

Stat 145, Thu 25-Feb-2021 -- Thu 25-Feb-2021
Biostatistics
Spring 2021

Thursday, February 25th 2021

Wk 4, Th
Topic:: Confidence intervals
Read:: Lock5 3.2

Point browser tab at
<https://pad.disroot.org/p/s145-25feb2021-gXX>

Choose XX to match group in which you find yourself:

- XX = 01 for Stob, Rudy, Anderson
- = 02 for Arthur, Ochiagha, Katje
- = 03 for Krikke, Triezenberg, Haveman
- = 04 for Pastoor, Ching, Latvaitis
- = 05 for Johnson, Cheek, Opalewski
- = 06 for Nedd, Bultje, Brink
- = 07 for Tanis, Schneider, Wolf
- = 08 for Aardema, Toldy
- = 09 for Steen, Sytsema, Lemon
- = 10 for Rai, Morren, Wakeman

Exercises to consider:

1.88

2.20 (b)

Please remember to "mark" which pages contain problems/problem parts as per
<http://scofield.site/courses/s143/videos/submittingWorkInGradescope.mp4>

Another video suggestion:

<http://scofield.site/courses/s143/videos/samplingDistsProportionsFirstLook.mp4>

Normal distributions

- Sampling distributions for sample means/proportions tend to look "normal" truer when sample size n is large
- Normal distribution calculator from StatKey
- 68-95-99.7% rule

In particular,

about 95% of values of \hat{p} lie within 2 standard deviations of p
that is, inside $[p - 2 \text{ SE}, p + 2 \text{ SE}]$

about 95% of values of \bar{x} lie within 2 standard deviations of μ
that is, inside $[\mu - 2 \text{ SE}, \mu + 2 \text{ SE}]$

call the amount added/subtracted the "margin of error" for 95% coverage

Q1: Use the Normal distribution calculator app from StatKey to

- (a) Plot a Normal distribution with mean 100 and std dev 20.
- (b) Find rel. freq of values from this distribution between 90 and 110
What is the Z-score of 110?
- (c) Find rel. freq of values from this distribution between 80 and 120
What is the Z-score of 80?
- (d) Find rel. freq of values from this distribution between 60 and 140
What is the Z-score of 140?
- (e) Find rel. freq of values from this distribution between 40 and 160
What is the Z-score of 40?

Q2: Say that a company fills its packages to an avg of 48 Kg with sd 2 Kg.

- (a) Plot a Normal distribution with these parameters
- (b) Find rel. freq of values from this distribution between 47 and 49
What is the Z-score of 49?
- (c) Find rel. freq of values from this distribution between 46 and 50
What is the Z-score of 46?
- (d) Find rel. freq of values from this distribution between 44 and 52
What is the Z-score of 52?
- (e) Find rel. freq of values from this distribution between 42 and 54
What is the Z-score of 42?

Idea of a 95% CI (Centered-interval construction method)

- get an estimate for population parameter
 - in case of μ (quant var), use \bar{x}
 - in case of p (binary categorical var), use \hat{p}
- get an estimate for margin of error (ME) for 95% coverage
most likely this involves estimating SE and doubling it
- construct centered interval

Q3: ~~ME~~ for 95% coverage is approximately ~~2 SE~~.

How should you get ~~SE for~~ ME for

99.7% coverage?

68% coverage?

100% coverage?

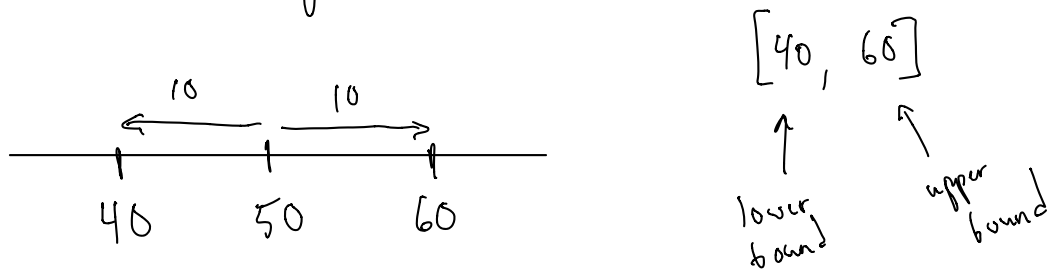
$2 * SE$

90% coverage?

① 68 - 95 - 99.7% rule for normal dists.

- about 68% of values in a normal dist. lie within 1 s.d.
- " 95% " " " " " " " 2 s.d.
- " 99.7% " " " " " " " 3 s.d.

② Centered intervals: an interval centered at 50 w/
margin of error 10 is

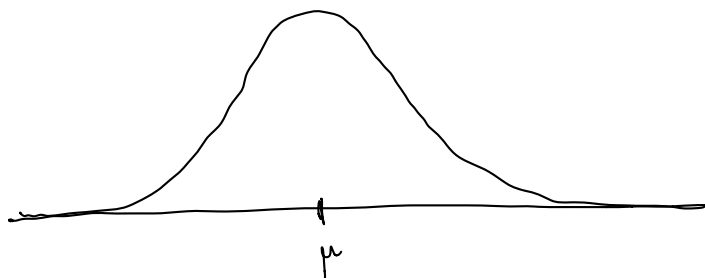


Join these ideas in discussing sampling dists.

Recall:

Sampling dist. for sample mean \bar{x}

- looks symmetric, bell-shaped (normal?)



By the 95% part. of the 68-95-99.7% rule,
about 95% of \bar{x} -values should lie within
 $2 \times (\text{SE Errors})$ of μ

In practice:

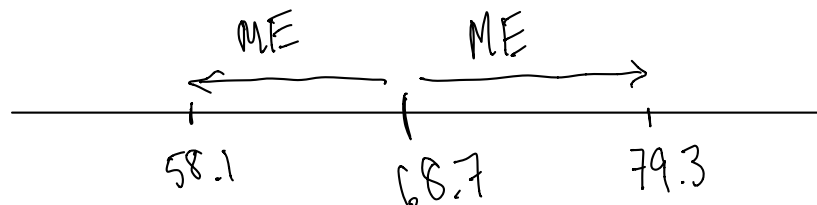
- Don't know μ
- Know \bar{x} from a sample collected
best point est. for μ : \bar{x}

A 95% confidence interval for μ : build a centered
interval with center \bar{x} , and margin of error
 $2 \cdot SE_{\bar{x}}$.

Ex.1 Say we want to estimate avg. height of a Calvin
female student. Suppose from a sample of 50
female students, get $\bar{x} = 68.7$ and
our $SE_{\bar{x}}$ is estimated to be 5.3.

Then a 95% CI

$$ME = 2(5.3) = 10.6$$



Centered - interval approach for confidence intervals

- always take the center to be the point est. from sample

\bar{x} when μ is desired

\hat{p} when p is desired

• If 95% coverage $ME = 2 * SE$

99.7% " $ME = 3 * SE$

68% " $ME = SE$

100% " $ME = +\infty$ (impossible)