Announcements

Resolved some conflicts between the website and my slides

Don’t know when I can have my UMD account… 🤷‍♀️

Piazza/something/Slack by the end of this week

Clarification: questions (8%)

At least 1 question per required paper

You can swap with an optional paper

Some questions in a week should be non-trivial

Clarification: presentation (8%)

~2 presentations; lemme know if you are auditing

if your paper is lengthy: 1

you can choose to present an optional paper

Note: sometimes I need to give a lecture; so I have to **push back** some presentations if there are three signups that week (e.g., 10/02)

Assignment Q.1

I finalized the wording

And decided to give you a pilot dataset

Lemme know if you have questions

Questions?

Readings & Data Vis

improvised presentation

visualize data of ACL papers

Readings

Top HCI conference

Rethinking Statistical Analysis Methods for CHI

In my very first first-author paper, a reviewer cited this when criticizing my interpretation of results

Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations

In my second year of phd, I found this paper when searching for CI materials; again in 2023 in a seminar class

Rethinking Statistical Analysis Methods for CHI

misinterpretation of p values;
prepare you for confidence intervals

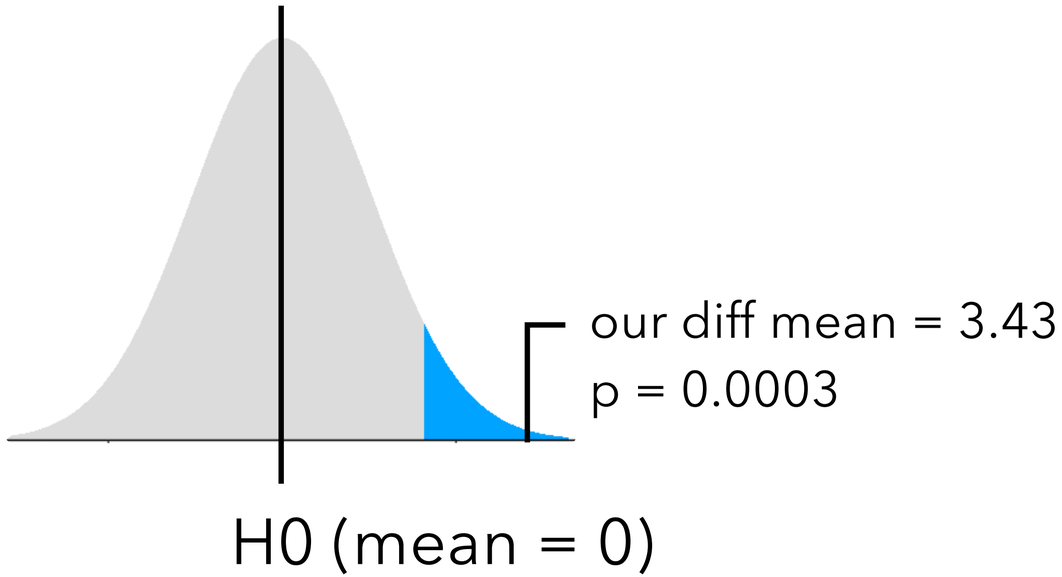
Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations

True or false game

P-values

P-values

The probability of you get the data, given the assumption that H_0 is true



True or false

P-values: true or false

A significant test result ($P < 0.05$) means that the test hypothesis is false or should be rejected.

false

A small *P* value simply flags the data as being unusual if all the assumptions used are correct. $P \leq 0.05$ only means that a discrepancy from the hypothesis prediction would be as large *or larger than* that observed no more than 5 % of the time.

P-values: true or false

The probability of getting certain results if the null is true

true

Textbook interpretation

P-values: true or false

A nonsignificant test result ($P > 0.05$) means that the test hypothesis is true or should be accepted.

false

A large *P* value only suggests that the data are *not* unusual if all the assumptions used to compute the *P* value (including the test hypothesis) were correct. The same data would also not be unusual under many other hypotheses.

P-values: true or false

A null-hypothesis P value greater than 0.05 means that no effect was observed, or that absence of an effect was shown or demonstrated.

false

Observing $P > 0.05$ for the null hypothesis only means that the null is one among the many hypotheses that have $P > 0.05$.
Unless the point estimate (observed association) equals the null value exactly, it is a mistake to conclude from $P > 0.05$ that a study found “no association” or “no evidence” of an effect.

P-values: true or false

The probability of that the null is true if we've obtained certain results

false

The fallacy of the transposed conditional – the p-value would quantify P(H0|D) - the probability that the null hypothesis (H0) is true, given the data (D) collected in the experiment.
The correct interpretation of the p-value is rather different: it quantifies P(D|H0) – the probability of the data given that H0 is true.

P-values: true or false

The P value is the probability that the test hypothesis is true; for example, if a test of the null hypothesis gave P = 0.01, the null hypothesis has only a 1 % chance of being true; if instead it gave P = 0.40, the null hypothesis has a 40 % chance of being true.

false

The *P* value *assumes* the test hypothesis is true—it is *not* a hypothesis probability. The *P* value simply indicates the degree to which the data conform to the pattern predicted by the test hypothesis and all the other assumptions used in the test

P-values: true or false

Statistical significance indicates a scientifically or substantively important relation has been detected.

false

When a study is large, very minor effects or small assumption violations can lead to statistically significant tests of the null hypothesis.
A small null *P* value simply flags the data as being unusual.

P-values: true or false

The P value is the chance of our data occurring if the test hypothesis is true; for example, P = 0.05 means that the observed association would occur only 5% of the time under the test hypothesis.

true

Textbook interpretation

P-values: true or false

Lack of statistical significance indicates that the effect size is small.

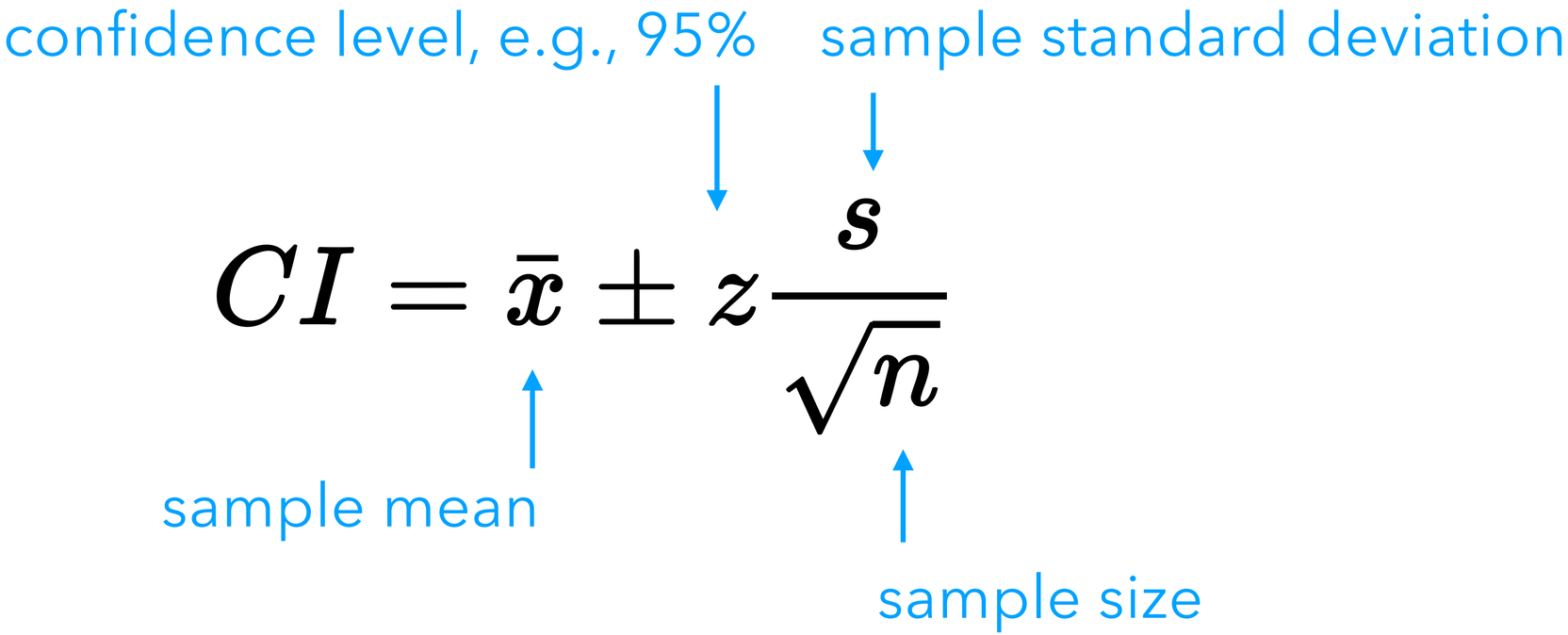
false

when a study is small, even large effects may be “drowned in noise” and thus fail to be detected as statistically significant by a statistical test

Confidence Intervals

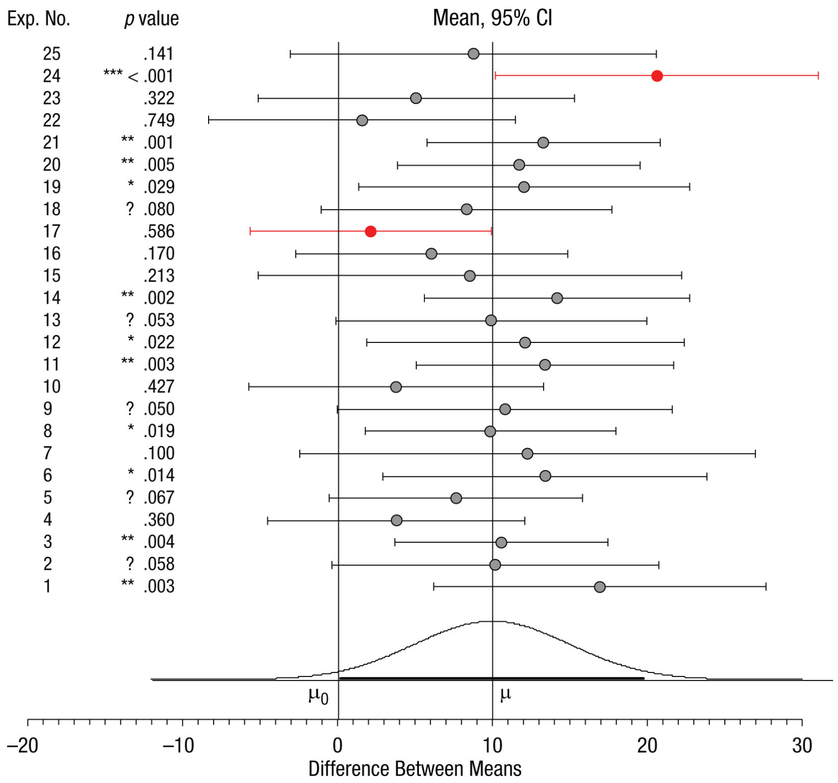
Confidence Intervals: Uncertainty

For example, mean & normal distribution



Confidence Intervals: CIs

If we repeat the experiments many times and calculate CIs in each experiment, how many of these CIs contain the true value falls.



True or false

Conference Intervals: true or false

The specific 95 % confidence interval presented by a study has a 95 % chance of containing the true effect size.

false

the 95 % refers only to how often 95 % confidence intervals computed from very many studies would contain the true size *if all the assumptions used to compute the intervals were correct*.

Conference Intervals: true or false

A confidence interval is randomly chosen from an infinite sequence, the dance of the CIs. In the long run 95% capture the true value, and 5% miss.

true

Textbook interpretation

Conference Intervals: true or false

On average, a 95% CI is an 83% prediction interval. There's a .83 chance that a 95% CI will capture the mean of a single replication experiment.

true

Textbook interpretation

Fewer ways to misinterpret confidence intervals...

Students questions