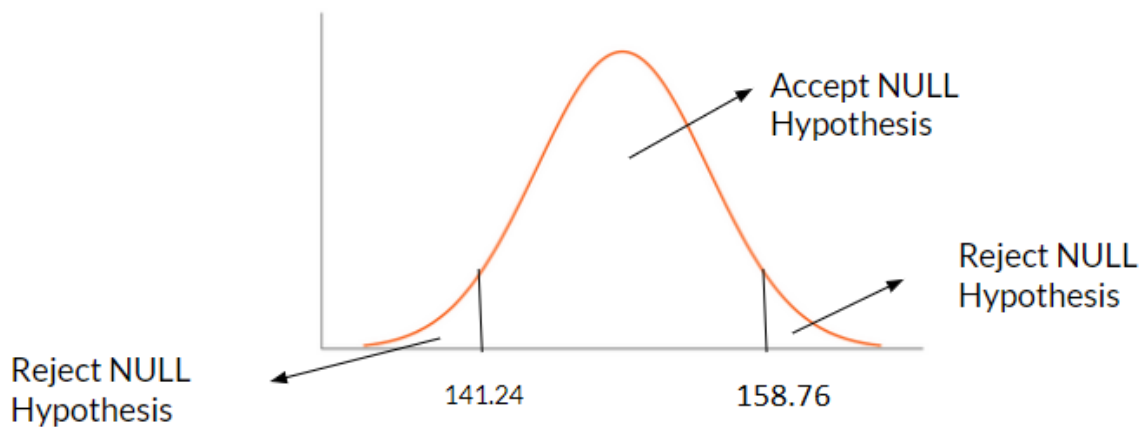
$$1 - 0.025 = 0.975$$

Therefore, from the z table(if population standard deviation is given)value is 1.96

$$z \alpha/2 = 1.96$$

$$\begin{aligned}\text{Lower Fence} &= \bar{x} - z \alpha/2 \left(\frac{\sigma}{\sqrt{n}} \right) \\ &= 150 - 1.96(100/\sqrt{500}) \\ &= 150 - 1.96 \times 4.47 \\ &= 141.24\end{aligned}$$

$$\begin{aligned}\text{Higher Fence} &= \bar{x} + z \alpha/2 \left(\frac{\sigma}{\sqrt{n}} \right) \\ &= 150 + 1.96(100/\sqrt{500}) \\ &= 150 + 1.96 \times 4.47 \\ &= 158.76\end{aligned}$$



The range for L T-Shirt is small as compared to the total sample drawn of $n = 500$

In the sample dataset, we have only 200 people who want L out of 500 people i.e. 40% of the sample dataset.

We can infer to reject the null hypothesis and accept the Alternative Hypothesis i.e. 60% of them don't want L T-Shirt

Thus, 60 % want XL T-Shirt and 40% want L T-Shirt

Therefore, out of 1,00,000 employees, we can order 60,000 XL T-Shirts and 40,000 L T-Shirts.