Two Independent Samples: Confidence Interval

Learning Objective: In a given context, carry out the inferential method for comparing groups and draw the appropriate conclusions.

Confidence Interval for $\mu_1 - \mu_2$ (Two-Sample t Confidence Interval)

So far we've discussed the two-sample t-test, which checks whether there is enough evidence stored in the data to reject the claim that $\mu_1 - \mu_2 = 0$ (or equivalently, that $\mu_1 = \mu_2$) in favor of one of the three possible alternatives.

If we would like to estimate $\mu_1 - \mu_2$ we can use the natural point estimate, $\overline{y_1} - \overline{y_2}$, or preferably, a 95% confidence interval which will provide us with a set of plausible values for the difference between the population means $\mu_1 - \mu_2$.

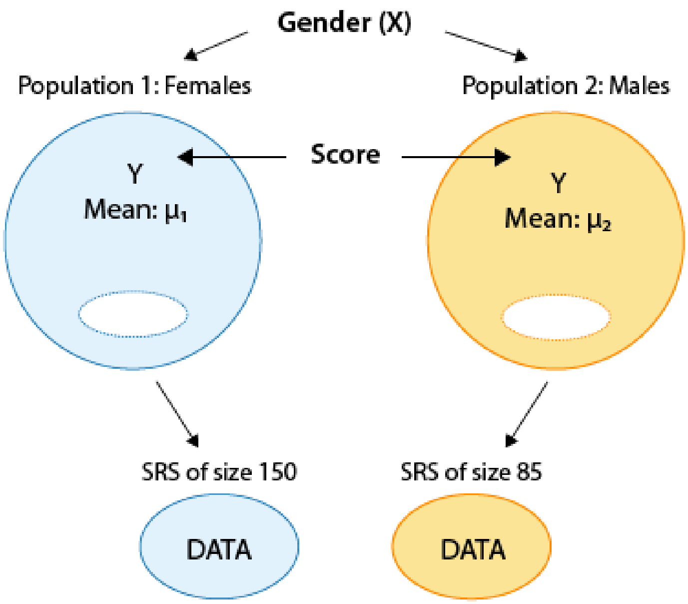
In particular, if the test has rejected $H_o : \mu_1 - \mu_2 = 0$, a confidence interval for $\mu_1 - \mu_2$ can be insightful since it quantifies the effect that the categorical explanatory variable has on the response.

Comment

We will not go into the formula and calculation of the confidence interval, but rather ask our software to do it for us, and focus on interpretation.

Example

Recall our leading example about the looks vs. personality score of females and males:



Here again is the output:

Two Sample T – Test and CI: Score (Y), Gender (X)

Gender (X)	n	Mean	Std. Dev.	Std. Err.
Female	**150**	**10.733334**	**4.254751**	0.347399
Male	**85**	**13.3294115**	**4.0189676**	0.43591824

Hypothesis test results:

μ_1 : mean of Score (Y) where X = Female

μ_2 : mean of Score (Y) where X = Male

$\mu_1 - \mu_2$: mean difference

$H_0 : \mu_1 - \mu_2 = 0$

$H_A : \mu_1 - \mu_2 \neq 0$

Difference	Sample Mean	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$	-2.5960784	0.55741435	182.97267	-4.657358	<0.0001

95% confidence interval results:

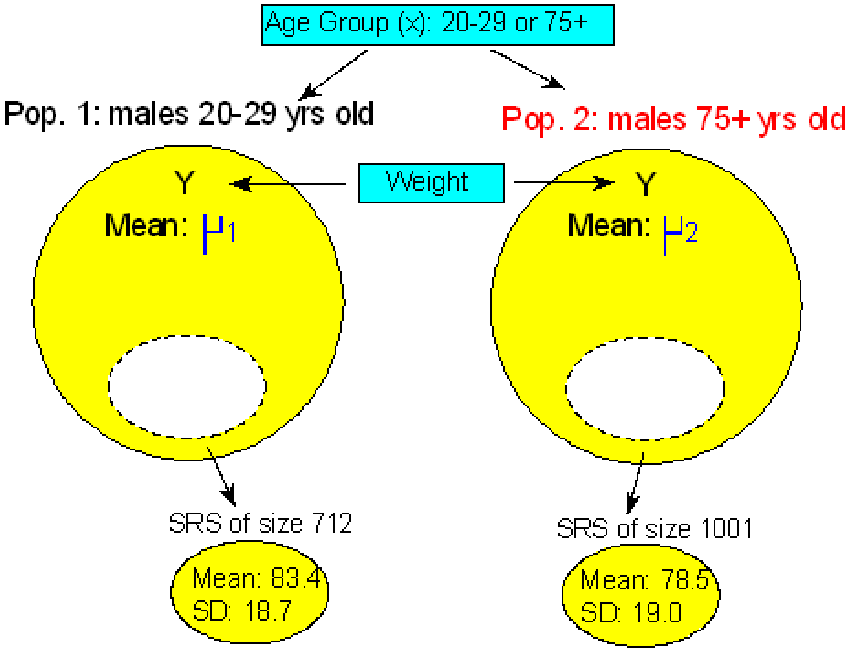
Difference	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
$\mu_1 - \mu_2$	-2.5960784	0.55741435	182.97267	**-3.6958647**	**-1.4962921**

Recall that we rejected the null hypothesis in favor of the two-sided alternative and concluded that the mean score of females is different from the mean score of males. It would be interesting to supplement this conclusion with more details about this difference between the means, and the 95% confidence interval for $\mu_1 - \mu_2$ does exactly that.

According to the output the 95% confidence interval for $\mu_1 - \mu_2$ is roughly (-3.7, -1.5). First, note that the confidence interval is strictly negative suggesting that μ_1 is lower than μ_2 . Furthermore, the confidence interval tells me that we are 95% confident that the mean "looks vs. personality score" of females (μ_1) is between 1.5 and 3.7 points lower than the mean looks vs. personality score of males (μ_2). The confidence interval therefore quantifies the effect that the explanatory variable (gender) has on the response (looks vs personality score).

Scenario: Weight by Mens Age

The purpose of this activity is to give you guided practice in interpreting a 95% confidence interval for $\mu_1 - \mu_2$ following a two-sample t-test that rejected H_o. Recall our second example:



Recall that we were testing

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_a : \mu_1 - \mu_2 > 0$$

and we found using statistical software that the test statistic was 5.31 with a p-value of 0.000. Based on the small p-value, we rejected H_o and concluded that males 20-29 years old weigh more, on average, than males 75+ years old. It would be interesting to follow up this conclusion and estimate how much more males 20-29 years old weigh, on average. The 95% confidence interval for $\mu_1 - \mu_2$ does exactly that, and is given by the formula

$$(\overline{Y}_1 - \overline{Y}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

As a reminder, here again is the summary of the study results:

	n	$\bar{Y}$	S
20-29 yrs old	712	83.4	18.7
75+ yrs old	1001	78.5	19.0

Learn By Doing

1/1 point (graded)

Plugging the study results into the above formula gives us a 95% confidence interval of (3.091, 6.709).

True or false? This means that with 95% confidence, we can say that males who are 20-29 years old weigh, on average, 3.1 to 6.7 kilograms less than males who are 75 and older.

- True
- False ✔

Answer

Correct:

This means that with 95% confidence, we can say that males who are 20-29 years old weigh, on average, 3.1 to 6.7 kilograms *more* than males who are 75 and older.

Submit