

Differences In The Means Of Two Groups

Differences May Be Statistically Significant Even With Overlapping Confidence Intervals

Andy Grogan-Kaylor

2023-12-03

Table of contents

1	Background	1
2	Demonstration	2
2.1	Set Up The Data	2
2.2	Confidence Intervals Overlap	3
2.3	t-test Of Mean Differences Is Statistically Significant	3

1 Background

Intuitively, if the *confidence intervals* of two variables overlap, we would expect that the two variables would not differ to a degree that is statistically significant.

However, as the brief example below illustrates, it is possible for two variables to have *overlapping confidence intervals*, yet to be different to a degree that is *statistically significant*.

```
library(Statamarkdown)
```

Stata found at /Applications/Stata/StataSE.app/Contents/MacOS/StataSE

The 'stata' engine is ready to use.

2 Demonstration

2.1 Set Up The Data

```
clear all // clear the workspace

set seed 3846 // set random seed

set obs 100 // 100 empty observations

generate x1 = rnormal(100, 10) // x1 has mean of 100, sd of 10

generate x2 = rnormal(102, 10) // x2 has mean of 102, sd of 10

list in 1/10 // list out some data

save demo.dta, replace
```

Number of observations (_N) was 0, now 100.

	x1	x2
1.	110.8965	102.8522
2.	85.56382	114.437
3.	104.4178	104.5644
4.	90.79031	110.5602
5.	108.6776	116.9658
6.	114.3565	116.0197
7.	87.86876	95.14593
8.	92.02374	117.2697
9.	103.8483	89.12561
10.	91.34591	95.71622

file demo.dta saved

2.2 Confidence Intervals Overlap

```
use demo.dta

ci means x1 x2 // confidence intervals of the two variables overlap
```

Variable	Obs	Mean	Std. err.	[95% conf. interval]	
x1	100	98.74361	1.040009	96.68001	100.8072
x2	100	101.9778	1.011382	99.971	103.9846

2.3 t-test Of Mean Differences Is Statistically Significant

```
use demo.dta

ttest x1 == x2 // t-test finds significant differences between x and x2
```

Paired t test

Variable	Obs	Mean	Std. err.	Std. dev.	[95% conf. interval]	
x1	100	98.74361	1.040009	10.40009	96.68001	100.8072
x2	100	101.9778	1.011382	10.11382	99.971	103.9846
diff	100	-3.234191	1.414134	14.14134	-6.04014	-.4282415
mean(diff) = mean(x1 - x2)				t = -2.2870		
H0: mean(diff) = 0				Degrees of freedom = 99		
Ha: mean(diff) < 0		Ha: mean(diff) != 0		Ha: mean(diff) > 0		
Pr(T < t) = 0.0122		Pr(T > t) = 0.0243		Pr(T > t) = 0.9878		