A Brief Introduction to One-sample and Two-sample t-tests

Nastaran Marzban

August 22, 2024

Have you ever wondered how to test if your data stands out or if two groups really differ? The one-sample and two-sample t-tests are simple yet powerful tools to answer this question. Let's dive into them to learn more.

1 One Sample t-test

A one-sample t-test is a statistical test used to determine whether the mean of a **single sample** differs significantly from a hypothesized population mean. In other words, we want to test the below hypothesis

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0.$$

To do this statistical test, after gathering some samples from the population we should calculate the Test Statistic (t) as below

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt(n)},$$

in which

- $\bar{x}$ is the sample mean.
- μ_0 is the hypothesized population mean.
- s is the sample standard deviation.
- n is the sample size.

Then we should compare its value to the critical t-value, obtained using a t-distribution table or statistical software. To use a t-distribution table, we find the critical t-value for $n - 1$ degrees of freedom at the 0.05 significance level (two-tailed). If the absolute value of the test statistic is greater than the critical t-value, we reject the null hypothesis.

Not to mention, the assumption of normality is essential, especially if the sample size is less than 30.

Example 1.1. *Imagine a nutritionist claiming that the average daily intake of calories for adults in an Erupean city is 2,500 calories. A statistician* *wants to test whether this claim is true or not. Meaning that the hypothesis test can be as follows*

$$H_0 : \mu = 2500$$
$$H_a : \mu \neq 2500.$$

2 Two-Sample t-test

The Two-Sample t-Test, is a statistical method used to determine whether there is a significant difference between the means of **two independent groups** or not. In another word, we want to test the below hypothesis

$$H_0 : \mu_1 = \mu_2$$
$$H_a : \mu_1 \neq \mu_2.$$

To do this statistical test, after gathering some samples from each population we should calculate the Test Statistic (t) as below

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

in which

- $\bar{x_1}$ and $\bar{x_2}$ are the sample means.
- s_1^2 and s_2^2 are the sample variances.
- n_1 and n_2 are the sample sizes.

Then we should compare this value to the critical t-value, obtained using a t-distribution table or statistical software. To use a t-distribution table, we find the critical t-value for $n_1 + n_2 - 2$ degrees of freedom at the 0.05 significance level (two-tailed). If the absolute value of the test statistic is greater than the critical t-value, we reject the null hypothesis.

Not to mention, the assumption of normality is essencial. Besides, observations in each group must be independent. Last but not least the variances of the two groups should be equal. This can be tested using **Levene's Test**. If the variances of the two groups are equal then $df = n_1 + n_2 - 2$

and if they are not equal we have $df = \dfrac{\left(\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1-1}+\dfrac{\left(\dfrac{s_2}{n_2}\right)^2}{n_2-1}}$. In the latter case we have **Welch Test**, a useful test when the two groups have unequal variances.

Example 2.1. *Imagine a techer claiming that his newly introduced method of teaching mathematics is much more helpful than his previous method. He asks a statistician to help him draw a valid conclusion.Thus, the statistician tests the below hypothesis*

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_a &: \mu_1 \neq \mu_2, \end{aligned}$$

and then he collects samples from each population to calculate the test statistic.

In a future post, I want to discuss what to do if we have more than two populations to compare their means. Does anyone have any suggestions?