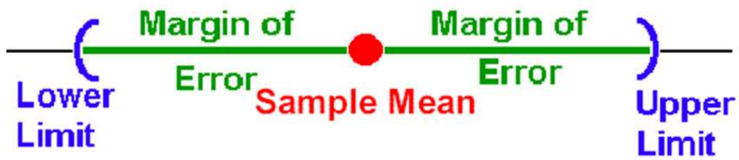
Confidence Intervals

Confidence Intervals

- One sample mean→ one *point estimate*: unlikely to be exactly correct
- To be correct with more certainty → use *interval estimates*

Anatomy of a Confidence Interval

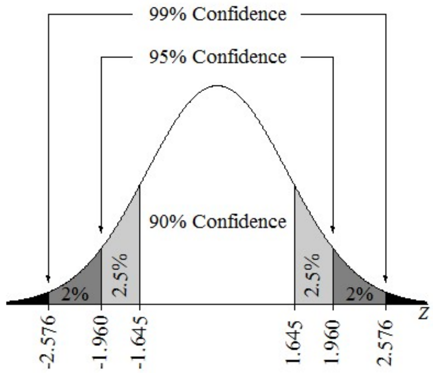


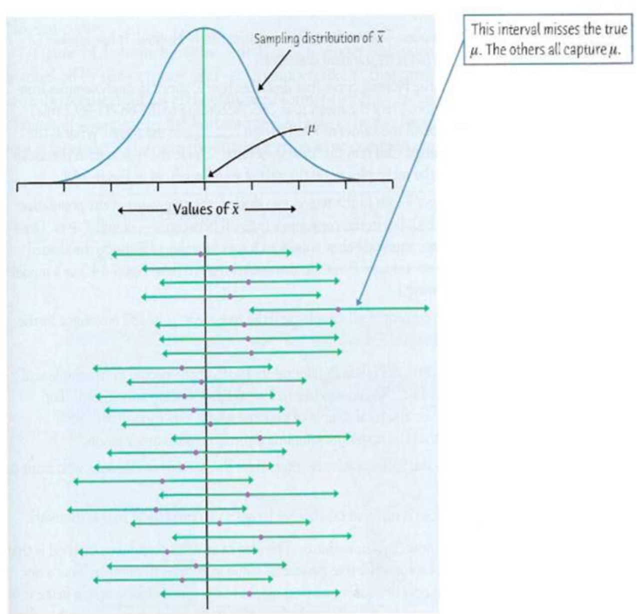
Level of Confidence

- Confidence interval → associated with a certain level of confidence
- 95% confidence interval: We are 95% sure that mean is in that interval, but there is a 5% chance that we are making an error
- We use alpha (α) to refer to this probability of error
- To be 100% sure, we would need an infinitely large confidence interval -- in theory, all estimates are possible, but some are extremely unlikely

Using CLT for Confidence Intervals

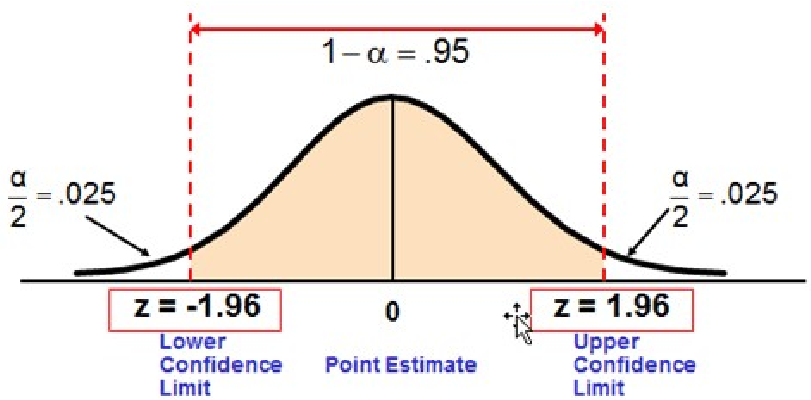
- Based on CLT, sampling distribution is normal
- For 95% of all samples, a sample mean will be within 2 SE from the population mean

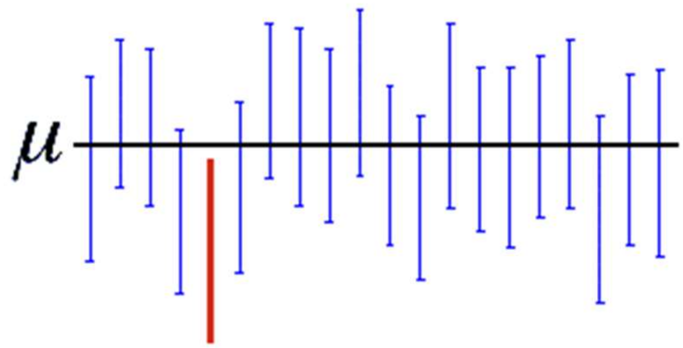




Typical Confidence Levels

- 95% confidence interval: z-score=+/-1.96 (within +/- 2 SE from the mean)
- 99% confidence interval: z-score=+/-2.576
- 90% confidence interval: z-score=+/-1.645





A 95% confidence interval indicates that 19 out of 20 samples (95%) from the same population will produce confidence intervals that contain the population parameter.

https://seeing-theory.brown.edu/frequentist-inference/index.html

https://rpsychologist.com/d3/CI/

Confidence Intervals for Means of Large Samples (n≥30)

1. Pick the desired level of confidence (e.g., 90%, 95%, 99%)
2. Divide the confidence level by 2, use that number as area value to enter Table B1 → find the corresponding z score
3. Calculate the two confidence interval limits, X_1 and X_2: $\overline{X} \pm z^*\sigma_{\bar{X}}$, where $\sigma_{\bar{X}} = {}^{s}/_{\sqrt{n}}$
4. Set the interval: Probability($X_1 \leq \mu \leq X_2$)=.95 (or .99, or .90, depending on the selected confidence level)

Example

Problem: Assume you want to have a 95% confidence interval for the mean. You take a random sample of n = 49, and get a sample mean $\overline{X}$ = 110 and standard deviation s = 14

1. Set level of confidence at CL = 95%

2. We want the middle 95% of the area under the curve. Half of it will be in the positive half of the curve → 95/2 = 47.50. Enter Table B1 with 47.50 → z=1.96

3. Calculate the two confidence interval limits: $\overline{X} \pm z\sigma_{\overline{X}}$ = 110 ± 1.96* (14/√49) = 110 ± 3.92 → X_1=106.08 and X_2=113.92.

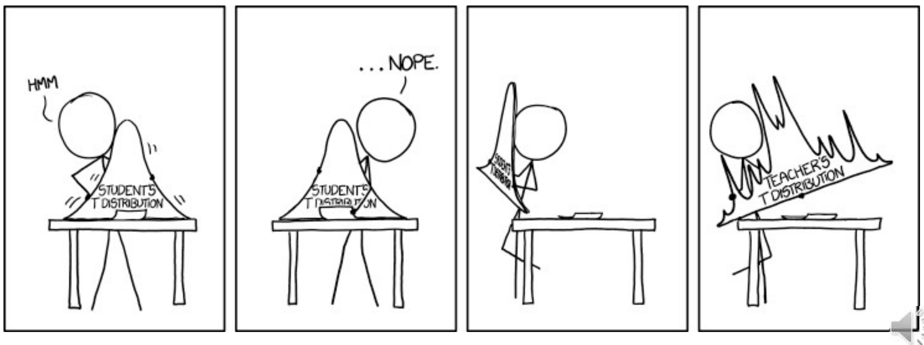
4. Answer: Probability($106.08 \leq \mu \leq 113.92$) = .95

In Words

- Probability($106.08 \leq \mu \leq 113.92$) = .95
- Based on our sample data, we are 95% confident that the “true” mean of this variable (i.e., population mean) is between 106.08 and 113.92.
- Alternatively: There is a 95% chance that the population mean is between 106.08 and 113.92.

Confidence Intervals: Small vs. Large Samples

- For large samples: based on normal curve
- For small samples (n < 30) – based on Student's t-distribution



Student's t Distribution

- DF = degrees of freedom = n-1

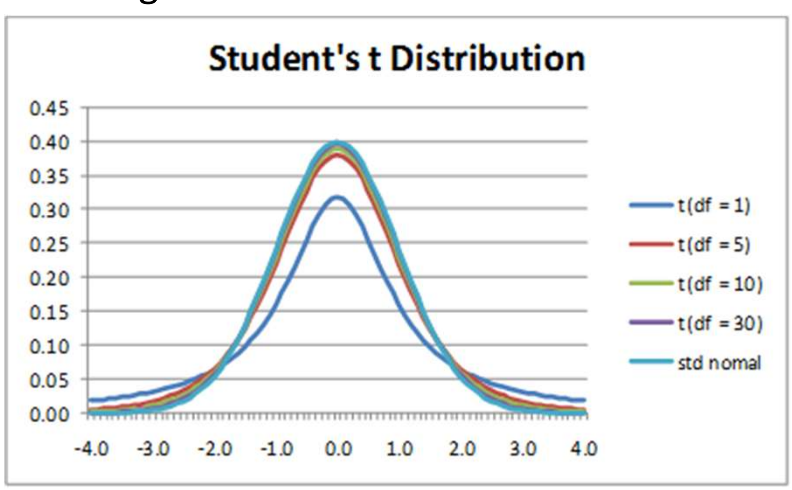


Table B2 (in Salkind)

- For confidence interval: only use the two-tailed test part
- Degrees of freedom, df = n – 1
- Use probability of error (alpha) rather than confidence level itself
- To get alpha: use the confidence level as proportion rather than percentage (e.g., .95 rather than 95) and subtract it from 1 (1 - .95 = .05)
- Use the column for this alpha and the row with appropriate df → t-value

Confidence Intervals for Means of Small Samples (n<30)

1. Pick the desired level of confidence (e.g., 90%, 95%, 99%)
2. Using the two-tailed test part of Table B2, find the t-value for the alpha level corresponding to your confidence level (1-CL/100) and your degrees of freedom (df=n-1)
3. Calculate the two confidence interval limits, X_1 and X_2: $\overline{X} \pm t^*\sigma_{\bar{X}}$, where $\sigma_{\bar{X}} = s/\sqrt{n}$
4. Set the interval: Probability($X_1 \leq \mu \leq X_2$)=.95 (or .99, or .90, depending on the selected confidence level)

Example

Problem: Assume you want to get a 95% confidence interval for the mean. You take a random sample of n=25 and get a sample standard deviation of s=10 and a sample mean of $\overline{X}$ =70.

1. Set CL at .95.

2. Obtain t. In Table B2, we look at the column in the two-tailed part of the table that corresponds to .05, since 1-.95=.05. We read down the column marked as .05 to the row with the degrees of freedom equal to 25-1 = 24. The entry is t=2.064.

3. Calculate the two confidence interval limits: $\overline{X} \pm t^{*}s/\sqrt{n}$
= 70 ± 2.064*(10/√25) = 70 ± 4.128 → X1=65.872 and X2=74.128.

4. Answer: Probability (65.872 ≤ μ ≤ 74.128) = .95

Confidence Intervals in Stata

`mean age`

```
Mean estimation                        Number of obs    =    1969
--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
         age |    48.1935   .3985971      47.41178    48.97522
--------------------------------------------------------------
```

Answer: Probability (47.41178 ≤ μ ≤ 48.97522) = .95

Based on our sample data, we are 95% confident that the “true” mean of age among Americans is between 47.41 and 48.98 years old.

Changing Confidence Level

`mean age, level(99)`

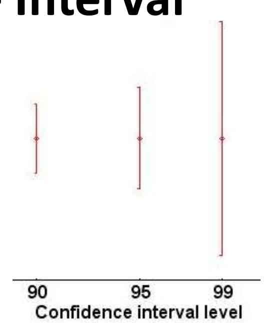
```
Mean estimation                          Number of obs     =     1969

--------------------------------------------------------------
             |       Mean   Std. Err.     [99% Conf. Interval]
-------------+------------------------------------------------
         age |    48.1935   .3985971      47.16578    49.22121
--------------------------------------------------------------
```

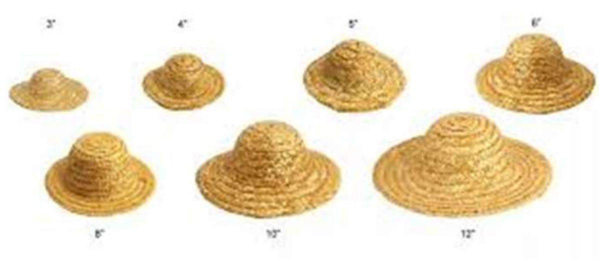
Answer: Probability $(47.16578 \leq \mu \leq 49.22121) = .99$

Based on our sample data, we are 99% confident that the "true" mean of age among Americans is between 47.17 and 49.22 years old.

What Makes a Confidence Interval Wider?

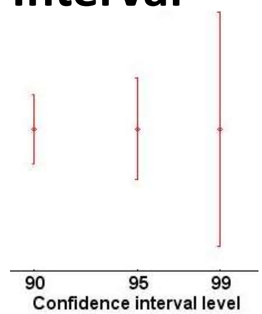
- Higher confidence level selected

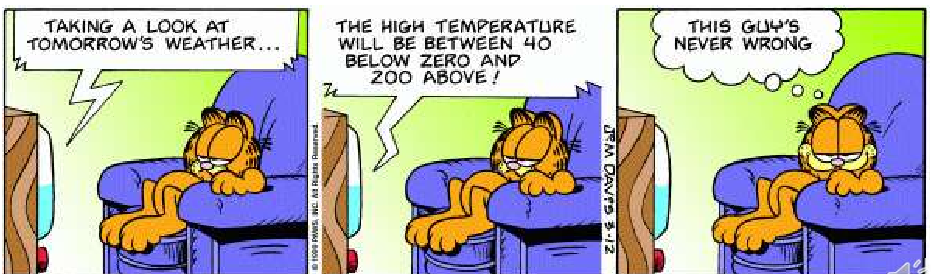




What Makes a Confidence Interval Wider?

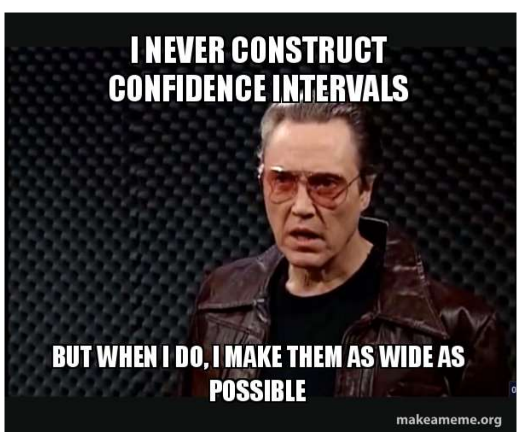
- Higher confidence level selected

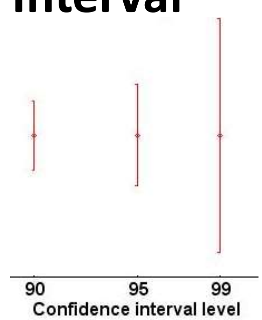




What Makes a Confidence Interval Wider?

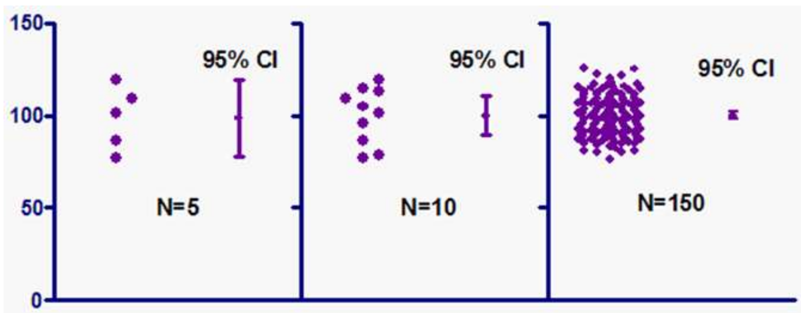
- Higher confidence level selected

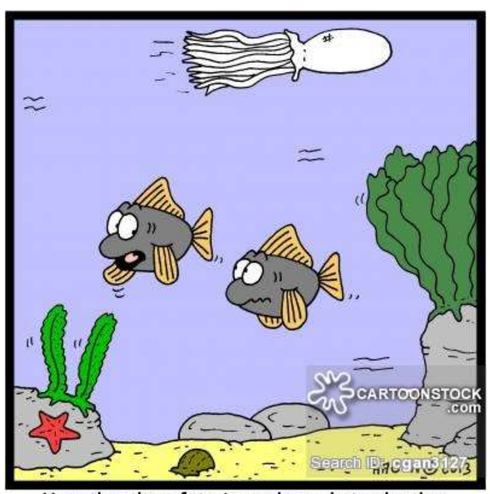




What Makes a Confidence Interval Wider?

- Higher confidence level selected
- Smaller sample size





Yes, there's safety in numbers, but only when the number is large: Right now, we have only a 50% chance of surviving a deadly attack...

What Makes a Confidence Interval Wider?

- Higher confidence level selected
- Smaller sample size
- Less homogenous population

Homogenous population:

Heterogenous (diverse) population:

What Makes a Confidence Interval Wider?

- Higher confidence level selected (larger z)
- Smaller sample size (larger n)
- Less homogenous population (larger s)

$$CI = \bar{x} \pm z\frac{s}{\sqrt{n}}$$

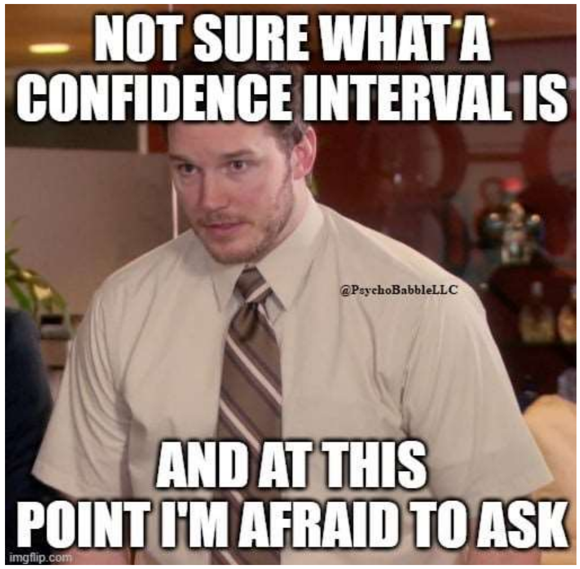
CI = confidence interval

$\bar{x}$ = sample mean

z = confidence level value

s = sample standard deviation

n = sample size


NOT SURE WHAT A
CONFIDENCE INTERVAL IS
@PsychoBabbleLLC
AND AT THIS
POINT I'M AFRAID TO ASK
imgflip.com