⚠️ Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit http://online.stanford.edu.

Course > Inference: Relationships C→Q > Two Independent Samples > Two Independent Samples: Example

🔖 **Bookmark this page**

Two Independent Samples: Example

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

Learning Objective: Specify the null and alternative hypotheses for comparing groups.

Let's look at another example, and then you'll do one yourself.

Example

According to the National Health And Nutrition Examination Survey (NHANES) sponsored by the U.S. government, a random sample of 712 males between 20 and 29 years of age and a random sample of 1,001 males over the age of 75 were chosen, and the weight of each of the males was recorded (in kg). Here is a summary of the results (Source: http://www.cdc.gov/nchs/data/ad/ad347.pdf):

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

Do the data provide evidence that the younger male population weighs more (on average) than the older male population? (Note that here the data are given in a summarized form, unlike the previous problem, where the raw data were given.)

Here is a figure that summarizes this example:
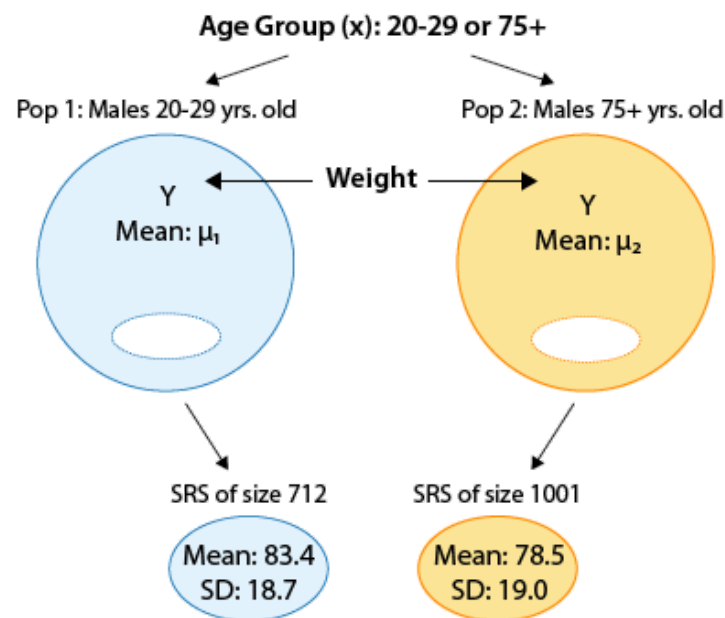
Note that we defined the younger age group and the older age group as population 1 and population 2, respectively, and μ_1 and μ_2 as the mean weight of population 1 and population 2, respectively.

Step 1:

Since we want to test whether the older age group (population 2) weighs less on average than the younger age group (population 1), we are testing:

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

or equivalently,

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

Step 2:

We can safely use the two-sample t-test in this case since:

1. The samples are independent, since each of the samples was chosen at random.

2. Both sample sizes are very large (712 and 1,001), and therefore we can proceed regardless of whether the populations are normal or not.

It is possible from these data to calculate the t-statistic of 5.31 and the p-value of 0.000. The t-value is quite large, and the p-value correspondingly small, indicating that our data are very different from what is claimed in the null hypothesis.

Step 3:

The p-value is essentially 0, indicating that it would be nearly impossible to observe a difference between the sample mean weights of 4.9 (or more) if the mean weights in the age group populations were the same (i.e., if H_o were true).

Step 4:

A p-value of 0 (or very close to it) indicates that the data provide strong evidence against H_o, so we reject it and conclude that the mean weight of males 20-29 years old is higher than the mean weight of males 75 years old and older. In other words, males in the younger age group weigh more, on average, than males in the older age group.