

Problems for Class 9

TRUE or FALSE problems

State whether you believe the given statement is TRUE or FALSE and provide a brief argument for your answer.

1. A confidence interval always includes the estimated population parameter.

FALSE

A $k\%$ confidence interval will include the estimated parameter with probability $k\%$.

2. Consider a 99% and a 90% confidence interval for the same population parameter estimated by the same estimator. Then the length of the 99% interval will always be greater than that of the 90% interval.

TRUE

The smaller the interval, the smaller the probability that it includes the population parameter.

3. We have a random sample from a population with pmf depending on the parameter θ . Let T_1 and T_2 be two sample statistics from this random sample. We have:

$$P[T_1 \leq \theta \leq T_2] = \theta^2, \text{ for } \theta \leq 1.$$

Then interval $[T_1, T_2]$ is called a $100\theta^2$ per cent confidence interval for θ .

FALSE

The probability must not depend on the parameter of the confidence interval.

4. When sampling from a normal population with unknown mean and variance, the expected value of the standardised sample mean is always zero.

TRUE

The standardised sample mean follows the t distribution which has an expected value of zero.

5. When sampling from a normal distribution, the expected value of the sample mean follows the t distribution.

FALSE

The expected value of a distribution, if it exists, is a constant.

Exercises

Exercises 1-5: NCT 8.10, 8.11, 8.16, 8.18, 8.25.

NCT 8.10

Calculate the margin of error to estimate the population mean

a. 98% confidence level; $n = 64$, variance = 144

$$ME = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.33 \left(\frac{12}{\sqrt{64}} \right) = 3.495$$

b. 99% confidence interval, $n=120$; standard deviation = 100

$$ME = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.58 \left(\frac{100}{\sqrt{120}} \right) = 23.552$$

NCT 8.11

Calculate the width to estimate the population mean, for

a. 90% confidence level, $n = 100$, variance = 169

$$\text{width} = 2ME = 2 \left[z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] = 2 \left[1.645 \left(\frac{13}{\sqrt{100}} \right) \right] = 4.277$$

b. 95% confidence interval, $n = 120$, standard deviation = 25

$$\text{width} = 2ME = 2 \left[z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] = 2 \left[1.96 \left(\frac{25}{\sqrt{120}} \right) \right] = 8.9461$$

NCT 8.16

a. $n = 16$, $\bar{x} = 4.07$, $\sigma = .12$, $z_{.005} = 2.58$

$$4.07 \pm 2.58(.12/4) = 3.9926 \text{ up to } 4.1474$$

b. narrower since the z score for a 95% confidence interval is smaller than the z score for the 99% confidence interval

c. narrower due to the smaller standard error

d. wider due to the larger standard error

NCT 8.18

Find the ME

a. $n = 20$, 90% confidence level, $s = 36$

$$ME = t_{\alpha/2} \frac{s}{\sqrt{n}} = ME = 1.729 \left(\frac{36}{\sqrt{20}} \right) = 13.9182$$

b. $n = 7$, 98% confidence level, $s = 16$

$$ME = t_{\alpha/2} \frac{s}{\sqrt{n}} = ME = 3.143 \left(\frac{16}{\sqrt{7}} \right) = 19.007$$

c. $n = 4$, 99% confidence level, $x_1 = 15$, $x_2 = 17$, $x_3 = 13$, $x_4 = 11$

$$\bar{x} = 14, \quad s = 2.58199 \quad ME = 5.841 \left(\frac{2.58199}{\sqrt{4}} \right) = 7.5407$$

NCT 8.25

$$n = 9, \quad \bar{x} = 157.82, \quad s = 38.89, \quad t_{8,0.25} = 2.306$$

$$\text{margin of error: } \pm 2.306(38.89/3) = \pm 29.8934$$

6. (a) For each population we must find t_1 and t_2 , s.t.

$$P(t_1 \leq z_i \leq t_2) = 0.90$$

$$\text{where } z_i = \frac{\bar{X}_i - \mu_i}{\frac{\sigma_i}{\sqrt{n}}}$$

$$\text{and where } z_i \sim N\left(\mu_i, \frac{\sigma_i^2}{\sqrt{n}}\right)$$

Using the standard normal cdf tables we find that:

$$t_1 = -1.65$$

$$t_2 = 1.65$$

Solving for μ_i we obtain:

$$P\left(\bar{X}_i - 1.65 \frac{\sigma_i}{\sqrt{n}} \leq \mu_i \leq \bar{X}_i + 1.65 \frac{\sigma_i}{\sqrt{n}}\right) = 0.90$$

Hence the 90% confidence intervals are:

$$\left[\bar{X}_i - 1.65 \frac{\sigma_i}{\sqrt{n}}, \bar{X}_i + 1.65 \frac{\sigma_i}{\sqrt{n}}\right] \quad \text{for } i=1,2,3$$

$$\text{with each having length: } 2 \cdot 1.65 \frac{\sigma_i}{\sqrt{n}}$$

We then have the following lengths:

$$\text{For } i=1: \sigma_1 = \sqrt{4} = 2 \quad \Rightarrow \text{length} = 2 \cdot 1.65 \frac{2}{4} = 2 \cdot 1.65 \frac{1}{2} = 1.65$$

$$\text{For } i=2: \sigma_2 = \sqrt{6} \quad \Rightarrow \text{length} = 2 \cdot 1.65 \frac{\sqrt{6}}{4} = 1.65 \frac{\sqrt{6}}{2} = 1.65 \cdot 1.22$$

$$\text{For } i=3: \sigma_3 = \sqrt{8} \quad \Rightarrow \text{length} = 2 \cdot 1.65 \frac{\sqrt{8}}{4} = 1.65 \frac{\sqrt{8}}{2} = 1.65 \cdot 1.41$$

The length of the interval therefore increases as σ_i increases. This is because, if we want to keep the probability that the random interval includes the estimated parameter constant, we must increase the length of the interval as the variance of our estimator increases.

(b)

We need to find t_1 and t_2 that satisfy:

$$P(t_1 \leq z_i \leq t_2) = 0.90$$

and are other than $t_1 = -1.65$ and $t_2 = 1.65$.

The 10% which is the complement of this probability can be split into two other alternative parts. For example, 1% in the upper tail and 9% in the lower tail, and vice versa.

For the case where we have 1% in the upper tail and 9% in the lower trail, we have:

$$t_2 = 2.33, t_1 = -1.34$$

For the case where we have 9% in the upper tail and 1% in the lower, we have:

$$t_2 = 1.34, t_1 = -2.33$$

Then, our confidence intervals become:

Confidence interval 1.

$$\begin{aligned} \left[\bar{X} - t_2 \frac{\sigma}{\sqrt{n}}, \bar{X} - t_1 \frac{\sigma}{\sqrt{n}} \right] &= \left[\bar{X} - 2.33 \cdot \frac{2}{4}, \bar{X} + 1.34 \cdot \frac{2}{4} \right] \\ &= [\bar{X} - 1.165, \bar{X} + 0.67] \end{aligned}$$

Confidence interval 2.

$$\begin{aligned} \left[\bar{X} - t_2 \frac{\sigma}{\sqrt{n}}, \bar{X} - t_1 \frac{\sigma}{\sqrt{n}} \right] &= \left[\bar{X} - 1.34 \cdot \frac{2}{4}, \bar{X} + 2.33 \cdot \frac{2}{4} \right] \\ &= [\bar{X} - 0.67, \bar{X} + 1.165] \end{aligned}$$

Hence both of these intervals have length: $1.165+0.67=1.835$, which is larger than what we found in (a), namely 1.65.

8. Let X be the expenditures on drinks during a week. We the have $X \sim N(\mu, 25)$, $n=49$

The general form of confidence intervals for sample means drawn from normal populations with known variance is:

$$\left[\bar{X} - t_2 \frac{\sigma}{\sqrt{n}}, \bar{X} - t_1 \frac{\sigma}{\sqrt{n}} \right]$$

For the standard normal cdf we find that:

$$t_1 = -1.96 \text{ and } t_2 = 1.96$$

We know that:

$$\sigma = \sqrt{25} = 5$$

$$\sqrt{n} = \sqrt{49} = 7$$

So the 95% confidence interval is:

$$\left[\bar{X} - 1.96 \frac{5}{7}, \bar{X} + 1.96 \frac{5}{7} \right] = [\bar{X} - 1.4, \bar{X} + 1.4]$$

The length of this interval is $2 \cdot 1.4 = 2.8$

For an interval with length 1 we have:

$$\begin{aligned} 2 \cdot t_1 \frac{\sigma}{\sqrt{n}} = 1 &\Rightarrow 2 \cdot 1.96 \frac{5}{\sqrt{n}} = 1 \Rightarrow 10 \cdot 1.96 = \sqrt{n} \Rightarrow 19.6 = \sqrt{n} \\ &\Rightarrow n = 19.6^2 = 384.16 \end{aligned}$$

So, we need a sample of 385 individuals.