Student Information

Name :Aybüke Aksoy

ID :2448090

Answer 1

Given a sample of data and a desired confidence level $(1-\alpha)$, we can construct a confidence interval [a,b] that will satisfy the condition:

$$P\{a \leq \theta \leq b\} = 1-\alpha$$

We can start by estimating θ.
Standardizing the unbiased estimator $\hat{\theta}$ with normal distribution, we get:

$$Z = \frac{\hat{\theta} - E(\hat{\theta})}{\sigma(\hat{\theta})} = \frac{\hat{\theta} - \theta}{\sigma(\hat{\theta})}$$

This variable falls between $-z_{\alpha/2}$ and $z_{\alpha/2}$ with the probability $(1-\alpha)$

$$P\{-z_{\alpha/2} \leq \frac{\hat{\theta} - \theta}{\sigma(\hat{\theta})} \leq z_{\alpha/2}\} = 1-\alpha$$

Solving the inequality for θ, we get:

$$a = \hat{\theta} - z_{\alpha/2}\sigma(\hat{\theta})$$

$$b = \hat{\theta} + z_{\alpha/2}\sigma(\hat{\theta})$$

While estimating the population mean $\hat{X}$:

$$\hat{\theta} = \hat{X}$$
$$a = \hat{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$
$$b = \hat{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

a)

$$Standard\ deviation : \sigma = 3$$

$$Sample\ size : n = 10$$

$$Sample\ mean : \hat{X} = \frac{20.1 + 12.8 + 18.9 + 16.4 + 20.3 + 10.1 + 15.4 + 12.4 + 24.7 + 18.5}{10} = 16.96$$

$$Confidence\ level1 : 1 - \alpha = 0.90\ ,\ \ \alpha = 0.1$$

$$Confidence\ level2 : 1 - \alpha = 0.99\ ,\ \ \alpha = 0.01$$

From the table:

$$z_{0.05} = 1.645$$

We can find a and b for $\%90$ confidence level as:

$$a = 16.96 - 1.645\frac{3}{\sqrt{10}} = 15.4$$

$$b = 16.96 + 1.645\frac{3}{\sqrt{10}} = 18.521$$

$$Confidence\ interval : [15.4, 18.521]$$

From the table:

$$z_{0.005} = 2.576$$

We can find a and b for $\%99$ confidence level as:

$$a = 16.96 - 2.576\frac{3}{\sqrt{10}} = 15.4$$

$$b = 16.96 + 2.576\frac{3}{\sqrt{10}} = 18.521$$

$$Confidence\ interval : [14.517, 19.403]$$

b)

$$Confidence\ level : 1 - \alpha = 0.98\ ,\ \ \alpha = 0.02$$

$$margin : z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1.55$$

Hence;

$$Sample\ size : n \geq (z_{\alpha/2}\frac{\sigma}{1.55})^2$$

From the table:

$$z_{0.01} = 2.326$$

$$n \geq (2.326\frac{3}{1.55})^2$$

$$n \geq 20.267$$

$$n = 21$$

Answer 2

a) No. The mean rating and sample size are not enough for statistics of a restaurant. We also need sample for standard deviation and variance.

b)

$$H_0 = \text{ the null hypothesis}$$

$$H_A = \text{ the alternative hypothesis}$$

Since we decide whether a restaurant will be included in our list of candidate restaurants or not from looking at the rating being lower than 7.5 with %5 level of significance, we can determine our null hypothesis as:

$$H_0 : \mu_0 = 7.5$$

$$H_A : \mu_0 < 7.5$$

A normally would not be in our list because its rating is 7.4; however, if we have enough evidence to reject the null hypothesis, which is also in favor of H_A, A can not be in our list. To decide, we can use the Z test: H_A is a left tail alternative. A level α test with a left-tail alternative should:

$$reject\ H_0, \quad if\ Z\ \leq -z_\alpha$$

$$accept\ H_0, \quad if\ Z\ > -z_\alpha$$

In this case, the rejection region is consists of small values of Z only.

$$A = [-z_\alpha, +\infty)$$

$$R = [-\infty, -z_\alpha)$$

For an estimator $\hat{\theta}$ for population parameter θ that has normal distribution:

$$Test\ Statistic : Z = \frac{\hat{\theta} - E(\hat{\theta})}{\sqrt{Var(\hat{\theta})}} = \frac{\hat{X} - \mu_0}{\sigma/\sqrt{n}}$$

$$Standard\ deviation : \sigma = 0.8$$

$$Sample\ size : n = 256$$

$$Sample\ mean : \hat{X} = 7.4$$

$$\alpha = 0.05$$

$$z_\alpha = 1.645$$

Hence,

$$Z = \frac{7.4 - 7.5}{0.8/\sqrt{256}} = -2$$

Since $Z < -z_\alpha$, we can reject the null hypothesis.
Thus, the restaurant A is not included in the list.

c) Using the information in part b with the only difference which is taking σ as 1.0 instead of 0.8 and calculating Z again:

$$Z = \frac{7.4 - 7.5}{1.0/\sqrt{256}} = -1.6$$

Since $Z > -z_\alpha$, we should accept the null hypothesis.
Thus, the restaurant A is included in the list.

d) Since the mean of user ratings for restaurant A=7.6 is already greater than $\mu_0 = 7.5$, our decision will not be affected by other parameters like standard deviation. Hence, we do not have to conduct any statistical test.

Answer 3

If we have enough evidence to reject the null hypothesis, which is also in favor of H_A, we can claim that there is not at least 90 minutes run-time improvement.
Hence we can take H_0 and H_A as:

$$H_0 : \mu_A - \mu_B = 90$$

$$H_A : \mu_A - \mu_B < 90$$

Computer A:

$$Sample\ Standard\ deviation : \sigma_A = 5.2$$

$$Sample\ size : n = 20$$

$$Sample\ mean : \hat{A} = 211$$

Computer B:

$$Sample\ Standard\ deviation : \sigma_B = 22.8$$

$$Sample\ size : m = 32$$

$$Sample\ mean : \hat{B} = 133$$

To decide, we can use the T test:

a)

$$reject\ H_0, \quad if\ t \leq -t_\alpha$$

$$accept\ H_0, \quad if\ t > -t_\alpha$$

While population variances are equal:

$$\alpha = 0.01$$

$$\sigma_A^2 = \sigma_b^2 = \sigma_2$$

$$Pooled\ Sample\ Variance: s_p^2 = \frac{(n-1)\sigma_A^2 + (m-1)\sigma_B^2}{n+m-2}$$

$$Degree\ of\ freedom: n+m-2$$

$$t = \frac{\hat{A} - \hat{B} - D}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

$$s_p^2 = \frac{19 * 5.2^2 + 31 * 22.8^2}{50} = 332.576$$

$$D = 50$$

$$t_\alpha = t_{0.01} = 2.403\ ,\ from\ the\ table$$

$$t = \frac{211 - 133 - 90}{\sqrt{332.576}\sqrt{\frac{1}{20} + \frac{1}{32}}} = -2.3084$$

Since $t > -t_\alpha$:
We should accept the null hypothesis.
Hence, B does provide 90 minutes or better improvement.

]b) While population variances are not equal:

$$T\ Ratio: \frac{(\hat{A} - \hat{B}) - (\mu_A - \mu_B)}{\sqrt{\frac{s_A^2}{n} + \frac{s_B^2}{m}}}$$

Using the information in part a:

$$T\ Ratio: \frac{(211 - 133) - 90}{\sqrt{\frac{5.2^2}{20} + \frac{22.8^2}{32}}} = -2.8606$$

$$Degrees\ of\ freedom: \frac{(\frac{\sigma_A^2}{n^2} + \frac{\sigma_B^2}{m^2})^2}{\frac{(\sigma_A^2/n)}{n-1} + \frac{(\sigma_B^2/m)^2}{m-1}} = 36$$

$$t_\alpha = t_{0.01} = 2.434\ ,\ from\ the\ table$$

Since $t < -t_\alpha$:
We should reject the null hypothesis.
Hence, B does not provide 90 minutes or better improvement.