Hypothesis Testing: Two-Sample Inference 8

Solved Problems

Two-Sample Inference

Hypothesis Testing for the Difference between Population Means

8.1 Introduction

A more frequently encountered situation is the two-sample hypothesis-testing problem. In this chapter (chapter 8), the appropriate methods of hypothesis testing for both the paired-samples and independent-samples situations are studied.

DEFINITION 8.1 In a **two-sample hypothesis-testing problem**, the underlying parameters of two different populations, *neither of whose values is assumed known*, are compared.

DEFINITION 8.4 Two samples are said to be **paired** when each data point in the first sample is matched and is related to a unique data point in the second sample.

DEFINITION 8.5 Two samples are said to be **independent** when the data points in one sample are unrelated to the data points in the second sample.

8.2 – 8.3 The Paired t Test and Interval Estimation

In this section, we discuss the estimation and hypothesis testing when the samples are dependent or related (paired samples). In this case, two data values x_{i1} and x_{i2} (*one in each sample*) for $i = 1, 2, \dots, n$ are collected from the same element (*unit or item*). Hence they are called paired or matched samples. Consider the difference:

$$d_i = x_{i2} - x_{i1} \ ; \ i = 1, 2, \dots, n$$

then the structure of the paired data takes the following form:

Element Number (i)	Sample 1	Sample 2	Difference (d_i)
1	x_{11}	x_{12}	$d_1 = x_{12} - x_{11}$
2	x_{21}	x_{22}	$d_2 = x_{22} - x_{21}$
.	.	.	
.	.	.	
.	.	.	
n	x_{n1}	x_{n2}	$d_n = x_{n2} - x_{n1}$

The differences $d_1, d_2, \dots, d_n$ represents a random sample of size n (*number of matched pairs*) with sample mean ($\bar{d}$) and sample standard deviation (S_d), where:

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} \text{ and } S_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - n(\bar{d})^2}{n-1}} = \sqrt{\frac{\left[\sum_{i=1}^{n} d_i^2 - \frac{\left(\sum_{i=1}^{n} d_i\right)^2}{n}\right]}{n-1}}$$

Sampling Distribution of $\bar{d}$

Let $d_1, d_2, \dots, d_n$ be a random sample of size n from $N(\mu_d, \sigma_d^2)$, that is, d_i is normally distributed with mean μ_d and unknown variance by σ_d^2. Then the sampling distribution of the sample mean ($\bar{d}$) is approximately normal with the following mean and standard deviation:

$$\mu_{\bar{d}} = \mu_d \quad \text{and} \quad \sigma_{\bar{d}} = \sqrt{\sigma_{\bar{d}}^2} = \frac{\sigma_d}{\sqrt{n}}$$

where

μ_d = mean of the paired differences for the population
σ_d = standard deviation of the paired differences for the population

Usually the sample size (n) is small and standard deviation (σ_d) is unknown in the case of paired data. This leads to the following test statistic for the mean μ_d:

$$t = \frac{\bar{d} - \mu_d}{S_d/\sqrt{n}} \sim t - \text{distribution with degrees of freedom} = n - 1$$

where S_d = sample standard deviation of the paired differences for the sample.

Now, based on the sampling distribution of the sample mean $(\bar{d})$ the $(1-\alpha)100\%$ confidence interval (CI) for μ_d and the hypothesis testing using the one-sample t test procedure called the paired t test can be obtained as follows:

(I) Interval Estimation for μ_d

The two-sided $(1-\alpha)100\%$ confidence interval (CI) for the true mean difference (Δ) or μ_d can be constructed as follows:

$$\text{CI} = \bar{d} \pm t_{(n-1,1-(\alpha/2))} \frac{S_d}{\sqrt{n}}$$

(II) Statistical Test (*Paired t Test*) for μ_d

The hypotheses and the rejection regions at level of significance α can be described as follows:

$H_0: \mu_d = 0 \ vs \ H_1: \mu_d > 0$ then reject H_0 if $t > t_{(n-1,1-\alpha)}$ otherwise Accept H_0.

$H_0: \mu_d = 0 \ vs \ H_1: \mu_d < 0$ then reject H_0 if $t < -t_{(n-1,1-\alpha)}$ otherwise Accept H_0.

$H_0: \mu_d = 0 \ vs \ H_1: \mu_d \neq 0$ then reject H_0 if $t > t_{(n-1,1-\alpha/2)}$ or $t < -t_{(n-1,1-\alpha/2)}$ Otherwise Accept H_0.

***p*-value**

The *p*-value for the two-sided paired t test can be computed as follows:

EQUATION 8.5

Computation of the *p*-Value for the Paired *t* Test

If $t < 0$,

$p = 2 \times$ [the area to the left of $t = \bar{d}/(s_d/\sqrt{n})$ under a t_{n-1} distribution]

If $t \geq 0$,

$p = 2 \times$ [the area to the right of t under a t_{n-1} distribution]

The computation of the *p*-value is illustrated in Figure 8.2.

Question (1)

The weights (in kgs) for a random sample of six patients selected from the Jordan University Hospital (JUH) before and after special exercise program are recorded in the following table:

Patient Number	1	2	3	4	5	6
Before	65	75	82	90	105	98
After	68	70	72	85	95	9

Answer the following:

(a) Construct the 95% confidence interval (CI) for the mean μ_d of the population paired differences?

Solution

Step (1)

The two-sided $(1-\alpha)100\%$ confidence interval (CI) for the true mean difference (Δ) or μ_d can be constructed as follows:

$$\text{CI} = \bar{d} \pm t_{(n-1,1-(\alpha/2))} \frac{S_d}{\sqrt{n}}$$

Step (2)

Patient No. $(\boldsymbol{i})$	**Before** $(\boldsymbol{x_{i1}})$	**After** $(\boldsymbol{x_{i2}})$	$\boldsymbol{d_i = x_{i2} - x_{i1}}$	$\boldsymbol{d_i^2}$
1	65	68	3	9
2	75	70	-5	25
3	82	72	-10	100
4	90	85	-5	25
5	105	95	-10	100
6	98	95	-3	9
Sum			-30	268

Step (3)

We calculate the sample mean ($\bar{d}$) and the sample standard deviation (S_d) as follows:

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} = \frac{-30}{6} = -5$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - \frac{\left(\sum_{i=1}^{n} d_i\right)^2}{n}}{n-1}} = \sqrt{\frac{268 - \frac{(-30)^2}{6}}{6-1}} = \sqrt{\frac{268 - 150}{5}} = 4.858$$

Step (4)

The critical value can be calculated as follows:

$$(1-\alpha)100\% = 95\%$$

$$1-\alpha = 0.95$$

$$\alpha = 0.05$$

$$t_{(n-1,\ 1-(\alpha/2))} = t_{(6-1,\ 1-(0.05/2))} = t_{(5,\ 0.975)} = 2.571$$

Step (5)

The (1 - α) × 100 % = 95% confidence interval for the mean μ_d of the population paired differences can be constructed as follows:

$$\text{Lower limit} = \bar{d} - t_{(n-1,\ 1-(\alpha/2))} \frac{S_d}{\sqrt{n}} = -5 - \left[(2.571)\left(\frac{4.858}{\sqrt{6}}\right)\right] = -5 - 5.099 = -10.099$$

$$\text{Upper limit} = \bar{d} + t_{(n-1,\ 1-(\alpha/2))} \frac{S_d}{\sqrt{n}} = -5 + \left[(2.571)\left(\frac{4.858}{\sqrt{6}}\right)\right] = -5 + 5.099 = 0.099$$

CI = (-10.099, 0.099)

Conclusion

We are 95% confident that the mean μ_d of the population paired differences of weights is between -10.099 and 0.099 kgs since.

(b) Can we conclude that there is a difference in weights of patients before and after the exercise program? Test using $\alpha = 0.01$?

Solution

Step (1)

The null and alternative hypothesis can be written as follows:

$$H_0: \mu_d = 0 \quad vs \quad H_1: \mu_d \neq 0 \ , \ \ \alpha = 0.01$$

Step (2)

We calculate the value of the test statistic t as follows:

$$t = \frac{\bar{d} - \mu_d}{S_d/\sqrt{n}} = \frac{-5 - 0}{4.858/\sqrt{6}} = -2.521$$

Step (3)

The rejection rule is given as follows: Rreject H_0 at level of significance α if

$$t > t_{(n-1,1-\alpha/2)} \text{ or } \quad t < -t_{(n-1,1-\alpha/2)}$$

Otherwise Accept H_0 $(|t| \leq t_{(n-1,1-\alpha/2)})$.

Step (4)

We get $t = -2.521 > -t_{(5,\ 1-(0.01/2))} = -t_{(5,\ 0.995)} = -4.032$

Then we Accept $H_0: \mu_d = 0$ at $\alpha = 0.01$ and conclude that the exercise program does not affect the weights of patients.

(c) Calculate the p-value for the test in (b)?

Solution

***p*-value**

The p-value for the two-sided paired t test can be computed as follows:

EQUATION 8.5

Computation of the p-Value for the Paired t Test

If $t < 0$,

$p = 2 \times$ [the area to the left of $t = \bar{d}/(s_d/\sqrt{n})$ under a t_{n-1} distribution]

If $t \geq 0$,

$p = 2 \times$ [the area to the right of t under a t_{n-1} distribution]

The computation of the p-value is illustrated in Figure 8.2.

$$p = 2\left[\text{the area to the write of t under } t_{(n-1,\ 1-(\alpha/2))}\right]$$
$$p = 2\,[P\left(t_{(n-1,\ 1-(\alpha/2))} \leq t\right)]$$
$$p = 2\,P(t_{(5,\ 0.995)} \leq\ -2.521)]$$

Example

If $t = 3.324 > 0$, then the p-value can be calculated using the following formula:

$$\begin{aligned} p &= 2 \times [\text{the area to the right of } t \text{ under a } t_{(n-1)} \text{ distribution}] \\ &= 2 \times P(t_{(n-1)} > t) \\ &= 2 \times [1- P(t_{(n-1)} \leq t)] \\ &= 2 \times [1- P(t_9 \leq 3.324)] \\ &= 2 \times [1- 0.995] \\ &= 2 \times [0.005] \\ &= 0.01 \end{aligned}$$

8.4 Two-Sample t Test for Independent Samples (Equal Variances)

Suppose that we have two populations which are normally distributed. If the first population has a mean μ_1 and a variance σ_1^2 (*or a standard deviation* $\sigma_1 = \sqrt{\sigma_1^2}$), and the second population has a mean μ_2 and a variance σ_2^2 (*or a standard deviation* $\sigma_2 = \sqrt{\sigma_2^2}$). Also, suppose that two independent random samples (*groups*) are drawn from these two populations. The first sample of size n_1 is drawn from the first population and has a sample mean ($\overline{X}_1$) and a sample variance (S_1^2). The second sample of size n_2 is drawn from the second population and has a sample mean ($\overline{X}_2$) and a sample variance (S_2^2). We want to test the hypothesis:

$$H_0: \mu_1 = \mu_2 \ \ vs\ \ H_1: \mu_1 \neq \mu_2$$

Assume that the underlying variances in the two populations are the same or equal (that is, $\sigma_1^2 = \sigma_2^2 = \sigma^2$). We know that $\overline{X}_1$ is normally distributed with mean μ_1 and variance σ^2/n_1 and $\overline{X}_2$ is normally distributed with mean μ_2 and variance σ^2/n_2.

It seems reasonable to base the significance test on the difference between the two sample means, $\overline{X}_1 - \overline{X}_2$ which is normally distributed with mean μ_1- μ_2 and variance $\sigma^2(1/n_1 + 1/n_2)$. In symbols, as follows:

EQUATION 8.7 $\bar{X}_1 - \bar{X}_2 \sim N\left[\mu_1 - \mu_2, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]$

Under H_0, we know that $\mu_1 - \mu_2$. Thus, Equation 8.7 reduces to

EQUATION 8.8 $\bar{X}_1 - \bar{X}_2 \sim N\left[0, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]$

If σ^2 were known, then $\overline{X}_1 - \overline{X}_2$ could be divided by $\sigma\sqrt{(1/n_1 + 1/n_2)}$. From Equation 8.8, we have:

EQUATION 8.9 $\dfrac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$

The test statistic in Equation 8.9 could be used as a basis for the hypothesis test. Unfortunately, σ^2 in general is unknown and must be estimated from the data. The best estimate of the population variance σ^2, which is denoted by S^2, is given by a weighted average of the two sample variances, where the weights are the number of df in each sample. In particular, S^2 will then have $(n_1 - 1)\ df$ from the first sample and $(n_2 - 1)\ df$ from the second sample, or:

$$[(n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2]\ df \text{ overall.}$$

EQUATION 8.10

The **pooled estimate of the variance** from two independent samples is given by

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

where $s = \sqrt{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right]/(n_1 + n_2 - 2)}$

Then S can be substituted for σ in Equation 8.9, and the resulting test statistic can then be shown to follow a t distribution with $n_1 + n_2 - 2\ df$ rather than a standard normal distribution, N(0, 1), distribution because σ^2 is unknown. Thus, the following test procedure is used:

EQUATION 8.11

Two-Sample *t* Test for Independent Samples with Equal Variances

Suppose we wish to test the hypothesis $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$ with a significance level of α for two normally distributed populations, where σ^2 is assumed to be the same for each population.

Compute the test statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $s = \sqrt{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right] / (n_1 + n_2 - 2)}$

If $t > t_{n_1+n_2-2,1-\alpha/2}$ or $t < -t_{n_1+n_2-2,1-\alpha/2}$

then H_0 is rejected.

If $-t_{n_1+n_2-2,1-\alpha/2}$ or $t \le t_{n_1+n_2-2,1-\alpha/2}$

then H_0 is accepted.

The acceptance and rejection regions for this test are shown in Figure 8.3.

Similarly, a p-value can be computed for the test. Computation of the p-value depends on whether $\bar{X}_1 \le \bar{X}_2$ $(t \le 0)$ or $\bar{X}_1 > \bar{X}_2$ $(t > 0)$. In each case, the p-value corresponds to the probability of obtaining a test statistic at least as extreme as the observed value t. This is given in Equation 8.12.

EQUATION 8.12

Computation of the p-Value for the Two-Sample *t* Test for Independent Samples with Equal Variances

Compute the test statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $s = \sqrt{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right] / (n_1 + n_2 - 2)}$

If $t \le 0$, $p = 2 \times$ (area to the left of t under a $t_{n_1+n_2-2}$ distribution).

If $t > 0$, $p = 2 \times$ (area to the right of t under a $t_{n_1+n_2-2}$ distribution).

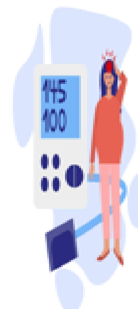
The computation of the p-value is illustrated in Figure 8.4.

EXAMPLE 8.10

Hypertension Suppose a sample of eight 35- to 39-year-old nonprenant, premenopausal OC users who work in a company and have a mean systolic blood pressure (SBP) of 132.86 mm Hg and sample standard deviation of 15.34 mm Hg are identified. A sample of 21 nonpregnant, premenopausal, non-OC users in the same age group are similarly identified who have mean SBP of 127.44 mm Hg and sample standard deviation of 18.23 mm Hg. What can be said about the underlying mean difference $(\mu_1 - \mu_2)$ in blood pressure between the two groups? Assess the statistical significance of the data using $\alpha = 0.05$?

Solution

Step (1)



Sample Number	Sample Size	Sample Mean	Sample Standard Deviation
Sample 1	$n_1 = 8$	$\bar{X}_1 = 132.86$	$S_1 = 15.34$
Sample 2	$n_2 = 21$	$\bar{X}_2 = 127.44$	$S_2 = 18.23$

Step (2) Define the two population means (μ_1) and (μ_2) as follows:

μ_1 = The mean blood pressures of the OC users.

μ_2 = The mean blood pressures of the non-OC users.

We want to test using $\alpha = 0.05$ the hypothesis:

$$H_0: \mu_1 = \mu_2 \;\; vs \;\; H_1: \mu_1 \neq \mu_2$$

Step (3)

The pooled estimate of the sample standard deviation (S) from the two independent samples is calculated as follows:

$$S = \sqrt{S^2} = \sqrt{\frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(8-1)(15.34)^2 + (21-1)(18.23)^2}{8 + 21 - 2}}$$

$$= \sqrt{\frac{1647.2092 + 6646.658}{27}} = \sqrt{\frac{8293.8672}{27}} = 17.527$$

Step (4)

The t-test statistic can be calculated as follows:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S * \sqrt{(1/n_1 + 1/n_2)}} = \frac{(132.86 - 127.44) - 0}{(17.527)\left(\sqrt{\frac{1}{8} + \frac{1}{21}}\right)} = \frac{5.42}{7.282} = 0.744 > 0$$

Step (5)

Since we have $t = 0.744 > 0$, then the rejection rule at level of significance α will be as follows:

$$Rule = \begin{cases} \text{Reject } H_0 \text{ if } t > t_{(n_1 + n_2 - 2\,,\, 1-(\alpha/2))} \\ \text{Accept } H_0 \text{ if } t \leq t_{(n_1 + n_2 - 2\,,\, 1-(\alpha/2))} \end{cases}$$

Step (7)

The critical value is obtained from Table 5 in the Appendix as follows:

$$\begin{aligned} & t_{(n_1 + n_2 - 2\,,\ 1-(\alpha/2))} \\ & = t_{(8 + 21 - 2\,,\ 1-(0.05/2))} \\ & = t_{(27,\ 0.975)} \\ & = 2.052 \end{aligned}$$

Step (8)

The decision will be as follows:
We get

$$t = 0.744 < t_{(27,\ 0.975)} = 2.052$$

it follows that H_0 is accepted using a two-sided t-test at the $\alpha = 5\%$ level.

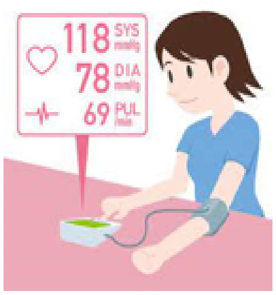
Conclusion

We conclude that the mean blood pressures of the OC users (μ_1) and the mean blood pressures of the non-OC users (μ_2) do not significantly differ from each other, that is, $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$.

p-value

To compute an approximate p-value, and because we have $t = 0.744 > 0$, then we will use the following rule:

$$\begin{aligned} p &= 2 \times [\text{the area to the right of } t \text{ under a } t_{(n_1 + n_2 - 2)} \text{ distribution}] \\ &= 2 \times P(t_{(n_1 + n_2 - 2)} > t) \\ &= 2 \times [1 - P(t_{(n_1 + n_2 - 2)} \leq t)] \\ &= 2 \times [1 - P(t_{27} \leq 0.744)] \\ &= 2 \times [1 - 0.75] \\ &= 2 \times [0.25] \\ &= 0.50 \end{aligned}$$



Now by using the p-value method we have:

$$p = 0.50 > \alpha = 0.05$$

then it follows that H_0 can be accepted using a two-sided Significance t test with α = 0.05.

Notation

The exact p-value obtained from MINITAB program is:

$$\begin{aligned} p &= 2 \times P(t_{27} > 0.744) \\ &= 0.46. \end{aligned}$$

8.5 Interval Estimation for the Comparison of Means from Two Independent Samples (Equal Variance Case)

In the previous section, methods of hypothesis testing for the comparison of means from two independent samples were discussed. It is also useful to compute the $(1 - \alpha) \times 100\%$ confidence intervals (CIs) for the true mean difference between the two groups (*or populations*) $(\mu1 - \mu2)$ as follows:

EQUATION 8.13

Confidence Interval for the Underlying Mean Difference $(\mu_1 - \mu_2)$ Between Two Groups (Two-Sided) $(\sigma_1^2 = \sigma_2^2)$

A two-sided $100\% \times (1 - \alpha)$ CI for the true mean difference $\mu_1 - \mu_2$ based on two independent samples with equal variance is given by

$$\left(\bar{x}_1 - \bar{x}_2 - t_{n_1+n_2-2,1-\alpha/2}\, s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}_1 - \bar{x}_2 + t_{n_1+n_2-2,1-\alpha/2}\, s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

where s^2 = pooled variance estimate given in Equation 8.12.

The derivation of this formula is provided in Section 8.11.

EXAMPLE 8.11

Hypertension Using the data in Examples 8.10, compute a 95% confidence interval (CI) for the true mean difference in systolic blood pressure (SBP) between 35- to 39-year-old OC users and non-OC users $(\mu_1 - \mu_2)$?

Solution

A confidence interval (CI) for the underlying mean difference $(\mu_1 - \mu_2)$ in SBP between the population of 35- to 39-year-old OC users and non-OC users can be calculated as follows:

Step (1)

$(1 - \alpha) \times 100\% = 95\%$

$1 - \alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$1 - \left(\frac{\alpha}{2}\right) = 1 - 0.95 = 0.975$

Step (2)

The critical value is obtained from Table 5 in the Appendix as follows:

$t_{(n_1 + n_2 - 2\ ,\ 1-(\alpha/2))}$

$= t_{(27,\ 0.975)}$

$= 2.052$

Step (3)

Using Equation 8.13, the lower and upper limits for the $(1 - \alpha) \times 100\% = 95\%$ confidence interval (CI) for the true mean difference in systolic blood pressure (SBP) between 35- to 39-year-old OC users and non-OC users $(\mu_1 - \mu_2)$ can be calculated as follows:

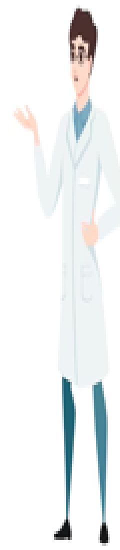
$$CI = \left(\bar{x}_1 - \bar{x}_2 - t_{n_1+n_2-2,1-\alpha/2} s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}_1 - \bar{x}_2 + t_{n_1+n_2-2,1-\alpha/2} s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \right)$$

$$\text{Lower Limit} = (\bar{X}_1 - \bar{X}_2) - t_{(n_1+n_2-2,\,1-(\alpha/2))}\, S\sqrt{(1/n_1 + 1/n_2)}$$

$$= (132.86 - 127.44) - \left[(2.052)(17.527)\left(\sqrt{\frac{1}{8}+\frac{1}{21}}\right)\right]$$

$$= 5.42 - 14.94$$

$$= -9.52$$



$$\text{Upper Limit} = (\bar{X}_1 - \bar{X}_2) + t_{(n_1+n_2-2,\,1-(\alpha/2))}\, S\sqrt{(1/n_1 + 1/n_2)}$$

$$= (132.86 + 127.44) + \left[(2.052)(17.527)\left(\sqrt{\frac{1}{8}+\frac{1}{21}}\right)\right]$$

$$= 5.42 + 14.94$$

$$= 20.36$$

Conclusion: CI =(-9.52, 20.36)

We are 95% confident that the true mean difference (μ_1 - μ_2) in SBP between the population of 35- to 39-year-old OC users and non-OC users is between -9.52 and 20.36. This interval is rather wide and indicates that a much larger sample is needed to accurately assess the true mean difference.

Exercises

Exercise (1)

An instructor wants to use two exams in her classes next year. This year, she gives both exams to the students. She wants to know if the exams are equally difficult and wants to check this by looking at the differences between scores. If the mean difference between scores for students is "close enough" to zero, she will make a practical conclusion that the exams are equally difficult. Use $\alpha = 0.05$. Here is the data:

Table 1: Exam scores for each student

Student	Exam 1 Score	Exam 2 Score	Difference
Bob	63	69	6
Nina	65	65	0
Tim	56	62	6
Kate	100	91	-9
Alonzo	88	78	-10
Jose	83	87	4
Nikhil	77	79	2
Julia	92	88	-4
Tohru	90	85	-5
Michael	84	92	8
Jean	68	69	1
Indra	74	81	7
Susan	87	84	-3
Allen	64	75	11
Paul	71	84	13
Edwina	88	82	-6

Answer:

$$\text{Standard Error} = \frac{s_d}{\sqrt{n}} = \frac{7.00}{\sqrt{16}} = \frac{7.00}{4} = 1.75$$

$$t = \frac{\text{Average difference}}{\text{Standard Error}} = \frac{1.31}{1.75} = 0.750$$

Exercise (2)

Paired t-test Example Solutions

A) Thirty sets of identical twins were enrolled in a study to measure the effect of home environment on certain social attitudes. One twin in each set was randomly assigned to a minority environment or a home environment. The twin assigned to the minority environment went to live with an African American family for a period of 1 year. At the end of the year, an attitudinal survey was administered. The data along with some descriptive statistics follow. Let alpha = 0.025 and test the hypothesis that living in the minority environment leads to higher scores on the attitudinal survey.

Table 1: Raw data for the attitude versus environment study

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Home	65	67	75	77	69	65	73	78	70	72	73	79	68	73	71
Minor	83	75	72	76	78	80	72	81	70	78	77	71	87	70	75
Diff	18	8	-3	-1	9	15	-1	3	0	6	4	-8	19	-3	4
ID	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Home	68	73	72	67	75	78	74	75	66	72	72	78	69	66	73
Minor	75	79	79	69	73	77	77	81	74	83	74	72	78	78	77
Diff	7	6	7	2	-2	-1	3	6	8	11	2	-6	9	12	4

Note: Σdiff = 138, SS_{diff} = 1236.6, Diff = Minor - Home

1) How can you tell that this is a paired experiment?

One clue that this is a Paired Experiment is that the investigator has used sets of twins. Typically, when this is done the analysis will be based on the differences between sets of scores rather than differences between the averages of one group versus the other.

2) Ho: $\mu_d \leq 0$
Ha: $\mu_d > 0$

3) $\alpha = 0.025$ df = 29 t-crit = 2.045

4) t-calc :

variance = $s^2 = 1236.6/29 = 42.64$

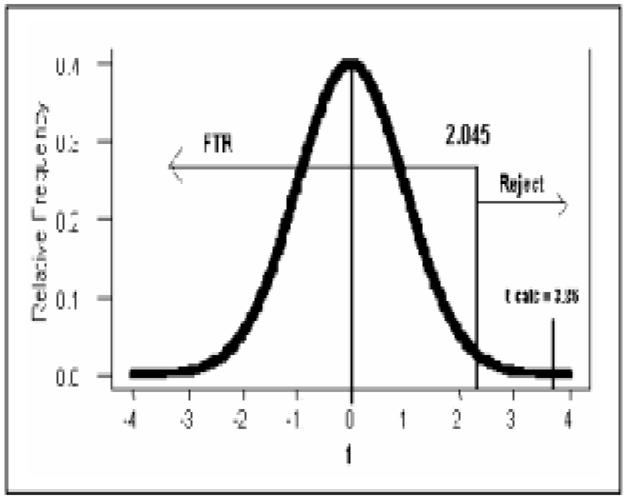
standard deviation = $\sqrt{42.64} = 6.53$

SE mean = $6.53/\sqrt{30} = 1.19$

$\bar{x} = {}^{138}\!/_{30} = 4.6$

t-calc = 4.6/1.19 = 3.86

5) The decision graphic is:



6) The statistical decision is:

Reject Ho

7) The English interpretation is:

At a significance level of 0.025 there is enough evidence to support the claim that living in a minority environment leads to higher scores on the attitudinal survey.

8) Construct a 99% CI for the true average difference in attitudinal scores achieved by subjects living in the two different environments.

$$\begin{aligned} \mu_d &= \bar{d} \pm t \cdot SE \\ &= 4.6 \pm 2.756(1.19) \\ &= 4.6 \pm 3.28 \\ &= (1.32, 7.88) \end{aligned}$$

We are 99% confident that the true average "attitude" difference between living environments is between 1.32 and 7.88. At a significance level of 0.01 we can say that living in a minority environment is associated with higher scores.

Exercise (3)

A psychologist was interested in exploring whether or not male and female college students have different driving behaviors. There were several 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 n_1 =24 male college students and a random n_2 =29 female college students. Here is a descriptive summary of the results of her survey:

Males *(X)*	**Females *(Y)***
[redacted]	
$\bar{x} = 105.5$ $s_x = 20.1$	$\bar{y} = 90.9$ $s_y = 12.2$

Therefore, the particular statistical question she framed was as follows:
Is there sufficient evidence at the $\alpha = 0.05$ level to conclude that the mean fastest speed driven by male college students (μ_1) differs from the mean fastest speed driven by female college students (μ_2)? Assume normal distributions?

Exercise (4)

Example

Mr Brown is the owner of a small bakery in a large town. He believes that the smell of fresh baking will encourage customers to purchase goods from his bakery. To investigate this belief, he records the daily sales for 10 days when all the bakery's windows are open, and the daily sales for another 10 days when all the windows are closed. The following sales, in £, are recorded.

Windows open	202.0 215.6	204.5 208.8	207.0 187.8	215.5 204.1	190.8 185.7
Windows closed	193.5 200.6	192.2 181.8	199.4 169.2	177.6 172.2	205.4 192.8

Assuming that these data may be deemed to be random samples from normal populations with the same variance, investigate the baker's belief.

Solution

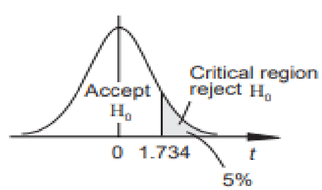
H_0: $\mu_1 = \mu_2$ (1 = open, 2 = closed)

H_1: $\mu_1 > \mu_2$ (one-tailed)

Significance level, $\alpha = 0.05$ (say)

Degrees of freedom, $\nu = 10 + 10 - 2 = 18$

Critical region is $t > 1.734$



Under H_0, the test statistic is

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sigma_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Calculation gives

$$\bar{x}_1 = 202.18, \quad \hat{\sigma}_1^{\,2} = 115.7284$$

and $$\bar{x}_2 = 188.47, \quad \hat{\sigma}_2^{\,2} = 156.6534$$

Hence $$\hat{\sigma}_p^{\,2} = \frac{9 \times 115.7284 + 9 \times 156.6534}{10 + 10 - 2}$$

$$= \frac{115.7284 + 156.6534}{2} \qquad \text{(mean when } n_1 = n_2\text{)}$$

so $$\hat{\sigma}_p = 11.67$$

Thus $$t = \frac{202.18 - 188.47}{11.67\sqrt{\frac{1}{10} + \frac{1}{10}}} = 2.63$$

This value does lie in the critical region so H_0 is rejected. Thus there is evidence, at the 5% level of significance, to suggest that the smell of fresh baking will encourage customers to purchase goods from Mr Brown's bakery.

Exercise (5)

A school mathematics teacher decides to test the effect of using an educational computer package, consisting of geometric designs and illustrations, to teach geometry. Since the package is expensive, the teacher wishes to determine whether using the package will result in an improvement in the pupils' understanding of the topic. The teacher randomly assigns pupils to two groups; a control group receiving standard lessons and an experimental group using the new package. The pupils are selected in pairs of equal mathematical ability, with one from each pair assigned at random to the control group and the other to the experimental group. On completion of the topic the pupils are given a test to measure their understanding. The results, percentage marks, are shown in the table.

Pair	1	2	3	4	5	6	7	8	9	10
Control	72	82	93	65	76	89	81	58	95	91
Experimental	75	79	84	71	82	91	85	68	90	92

Assuming percentage marks to be normally distributed, investigate the claim that the educational computer package produces an improvement in pupils' understanding of geometry.

Solution

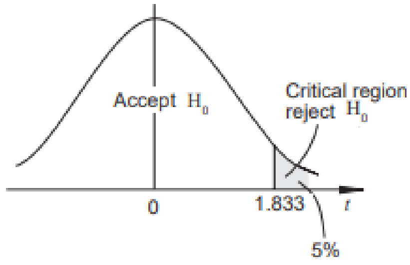
H_0: $\mu_d = 0$ Difference = Experimental – Control

H_1: $\mu_d > 0$ (one-tailed)

Significance level, $\alpha = 0.05$ (say)

Degrees of freedom, $v = 10 - 1 = 9$

Critical region is $t > 1.833$



Under H_0, the test statistic is

$$t = \frac{\bar{d}}{\frac{\hat{\sigma}_d}{\sqrt{n}}}$$

The 10 differences (Experimental - Control) are

d: 3 −3 −9 6 6 2 4 10 −5 1

Hence $\sum d = 15$ and $\sum d^2 = 317$

so $\bar{d} = 1.5$ and $\hat{\sigma}_d = 5.72$

Thus $$t = \frac{1.5}{\frac{5.72}{\sqrt{10}}} = 0.83$$

This value does not lie in the critical region so H_0 is not rejected. Thus there is no evidence, at the 5% level of significance, to suggest that the educational computer package produces an improvement in pupils' understanding of geometry.

Exercise (6)

A random sample of eleven students sat a Chemistry examination consisting of one theory paper and one practical paper. Their marks out of 100 are given in the table below.

Student	A	B	C	D	E	F	G	H	I	J	K
Theory mark	30	42	49	50	63	38	43	36	54	42	26
Practical mark	52	58	42	67	94	68	22	34	55	48	17

Assuming differences in pairs to be normally distributed, test, at the 5% level of significance, the hypothesis of no difference in mean mark on the two papers. (AEB)
