8.3 CI for μ, σ NOT known (old 8.4)

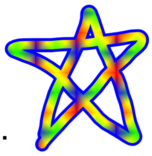
GOALS:

1. Learn the properties of the *student-t* distribution and the *t-curve*.
2. Understand how degrees of freedom, *df,* relates to *t-curves.*
3. Recognize that t-curves approach the SNC as *df* increases.
4. Perform the *t-interval* procedure to find the confidence interval when σ is not known.

Study 8.3,# 109-113,119,123-127(75-91*), 133(~97*) *old 8.4

Statistics Home Page Class Notes Homework

8.1 Estimating the Population Mean, μ

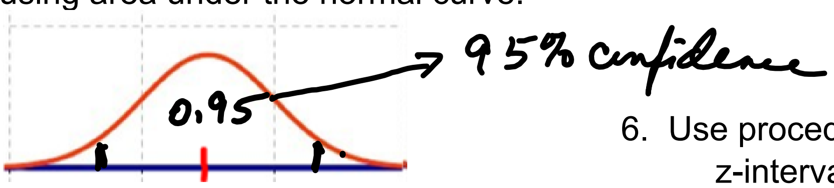


Overview

1. Want to find unknown mean from unknown population.
2. Find a sample mean.
3. Use sample mean to estimate population mean.

 POINT ESTIMATE

4. Know that sample mean is not expected to = population mean.
5. If n.d., can find an interval with the level of confidence wanted using area under the normal curve.



6. Use procedure:
 z-interval
 t-interval

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

z-interval	t-interval
SNC	*t* - curve
standardized $\bar{x}$	studentized $\bar{x}$
$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$	$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
n.d. or large *n*	n.d. or large *n*
σ known	σ **not** known
same for all n	different curve for each *n*

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

test statistic

estimating

1 variable: $\bar{x}$ 2 variables: $\bar{x}$ s

σ

8.3 CI for μ, σ NOT known (old 8.4)

Not covering because σ is rarely known

z-interval

☆ Find a CI for μ ☆

Assumptions: 1. Simple Random Sample
2. nd or large n
3. σ known

Procedure

1. For CL of 1- α find $z_{\alpha/2}$ from Table II or calculator
2. Find CI:

$$\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

3. Intepret:
   a) If n.d., CI precise
   b) If not n.d., n large, CI approximate

Covering this, does not require σ

t-interval

☆ Find a CI for μ: *t* - interval ☆

Assumptions: 1. Simple Random Sample
2. nd or large n
3. σ **unknown**

Procedure

1. For CL of 1- α find $t_{\alpha/2}$ from Table IV or **invT(area, df)** df = n - 1, where n = sample size
2. Find CI:

$$\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

3. Intepret:
   a) If n.d., CI precise
   b) If not n.d., n large, CI approximate

8.3 CI for μ, σ NOT known (old 8.4)

In more realistic situations, σ is NOT known.
Need to use sample *s* instead of σ

But, can NOT use standardized version of $\overline{X}$
ie: no z score

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

Wikipedia: Why is this procedure called the ***student t-test ?***

The *t*-statistic was introduced in 1908 by William Sealy Gosset, a chemist working for the Guinness brewery in Dublin, Ireland. "Student" was his pen name. [1][2][3][4]

Gosset had been hired owing to Claude Guinness's policy of recruiting the best graduates from Oxford and Cambridge to apply biochemistry and statistics to Guinness's industrial processes.[2] Gosset devised the *t*-test as an economical way to monitor the quality of stout. The *t*-test work was submitted to and accepted in the journal *Biometrika* and published in 1908.[5] Company policy at Guinness forbade its chemists from publishing their findings, so Gosset published his statistical work under the pseudonym "Student" (see Student's *t*-distribution for a detailed history of this pseudonym, which is not to be confused with the literal term *student*).

σ

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

SNC	*t* - curve
only 1 variable	2 variables
$\bar{x}$	$\bar{x}$ s

When σ is known, there is only one parameter to estimate, the population *μ*. Therefore, there is only one variable, $\bar{x}$

When σ is NOT known, there are two parameters to estimate: the population *μ,* and the population standard deviation, σ. Therefore, there are 2 variables, $\bar{x}$ and s

SNC	*t* - curve
standardized $\bar{x}$	studentized $\bar{x}$
$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$	$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
n.d. or large *n*	n.d. or large *n*
σ known	σ not known
same for all n	different curve for each *n*

Many t-curves- different curve for each n.
When you change n, change $\sigma_{\bar{x}}$ and change shape.

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

calculator demo **(not required)**:

1. y_1= normalpdf(x,0,1)
2. y_2= tpdf(x,1)
3. use stat to enter 2, 4, 20 into L1
4. y_3= tpdf(x,L1)
5. y_4= tpdf(x,100)

window: -3<x<3, 0<y<0.4

not required

geogebra t-curve demo

df = n-1

For df =1, the t curve is wider and shorter than SNC. What does this tell us about the two curves?(Hint: shape)

As df increases, t- curves --> snc

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

Properties of *t - curve*

1. Total area under curve = 1.
2. Approaches horizontal axis as asymptote
3. Symmetric about 0.
4. As the df increases, ***t***-curves --> SNC

} same as the SNC

8.3 CI for μ, σ NOT known (old 8.4)

Degrees Of Freedom

The number of values in a study that are free to vary.

eg: If have 5 pieces of fruit in a bowl, and you eat one

each day.

On Day 1, you have a choice of	5
Day 2	4
Day 3	3
Day 4	2
Day 5	NO choice

For n = 5, free to choose 4 times:

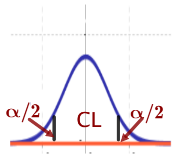
df = n - 1 = 4

Degrees of Freedom refers to the maximum number of logically independent values, which are values that have the **freedom to vary**, in the data sample.

8.3 CI for μ, σ NOT known (old 8.4)

☆ Find a CI for μ: *t* - interval ☆

Assumptions: 1. Simple Random Sample
2. nd or large n
3. σ **un**known



Procedure

1. For CL of 1- α
   df = n - 1, where n = sample size

2. Find CI:

$$\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

3. Intepret:
   a) If n.d., CI precise
   b) If not n.d., n large, CI approximate

We will use the t-interval procedure on the calculator. So, do **not** need to specify the $\mathbf{t}_{\alpha/2}$ value

DO need to specify the $\alpha/2$ value.

Class Notes: Prof. G. Battaly, Westchester Community College, NY

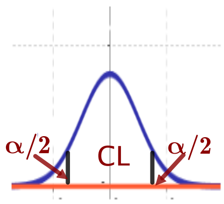
Statistics Home Page Class Notes Homework

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

8.3 CI for μ, σ NOT known (old 8.4)

☆ Find a CI for μ: *t* - interval ☆

Assumptions: 1. Simple Random Sample
2. nd or large n
3. σ **un**known



Procedure

1. For CL of 1- α
   df = n - 1, where n = sample size

2. Find CI:

$$\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Standard Error

3. Intepret:
   a) If n.d., CI precise
   b) If not n.d., n large, CI approximate

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

While some hawk species have been declining in the Northeast, Red-shouldered Hawks (RS) seem to have increased. Data from hawk watch sites for the 10 years from 2014 to 2023 have a mean of 3729.8 RS migrating across the Northeast, with a standard deviation of 1286.43 RS.

Find the 95% Confidence Interval for the mean number of RS hawks that migrate across the Northeast. (Assume random sampling. A NPP of the data shows an approximately straight line.)

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of $\alpha/2$
5. STAT/TESTS/TInterval
6. Result as an interval
   _____ ≤ μ ≤ _____
7. Interpretation

Switch steps 4 and 5 if have data.

☆ Find a CI for μ: *t* - interval ☆

Assumptions: 1. Simple Random Sample
2. nd or large n
3. σ **unknown**

Procedure

1. For CL of 1- α , df = n - 1, where n = sample si
2. Find CI:

$$\bar{x} - t_{\alpha/2}\cdot\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$$

3. Intepret:
   a) If n.d., CI precise
   b) If not n.d., n large, CI approximate

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

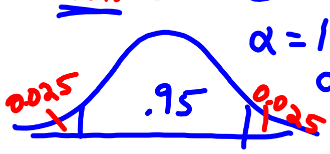
8.3 CI for μ, σ NOT known (old 8.4)

G: n.d., srs., no σ (A: ✓ ✓ ✓)

$\bar{x} = 3729.8$, $s = 1286.43$, $n = 10$

F: CI → t-interval
95% C.I.

CL = 0.95
α = 1 − CL = 1 − 0.95
α = 0.05
α/2 = 0.025



$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$ df = 10 − 1 = 9

$3729.8 \pm t_{0.025}\cdot\frac{1286.43}{\sqrt{10}}$

$2809.5 \le \mu \le 4650.1$

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of $\alpha/2$
5. STAT/TESTS/TInterval
6. Result as an interval
   _____ ≤ μ ≤ _____
7. Interpretation

Switch steps 4 and 5 if have data.

☆ Find a CI for μ: *t* - interval

Assumptions: 1. Simple Random Sample
2. nd or large n
3. σ **unknown**

Procedure

1. For CL of 1- α find $t_{\alpha/2}$ from Table IV
   invT(area, df) df = n - 1, where n = sample size
   Word Problem
2. Find CI:

$$\bar{x} - t_{\alpha/2}\cdot\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$$

3. Intepret:
   a) If n.d., CI precise
   b) If not n.d., n large, CI approximate

Concl: We have 95% confidence that the mean RS hawks migrating over the northeast lies between 2809.5 and 4650.1.

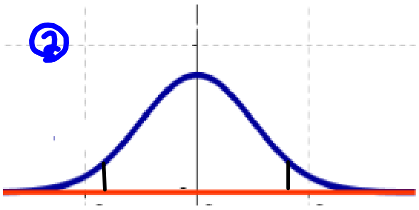
Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

① srs, n large, σ not known

G: srs, $\bar{x}$ = 25, n=36, s=3, CL=95% F: CI

② 

CL =_____

$\alpha = 1 -$ _____ = _____

$\alpha/2 =$ _____

$t_{\alpha/2} = t$ ☐

③ → df = n -1 =_______

Margin of Error

$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

Standard Error

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
   _____ ≤ μ ≤ _____
7. Interpretation

Switch steps 4 and 5 if have data.

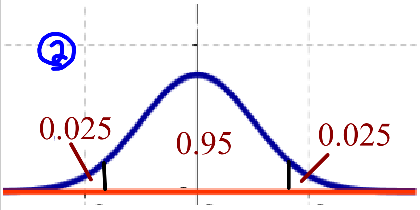
Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

① srs, n large, σ not known

G: srs, $\bar{x}$ = 25, n=36, s=3, CL=95% F: CI

② 

CL = 0.95

$\alpha = 1 - 0.95 = 0.05$

$\alpha/2 = 0.025$

$t_{\alpha/2} = t_{0.025}$

df = n -1 = 35

③ $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

④ $25 \pm t_{0.025} \frac{3}{\sqrt{36}}$

⑤ calculator:
STAT / TESTS
8: TInterval
Inpt: STAT
$\bar{x}$: 25
S_x: 3
n: 36
C-Level: .95
Calculate

Result: (23.985,26.015)

⑥ $23.99 \le \mu \le 26.02$

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
   _____ ≤ μ ≤ _____
7. Interpretation

Switch steps 4 and 5 if have data.

⑦ Conclude: We have 95% confidence that the population mean lies within the interval from 23.99 to 26.02

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

Calculator details

calculator:
STAT / TESTS
8: TInterval
Inpt: STAT <- Could be DATA
$\bar{x}$: 25
S_x: 3
n: 36
C-Level: .95
Calculate

Result: (23.985,26.015)

$\bar{x}$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

8.3 CI for μ, σ NOT known (old 8.4)

G: A random sample of 16 batteries resulted in a mean weight of 55 gm and a standard deviation of 5 gm. If the weight is known to be normally distributed, estimate the population mean with a 99% confidence interval.

G: ____________________________ F: _______

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
_____ ≤ μ ≤ _____
7. Interpretation
Switch steps 4 and 5 if have data.

51.32 ≤ μ ≤ 58.68

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

$\bar{x}$

8.3 CI for μ, σ NOT known (old 8.4)

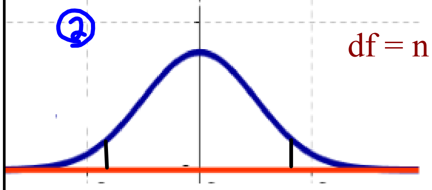
G: A random sample of 16 batteries resulted in a mean weight of 55 gm and a standard deviation of 5 gm. If the weight is known to be normally distributed, estimate the population mean with a 99% confidence interval.

$\overline{x}$

G: srs, nd, $\overline{x}$ = 55, n=16, s=5, CL=99% F: CI

① A: srs, nd, σ not known

②



③

df = n - 1 = ______

CL = _____

α = 1 - _____ = _____

α/2 = _____

$t_{\alpha/2}$ = t ▭

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
_____ ≤ μ ≤ _____
7. Interpretation
Switch steps 4 and 5 if have data.

$$\overline{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

51.32 ≤ μ ≤ 58.68

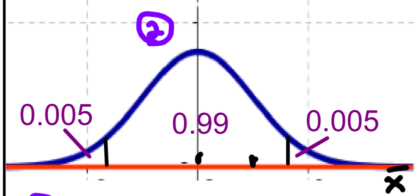
8.3 CI for μ, σ NOT known (old 8.4)

G: A random sample of 16 batteries resulted in a mean weight of 55 gm and a standard deviation of 5 gm. If the weight is known to be normally distributed, estimate the population mean with a 99% confidence interval.

55
60

G: srs, nd, $\overline{x}$ = 55, n=16, s=5, CL=99% F: CI

②



① Assumptions met: 1)srs 2)nd 3)σ not known

CL = 0.99

α = 1 - 0.99 = 0.01

α/2 = 0.005

$t_{\alpha/2}$ = $t_{0.005}$

df = n - 1 = 15

③ $$\overline{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

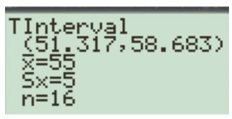
④ $$55 \pm t_{0.005} \frac{5}{\sqrt{16}}$$

⑥ 51.3mg ≤ μ ≤ 58.7mg

⑦ Conclude: Have 99% confidence that the population mean battery weight lies between 51.2 mg and 58.7 mg.

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
_____ ≤ μ ≤ _____
7. Interpretation
Switch steps 4 and 5 if have data.

⑤



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

51.32 ≤ μ ≤ 58.68

p.383 8.136 The following data represent the age (in weeks) at which babies first crawl based on a randomized survey of 12 mothers.

52	30	44	35
47	37	56	26
52	47	52	26

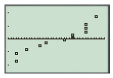
Find the 95% confidence interval for the mean number of weeks to a baby's first crawl.

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
   _____ ≤ μ ≤ _____
7. Interpretation

Switch steps 4 and 5 if have data.

Switch steps 4 and 5 if have data.

Numeric results:
35.1 ≤ μ ≤ 48.9 weeks
x bar=42, s=10.79



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

p.383 8.136 The following data represent the age (in weeks) at which babies first crawl based on a randomized survey of 12 mothers.

52	30	44	35
47	37	56	26
52	47	52	26

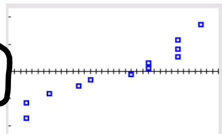
Find the 95% confidence interval for the mean number of weeks to a baby's first crawl.

G: n=12 F: 95% C.I.

~n.d, NPP ✓

∴ A: srs, n.d., no σ

t-interval



$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$42 \pm t_{0.025} \cdot \frac{10.8}{\sqrt{12}}$$

(35.1, 48.9)

35.1 wks ≤ μ ≤ 48.9 weeks

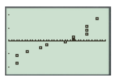
Conclude: Have 95% confidence that popul. mean lies betw. 35.1 and 48.9 weeks.

REQUIRED:
1. Check assumptions
2. **sketch showing both CL and α/2
3. Write formula and df
4. Substitute into equation showing subscript on t as value of α / 2
5. STAT/TESTS/TInterval
6. Result as an interval
   _____ ≤ μ ≤ _____
7. Interpretation

Switch steps 4 and 5 if have data.

Switch steps 4 and 5 if have data.

Numeric results:
35.1 ≤ μ ≤ 48.9 weeks
x bar=42, s=10.79



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework