For the two-sided hypothesis

$$H_0\colon\ \mu_1 = \mu_2,$$
$$H_1\colon\ \mu_1 \neq \mu_2,$$

we reject H_0 at significance level α when the computed t-statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}},$$

where

$$s_p^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

exceeds $t_{\alpha/2,n_1+n_2-2}$ or is less than $-t_{\alpha/2,n_1+n_2-2}$.

Recall from Chapter 9 that the degrees of freedom for the t-distribution are a result of pooling of information from the two samples to estimate σ^2. One-sided alternatives suggest one-sided critical regions, as one might expect. For example, for $H_1\colon\ \mu_1 - \mu_2 > d_0$, reject $H_1\colon\ \mu_1 - \mu_2 = d_0$ when $t > t_{\alpha,n_1+n_2-2}$.

An experiment was performed to compare the abrasive wear of two different laminated materials. Twelve pieces of material 1 were tested by exposing each piece to a machine measuring wear. Ten pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average (coded) wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 81 with a sample standard deviation of 5. Can we conclude at the 0.05 level of significance that the abrasive wear of material 1 exceeds that of material 2 by more than 2 units? Assume the populations to be approximately normal with equal variances.

Let μ_1 and μ_2 represent the population means of the abrasive wear for material 1 and material 2, respectively.

1. H_0: $\mu_1 - \mu_2 = 2$.
2. H_1: $\mu_1 - \mu_2 > 2$.
3. $\alpha = 0.05$.
4. Critical region: $t > 1.725$, where $t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$ with $v = 20$ degrees of freedom.
5. Computations:

$$\bar{x}_1 = 85, \qquad s_1 = 4, \qquad n_1 = 12,$$
$$\bar{x}_2 = 81, \qquad s_2 = 5, \qquad n_2 = 10.$$

Hence

$$
\begin{aligned}
s_p &= \sqrt{\frac{(11)(16) + (9)(25)}{12 + 10 - 2}} = 4.478, \\
t &= \frac{(85 - 81) - 2}{4.478\sqrt{1/12 + 1/10}} = 1.04, \\
P &= P(T > 1.04) \approx 0.16. \qquad \text{(See Table A.4.)}
\end{aligned}
$$

6. Decision: Do not reject H_0. We are unable to conclude that the abrasive wear of material 1 exceeds that of material 2 by more than 2 units.

10.28 According to *Chemical Engineering*, an important property of fiber is its water absorbency. The average percent absorbency of 25 randomly selected pieces of cotton fiber was found to be 20 with a standard deviation of 1.5. A random sample of 25 pieces of acetate yielded an average percent of 12 with a standard deviation of 1.25. Is there strong evidence that the population mean percent absorbency is significantly higher for cotton fiber than for acetate? Assume that the percent absorbency is approximately normally distributed and that the population variances in percent absorbency for the two fibers are the same. Use a significance level of 0.05.

- the sample mean of cotton fiber is $\overline{x}_1 = 20$
- sample standard deviation of cotton fiber is $s_1 = 1.5$
- sample size of cotton fiber is $n_1 = 25$

- the sample mean of acetate is $\overline{x}_2 = 12$
- sample standard deviation of acetate is $s_2 = 1.25$
- sample size of acetate is $n_1 = 25$

The null hypotheses is

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

The alternative hypotheses is

$$H_1 : \mu_1 - \mu_2 > 0$$

We also have $\alpha = 0.05$

Lets find degrees of freedom:

$$\begin{aligned}\text{deegres of freedom} &= n_1 + n_2 - 2 \\ &= 25 + 25 - 2 \\ &= 48\end{aligned}$$

⋆ From the t -distribution table, we see that the critical value at 0.05 level for 48 degrees of freedom is 1.677

- The pooled standard deviation is

$$\begin{aligned} s_p &= \sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}} \\ &= \sqrt{\frac{(25 - 1) \cdot 1.5^2 + (25 - 1) \cdot 1.25^2}{25 + 25 - 2}} \\ &= \sqrt{\frac{24 \cdot 2.25 + 24 \cdot 1.56}{48}} \\ &= \sqrt{1.906} \\ &= 1.381 \end{aligned}$$

- We will reject the null hypothesis if $t > t_\alpha$

Lets now find test statistic as follows:

$$
\begin{aligned}
t &= \frac{\overline{x}_1 - \overline{x}_2}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\
&= \frac{20 - 12}{1.381 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}}} \\
&= \frac{8}{1.381 \cdot \sqrt{\frac{2}{25}}} \\
&= 20.48
\end{aligned}
$$

Therefore, $t > t_\alpha$, so we reject the null hypothesis.

$\star$ Hence, we conclude that the population mean percent absorbency for cotton fiber is significantly higher than the mean for acetate.

The population mean percent absorbency for cotton fiber is significantly higher than the mean for acetate.

H_0	Value of Test Statistic	H_1	Critical Region
$\mu = \mu_0$	$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ known	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu = \mu_0$	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n - 1$, σ unknown	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$; σ_1 and σ_2 known	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$; $v = n_1 + n_2 - 2$, $\sigma_1 = \sigma_2$ but unknown, $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t' = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$; $v = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$, $\sigma_1 \neq \sigma_2$ and unknown	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t' < -t_\alpha$ $t' > t_\alpha$ $t' < -t_{\alpha/2}$ or $t' > t_{\alpha/2}$
$\mu_D = d_0$ paired observations	$t = \dfrac{\bar{d} - d_0}{s_d/\sqrt{n}}$; $v = n - 1$	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

The **Population** Standard Deviation: $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$

The **Sample** Standard Deviation: $s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$

The following data represent the running times of films produced by two motion-picture companies:

Company	**Time (minutes)**						
1	102	86	98	109	92		
2	81	165	97	134	92	87	114

Test the hypothesis that the average running time of films produced by company 2 exceeds the average running time of films produced by company 1 by 10 minutes against the one-sided alternative that the difference is less than 10 minutes. Use a 0.1 level of significance and assume the distributions of times to be approximately normal with unequal variances.

- Let $\overline{x}_1$ denotes sample average run time of films produced by company 1.
- Let μ_1 denotes true average run time of films produced by company 1.
- Let n_1 denote the sample size for films produced by company 1.
- Let s_1^2 denotes sample variance of running time of films produced by company 1.

- Let $\overline{x}_2$ denotes sample average run time of films produced by company 2.
- Let μ_2 denotes true average run time of films produced by company 2.
- Let n_2 denote the sample size for films produced by company 2.
- Let s_2^2 denotes the sample variance of running time of films produced by company 2.

⋆ We need to determine whether the average running time of films produced by company 2 exceeds the average running time of films produced by company 1 by 10 minutes.

⋆ We have:
$n_1 = 5$,
$n_2 = 7$.

- The null hypothesis is

$$H_0 : \mu_1 - \mu_2 = -10$$

- The alternative hypothesis is

$$H_1 : \mu_1 - \mu_2 < -10$$

- The level of signnificant is $\alpha = 0.1$.

- The test statistics is

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (1),$$

where $d_0 = \mu_1 - \mu_2 = -10$.

Let's find $\overline{x}_i$ and s_i^2, $i \in \{1, 2\}$.

We have:

$$\overline{x}_1 = \frac{1}{5}(102 + 86 + 98 + 109 + 92) = 97.4$$

$$\overline{x}_2 = \frac{1}{7}(81 + 165 + 97 + 134 + 92 + 87 + 114) = 110$$

$$s_1 = \frac{1}{7}\Big((102 - 97.4)^2 + (86 - 97.4)^2 + ... + (92 - 97.4)^2\Big)$$
$$= 78.8$$

$$s_2 = \frac{1}{7}\Big((81 - 110)^2 + (165 - 110)^2 + ... + (114 - 110)^2\Big)$$
$$= 913.3$$

Hence, using equation (1), and given data, we have:

$$\begin{aligned} t &= \frac{(\overline{x}_1 - \overline{x}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\ &= \frac{(97.4 - 110) - (-10)}{\sqrt{\frac{78.8}{5} + \frac{913.3}{7}}} \\ &= \frac{-2.6}{12.09} \\ &\approx -0.21 \end{aligned}$$

We can find degrees of freedom by formula:

$$v = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2 / (n_1 - 1) + \left(s_2^2/n_2\right)^2 / (n_2 - 1)}$$

Therefore, the degrees of freedom is:

$$v = \frac{\left(\frac{78.8}{5} + \frac{913.3}{7}\right)^2}{\frac{(78.8/5)^2}{5-1} + \frac{(913.3/7)^2}{7-1}} = 7$$

Using Student t-Distribution Probability Table, we see that the t-value at 0.1 level of significance, with 7 degrees of freedom is

$$t_{0.01} = 1.415$$

Because observed value (which is -0.21) is between -1.415 and 1.415 (i.e. observed value is not in the critical region), we fail to reject the null hypothesis and conclude that the average running time of films produced by company 2 exceeds the average running time of films produced by company 1 by 10 minutes.