EIN 5226

Comparison Tests Continuous - Means

Chapter 19 Sections 19.1-5,19.8-14

Karen E. Schmahl Ph.D., P.E.

1

Introduction

In previous chapters, we learned how to construct confidence intervals and perform hypothesis tests for a single mean compared to a defined value.

Often, we wish to compare the differences between two samples.

- Large sample sizes
- Small sample sizes, variances assumed equal
- Small sample sizes, variances assumed not equal
- Paired comparisons

2

Examples – Comparison Tests Continuous Response

- Is there a difference between two vendors products?
- Did the new tooling activity really improve the process?
- Do the two machines have the same capability
- Is the time to failure one electrical component the same as another?
- Is the strength of the two welds the same?
- Is the service time for customers the same for morning lunch shifts in a fast food restaurant?

3

19.2 Comparing Continuous Data Responses

- Mean
  - Sample outputs can be compared to determine if a difference is large enough to be significant.
  - Null hypothesis for the comparison test is that there is no difference, while the alternative hypothesis is that there is a difference.
  - Comparison test of means is robust to the shape of the underlying distribution not being normal.

4

Statistical Inference - Using p values

Steps in Hypothesis Testing

1. Identify your objective
2. State the null hypothesis, H_0
3. State the alternative hypotheses, H_a.
4. Calculate the appropriate test statistic
5. Compute the p value of the test statistic
6. Determine the acceptable risk
7. Compare the p value of test statistic to the to the acceptable risk.

5

Tests comparing Two Means

Large Sample Sizes (n≥30, or σ known)

	Two-tailed	One-tailed
Null hypothesis	$H_0: \mu_1=\mu_2$	
Alternate hypothesis	$H_a: \mu_1\neq\mu_2$	$\mu_1>\mu_2$ or $\mu_1<\mu_2$
Test statistic	$z_{test} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$	
p (reject if p<α)	sum the areas in the tails, cut off by Z_{test}	Area to right (>) or left (<) of Z_{test}

6

Vendor Comparison

The product of two vendors are being compared. The specification on a critical characteristic is 60 ± 5. Sample data for the vendors is shown below.

Vendor	n	average	Std. Dev.
Big Sky Co.	41	61.7	1.1
Mellow Yellow Co.	36	62.1	0.8

With 95% confidence, can it be concluded that there is a difference between the averages of the vendors' processes?

What are the null and alternate hypotheses for this problem?

A) H_0: $\mu = 60$ H_a: $\mu \neq 60$

B) H_0: $\mu_B = \mu_M$ H_a: $\mu_B \neq \mu_M$

C) H_0: $\mu_M = \mu_B$ H_a: $\mu_M \neq \mu_B$

D) Either B) or C) could be used.

7

Vendor Comparison

$$z_{test} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Vendor	n	average	Std. Dev.
Big Sky Co.	41	61.7	1.1
Mellow Yellow Co.	36	62.1	0.8

With 95% confidence, can it be concluded that there is a difference between the averages of the vendors' processes?

The Z_{test} value for the problem is

A) 1.84   B) 1.81   C) 8.45   D) 3.68

The p value for the problem is

A) .0500   B) .0329   C) .0658   D) .0702

8

Vendor Comparison

Vendor	n	average	Std. Dev.
Big Sky Co.	41	61.7	1.1
Mellow Yellow Co.	36	62.1	0.8

With 95% confidence, can it be concluded that there is a difference between the averages of the vendors' processes?

This analysis would result in the following conclusion:

T / F p is less than α, therefore reject the null hypothesis.

T / F With 95% confidence we cannot conclude there is a difference between the vendor averages.

9

Two-Sample T-Test and CI

```
Sample   N    Mean  StDev  SE Mean
1       41   61.70   1.10     0.17
2       36  62.100  0.800     0.13

Difference = μ (1) - μ (2)
Estimate for difference:  -0.400
95% CI for difference:  (-0.834, 0.034)
T-Test of difference = 0 (vs ≠): T-Value = -1.84
P-Value = 0.070  DF = 72
```

Minitab – does not do a 2 sample Z

Theoretically we do not know the variances for the processes, so the 2 sample t is more precise.

10

Tests comparing Two Means
Small Sample Sizes (n<30,) σ equal

Variances unknown and assumed equal	Two-tailed	One-tailed
Null hypothesis	$H_0: \mu_1 = \mu_2$	
Alternate hypothesis	$H_a: \mu_1 \neq \mu_2$	$\mu_1 > \mu_2$ or $\mu_1 < \mu_2$
Critical value ($v = n_1 + n_2 - 2$ degrees of freedom)	$\pm t_{\alpha/2,v}$	$+ t_{\alpha,v}$ or $- t_{\alpha,v}$
Test statistic $t_{test} = \dfrac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	Where pooled standard deviation is $s_p = \sqrt{\dfrac{(n_1-1)s_1^{\,2} + (n_2-1)s_2^{\,2}}{n_1+n_2-2}}$	
p (reject if p<α)	sum the areas in the tails	Area to right (>) or left (<)

11

Practice

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky? (Assume equal variances)

What are the null and alternate hypotheses for this problem?

A) H_0: $\mu_M = \mu_B$ H_a: $\mu_M \neq \mu_B$

B) H_0: $\mu_M = \mu_B$ H_a: $\mu_M < \mu_B$

C) H_0: $\mu_M = \mu_B$ H_a: $\mu_M > \mu_B$

12

Practice

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky? (Assume equal variances)

What is the estimate of the common standard deviation for the data sets?

$$s_p = \sqrt{\frac{(n_1-1)s_1{}^2 + (n_2-1)s_2{}^2}{n_1+n_2-2}} = \sqrt{\frac{(18-1)1.1^2 + (24-1).08^2}{18+24-2}} = 0.9240$$

What degrees of freedom should be used for the problem?

A) 24 B) 42 C) 41 D) 40

13

Practice

$$t_{test} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

s_p=.9240
ν =40

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky? (Assume equal variances)

What is the value of the test statistic for the problem?

A) -1.362 B) -4.288 C) -1.645 D) 1.960

14

Practice

$$t_{test} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

s_p=.9240
ν =40

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky? (Assume equal variances)

What is the value of t at the critical point?

A) $-t_{.05,40}$ B) -1.684 C) $t_{.10,40}$ D) -1.303 E) C and D

15

Practice

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

s_p=.9240
v =40

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky? (Assume equal variances)

T / F $|t_{test}| > |t_{critical}|$ therefore reject the null hypothesis

T / F At a 90% level of confidence it can be concluded that the mean of the Mellow Yellow process is larger than the mean of the Big Sky process.

16

Tests comparing Two Means
Small Sample Sizes (n<30,) σ not equal

Variances unknown and NOT equal	Two-tailed	One-tailed
Null hypothesis	$H_0: \mu_1=\mu_2$	
Alternate hypothesis	$H_a: \mu_1\neq\mu_2$	$\mu_1>\mu_2$ or $\mu_1<\mu_2$
Critical value	$\pm t_{\alpha/2, v}$	$+ t_{\alpha, v}$ or $- t_{\alpha, v}$
Test statistic $t_{test} = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$	$v = \frac{\left[\left(s_1^2/n_1\right)+\left(s_2^2/n_2\right)\right]^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}}$ (always round down)	
p (reject if p<α)	sum the areas in the tails	Area to right (>) or left (<)

17

Practice

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky?

(Assume unequal variances)

$$v = \frac{\left[\left(s_1^2/n_1\right)+\left(s_2^2/n_2\right)\right]^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}} = 24.78 \gg 24$$

18

Practice

$$t_{test} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}}$$

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky?

v=24 (Assume unequal variances)

What is the value of the test statistic for the problem?

A) -1.263 B) -4.260 C) -3.991 D) 1.291

19

Practice

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky?

v=24 (Assume unequal variances)

What is the value of t at the critical point?

A) $-t_{.05,24}$ B) 1.735 C) $t_{.010,24}$ D) -1.318 E) C and D

20

Practice

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky?

v=24 (Assume unequal variances)

H_0: $\boldsymbol{\mu}_M = \boldsymbol{\mu}_B$ H_a: $\boldsymbol{\mu}_M > \boldsymbol{\mu}_B$

t_{test} = -1.263 $t_{crit} = -t_{.10,24}$ = -1.318

$|t_{test}|\ not > |t_{critical}|$ therefore cannot reject the null hypothesis

21

Vendor	n	average	Std. Dev.
Big Sky Co.	16	61.7	1.1
Mellow Yellow Co.	26	62.1	0.8

With 90% confidence, can it be concluded that Mellow Yellow has a greater process average than Big Sky?

H_0: $\mu_M = \mu_B$ H_a: $\mu_M > \mu_B$

Assume equal variances
$t_{test} = -1.362$ $t_{crit} = -t_{.10,40} = -1.303$
$|t_{test}| > |t_{critical}|$
therefore reject the null hypothesis

Assume unequal variances
$t_{test} = -1.263$ $t_{crit} = -t_{.10,24} = -1.318$
$|t_{test}|$ $not > |t_{critical}|$
therefore cannot reject the null hypothesis

22

Paired Comparison Test

- When sample groups not independent
- Examples: Before and after type studies
  Product comparisons
- Comparison is between average difference between the pairs
- Always done as a t-test

23

Paired Comparison Test
On two population means

	Two-tailed	One-tailed
Null hypothesis	$H_0: \mu_d = \mu_0$	
Alternate hypothesis	$H_a: \mu_d \neq \mu_0$	$\mu_d > \mu_0$ or $\mu_d < \mu_0$
Critical value (v = n-1 degrees of freedom)	$\pm t_{\alpha/2, v}$	$+ t_{\alpha, v}$ or $- t_{\alpha, v}$
Test statistic $t_{test} = \frac{\bar{D} - \mu_0}{S_d / \sqrt{n}}$	where $\bar{D}$ = average difference, S_d = standard deviation of the differences	
p (reject if p<α)	sum the areas in the tails	Area to right (>) or left (<)

Note: n = # of pairs

24

A company prided itself on its continuous improvement program. Over a period of several months, numerous small improvements were made in the shop

To assess the impact of the improvements on time to completion, 20 items were selected at random and times collected both before and after the improvement efforts.

The before and after times are listed in the table.

Item #	Date 1	Time 1	Date 2	Time 2
229	10/30/2016	47.62	2/14/2017	49.62
225	10/31/2016	48.45	2/15/2017	47.53
216	11/1/2016	45.08	2/16/2017	44.92
214	11/2/2016	42.63	2/17/2017	44.88
210	11/3/2016	44.78	2/18/2017	33.83
630	11/4/2016	48.40	2/19/2017	38.42
201	11/5/2016	42.62	2/20/2017	36.67
199	11/5/2016	42.58	2/21/2017	46.60
195	11/6/2016	55.13	2/22/2017	54.08
186	11/7/2016	53.70	2/23/2017	56.90
184	11/7/2016	78.70	2/24/2017	51.82
180	11/8/2016	40.15	2/25/2017	38.18
171	11/9/2016	46.73	2/26/2017	37.75
169	11/10/2016	38.52	2/27/2017	44.72
165	11/11/2016	38.88	2/28/2017	44.42
150	11/14/2016	50.57	3/1/2017	40.15
156	11/15/2016	56.70	3/2/2017	41.58
154	11/16/2016	35.15	3/2/2017	35.58
139	11/17/2016	45.07	3/3/2017	31.00
123	11/18/2016	49.27	3/3/2017	47.12

25

i	Item #	Time 1	Time 2	Difference
1	229	47.62	49.62	-2.00
2	225	48.45	47.53	0.92
3	216	45.08	44.92	0.17
4	214	42.63	44.88	-2.25
5	210	44.78	33.83	10.95
6	630	48.40	38.42	9.98
7	201	42.62	36.67	5.95
8	199	42.58	46.60	-4.02
9	195	55.13	54.08	1.05
10	186	53.70	56.90	-3.20
11	184	78.70	51.82	26.88
12	180	40.15	38.18	1.97
13	171	46.73	37.75	8.98
14	169	38.52	44.72	-6.20
15	165	38.88	44.42	-5.53
16	150	50.57	40.15	10.42
17	156	56.70	41.58	15.12
18	154	35.15	35.58	-0.43
19	139	45.07	31.00	14.07
20	123	49.27	47.12	2.15

A paired comparison test is appropriate as the times are not independent.

With 95% confidence, can you conclude that the current times are lower than the previous times.

$H_0: \mu_d = 0$ $H_a: \mu_d > 0$

$n=20$

$\overline{D} = 4.25$
$S_d = 8.42$

26

With 95% confidence, can you conclude that the current times are lower than the previous times.

$$t_{test} = \frac{\overline{D} - u_0}{S_d / \sqrt{n}}$$

$H_0: \mu_d = 0$ $H_a: \mu_d > 0$

$\overline{D} = 4.25$ $S_d = 8.42$ $n=20$

What is the value of the test statistic for the problem?

A) 1.45 B) 2.15 C) 3.32 D) 2.26

What degrees of freedom should be used for the problem?
A) 20 B) 19 C) 40 D) 41

What is the value of $t_{critical}$ for the hypothesis test?

A) 1.65 B) 2.85 C) 3.94 D) 2.26

27

With 95% confidence, can you conclude that the current times are lower than the previous times.

$$t_{test} = \frac{\overline{D} - u_0}{S_d / \sqrt{n}}$$

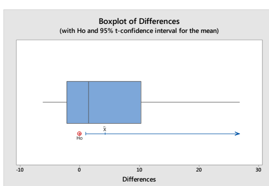
$H_0: \mu_d = 0$ $H_a: \mu_d > 0$

$\overline{D} = 4.25$ $S_d = 8.42$ $n=20$

T / F $|t_{test}| > |t_{critical}|$ therefore reject the null hypothesis

T / F At a 95% level of confidence it can be concluded that the continuous improvement program was effective at lowering processing times.

28




```
Paired T for initial - now

             N   Mean  StDev  SE Mean
initial     20  47.54   9.24     2.07
now         20  43.29   6.90     1.54
Difference  20   4.25   8.42     1.88


95% lower bound for mean difference: 0.99
T-Test of mean difference = 0 (vs > 0): T-Value = 2.26  P-Value = 0.018
```

29

i	Exam 1	2nd chance	Difference
1	81	65	-16
2	55	85	30
3	82	90	8
4	79	90	11
5	58	90	32
6	82	88	6
7	55	60	5
8	61	85	24
9	70	80	10
10	85	90	5
11	73	80	7
12	76	90	14
13	67	62	-5
14	64	47	-17
15	52	85	33
16	79	90	11
17	67	65	-2
18	82	90	8
19	64	80	16
20	79	75	-4
21	55	90	35
22	88	80	-8
23	64	90	26
24	67	80	13
25	64	77	13
26	61	80	19
27	52	62	10

<u>Paired Comparison</u>

Grades 2nd Chance Exam vs Exam 1

(Students scoring >90 on Exam 1 omitted)

With 90% confidence, can the taking of a 2nd chance exam improve a score by at least 5 points.

$H_0: \mu_d = 0$ $H_a: \mu_d > 5$

$\overline{D} = 10.52$ $S_d = 13.89$ $n=27$

30

Paired Comparison
Grades 2nd Chance Exam vs Exam 1
(Students scoring >90 on Exam 1 omitted)

$$t_{test} = \frac{\overline{D} - u_0}{S_d / \sqrt{n}}$$

$H_0: \mu_d = 0$ $H_a: \mu_d > 5$ 90% confidence

$\overline{D} = 10.52$ $S_d = 13.89$ $n=27$

What is the value of the test statistic for the problem?

A) 1.55 B) 2.06 C) 3.94 D) 10.73

What is the value of $t_{critical}$ for the hypothesis test?

A) 1.315 B) 1.314 C) 1.706 D) 1.703

T / F At a 90% level of confidence it can be concluded that the students would gain at least 5 points with a 2nd chance exam..

31

Related Assignments

Please see Blackboard for related assignments

32

COLLEGE OF ENGINEERING