18

STATISTICAL SIGNIFICANCE

(Confidence Interval)

KSU COLLEGE OF MEDICINE
2019 - 2020

ACKNOWLEDGMENTS

TEAM MEMBERS

AHMED ALANZI

REHAB ALANZI

REVIEWER

ASEEL BADUKHON

Special thanks to SARAH ALENEZY & 436 TEAM

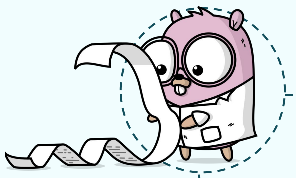
TABLE OF **CONTENTS**

LECTURE **OBJECTIVES**

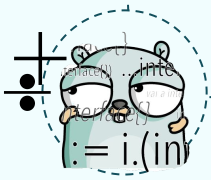
By the end of this lecture, I am able to understand:

- Able to understand the concept of confidence intervals.
- Able to apply the concept of statistical significance using confidence intervals in analyzing the data.
- Able to interpret the concept of 95% confidence intervals in making valid conclusions.

OVERVIEW

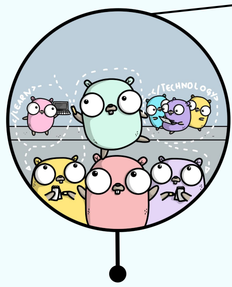


Hypothesis Testing ---- P-Value

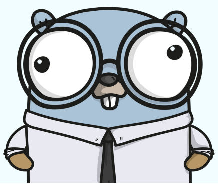


Estimation ---- Confidence Interval (CI)

Sample is assumed to be representative to the population. In research: measurement are always done in the sample, the results will be applied to population.



- An observed value drawn from the sample is called a statistic
- The corresponding value in population is called a parameter
- We measure, analyze, etc statistics and translate them as parameters



There is no need to do the whole population in the study. We depend on a sample and from the sample calculations we estimate population parameters

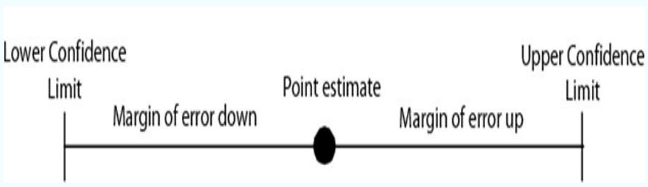
Statics	Sample
Parameter	Population

CONFIDENCE INTERVAL

Estimation

Two forms of estimation:

- **Point estimation** = single value, e.g., (mean, proportion, difference of two means, difference of two proportions, OR, RR etc.,)
- **Interval estimation** = range of values ⇒ confidence interval (CI). A confidence interval consists of:
- E.g. Class mean marks, number of students who do daily exercise, prevalence of diabetes and incidents of hypertension > all are one value so considered point estimates



interval estimation : contant 3 thing, margin error down ,point estimate,margin of error up

CONFIDENCE INTERVAL

- P values give no indication about the clinical importance of the observed association
- Relying on information from a sample will always lead to some level of uncertainty.
- Confidence interval is a range of values that tries to quantify this uncertainty:
  For example , 95% CI means that under repeated sampling 95% of CIs would contain the true population parameter
- P value will tell you whether the results are statically significant or not, whether the risk factor is associated with outcome or not but it won't tell you the direction and it won't tell you the magnitude

COMPUTING CONFIDENCE INTERVALS (CI)

General Formula

(Sample statistic) ± [(confidence level) X (measure of how high the sampling variability is)]

- Sample statistic: observed magnitude of effect or association (e.g., odds ratio, risk ratio, single mean, single proportion, difference in two means, difference in two proportions, correlation, regression coefficient, etc.,) anything that calculated from the sample
- Confidence level: varies – 90%, 95%, 99%. For example, to construct a 95% CI, $Z_{a/2}$ =1.96 always we use 95%. it means how much error will happen when we take different samples
- Sampling variability: Standard error (S.E.) of the estimate is a measure of variability

Don't get confused with the terms of **STANDARD DEVIATION** (Variability within a sample) and **STANDARD ERROR** (Variability among samples (more than one group))

Example:

Data: X = {6, 10, 5, 4, 9, 8}; N = 6

Mean:

$$\bar{x} = \frac{\Sigma x}{N} = \frac{42}{6} = 7$$

Variance:

$$s^2 = \frac{\sum(\bar{x} - x)^2}{N} = \frac{28}{6} = 4.67$$

Standard Deviation:

$$s = \sqrt{s^2} = \sqrt{4.67} = 2 \cdot 16$$

on average the 6 value are deviating from the mean by 2.16

x	$x - \bar{x}$	$(x - \bar{x})^2$
1	-1	6
9	3	10
4	-2	5
9	-3	4
4	2	9
1	1	8
Total: 28		Total:42

COMPUTING CONFIDENCE INTERVALS (CI)

Statistical Inference is based on Sampling Variability:

Sample Statics	Sampling Variability	Standard Error
We summarize a sample into one number; e.g., could be a mean, a difference in means or proportions, an odds ratio, or a correlation coefficient .what you calculated E.g.: average blood pressure of a sample of 50 Saudi men. E.g.: the difference in average blood pressure between a sample of 50 men and a sample of 50 women.	If we could repeat an experiment many, many times on different samples with the same number of subjects, the resultant sample statistic would not always be the same (because of chance!).: measure of sampling variability is nothing but standard error.	A measure of the sampling variability among different samples

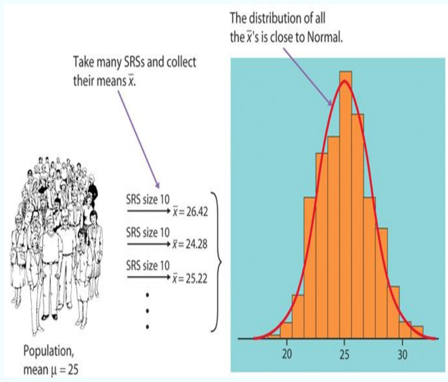
Standard error of the mean:

- Standard error of the mean (sem):

$$s_{\bar{x}} = sem = \frac{s}{\sqrt{n}}$$

Comments:

- n = sample size
- Even for large s, if n is large, we can get good precision for sem
- Always (SE) smaller than standard deviation (s)



COMPUTING CONFIDENCE INTERVALS (CI)

Example

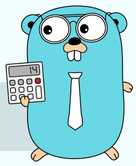
In a representative sample of 100 observations of heights of men, drawn at random from a large population, suppose the sample mean is found to be 175 cm (sd=10cm) .
- Can we make any statements about the population mean ?

Ans:

- We cannot say that population mean is 175 cm because we are uncertain as to how much sampling fluctuation has occurred. Because if you take other 100 sample you'll get different value
- What we do instead is to determine a range of possible values for the population mean, with 95% degree of confidence.
- This range is called the 95% confidence interval and can be an important adjuvant to a significance test.

In the example, n =100 ,sample mean = 175, S.D., =10, and the S. Error =10/√100= 1

- Using the general format of confidence interval* :
  Therefore, the 95% confidence interval is, 175 ± 1.96 x 1 = 173 to 177
  1.96 is Z value if you repeat on another sample you won't get 175 but it'll be in this interval 173 to 177
- That is, if numerous random sample of size 100 are drawn and the 95% confidence interval is computed for each sample, the population mean will be within the computed intervals in 95% of the instances.



*Statistic ± confidence factor x Standard Error of statistic

Confidence interval

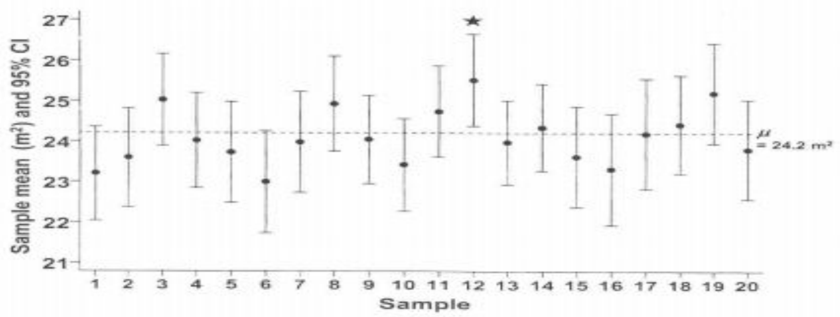


Fig. 6.2 Mean sprayable areas, with 95% confidence intervals, from 20 samples of 100 houses in a rural area. The star indicates that the CI does not contain the population mean

This is called "forest plot":
1 line doesn't cross population mean
19 lines cross = 95%
. So this is the concept of confidence interval here we repeated 20 times and 19 of them contain population value only one time doesn't contain.

Horizontal dots line = population mean
Black dots = sample mean
Vertical line = CI

COMPUTING CONFIDENCE INTERVALS (CI)

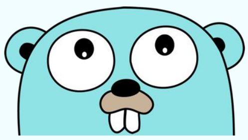
Confidence interval

- The previous picture shows 20 confidence intervals for μ
- Each 95% confidence interval has fixed endpoints where μ might be in between (or not).
- There is no probability of such an event! It's by Repetition

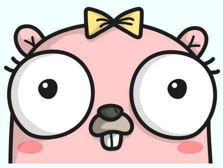
- Suppose $\alpha = 0.05$, we cannot say: "with probability 0.95 the parameter μ lies in the confidence interval."
- We only know that by repetition, 95% of the intervals will contain the true population parameter (μ)
- In 5 % of the cases however it doesn't. And unfortunately we don't know in which of the cases this happens
- That's why we say: with confidence level $100(1 - \alpha)$ % μ lies in the confidence interval."

Different Interpretations of the 95% confidence interval:

We are 95% sure that the TRUE parameter value is in the 95% confidence interval



If we repeated the experiment many many times, 95% of the time the TRUE parameter value would be in the interval



Most commonly used CI:

90% CI **90%** corresponds to α `0.10`

95% CI **95%** corresponds to α `0.05`

99% CI **99%** corresponds to α `0.01`

NOTE:

- p value: only for analytical studies when you compare 2 or more groups
- CI: for descriptive and analytical studies

How to calculate CI:

General Formula: $CI = p \pm Z\alpha \times SE$
p = point of estimate, a value drawn from sample (a statistic)
• $Z\alpha$ = standard normal deviate for α, if $\alpha = 0.05 \rightarrow Z\alpha = 1.96$ (~ 95% CI).
SE=standard error

EXAMPLES

Example 1

100 KKUH students → 60 do daily exercise (p=0.6)
What is the proportion of students do daily exercise in the KSU ?

$$SE(p)CI = \sqrt{\frac{pq}{n}}$$

$$95\%CI = 0 \cdot 6 \pm 1.96\sqrt{\frac{0.6x0.4}{100}}$$

$$0.6 \pm 1.96x0.5\backslash 10$$

$$0.6 \pm 0.1 = 0.5; 0.7$$

To get parameter value of KSU students I've to calculate CI
P: who do exercise = 0.6
q: who don't do exercise = 0.4

In this example we don't compare anything b\c it's descriptive

Example 2

CI of the mean:

100 newborn babies, mean BW = 3000 (SD = 400) grams,

what is 95% CI? 95% CI = x ± 1.96 (SEM)

$$SEM = SD/\sqrt{n}$$

$$95\%CI = 3000 \pm 1.96\left(\frac{400}{\sqrt{100}}\right)$$

$$= 3000 \pm 80 = (3000 - 80);(3000 + 80)$$

$$= 2920; 3080$$

Example 3

CI of difference between proportions (p1-p2): Comparing two groups A and B

50 patients with drug A, 30 cured (p1=0.6)

50 patients with drug B, 40 cured (p2=0.8) better than A

$$95\%CI(p_1 - p_2) = (p_1 - p_2) \pm 1.96xSE(p_1 - p_2)$$

$$SE(p_1 - p_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

$$= \sqrt{\frac{(0.6 \times 0.4)}{50} + \frac{(0.8 \times 0.2)}{50}} = \sqrt{0.008} = 0.09$$

$$\Rightarrow 95\%CI(p_1 - p_2) = [0.2 - (0.09 * 1.96)];[0.2 + (0.09 * 1.96)]$$

$$= 0.024, 0.3764 = 2.4\% to 37.6\%$$

Is the difference of 20% statically significant ? The null hypothesis is 0 so are they getting 0 in this example? Minimum is 2.4 and maximum is 37.6 so at any case we won't get 0 so this study is statically significant. Interpretation: if someone do this type of study 100 times then 95 times will be between 2.4% to 37.6% Worst scenario they'll get 2.4 not zero.

It doesn't include the null hypothesis > 0

EXAMPLES

Example 4

CI for difference between 2 means:

Mean systolic BP:

- 50 smokers = 146.4 (SD 18.5) mmHg
- 50 non-smokers = 140.4 (SD 16.8) mmHg
  - x_1-x_2 = 6.0 mmHg
  - 95% CI(x_1-x_2) = (x_1-x_2) ± 1.96 x SE (x_1-x_2)
  - SE(x_1-x_2) = S x (1/n_1 + 1/n_2)

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}}$$

$$s = \sqrt{\frac{(49 \times 18.6) + 49 \times 16.2}{98}} = 17.7$$

$$SE(x_1 - x_2) = 17.7 \times \sqrt{\left(\frac{1}{50} + \frac{1}{50}\right)} = 3.53$$

$$95\%CI = 6.0 \pm (1.96X3.53) = -1.0;13.0$$

Is this study (comparing between smokers and non-smokers) statically significant? No Because it includes 0. So if someone repeat he may get -1 , he may get 0 or he may get 13 so 0 is included so it's not statically significant.

CONFIDENCE INTERVAL

Other Commonly supplied CI

- Relative Risk (RR)
- Odds Ratio (OR)
- Sensitivity Specificity (SE SP)
- Likelihood Ratio (LR)
- Relative Risk Reduction (RRR)
- Number Needed to Treat (NNT)

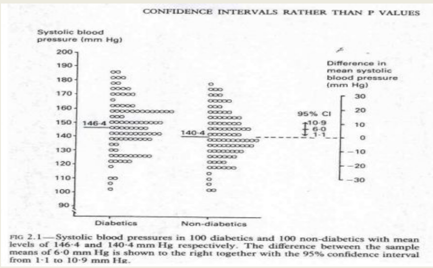
CONFIDENCE INTERVAL

CHARACTERISTICS OF CI'S

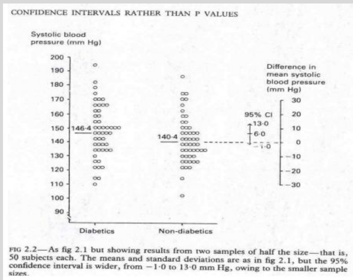
- The (im) precision (Reliability, repeatability, reproducibility) of the estimate is indicated by the width of the confidence interval.
- The wider the interval the less precision

The width of C.I. Depends on:

Sample size	Variability	Degree of confidence



Normal distribution bell shaped curve.
It's statistically significant because it includes 0
Upper limit= 10.9,
Difference= 6, Lower limit = 1.1



By reducing the sample size to half the bell shape is gone, high variability even though the accurate mean difference is same, statistical significance disappeared and the CI increase width
n= 50 So when we decrease the sample size the width will increase

EFFECT OF VARIABILITY

- **Properties of error;**

1. Error increases with smaller sample size For any confidence level, large samples reduce the margin of error
2. Error increases with larger standard Deviation As variation among the individuals in the population increases, so does the error of our estimate
3. Error increases with larger z values Tradeoff between confidence level and margin of error

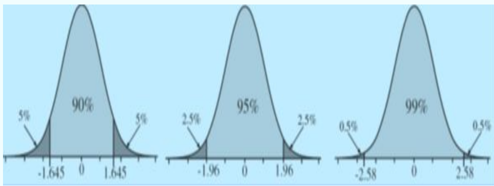
Not only 95%....

- 90% confidence interval: NARROWER than 95% $\bar{x} \pm 1.65 sem$
- 99% confidence interval: WIDER than 95% $\bar{x} \pm 2.58 sem$

CONFIDENCE INTERVAL

Common Levels of Confidence

Confidence level $1-\alpha$	Alpha level α	Z value $z_{1-(\alpha/2)}$
.90	.10	1.645
.95	.05	1.960
.99	.01	2.576



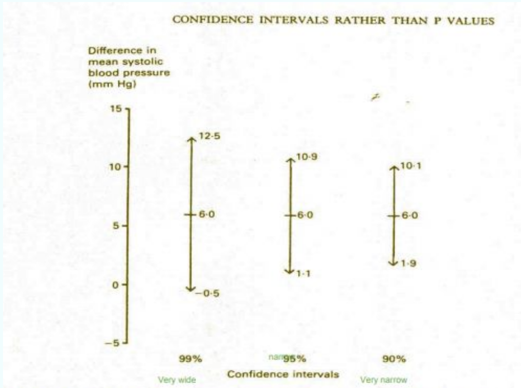


FIG 2.3—Confidence intervals associated with differing degrees of "confidence" using the same data as in fig 2.1.

Application of confidence intervals

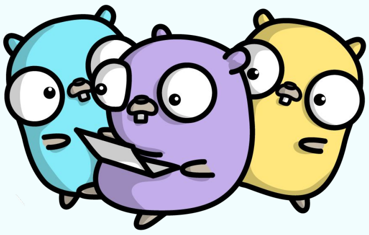
Example: The following finding of non-significance in a clinical trial on 178 patients.

- Chi-square value = 1.74 (p > 0.1) (non –significant)

i.e. there is no difference in efficacy between the two treatments

- The observed difference is: 75% - 66% = 9% and the 95% confidence interval for the difference is: - 4% to 22%
- This indicates that compared to treatment B, treatment A has, at best an appreciable advantage (22%) and at worst , a slight disadvantage (- 4%).
- This inference is more informative than just saying that the difference is non significant.

Treatment	Success	Failure	Total
A	76 (75%)	25	101
B	51(66%)	26	77
Total	127	51	178



CONFIDENCE INTERVAL

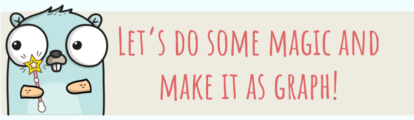
Interpretation of Confidence intervals

Width of the confidence interval (CI):

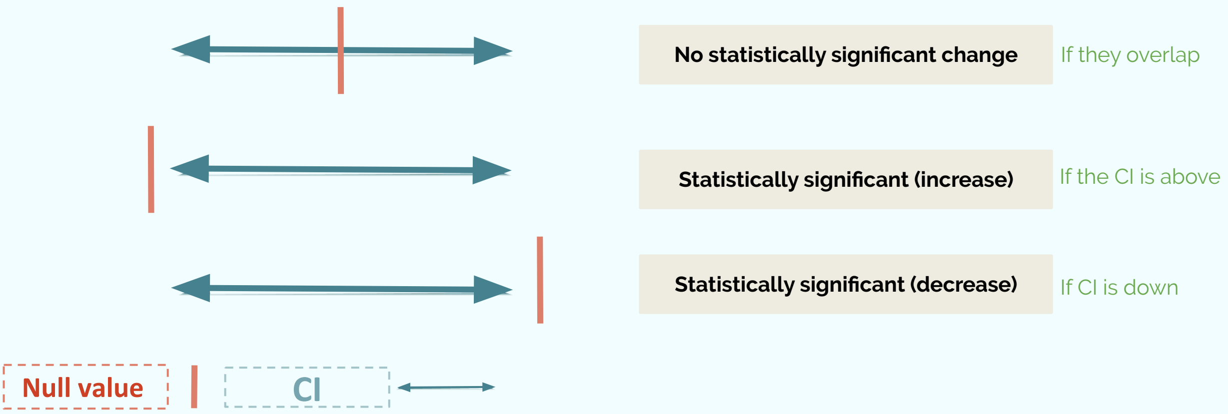
- A narrow CI implies high precision
- A wide CI implies poor precision (usually due to inadequate sample size)

Does the interval contain a value that implies no change or no effect or no association?

- CI for a difference between two means: Does the interval include 0 (zero)? no significant difference
- CI for a ratio (e.g, OR, RR): Does the interval include 1? no association



Interpretation of Confidence intervals Important



Duality between P-values and CIs:

- If a 95% CI includes the null effect, the P-value is >0.05 (and we would fail to reject the null hypothesis)
- If the 95% CI excludes the null effect, the P-value is <0.05 (and we would reject the null hypothesis)

Comparison of p values and confidence interval:

P-Value	Confidence interval (CI)
Gives you the probability that the result is merely caused by chance or not by chance, it does not give the magnitude and direction of the difference	Indicates estimate of value in the population given one result in the sample, it gives the magnitude and direction of the difference
Answers the question: "Is there a statistically significant difference between the two treatments?" (or two groups)	The **point estimate and its confidence interval** answers the question… "What is the size of that treatment difference?", and "How precisely did this trial determine or estimate the treatment difference?"

EXAMPLES

Interpretation of Confidence intervals

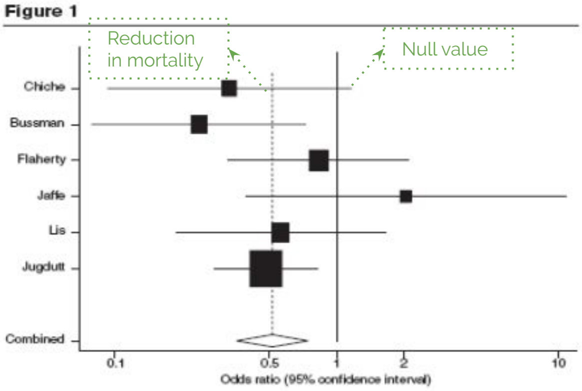
Trial	Number dead / randomized: Intravenous nitrate	Number dead / randomized: Control	Risk Ratio	95% C.I.	P value
Chiche	3/50	8/45	0.33	(0.09,1.13)	0.08
Wide interval: suggests reduction in mortality of 91% and an increase of 13%					
Flaherty	11/56	11/48	0.83	(0.33,2.12)	0.70
Jaffe	4/57	2/57	2.04	(0.39,10.71)	0.40
Reduction in mortality as little as 18%, but little evidence to suggest that IV nitrate is harmful					
Jugdutt	24/154	44/156	0.48	(0.28, 0.82)	0.007

1-0.09=0.91=91%

Table adapted from Whitley and Ball. Critical Care; 6(3):222-225, 2002

Non-statically 1-0.09=0.91=91%
significance because RR less than 1
-Remember: in RR we look for 1 not 0

Figure 1



Individual and combined odds ratios and 95% confidence intervals for six intravenous nitrate trials.

EXAMPLE:

Which of the following odds ratios for the relationship between various risk factors and heart disease are statistically significant at the .05-significance level? Which are likely to be clinically significant?

A. Odds ratio for every 1-year increase in age: 1.10 (95% CI: 1.01—1.19)

B. Odds ratio for regular exercise (yes vs. no): 0.50 (95% CI: 0.30—0.82)

C. Odds ratio for high blood pressure (high vs. normal): 3.0 (95% CI: 0.90—5.30)

Interpretation of odds ratio OR: the odds of having a heart disease is 3 times more in patients who does not maintain their blood pressure in comparison to those who do maintain it.

D. Odds ratio for every 50-pound increase in weight: 1.05 (95% CI: 1.01—1.20)

Statistically significant?	Clinically significant?
✓	✓
✓	✓ Can you tell your patients to do regular exercise
It includes 1	✓
✓	

Summary of key points

- **A P-value** is a probability of obtaining an effect as large as or larger than the observed effect, assuming null hypothesis is true
  - Provides a measure of strength of evidence against the Ho
  - Does not provide information on magnitude of the effect
  - Affected by sample size and magnitude of effect: interpret with caution!
- **Confidence interval quantifies:**
  - How confident are we about the true value in the source population
  - Better precision with large sample size
  - Much more informative than P-value
- **Keep in mind clinical importance when interpreting statistical significance!**