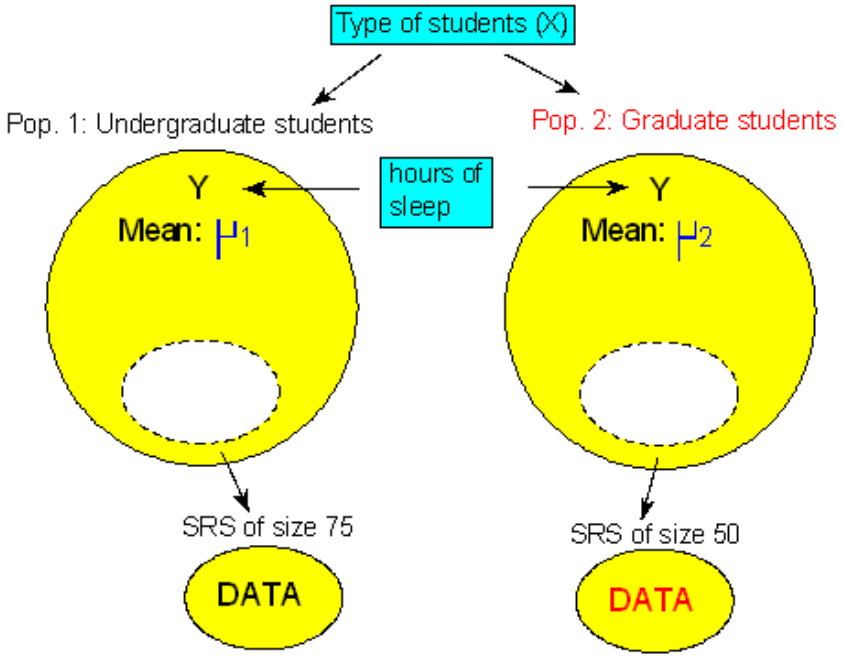
Statistics Package Exercise: Carrying Out the Two-Sample t-test

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

The purpose of this activity is to give you guided practice in carrying out the two-sample t-test, and to show you how to use software to aid in the process.

Background

A study was conducted at a large state university in order to compare the sleeping habits of undergraduate students to those of graduate students. Random samples of 75 undergraduate students and 50 graduate students were chosen and each of the subjects was asked to report the number of hours he or she sleeps in a typical day. The thought was that since undergraduate students are generally younger and party more during their years in school, they sleep less, on average, than graduate students. Do the data support this hypothesis? The following figure summarizes the problem:



Note that we defined:

μ_1—the mean number of hours undergraduate students sleep in a typical day

μ_2—the mean number of hours graduate students sleep in a typical day

---

Comment: Before we move on to carry out the test, it is important to realize that in the two-sample problem, the data can be provided in three possible ways:

(i) Sample data in one column, and another column that indicates which sample the observation belongs to. Recall that this is the way the data were given in our leading example (looks vs. personality score and gender):

Score (Y)	Gender (X)
15	Male
13	Female
10	Female
12	Male
14	Female
14	Male
6	Male
17	Male
...	...

Note that essentially, one column contains the explanatory variable, and one contains the response.

(ii) Sample data in different columns—data from each of the two samples appear in a column dedicated to that category. As you'll see, this is the way the data are provided in this example:

Undergraduate	Graduate
6	8
5	5
3	6
6	6
...	...

(iii) Summarized data—we are not given the actual data, but just the data summaries: sample sizes, sample means and sample standard deviations of both samples. Recall that in our second example, the data were given in this format.

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

- **R** **StatCrunch** **TI Calculator** **Minitab** **Excel**

R Instructions

To carry out the test, open R with the data set preloaded by right-clicking here and choosing "Save Target As" to download the file to your computer. Then find the downloaded file and double-click it to open it in R.

The data have been loaded into the data frame

```
sleep
```

. The two variables in the data frame are

```
undergraduate
```

and

```
graduate
```

.

To carry out the t-test, enter the command:

- ```
  t.test(sleep$undergraduate,sleep$graduate,alternative = "less")
  ```
  ```

Note: Using R, when we used

```
t.test()
```

for a one-sample t-test in a previous activity, we specified a one-sample data set, a hypothetical mean, and an alternative hypothesis. To perform a two-sample t-test, we use the same command,

```
t.test()
```

, but specify two sample data sets and an alternative hypothesis. If the data set were structured so the sample data is in one column (called

```
sleep
```

) and another column that indicates which sample the observation belongs to (called

```
student.type
```

), then the command would be

```
t.test(sleep~student.type, alternative="less")
```

.

Learn By Doing (1/1 point)

Carry out the test and report the test statistic and p-value.

Your Answer:

t = -1.2304
p-value = 0.1106

Our Answer:
RStatCrunch TI CalculatorMinitabExcel R Here is the R output: The test statistic is t = -1.2304 and the p-value is 0.1106. StatCrunch Here is the StatCrunch output: TI Calculator Here is the output: The test statistic is t = -1.2304 and the p-value is 0.1106. Minitab Here is the output. The test statistic and p-value are highlighted. Excel Here is the output. The test statistic and p-value are highlighted.

Resubmit Reset

Learn By Doing (1/1 point)

Draw your conclusions in context.

Your Answer:

Assuming Ho is true, there was an 11.06% possibility for us to get (mu1), which were -1.23 standard errors away from the (mu2). Since this is higher than 0.05, it can be deemed not stastistically significant, so we reject Ho, and accept Ha.

Our Answer:

The p-value is not small (in particular, it is larger than 0.05), indicating that it is still reasonably likely (probability 0.111) to get data like those observed, or even more extreme data, under the null hypothesis (i.e., assuming that undergraduate and graduate students have the same mean sleeping hours). Therefore, the data do not provide evidence to reject Ho, and we cannot conclude that undergraduate students sleep less, on average, than graduate students.

Resubmit Reset