Solution Section 4.2 – Inferences about Two Means: Dependent

Exercise

Listed below are the time intervals (in minutes) before and after eruptions of the Old Faithful geyser. Find the values of $\bar{d}$ and s_d. In general, what does μ_d represent?

Time interval before eruption	98	92	95	87	96
Time interval after eruption	92	95	92	100	90

Solution

The difference values are:

	98	92	95	87	96
	92	95	92	100	90
Difference = d	6	−3	3	−13	6
d^2	36	9	9	169	36

$n=5; \qquad \sum d = 6-3+3-13+6=-1; \quad \sum d^2 = 36+9+9+169+36=259$

$$\bar{d} = \frac{\sum d}{n} = \frac{-1}{5} = -0.2 \text{ min}$$

$$s_d^2 = \frac{n\sum d^2 - \left(\sum d\right)^2}{n(n-1)}$$

```
(5*259-(-1)²)/(5
*4
          64.7000
```

$$= \frac{5(259)-(-1)^2}{5(4)}$$

$$= 64.7$$

$$s_d = \sqrt{64.7} = 8.0 \text{ min}$$

In general, μ_d represents the true mean of the differences from the population of matched pairs (which is mathematically equivalent to the true of the difference between the means of the two populations).

Exercise

Listed below are measured fuel consumption amount (in miles/gal) from a sample of cars.

City fuel consumption	18	22	21	21
Highway fuel consumption	26	31	29	29

Assume that wou want to use a 0.05 significance level to test the claim that the paired sample data come from a population for which the mean differecne is $\mu_d = 0$. Find

a) $\bar{d}$

b) s_d

c) The t test statistic

d) The critical values.

Solution

The difference values are:

	18	22	21	21
	26	31	29	29
Difference = d	-8	-9	-8	-8
d^2	64	81	64	64

$n = 4; \quad \sum d = -8 - 9 - 8 - 8 = -33; \quad \sum d^2 = 64 + 81 + 64 + 64 = 273$

a) $\bar{d} = \frac{\sum d}{n} = \frac{-33}{4} \underline{= -8.3\ mpg|}$

b) $s_d^2 = \frac{n\sum d^2 - (\sum d)^2}{n(n-1)} = \frac{4(273) - (-33)^2}{4(3)} = 0.25$

$s_d = \sqrt{0.25} \underline{= .5\ mpg|}$

c) $t_{\bar{d}} = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-8.25 - 0}{\frac{0.5}{\sqrt{4}}} = -33.00$

d) With $df = 3 \quad and \quad \alpha = 0.05$, the critical values are $t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 3.182$

Exercise

Listed below are predicted high temperatures that were forecast different days.

Predicted high temperatures forecast 3 *days ahead*	79	86	79	83	80
Predicted high temperatures forecast 5 *days ahead*	80	80	79	80	79

Assume that wou want to use a 0.05 significance level to test the claim that the paired sample data come from a population for which the mean differecne is $\mu_d = 0$. Find

a) $\bar{d}$

b) s_d

c) The t test statistic

d) The critical values.

Solution

The difference values are:

	79	83	79	83	80
	80	80	79	80	79
Difference = d	–1	6	0	3	1
d^2	1	36	0	9	1

$n = 4; \qquad \sum d = -1+6+0+3+1 = 9; \quad \sum d^2 = 1+36+0+9+1 = 47$

a) $\bar{d} = \dfrac{\sum d}{n} = \dfrac{9}{5} \underline{= 1.8\ °F}$

b) $s_d^2 = \dfrac{n\sum d^2 - \left(\sum d\right)^2}{n(n-1)} = \dfrac{5(47)-(9)^2}{5(4)} = 7.7$

$s_d = \sqrt{7.7} \underline{= 2.8\ °F}$

c) $t_{\bar{d}} = \dfrac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \dfrac{1.8-0}{\frac{2.7749}{\sqrt{5}}} = 1.45$

d) With $df = 4 \quad and \quad \alpha = 0.05$, the critical values are $t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.776$

Exercise

Listed below are body mass indices (BMI). The BMI of each student was measured in September and April of the freshman year.

a) Use a 0.05 significance level to tet the claim that the mean change in BMI for all students is equal to 0. Does BMI appear to change during freshman year?

b) Construct a 95% confidence interval estimate of the change in BMI during freshman year. Does the confidence interval include 0, and what does that suggest about BMI during freshman year?

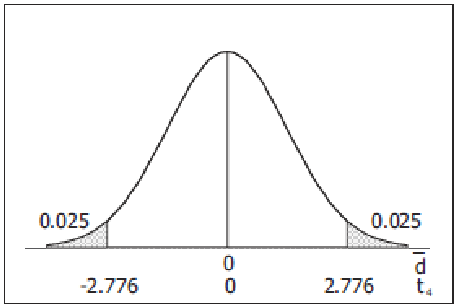
April BMI	20.15	19.24	20.77	23.85	21.32
September BMI	20.68	19.48	19.59	24.57	20.96

Solution

	20.15	19.24	20.77	23.85	21.32
	20.68	19.48	19.59	24.57	20.96
Difference = d	–0.53	–0.24	1.18	–0.72	0.36
d^2	.2809	.0576	1.3924	.5184	.1296

$n=5; \quad \sum d = -53 - .24 + 1.18 - .72 + .36 = .05; \quad \sum d^2 = 2.3789$

a) Original claim: $\mu_d = 0$



$H_0: \mu_d = 0$

$H_1: \mu_d \neq 0$

$df = 4 \quad and \quad \alpha = 0.05$

$t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.776$

$$s_d = \sqrt{\frac{n\sum d^2 - (\sum d)^2}{n(n-1)}}$$

$$= \sqrt{\frac{5(2.3789) - (0.05)^2}{5(4)}}$$

$$= 0.7711$$

$$t_{\bar{d}} = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{0.01 - 0}{\frac{0.7711}{\sqrt{5}}} = 0.029$$

$P - value = 2 \cdot tcdf\left(0.029,\ 99,\ 4\right) = 0.9783$

Conclusion:

Do not reject H_0; there is not sufficient evidence to reject the claim $\mu_d = 0$. There is not sufficient evidence to reject the claim that the mean change in BMI for all students is equal to 0.

No; BMI does not appear to change during the freshman year.

b) $\bar{d} - E < \mu_d < \bar{d} + E$

$$\bar{d} - t_{\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{\alpha/2}\frac{s_d}{\sqrt{n}}$$

$$0.01 - 2.776\left(\frac{0.771}{\sqrt{5}}\right) < \mu_d < 0.01 + 2.776\left(\frac{0.771}{\sqrt{5}}\right)$$

$$-0.947 < \mu_d < 0.967$$

Yes; the confidence interval includes 0, which suggests that the mean of the differences could be 0 and that there is no change in BMI during the freshman year

Exercise

Listed below are body temperature (in °F) of subjects measured at 8:00 AM and at 12:00 AM. Construct a 95% confidence interval estimate of the difference between the 8:00 AM temperatures and the 12:00 AM temperatures. Is body temperature basically the same at both times?

8:00 AM	97.0	96.2	97.6	96.4	97.8	99.2
12:00 AM	98.0	98.6	98.8	98.0	98.6	97.6

Solution

	97.0	96.2	97.6	96.4	97.8	99.2
	98.0	98.6	98.8	98.0	98.6	97.6
Difference = d	−1.0	−2.4	−1.2	−1.6	−0.8	1.6
d^2	1	5.76	1.44	2.56	.64	2.56

$n = 6;\quad \sum d = -5.4;\quad \sum d^2 = 13.96\quad \bar{d} = \frac{\sum d}{n} = \frac{-5.4}{6} = .9$

$$s_d^2 = \frac{n\sum d^2 - (\sum d)^2}{n(n-1)} = \frac{6(13.96) - (-5.4)^2}{6(5)} = 1.82$$

$$s_d = \sqrt{1.82} = 1.349$$

$df = 5 \quad and \quad \alpha = 0.05$

$$t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.571$$

$$\bar{d} - t_{\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{\alpha/2}\frac{s_d}{\sqrt{n}}$$

$$-0.90 - 2.571\left(\frac{1.349}{\sqrt{6}}\right) < \mu_d < -0.90 + 2.571\left(\frac{1.349}{\sqrt{6}}\right)$$

$$-2.32\ °F < \mu_d < 0.52\ °F$$

Yes; since the confidence intervals includes 0, body temperature is basically the same at both times.

Exercise

Listed below are systolic blood pressure measuremetns (mm Hg) taking rom the right and left arms of the same woman. Use a 0.05 significance level to test for a difference the measurements from the two arms. What do you conclude?

Right arm	102	101	94	79	79
Left arm	175	169	182	146	144

Solution

	102	101	94	79	79
	175	169	182	146	144
Difference = d	−73	−68	−88	−67	−65
d^2	5329	4624	7744	4489	4225

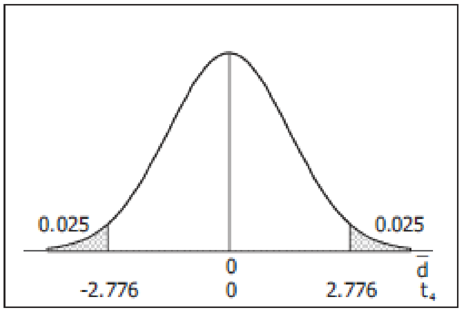
$n = 5; \qquad \sum d = -361; \quad \sum d^2 = 26,411 \quad \bar{d} = \frac{\sum d}{n} = \frac{-361}{5} = -72.2$

$$s_d^2 = \frac{n\sum d^2 - (\sum d)^2}{n(n-1)} = \frac{5(26,411) - (-361)^2}{5(4)} = 86.6947$$

$$s_d = \sqrt{86.6947} \underline{= 9.311}$$

$df = 4 \quad and \quad \alpha = 0.05$

$t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.776$



$$t_{\bar{d}} = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

$$= \frac{-72.2 - 0}{\frac{9.311}{\sqrt{5}}}$$

$$\underline{= -17.339}$$

$P - value = 2 \cdot tcdf\,(-99, \;\; -17.338, \;\; 4) \underline{= 6.488E - 5 = 0.00006}$

Conclusion:

Reject H_0; there is sufficient evidence to conclude that $\mu_d \neq 0$ (in fact, that $\mu_d < 0$). There is sufficient evidence to support the claim that there is a difference in measurements between the two arms. The statistical conclusion is that the right arm. Since the right and left arms should yield the same measurements, the practical conclusion is that a mistake has been made. The most obvious explanation is that diastolic (and not the systolic) values were mistakenly recorded for the right arm. Further investigation is definitely in order.

Exercise

As part of the National Health and Nutrition Examination Survey, the Department of Health and Human Services obtained self-reported heights and measured heights for males ages 12 – 16. All measurement are in inches. Listed below are sample results

Reported height	68	71	63	70	71	60	65	64	54	63	66	72
Measured height	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67.0	55.6	74.2	65.0	70.8

a) Is there sufficient evidence to support the claim that there is a difference between self-reported heights and measured heights of males? Use a 0.05 significance level.

b) Construct a 95% confidence interval estimate of the man difference between reported heights and measured heights. Interpret the resulting confidence interval, and comment on the implications of whether the confidence interval limits contain 0.

Solution

a)

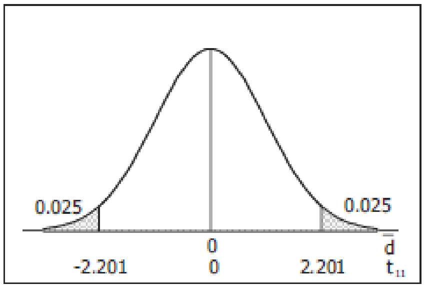
	68	71	63	70	71	60	65	64	54	63	66	72
	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67.0	55.6	74.2	65.0	70.8
Difference = d	0.1	1.1	–1.9	1.7	0.7	–0.6	0.5	–3.0	–1.6	–11.2	1.0	1.2
d^2	.01	1.21	3.61	2.89	.49	.36	.25	9	2.56	125.44	1	1.44

$n=12;\qquad \sum d=-12.0;\quad \sum d^2=148.26\quad \bar{d}=\dfrac{\sum d}{n}=\dfrac{-12}{12}=-1.0$

$$s_d=\sqrt{\frac{n\sum d^2-\left(\sum d\right)^2}{n(n-1)}}$$

$$=\sqrt{\frac{12(148.26)-(-12)^2}{12(11)}}$$

$$\underline{=3.52}$$



Original claim: $\mu_d\neq 0$ *inches*

$$H_0:\ \mu_d=0$$

$$H_1:\ \mu_d\neq 0$$

$df=11\quad and\quad \alpha=0.05$

$t=\pm t_{\alpha/2}=\pm t_{0.025}=\pm 2.201$

$$t_{\bar{d}}=\frac{\bar{d}-\mu_d}{\frac{s_d}{\sqrt{n}}}$$

$$=\frac{-1.0-0}{\frac{3.52}{\sqrt{12}}}$$

$$\underline{=-0.984}$$

$P-value = 2 \cdot tcdf\,(-99,\ -0.984,\ 11)$

$\qquad = 0.3461$

Conclusion:

Do not reject H_0; there is not sufficient evidence to reject the claim $\mu_d \neq 0$. There is not sufficient evidence to support the claim that there is a difference between self-reported heights and measured height of such males.

b) $df = 11 \quad and \quad \alpha = 0.05$

$$\bar{d} - t_{\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{\alpha/2}\frac{s_d}{\sqrt{n}}$$

$$-1.0 - 2.201\left(\frac{3.52}{\sqrt{12}}\right) < \mu_d < -1.0 + 2.201\left(\frac{3.52}{\sqrt{12}}\right)$$

$$-3.2\ in < \mu_d < 1.2\ in$$

Since the confidence interval contains 0, there is no significant difference between the reported and measured heights.

Exercise

Listed below are combined city – highway fuel consumption ratings (in miles/gal) for different cars measured under both the old rating system and a new rating system introducing in 2008. The new ratings were implemented in response to complaints that the old ratings were too high. Use a 0.01 significance level to test the claim the old ratings are higher than the new ratings.

Old rating	16	18	27	17	33	28	33	18	24	19	18	27	22	18	20	29	19	27	20	21
New rating	15	16	24	15	29	25	29	16	22	17	16	24	20	16	18	26	17	25	18	19

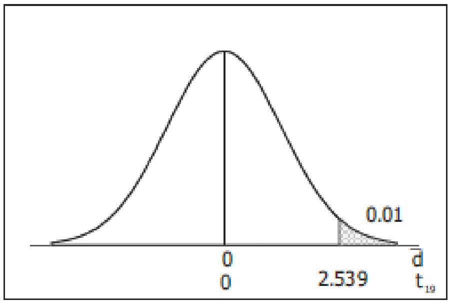
Solution

	16	18	27	17	33	28	33	18	24	19	18	27	22	18	20	29	19	27	20	21
	15	16	24	15	29	25	29	16	22	17	16	24	20	16	18	26	17	25	18	19
Diff = d	1	2	3	2	4	3	4	2	2	2	2	3	2	2	2	3	2	2	2	2
d^2	1	4	9	4	16	9	16	4	4	4	4	9	4	4	4	9	4	4	4	4

$n = 20; \qquad \sum d = 47; \quad \sum d^2 = 121 \quad \bar{d} = \frac{\sum d}{n} = \frac{47}{20} = 2.35$

$$s_d = \sqrt{\frac{n\sum d^2 - \left(\sum d\right)^2}{n(n-1)}} = \sqrt{\frac{20(121) - (47)^2}{20(19)}}$$

$\quad = 0.745$



Original claim: $\mu_d > 0$ *mpg*

$$H_0 : \mu_d = 0$$

$$H_1 : \mu_d > 0$$

$df = 19 \quad and \quad \alpha = 0.01$

$$t = t_{\alpha} = t_{0.01} = 2.539$$

$$t_{\bar{d}} = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

$$= \frac{2.35 - 0}{\frac{0.745}{\sqrt{20}}}$$

$$\underline{= 14.104}$$

$$P - value = tcdf\left(14.104,\ 99,\ 19\right)$$

$$\underline{= 9.093E - 12\ \approx 0}$$

Conclusion:

Reject H_0; there is sufficient evidence to conclude that $\mu_d > 0$. There is sufficient evidence to support the claim that the old ratings are higher than the new ratings.

Exercise

Listed below are 2 tables. Construct a 95% confidence interval estimate of the mean of the differences between weights of discarded paper and weights of discarded plastic. Which seems to weigh more: discarded paper or discarded plastic?

Paper

2.41	7.57	9.55	8.82	8.72	6.96	6.83	11.42	16.08	6.38	13.05	11.36	15.09
2.80	6.44	5.86	11.08	12.43	6.05	13.61	6.98	14.33	13.31	3.27	6.67	17.65
12.73	9.83	16.39	6.33	9.19	9.41	9.45	12.32	20.12	7.72	6.16	7.98	9.64
8.08	10.99	13.11	3.26	1.65	10.00	8.96	9.46	5.88	8.26	12.45	10.58	5.87
8.78	11.03	12.29	20.58	12.56	9.92	3.45	9.09	3.69	2.61			

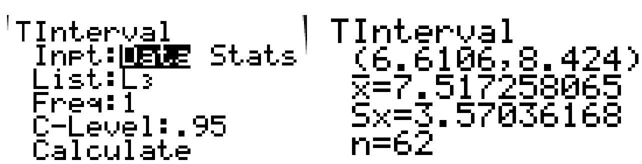
Plastic

0.27	1.41	2.19	2.83	2.19	1.81	0.85	3.05	3.42	2.10	2.93	2.44	2.17
1.41	2.00	0.93	2.97	2.04	0.65	2.13	0.63	1.53	4.69	0.15	1.45	2.68
3.53	1.49	2.31	0.92	0.89	0.80	0.72	2.66	4.37	0.92	1.40	1.45	1.68
1.53	1.44	1.44	1.36	0.38	1.74	2.35	2.30	1.14	2.88	2.13	5.28	1.48
3.36	2.83	2.87	2.96	1.61	1.58	1.15	1.28	0.58	0.74			

Solution

Using Ti-84, store paper into List 1 and plastic in List 2

To create list 3: **[2nd] 1** (L1) **– [2nd] 2** (L2) [***STO→***] **[2nd] 3** (L3)



$6.6107\ lbs < \mu_d < 8.424\ lbs$

Since the confidence interval includes only positive values, there discarded paper appears to weigh more than the discarded plastic.

Exercise

Suppose you wish to test the claim that μ_d, the mean value of the differences d for a population of paired data, is different from 0. Given a sample of $n = 23$ and a significance level of $\alpha = 0.05$, what criterion would be used for rejecting the null hypothesis?

Solution

Given: $n = 23 \Rightarrow df = 23 - 1 = 22$ and $\alpha = 0.05$

Degrees of Freedom	Area in Two Tails 0.01	0.02	0.05	0.10	0.20
22	**2.819**	2.508	**2.074**	1.717	1.321

To reject null hypothesis if test statistic is: $|t| > 2.074$ or $< -2.074 \quad or \quad > 2.074$

Exercise

Assume that he paired data came from a population that is normally distributed. Using a 0.05 significance level, find $\bar{d}$, s_d, the t test statistic, and the critical values to test the claim that $\mu_d = 0$

x	14	8	4	14	3	12	4	13
y	15	8	7	13	5	11	6	15

Solution

	14	8	4	14	3	12	4	13	
	15	8	7	13	5	11	6	15	
Difference = d	−1	0	−3	1	−2	1	−2	−2	
$(d - \bar{d})^2$	0	1	4	4	1	4	1	1	16

$$\bar{d} = \frac{-1+0-3+1-2+1-2-2}{8} = -1$$

$$s_d = \sqrt{\frac{\sum(d-\bar{d})^2}{n-1}} = \sqrt{\frac{16}{7}} \approx 1.5119$$

$$t_{\bar{d}} = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

$$= \frac{-1-0}{\frac{1.5119}{\sqrt{8}}}$$

$$= -1.8708$$

$df = 7 \quad and \quad \alpha = 0.05$

$t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.365$

Exercise

Assume that he paired data came from a population that is normally distributed. Using a 0.05 significance level, find $\bar{d}$, s_d, the t test statistic, and the critical values to test the claim that $\mu_d = 0$

x	12	5	1	20	3	16	12	8
y	7	10	5	15	7	14	10	13

Solution

	12	5	1	20	3	16	12	8
	7	10	5	15	7	14	10	13
Difference = d	5	−5	−4	5	−4	2	2	−5
$\left(d-\bar{d}\right)^2$	25	25	16	25	16	4	4	25

$$\bar{d} = \frac{5-5-4+5-4+2+2-5}{8} = -0.5$$

$$s_d = \sqrt{\frac{n\sum d^2 - \left(\sum d\right)^2}{n(n-1)}} = \sqrt{\frac{8(140)-(-0.5)^2}{7}} \approx 4.440$$

$$t_{\bar{d}} = \frac{\bar{d}-\mu_d}{\frac{s_d}{\sqrt{n}}}$$

$$= \frac{-0.5-0}{\frac{4.440}{\sqrt{8}}}$$

$$= -0.319$$

$df = 7 \quad and \quad \alpha = 0.05$

$t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.365$