proprep™

Workbook

Table of Contents

Confidence Intervals for Sample Means

The Big Idea

The sample average is an estimator for the population average. We estimate the population average using the sample average and build a confidence interval around it.

A 95% confidence interval means that there is a 95% probability that the parameter, μ, is included within that interval.

Example (Solution in the recording)

A researcher randomly samples 25 students who took an IQ test.
She built a 95% confidence interval for the average exam score
and obtained the interval from 510 to 590.
What does this means?

Known Population Variance

Theory

Suppose we want to build a confidence interval for the mean μ, and that we know the population variance, or σ^2. We try to estimate the parameter μ, using the statistic $\overline{x}$.

Conditions for building the confidence interval

1) $X \sim N$ or $n \geq 30$.
2) σ^2 (the population variance) is known.

Formula for the confidence interval: $\overline{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$.

Example (solution in the recording)

According to the manufacturer's figures, a battery's lifespan follows a normal probability distribution with a standard deviation of 1 hour. We want to estimate a battery's average lifespan. 4 batteries are randomly sampled, with an average lifespan of 13.5 hours.
Build a 95% confidence interval for the battery's expected lifespan.

Maximum Estimation Error

$$\varepsilon = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

ε is the maximum estimation error, also called the statistical error or the sampling error.

Example (solution in the recording)

What can be stated with a 95% level of certainty about the estimation error?

Mathematical connections with the confidence interval

- The length of the confidence interval is double the maximum estimation error: $L = 2\varepsilon$.
- The sample average always falls exactly in the middle of the confidence interval: $\bar{X} = \dfrac{A+B}{2}$.
- When the number of observations increases, we have more information. This means that our estimation will be more accurate, and the confidence interval will be smaller.
- When the confidence level is higher, the confidence interval will be bigger.

Questions

1) A researcher is estimating the average salary in Wisconsin.
Based on a sample, she determines with a 95% level of certainty that the average weekly salary is between $\$1,000$ and $\$1,400$.

a. What is the length of the confidence interval?
b. What are the chances that the sample error is greater than $\$300$?

2) The average output of a factory is 4,950 products per day over a random set of 100 days. Assume that the population standard deviation is known to be 150 products per day. Estimate the average daily output of a certain factory at a 95% confidence level.

3) We want to estimate a battery's average lifespan. Assume that the battery's lifespan follows a normal probability distribution with a standard deviation of 20 hours. We sample 25 devices and find their average lifespan to be 230 hours.

a. Construct a 90% confidence interval for the battery's average lifespan.
b. Will a 95% confidence interval be shorter or longer? Explain.

4) The average monthly salary of 200 employees in Ohio is found to be $\$5,000$.
Assume that the population standard deviation of salaries is $\$1,000$.

a. Construct a 95% confidence interval for the actual average salary.
b. What sample size is needed to shorten the confidence interval by 50%?
c. If we expand the sample size from 200 and construct a 95% confidence interval, what happens to the interval?

5) A company is studying the recovery time of a new drug. 60 people participated in the study, with an average recovery of 4 days. The population standard deviation is 2 days.

a. Construct a confidence interval for the mean recovery time at a 90% level of confidence.
b. What happens to the length of the confidence interval, if the sample were four times as large? Explain.
c. What happens to the length of the confidence interval, if we use a higher level of confidence? Explain.

6) A researcher constructs a confidence interval for an average, using a sample of 16 observations and obtains $82 < \mu < 92$.
Assume that the variable follows a normal distribution and that its standard deviation is 10.

a. What is the sample average?
b. What is the level of confidence for this confidence interval?
c. What are the chances that the estimation error is greater than 5?

7) Which of the following factors does not affect the length of the confidence interval? (when the population variance is known)

a. The confidence level.
b. The population standard deviation.
c. The sample size.
d. The sample standard deviation.

8) A researcher constructs a confidence interval for the average and obtains the following confidence interval: $63 < \mu < 83$.
Assume a known population standard deviation, and a sample size of 40.

a. What sample size is needed for a confidence interval of length 10?
b. The original confidence level was 95%. Construct a confidence interval at a 98% level of confidence.

Answer Key

1) a. $\$400$; b. $\$200$

2) $(4920.6,\ 4979.4)$

3) a. $(223.42,\ 236.58)$ b. Longer.

4) a. $(4861.4,\ 5138.6)$ b. $n = 800$ c. Interval will get smaller.

5) a. $(3.5753,\ 4.4247)$ b. Interval will be halved. c. Interval will be longer.

6) a. 87 b. 95.45% c. 4.55%

7) The sample mean.

8) a. $n = 160$ b. Solution in the recording.

Unknown Population Variance

Theory

The t probability distribution is a bell-shaped symmetric probability distribution with a mean of 0. The t probability distribution is similar to the Z probability distribution, but since is wider, its corresponding values are larger.

The t probability distribution involves on a concept called degrees of freedom.
The degrees of freedom are $df = n-1$.
As the degrees of freedom increase, the probability distribution becomes higher and narrower.
As the degrees of freedom tend towards infinity, the t probability distribution converges to the Z probability distribution.

Example (solution in the recording)

The time it takes 8^{th} grade students to solve a given arithmetic question has a normal probability distribution. In order to estimate the expectation of the time needed to find the solution, four 8^{th} grade students were sampled.
The following are the results obtained in minutes: 4.7, 5.2, 4.6, 5.3.
Construct a confidence interval at a 95% level of significance for the average time taken by 8^{th} grade students to solve a question.

Questions

1) A study investigates how a certain drug affects pulse rate. A sample of 5 participants measured their pulse and recorded the number of beats per minute: 89, 79, 84, 88, 84. Assume that pulse rate is approximately normal.

a. Construct a 95% confidence interval for the expected pulse rate among all users of this drug.
b. Assuming that the average pulse rate for people who do not take the drug is 70, does the drug affect pulse rate, at a 95% level of certainty?
c. In continuation of part a: if we were to construct a confidence interval at a 99% level of certainty instead, what will happen to the confidence interval?

2) In a sample of 25 college students, the average height was 178cm, with a standard deviation of 13cm.
Create a confidence interval at a 90% level of confidence for the expected height of college students.

3) Steve wants to estimate the average time (in minutes) that it takes him to get to work. He samples his commute time for five days, with the following results: 27, 34, 32, 40, 30.

a. Estimate the average travel time at a 95% level of certainty.
b. How would the size of the confidence interval change, if Steve sampled more days?

4) Scores on an intelligence test follow a normal probability distribution. The scores of 25 people averaged 102, with a sample standard deviation of 13.

a. Construct a confidence interval for the population average score at a 95% level of certainty.
b. Repeat part a, assuming that the obtained standard deviation of 13 is the population standard deviation.
c. Explain the differences in the answers to parts a and b.

5) 60 babies were weighed at birth, with an average of 7.7 lbs and a sample standard deviation of 1 lb.
Construct a confidence interval for birth weight at 95% confidence.
Explain what this means.

6) Two statisticians constructed 95% confidence intervals for the same parameter μ.
Each statistician had a different sample of 10 observations.
Statistician A assumed that $\sigma = 20$.
Statistician B calculated the sample standard deviation, and found that $\sigma = 20$.
Which of the statisticians will have a longer confidence interval (select the correct answer)?

a. Statistician A.
b. Statistician B.
c. Both statisticians will have confidence intervals of the same length.
d. It depends on the sample results of each statistician.

Answer Key

1) a. $(79.88,\ 89.72)$; b. Yes. c. Interval will be longer.

2) $(173.55,\ 182.45)$

3) a. $(26.543,\ 38.657)$ b. Interval will be shorter.

4) a. $(96.634,\ 107.37)$ b. Interval will be shorter.

5) $(7.4417,\ 7.9583)$

6) b. Statistician B.