Homework assignment #8: Solutions

Milica Cudina

2021-11-02

Problem #1 (5 points)

Consider a normal population with mean 10 and standard deviation 1. What is the probability that a sample mean of a sample of size 16 falls into the interval $[9.525, 9.9]$?

Solution: The sample mean of a sample of size 16 has the distribution

$$\bar{X} \sim Normal\left(mean = 10, sd = \frac{1}{\sqrt{16}} = \frac{1}{4}\right).$$

We are looking for the probability

$$\mathbb{P}[\bar{X} \in [9.525, 9.9]] = \mathbb{P}[9.525 \le \bar{X} \le 9.9] = \mathbb{P}[\bar{X} \le 9.9] - \mathbb{P}[\bar{X} < 9.525].$$

Rewriting $\bar{X}$ in standard units (with $Z \sim N(0, 1)$), we get

$$\mathbb{P}[\bar{X} < 9.525] = \mathbb{P}\left[\frac{\bar{X} - 10}{\frac{1}{4}} < \frac{9.525 - 10}{\frac{1}{4}}\right] = \mathbb{P}[Z < -1.9],$$

$$\mathbb{P}[\bar{X} < 9.9] = \mathbb{P}\left[\frac{\bar{X} - 10}{\frac{1}{4}} < \frac{9.9 - 10}{\frac{1}{4}}\right] = \mathbb{P}[Z < -0.4].$$

Using the standard normal tables, we get our answer:

$$\mathbb{P}[Z < -0.4] - \mathbb{P}[Z < -1.9] = 0.3446 - 0.0287 = 0.3159.$$

Alternatively, in `R`, we have

```
pnorm(9.9, 10, 0.25) - pnorm(9.525, 10, 0.25)
## [1] 0.3158617
```

Problem #2 (5 points)

A particular type of wool for clothes manufacturing has to have a specific tensile strength in order to be used in weaving machines without breaking. We model its tensile strength as normally distributed with standard deviation 0.4 MPa. How is the variance of the sample mean changed when the sample size decreases from 64 to 36?

Solution: The variance of the sample mean $\bar{X}_{64}$ for the sample of size 64 is

$$Var[\bar{X}_{64}] = \frac{(0.4)^2}{64} = 0.0025.$$

The variance of the sample mean $\bar{X}_{36}$ for the sample of size 36 is

$$Var[\bar{X}_{36}] = \frac{(0.4)^2}{36} = 0.0044.$$

Problem #3 (5 points)

To estimate a population mean, our resident statistician Martyn Rivera plans to pick two independent simple random samples, each of size 100, from the population. He also plans to calculate the confidence interval with level C for each sample. What is the probability that at least one of his confidence intervals will cover the population mean?

Solution: The probability that neither of the two confidence intervals contains the true population mean is $(1-C)^2$. So, the probability we are seeking is $1 - (1-C)^2 = C(2-C)$.

Problem #4 (5 points)

Consider the following statement:

"At the 95% confidence level, we estimate that the true population mean of the tail length of rhesus macaques is between 20.7 cm and 22.9 cm. If we were to construct a confidence interval at the 99% confidence level, the resulting interval would be narrower."

Are there any errors in the above two statements? Justify your answer.

  (a) This interpretation is correct. There are no errors.

  (b) In the first sentence, the phrase "true population mean" should be replaced with "sample mean."

  (c) In the second sentence, "narrower" should be replaced with "wider."

  (d) In the first sentence, the word "estimate" should be replaced with "know for certain."

  (e) None of the above.

Solution: The correct answer is **(c)** since the critical value associated with the 99% confidence level is larger than the critical value associated with the 95% confidence level.

Problem #5 (5 points)

The Chicago Tribune and the Los Angeles Times conducted separate national polls where they asked full-time employees how many hours they work per week, and reported confidence intervals at the 95% confidence level. The Chicago Tribune surveyed 500 people, and the Los Angeles Times surveyed 300 people. Which paper reported a larger margin of error? Assume the standard deviations of the two populations were equal and justify your answer.

  (a) The Chicago Tribune

  (b) The Los Angeles Times

  (c) The margins of error are the same

  (d) The margins of error will depend on the means of the samples

  (e) There is not enough information to answer this question

Solution: The correct answer is **(b)** since the sample size for the LA Times survey is smaller.

Problem #6 (2 points)

Source: Problem 6.21 from the Moore/McCabe/Craig.

Consider the following two scenarios:

- Take a simple random sample of 100 sophomore students at your college or university.

- Take a simple random sample of 100 students at your college or university.

For each of these samples you record the amount spent on textbooks used for classes during the fall semester. You should suspect that the first sample should have the smaller margin of error. *True or false? Why?*

Solution: The answer is **TRUE**. The reason is that the subpopulation of sophomores is more homogeneous than the entire population.

Problem #7 $(4 + 4 + 2 = 10$ **pts**$)$

Source: "Probability and Statistics for Engineers and Scientists" by Walpole, Myers, Myers, and Ye.

The average zinc concentration recovered from a sample of measurements taken at 36 different locations in a river is found to be 2.6 grams per milliliter. Assume the normal distribution for the concentration and let the population standard deviation be 0.3 gram per milliliter. Find the

(i) 95% confidence interval, and

(ii) 99% confidence interval

for the mean zinc concentration in the river.

Moreover, how large a sample is required if we want to be 95% confident that our estimate of the mean parameter is off by less than 0.05?

Solution: The point estimate of the population mean is $\bar{x} = 2.6$. The standard error is $\frac{\sigma}{\sqrt{n}} = \frac{0.3}{\sqrt{36}} = 0.05$.

(i) For a 95% confidence interval, the critical value is $z^* = \Phi^{-1}(0.025) = 1.96$. So, the 95%−confidence interval is

$$2.6 - (1.96)(0.05) < \mu < 2.6 + (1.96)(0.05),$$

i.e., 2.6 ± 0.098, i.e., $(2.502, 2.698)$.

(ii) For a 99% confidence interval, the critical value is $z^* = \Phi^{-1}(0.005) = 2.575$. So, the 99%−confidence interval is

$$2.6 - (2.575)(0.05) < \mu < 2.6 + (2.575)(0.05),$$

i.e., 2.6 ± 0.12875, i.e., $(2.47125, 2.72875)$.

As for the required sample size so that we are 95% confident that our estimate of the mean parameter is off by less than 0.05, the condition is

$$(1.96)\left(\frac{0.3}{\sqrt{n}}\right) \leq 0.05 \quad \Rightarrow \quad n \geq \left(\frac{1.96(0.3)}{0.05}\right)^2 = 138.2976.$$

Since n must be an integer, the condition is $n \geq 139$.

Problem #8 (5 points)

Source: "Mathematical Statistics with Applications" by Ramachandran and Tsokos.

Fifteen vehicles were observed at random for their speeds (in mph) on a highway with a speed limit posted as 70 mph. It was found that the observed sample average was 73.3 mph. Suppose that from the past experience we can assume that vehicle speeds are normally distributed with the standard deviation of the speed equal to 3.2 mph. Construct a 90% confidence interval for the true mean speed μ of the vehicles on this highway.

Solution: We are given that the point estimate for the mean speed is $\bar{x} = 73.3$. The population standard deviation is given to be $\sigma = 3.2$. The critical value of the standard normal distribution at the confidence level of 90% is $z^* = \Phi^{-1}(0.95) = 1.645$. So, the confidence interval is

$$73.3 - (1.645)\left(\frac{3.2}{\sqrt{15}}\right) < \mu < 73.3 + (1.645)\left(\frac{3.2}{\sqrt{15}}\right)$$

i.e., 73.3 ± 1.359159, i.e., $(71.94084, 74.65916)$.

Problem #9 (3 points)

The standard deviation of the zinc concentration in a certain river is given to be 0.45 grams per milliliter. Assuming the normal distribution as the model for the zinc concentration, you want to create a 90%-confidence interval for the mean zinc concentration. Your goal is to have the margin of error of at most 0.04. How large should your sample be?

Solution: We need to have

$$n \geq \left(\frac{1.645(0.45)}{0.04}\right)^2 = 342.481. \tag{1}$$

So, $n \geq 343$.

Problem #10 (5 points)

For your own RT, go to: This website. You will also see that the distribution is not really well-modelled as normal.

You want to study the reaction time (RT) of undergraduate students. From past research, you are comfortable modelling the RT using the normal distribution with a standard deviation of 30 ms. You gather a simple random sample of 225 undergraduates. Their sample average is 215 ms. What is the 80%-confidence interval?

Solution: The critical value z^* corresponding to the confidence level of 80% is $z^* = \Phi^{-1}(0.90) = 1.28$. The standard deviation of the sample mean with the given $\sigma = 30$ and the sample size $n = 225$ is

$$\frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{225}} = \frac{30}{15} = 2.$$

So, the margin of error equals

$$z^*\left(\frac{\sigma}{\sqrt{n}}\right) = 1.28(2) = 2.56.$$

So, the 80%-confidence interval is $(212.44, 217.56)$.