MATHEMATICS 360-255-LW
Quantitative Methods II
Martin Huard
Fall 2009

XV – Inferences for Dependent Samples

SOLUTIONS

1. A company sent seven of its employees to attend a course on building self-confidence. These employees were evaluated for their self-confidence before and after attending this course. The following table gives the scores (on a scale of 1 to 15, 1 being the lowest and 15 being the highest score) of these employees before and after they attended the course. Assume that the population of paired differences has a normal distribution

Before	8	5	4	9	6	8	5
After	10	7	5	11	6	7	9
A-B	2	2	1	2	0	-1	4

a) Construct a 95% confidence interval for the mean of the population paired difference.

Step 1 Assumptions: The sampled populations are normally distributed.

Step 2 a) Test statistic: t with $df = n-1 = 6$

b) Level of confidence: $1-\alpha = 0.95$ or $\alpha = 0.05$

Step 3 Point estimate: $\bar{d} = 1.429$

Step 4 a) $t_{\left(df,\frac{\alpha}{2}\right)} = t_{(6,0.025)} = 2.447$

b) $E = t_{\left(df,\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}} = 2.447 \frac{1.618}{\sqrt{7}} = 1.50$

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

$1.43 - 1.50 < \mu_d < 1.43 + 1.50$

$-0.07 < \mu_d < 2.93$

Step 5 The 95% confidence interval for the mean difference in the self-confidence score is -0.07 to 2.93.

b) Test at the 1% significance level if attending this course increases then mean score of employees. Try with both approaches, the classical and the *p*-value.

Classical Approach

Step 1 Assumptions: The populations are normally distributed.

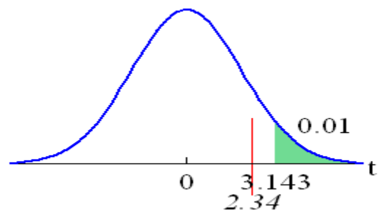
Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n - 1 = 6$

b) Right-tailed test with $\alpha = 0.01$

c) $t_{(df,\alpha)} = t_{(6,0.01)} = 3.143$



Step 4 $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{1.429 - 0}{\frac{1.618}{\sqrt{7}}} = 2.34$

Step 5 a) t is not in the critical region

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 1% level of significance to conclude that attendance to this course increases the mean score of employees.

p-value approach

Step 1 Assumptions: The populations are normally distributed

Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n - 1 = 6$

b) Right-tailed test with $\alpha = 0.01$

Step 4 a) $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{1.429 - 0}{\frac{1.618}{\sqrt{7}}} = 2.34$

b) $0.027 < p\text{-value} < 0.031$

Step 5 a) $p\text{-value} > 0.027 > \alpha = 0.01$

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 1% level of significance to conclude that attendance to this course increases the mean score of employees.

2. A private agency claims that the crash course it offers significantly increases the writing speed of secretaries. The following table gives the scores of eight secretaries before and after they attended the course.

Before	81	75	89	91	65	70	90	69
After	97	72	93	110	78	69	115	75
A-B	16	-3	4	19	13	-1	25	6

Assume that the writing speeds before and after attending the course are normally distributed.

a) Construct a 90% confidence interval for the mean increase in writing speed of secretaries.

Step 1 Assumptions: The sampled populations are normally distributed.

Step 2 a) Test statistic: t with $df = n-1 = 7$

b) Level of confidence: $1-\alpha = 0.90$ or $\alpha = 0.10$

Step 3 Point estimate: $\bar{d} = 9.875$

Step 4 a) $t_{\left(df,\frac{\alpha}{2}\right)} = t_{(7,0.05)} = 1.895$

b) $E = t_{\left(df,\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}} = 1.895 \frac{9.949}{\sqrt{7}} = 6.664$

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

$9.87 - 6.66 < \mu_d < 9.87 + 6.66$

$3.21 < \mu_d < 16.55$

Step 5 The 90% confidence interval for the mean increase in writing speed of secretaries is 3.21 to 16.55.

b) Using a 5% level of significance, can you conclude if attending this course increases the writing speed of secretaries? Try with both approaches, the classical and the *p*-value.

Classical Approach

Step 1 Assumptions: The populations are normally distributed.

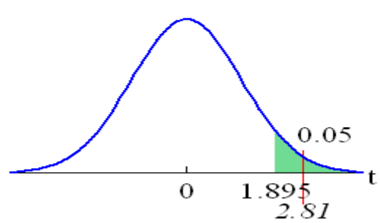
Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n-1 = 7$

b) Right-tailed test with $\alpha = 0.05$

c) $t_{(df,\alpha)} = t_{(7,0.05)} = 1.895$



Step 4 $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{9.875 - 0}{\frac{9.949}{\sqrt{8}}} = 2.81$

Step 5 a) t is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that attendance to this course increases the mean writing speed of secretaries.

p-value approach

Step 1 Assumptions: The populations are normally distributed

Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n - 1 = 7$

b) Right-tailed test with α = 0.05

Step 4 a) $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{9.875 - 0}{\frac{9.949}{\sqrt{8}}} = 2.81$

b) $0.011 < p\text{-value} < 0.013$

Step 5 a) $p\text{-value} < 0.013 < \alpha = 0.05$

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that attendance to this course increases the mean writing speed of secretaries.

3. In a study on human behavior, a psychologist studied the incidence of suicide in five randomly selected communities of moderate size both before and after publicity was given to the suicide of a famous singer. Assume that the number of suicides in a community is normally distributed.

Before	3	4	9	7	5
After	6	7	10	9	8
A-B	3	3	1	2	3

a) Construct a 95% confidence interval for the difference in the number of suicides between before and after the publicity.

Step 1 Assumptions: The sampled populations are normally distributed.

Step 2 a) Test statistic: t with $df = n - 1 = 4$

b) Level of confidence: $1 - \alpha = 0.95$ or $\alpha = 0.05$

Step 3 Point estimate: $\bar{d} = 2.400$ suicides

Step 4 a) $t_{\left(df, \frac{\alpha}{2}\right)} = t_{(4, 0.025)} = 2.776$

b) $E = t_{\left(df, \frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}} = 2.776 \frac{0.8944}{\sqrt{5}} = 1.111$

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

$2.40 - 1.11 < \mu_d < 2.40 - 1.11$

$1.29 < \mu_d < 3.51$

Step 5 The 95% confidence interval for the mean difference in the number of suicides before and after the publicity is 1.29 to 3.51 suicides.

b) Using a 5% level of significance, can you conclude that publicity has an effect on suicide? Try with both approaches, the classical and the *p*-value.

Classical Approach

Step 1 Assumptions: The populations are normally distributed.

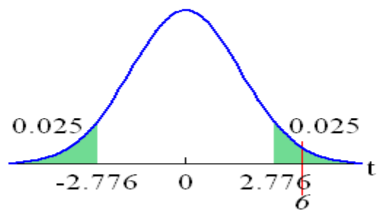
Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d \neq 0$

Step 3 a) Test statistic: t with $df = n - 1 = 4$

b) Two-tailed test with $\alpha = 0.05$

c) $t_{(df, \frac{\alpha}{2})} = t_{(7, 0.025)} = 2.776$



Step 4 $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{2.400 - 0}{\frac{0.8944}{\sqrt{5}}} = 6.00$

Step 5 a) t is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that publicity has an effect on suicide.

p-value approach

Step 1 Assumptions: The populations are normally distributed

Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d \neq 0$

Step 3 a) Test statistic: t with $df = n - 1 = 4$

b) Two-tailed test with $\alpha = 0.05$

Step 4 a) $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{2.400 - 0}{\frac{0.8944}{\sqrt{5}}} = 6.00$

b) $p\text{-value} < 2 \cdot 0.008 = 0.016$

Step 5 a) $p\text{-value} < 0.016 < \alpha = 0.05$

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that publicity has an effect on suicide.

4. A psychologist wants to examine the effectiveness of family therapy as a treatment for anorexia. Seventeen girls were selected at random, and were weighed before and after the treatments. The weights of the girls, in kg, are given below. Assume that the weights of girls are normally distributed.

Before	42.3	42.1	43	41.1	43.2	38.7	37.6	46.3	39.4
After	47.3	46.3	43.2	46.3	50.1	36.9	37.3	49.8	37.2
A-B	5	4.2	0.2	5.2	6.9	-1.8	-0.3	3.5	-2.2

Before	39.9	40.8	41.3	38.3	41.5	52.3	42.0	35.6
After	42.2	43.5	50.2	37.5	39.2	52.4	41.8	39.6
A-B	2.3	2.7	8.9	-0.8	-2.3	0.1	-0.2	4

a) Construct a 99% confidence interval for the mean weight gain by girls following the family therapy.

Step 1 Assumptions: The sampled populations are normally distributed.

Step 2 a) Test statistic: t with $df = n-1 = 16$

b) Level of confidence: $1-\alpha = 0.99$ or $\alpha = 0.01$

Step 3 Point estimate: $\bar{d} = 2.082$ kg

Step 4 a) $t_{\left(df,\frac{\alpha}{2}\right)} = t_{(16,0.005)} = 2.921$

b) $E = t_{\left(df,\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}} = 2.921 \frac{3.334}{\sqrt{17}} = 2.361$

c) $$\bar{d} - E < \mu_d < \bar{d} + E$$
$$2.08 - 2.36 < \mu_d < 2.08 + 2.36$$
$$-0.28 < \mu_d < 4.44$$

Step 5 The 99% confidence interval for the mean weight gain by girls who follow the family therapy is -0.28 kg to 4.44 kg.

b) Using a 10% level of significance, can you conclude that family therapy is an effective treatment for anorexia? Try with both approaches, the classical and the *p*-value.

Classical Approach

Step 1 Assumptions: The populations are normally distributed.

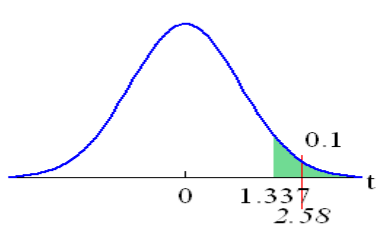
Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n-1 = 16$

b) Two-tailed test with $\alpha = 0.10$

c) $t_{(df,\alpha)} = t_{(16,0.1)} = 1.337$



Step 4 $$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{2.082 - 0}{\frac{3.334}{\sqrt{17}}} = 2.58$$

Step 5 a) t is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that family therapy is an effective treatment for anorexia.

p-value approach

Step 1 Assumptions: The populations are normally distributed

Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n - 1 = 16$

b) Two-tailed test with α = 0.10

Step 4 a) $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{2.082 - 0}{\frac{3.334}{\sqrt{17}}} = 2.58$

b) $0.009 < p\text{-value} < 0.012$

Step 5 a) $p\text{-value} < 0.012 < \alpha = 0.10$

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that family therapy is an effective treatment for anorexia.

5. Many mothers say that they can recognize the cries of their own babies and that they can distinguish between a cry of pain and a hunger cry. A psychologist studied this phenomenon using the following experiment. A random sample of mothers listened to tape-recorded sets of five cries from different babies, one of which was their own. They then had to decide which baby was theirs. Each mother heard 20 such sets, in which 10 were cries of hungry babies and 10 were cries produced by a slight pin prick on a foot. The results are presented in the following table, where the correct number of identifications (out of 10) is shown. The psychologist claims that mothers are more successful in picking out their own babies when a hunger cry is involved, since the mothers have more experience with that situation. Assume that the number of correct identifications is normally distributed.

Hungry Cry	6	6	6	5	3	7	9	2
Pain Cry	5	4	7	3	2	6	4	3
H - C	1	2	-1	2	1	1	5	-1

a) Construct a 90% confidence interval for the mean difference in the number of correct identification between the Hungry cry and the Pain cry.

Step 1 Assumptions: The sampled populations are normally distributed.

Step 2 a) Test statistic: t with $df = n - 1 = 7$

b) Level of confidence: $1 - \alpha = 0.90$ or $\alpha = 0.10$

Step 3 Point estimate: $\bar{d} = 1.25$

Step 4 a) $t_{\left(df, \frac{\alpha}{2}\right)} = t_{(7, 0.05)} = 1.895$

b) $E = t_{\left(df, \frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}} = 1.895 \frac{1.909}{\sqrt{8}} = 1.278$

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

$1.25 - 1.29 < \mu_d < 1.25 + 1.29$

$-0.04 < \mu_d < 2.54$

Step 5 The 90% confidence interval for the mean difference in the number of correct identification between the Hungry and the Pain cry is -0.04 to 2.54.

b) Test the psychologists claim using a 10% level of significance? Try with both approaches, the classical and the *p*-value.

Classical Approach

Step 1 Assumptions: The populations are normally distributed.

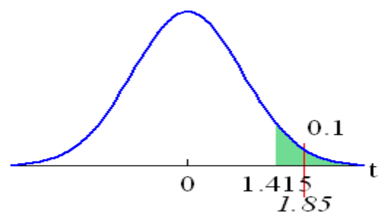
Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n - 1 = 7$

b) Two-tailed test with $\alpha = 0.10$

c) $t_{(df,\alpha)} = t_{(7,0.1)} = 1.415$



Step 4 $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{1.25 - 0}{\frac{1.909}{\sqrt{8}}} = 1.85$

Step 5 a) t is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that mothers are more successful in picking out their own babies when a hunger cry is involved.

p-value approach

Step 1 Assumptions: The populations are normally distributed

Step 2 $H_O: \mu_d = 0$

$H_A: \mu_d > 0$

Step 3 a) Test statistic: t with $df = n - 1 = 7$

b) Two-tailed test with $\alpha = 0.10$

Step 4 a) $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{1.25 - 0}{\frac{1.909}{\sqrt{8}}} = 1.85$

b) $0.050 < p\text{-value} < 0.057$

Step 5 a) $p\text{-value} < 0.057 < \alpha = 0.10$

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that mothers are more successful in picking out their own babies when a hunger cry is involved.