Stat 131 (Mathematical Statistics III)

Lesson 4.4. Independent samples t test

Learning Outcomes

At the end of the lesson, students should be able to

1. articulate the test procedure for testing hypothesis about the difference of two population means; and
2. apply the independent samples t test to compare the means of two independent (normal) populations.

Discussion

A second application of the t distribution is in constructing a small-sample test to compare the means of two normal populations. Suppose $Y_{11}, Y_{12}, \cdots, Y_{1n_1}$ and $Y_{21}, Y_{22}, \cdots, Y_{2n_2}$ are independent random samples from normal distributions with unknown means μ_1 and μ_2 and unknown variances σ_1^2 and σ_1^2. Let $\bar{Y}_i$, S_i^2, and n_i, for $i = 1, 2$, be the corresponding sample means, sample variances, and sample sizes. If we can show (thru an F test) that $\sigma_1^2 = \sigma_2^2 = \sigma^2$, then

$$T = \frac{(\overline{Y}_1 - \overline{Y}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

follows a t distribution with $df = n_1 + n_2 - 2$, where

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

is the pooled sample standard deviation and is an unbiased estimator of σ.

If we are interested in testing $H_0 : \mu_1 = \mu_2$ against a suitable alternative hypothesis, we can use the test statistic

$$T = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

for which when H_0 is true reduces to

$$T = \frac{\overline{Y}_1 - \overline{Y}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

The rejection regions for each type of alternative hypothesis are given in the table below.

Alternative Hypothesis	Rejection Region
$H_1 : \mu_1 > \mu_2$	$RR = \{t : t \geq t_{\alpha,df}\}$
$H_1 : \mu_1 < \mu_2$	$RR = \{t : t \leq -t_{\alpha,df}\}$
$H_1 : \mu_1 \neq \mu_2$	$RR = \{t : \lvert t \rvert \geq t_{\alpha/2,df}\}$

The above test statistic is popularly known as the **Student's t-test**. Let's take a look at examples.

Example 4.4.1

In a packing plant, a machine packs cartons with jars. It is supposed that a new machine will pack faster on the average than the machine currently used. To test that hypothesis, the times it takes each machine to pack ten cartons are recorded and the summary statistics are given below.

	New Machine	Old Machine
n	10	10
$\bar{y}$	42.14	43.23
s	0.683	0.750

Do the data provide sufficient evidence to conclude that, on the average, the new machine packs faster?

SOLUTION

Let Y_1 and Y_2 denote the time required to pack cartons by the new and old machine, respectively and assume that $Y_1 \sim N(\mu_1, \sigma_1^2)$ and $Y_2 \sim N(\mu_2, \sigma_2^2)$.

For this problem the appropriate hypotheses are:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

For now, let us assume that the populations variances are equal, that is $\sigma_1^2 = \sigma_2^2 = \sigma^2$. The pooled sample standard deviation is

$$\begin{aligned}
s_p &= \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} \\
&= \sqrt{\frac{(10-1)0.683^2+(10-1)0.750^2}{10+10-2}} \\
&\approx 0.717
\end{aligned}$$

Hence, the value of the test statistic is

$$\begin{aligned}
t &= \frac{\overline{Y}_1-\overline{Y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \\
&= \frac{42.14-43.23}{0.717\sqrt{\frac{1}{10}+\frac{1}{10}}} \\
&\approx -3.399
\end{aligned}$$

This test statistic has a t distribution with 18 degrees of freedom. Thus, at $\alpha = 0.05$, we reject H_0 in favor of H_1 if $\{t < -t_{0.05,18} = -1.734\}$. The critical value was obtained from a T table or using the R command **qt(0.05,18)**.

The p-value associated with $t = -3.399$ is approximately 0.0016 which can be obtained using the R command **pt(-3.399, 18)**.

Using either the rejection region or the p-value, we reject H_0.

Therefore, at $\alpha = 0.05$, the data provide sufficient evidence to conclude that, on the average, the new machine packs faster than the old machine.

When the population variances are not equal, the alternative test statistic for testing H_0 against a suitable alternative hypothesis is given by

$$T = \frac{(\bar{Y}_1-\bar{Y}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

which follows a t distribution with adjusted degrees of freedom determined by the equation:

$$\nu = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}}$$

Usually, the value of ν is rounded off to the nearest integer.

Under the null hypothesis, the above test statistic reduces to

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

The above test statistic is popularly known as the **Welch's t-test**.

Example 4.4.2

Let us revisit *Example 4.4.1* and assume that the population variances are not equal.

SOLUTION

Let Y_1 and Y_2 denote the time required to pack cartons by the new and old machine, respectively and assume that $Y_1 \sim N(\mu_1, \sigma_1^2)$ and $Y_2 \sim N(\mu_2, \sigma_2^2)$.

For this problem the appropriate hypotheses are:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

Assuming $\sigma_1^2 \neq \sigma_2^2$. Then, the value of the test statistic is

$$\begin{aligned} t &= \frac{\overline{Y}_1 - \overline{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\ &= \frac{42.14 - 43.23}{\sqrt{\frac{0.683^2}{10} + \frac{0.750^2}{10}}} \\ &\approx -3.398 \end{aligned}$$

This test statistic has a t distribution with degrees of freedom approximately equal to

$$\begin{aligned} \nu &= \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \\ &= \frac{\left(\frac{0.683^2}{10} + \frac{0.750^2}{10}\right)^2}{\frac{\left(\frac{0.683^2}{10}\right)^2}{10-1} + \frac{\left(\frac{0.750^2}{10}\right)^2}{10-1}} \\ &\approx 18 \end{aligned}$$

Thus, at $\alpha = 0.05$, we reject H_0 in favor of H_1 if $\{t < -t_{0.05,18} = -1.734\}$. The critical value was obtained from a T table or using the R command **qt(0.05,18)**.

The p-value associated with $t = -3.398$ is approximately 0.0016 which can be obtained using the R command **pt(-3.399, 18)**.

Using either the rejection region or the p-value, we reject H_0.

Therefore, at $\alpha = 0.05$, the data provide sufficient evidence to conclude that, on the average, the new machine packs faster than the old machine.

Remark

In actual applications, we have to verify the assumption that the population variances are equal using an F test before we decide which between the Student's t- or Welch's t-test to use. This is illustrated in the following example.

Example 4.4.3

A psychologist was interested in exploring whether or not male and female college students have different driving behaviors. There were a number of ways that she could quantify driving behaviors. She opted to focus on the fastest speed ever driven by an individual. She conducted a survey of a random 34 male college students and a random 29 female college students. Here is a descriptive summary of the results of her survey:

	Male	Female
n	34	29
$\overline{y}$	105.5	90.9
s	20.1	12.2

Is the mean fastest speed driven by male college students different than the mean fastest speed driven by female college students?

SOLUTION

Let Y_1 and Y_2 denote the driving speed of male and female students, respectively, and assume that $Y_1 \sim N(\mu_1, \sigma_1^2)$ and $Y_2 \sim N(\mu_2, \sigma_2^2)$.

We are interested in testing $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$. Before we can decide the test statistic for this testing problem, let us verify first if the population variances are statistically equal or not.

A. F test for equality of population variances

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

The test statistic is

$$F = \frac{S_1^2}{S_2^2} = \frac{20.1^2}{12.2^2} \approx 2.714$$

At $\alpha = 0.05$, the rejection region is $RR = \{F : F \geq F_{0.05/2,33,28} = 2.089\}$. The p-value associated with $F = 2.714$ is 0.0086.

Based on the rejection region or the p-value, we reject H_0.

Therefore, at $\alpha = 0.05$, there is sufficient evidence in the sample that the population variances are not equal.

B. T test for equality of population means

$$
\begin{aligned}
H_0 &: \mu_1 = \mu_2 \\
H_1 &: \mu_1 \neq \mu_2
\end{aligned}
$$

Based on the results of the F test, we shall use the Welch's t test.

$$
\begin{aligned}
t &= \frac{\overline{Y}_1 - \overline{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\
&= \frac{105.5 - 90.9}{\sqrt{\frac{20.1^2}{34} + \frac{12.2^2}{29}}} \\
&\approx 3.539
\end{aligned}
$$

This test statistic has a t distribution with approximate degrees of freedom given by

$$
\begin{aligned}
\nu &= \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \\
&= \frac{\left(\frac{20.1^2}{34} + \frac{12.2^2}{29}\right)^2}{\frac{\left(\frac{20.1^2}{34}\right)^2}{34-1} + \frac{\left(\frac{12.2^2}{29}\right)^2}{29-1}} \\
&\approx 55
\end{aligned}
$$

Thus, at $\alpha = 0.05$, the rejection region is $RR = \{t : |t| > t_{0.025,55} = 2.004\}$. In addition, the p-value associated with $t = 3.539$ is 8×10^{-4}.

Hence, based on either the rejection region or the p-value, we reject H_0.

Therefore, at $\alpha = 0.05$, there is sufficient evidence that the mean fastest speed driven by male college students is different than the mean fastest speed driven by female college students.

Learning Tasks

Instruction: Answer the following as indicated.

1. Two methods for teaching reading were applied to two randomly selected groups of elementary schoolchildren and then compared on the basis of a reading comprehension test given at the end of the learning period. The sample means and variances computed from the test scores are shown in the accompanying table. Do the data present sufficient evidence to indicate a difference in the mean scores for the populations associated with the two teaching methods?

	Method 1	Method 2
n	11	14
$\bar{y}$	64	69
s	52	71

2. Medical researchers are interested to know if under normal conditions men have lower body temperatures (°F) than women. They collected data from a large number of men and women, and random samples from that data are presented in the accompanying table. Is the data sufficient to indicate that the mean body temperature of men is lower than that of women?

Men	Women
96.9	97.8
97.4	98.0
97.5	98.2
97.8	98.2
97.8	98.2
97.9	98.6
98.0	98.8
98.6	99.2
98.8	99.4