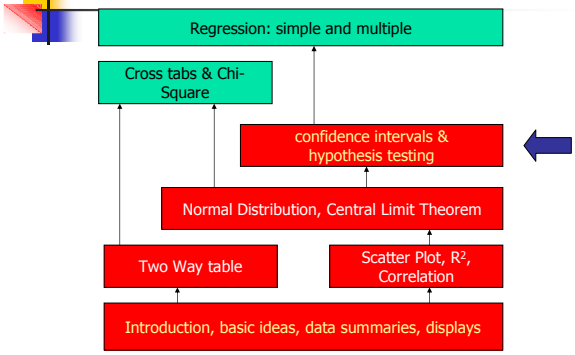
PUB – POS 316
Week 8b

Confidence intervals

Navid Ghaffarzadegan
navidg@gmail.com
Last updated – Jan 1, 10

Course Road Map



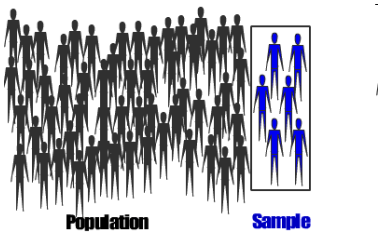
Agenda

- Introduction
- Confidence intervals for sample mean
- Confidence intervals for proportion

Introduction

- Inference
  - We want to draw conclusion from data
  - In a systematic and precise way
  - Use of probability calculations
- Example: we want to know the average monthly expenditure of SUNY Albany students. What should we do?

Population vs Sample



Population	Sample
quantity (count) = N	quantity (count) = n
mean = μ	mean = $\bar{x}$
variance = σ^2	variance = s^2
standard deviation = σ	standard deviation = s

Introduction

- Example: we want to know the average monthly expenditure of SUNY Albany students. What should we do?
- Let's say from a sample of 1000 people we get that the average expenditure is $525/month.
- If I conduct another sample, will I have the same results?
- How confident are we in our result? Margin of error? Interval for the answer?

Confidence intervals for sample mean

Confidence interval: An interval containing the true value of the parameter with some probability.

- Price of a medium size cup of coffee in Albany is:
  - $ 1.85 ±$0.35 (with 95% confidence)
- 1. There is a true value out there (parameter).
- 2. We don't know what that is.
- 3. Our sample study says it is $1.85.
- 4. We expect the results to change less than $0.35 if we conduct this study again and again
- 5. More accurate: We expect the results change less than $0.35 in 95% of similar studies.
- 6. So, we are 95% sure that the average price of coffee (the true parameter), is around $1.5 and $2.2.

Confidence intervals for sample mean

Confidence interval: An interval containing the true value of the parameter with some probability.

- The average monthly expenditure of a UAlbany student is:
  - $ 525 ±$28 (with 80% confidence)
- 1. There is a true value out there (parameter).
- 2. We don't know what that is.
- 3. Our sample study says it is $525.
- 4. We expect the results to change less than $28 if we conduct this study again and again
- 5. More accurate: We expect the results change less than $28 in 80% of similar studies.
- 6. So, we are 80% sure that the average monthly exp. (the true parameter), is around $497 and $553.

Confidence intervals for sample mean

We can state it in different ways:

- The average monthly expenditure of a UAlbany student is $ 525 ±$28 (with 80% confidence)
- The average monthly expenditure of a UAlbany student is $ 525 with $28 margin of error (with 80% confidence)
- The average monthly expenditure of a UAlbany student is $ 525 with 5% margin of error (with 80% confidence)

Or: a statement for 95% confidence. [more common]

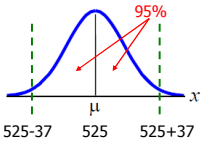
- The average monthly expenditure of a UAlbany student is $ 525 ±$37 (with 95% confidence)
- The average monthly expenditure of a UAlbany student is $ 525 with $37 margin of error (with 95% confidence)

Confidence intervals for sample mean

How to calculate the margin of error:

- The average monthly expenditure of a UAlbany student is $ 525 ±$37 (with 95% confidence)

If we assume a normal distribution for our sampling results:

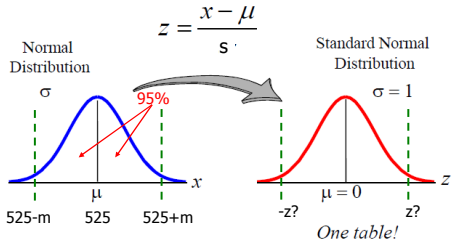


How to find the margin?
±$37

Confidence intervals for sample mean

We have a table for z-score!

$$z = \frac{x-\mu}{s}$$



One table!

$$z = \frac{x-\mu}{s} \Rightarrow x = \mu + z.s \Rightarrow x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

The average monthly expenditure of a UAlbany student is $525 ±$37 (with 95% confidence)

X= $525 ±$37

Do you see how the margin of error was calculated?!
They knew z. Do we know z?
They knew s. Do we know s?

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

Based on a survey we know that the price of a medium size cup of coffee in Albany is $1.85. This survey included 36 coffee shops in Albany. We know that the standard deviation for price of a cup of coffee is $0.6.

Find the margin of error in this report for 95% confidence interval.

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

Based on a survey we know that the UAlbany students study 2.5 hours every day. This study included 100 students. We know that the population standard deviation is 10.

What do you think the true value for hours of study in our population is? with 95% confidence interval.

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

Based on a survey we know that the UAlbany students study 2.5 hours every day. This study included 10,000 students. We know that the population standard deviation is 10.

What do you think the true value for hours of study in our population is? with 95% confidence interval.

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

Based on a survey we know that the UAlbany students study 2.5 hours every day. This study included 100 students. We know that the population standard deviation is 10.

What do you think the true value for hours of study in our population is? with 80% confidence interval.

Compare the previous results.

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

Looking at my last month lunch plan (30 days), I found that I usually pay around $6.25 for my lunch. Let's assume the standard deviation for lunch price is $3.

How much do you think (estimate!) I will spend for my lunch tomorrow. (with 95% confidence interval)

Confidence intervals for sample mean

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$$

The procedure:

1. Write $x = \mu \pm z.s$
2. Write $x = \mu \pm z.\frac{\sigma}{\sqrt{n}}$
3. Find z (you usually know it!), σ, and n

Agenda

- Introduction
- Confidence intervals for sample mean
- Confidence intervals for proportion

Confidence intervals for proportion

Remember: Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\sqrt{\frac{p(1-p)}{n}}$$

The procedure:

1. Write $x = \mu \pm z.s$
2. Write $x = \mu \pm z.\sqrt{\frac{p(1-p)}{n}}$
3. Find z (you usually know it!), p, and n

The only difference is the ***s***

Confidence intervals for proportion

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\sqrt{\frac{p(1-p)}{n}}$$

Based on a survey we know that 52% of UAlbany students drink more that once in a week. This survey included 64 students in UAlbany.

Report your estimation of the true parameter with 95% confidence.

Confidence intervals for proportion

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\sqrt{\frac{p(1-p)}{n}}$$

Based on a survey we know that 52% of UAlbany students drink more that once in a week. This survey included 1000 students in UAlbany.

Report your estimation of the true parameter with 95% confidence.

Confidence intervals for proportion

Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow x = \mu \pm z.\sqrt{\frac{p(1-p)}{n}}$$

Based on a survey we know that 52% of UAlbany students drink more that once in a week. This survey included 1000 students in UAlbany.

Report your estimation of the true parameter with 80% confidence.

Summary

Remember: Only remembering the following equation is enough

$$x = \mu \pm z.s \Rightarrow \begin{cases} x = \mu \pm z.\dfrac{\sigma}{\sqrt{n}} \\ x = \mu \pm z.\sqrt{\dfrac{p(1-p)}{n}} \end{cases}$$

The procedure:

1. Write $x = \mu \pm z.s$
2. Write the proper extension

3. Find z (you usually know it!), σ or p, and n