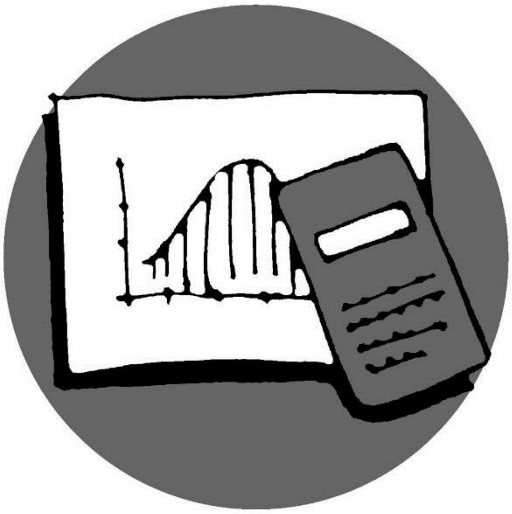
Library, Teaching and Learning

Least Significant Differences (LSDs) and Contrast Sums of Squares

QMET201



Lincoln University
Te Whare Wānaka o Aoraki
CHRISTCHURCH • NEW ZEALAND
New Zealand's specialist land-based university

MULTIPLE COMPARISONS

Having established that, for a number of treatments, there is a significant difference between at least one pair of means, the next process is to find which pair(s).

Least Significant Differences To find the LSD

- Calculate MS_{error} (= EMS = s_p^2) or read from ANOVA table.
- Calculate $SED = \sqrt{EMS \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

  where n_1, n_2 are the number of values used to calculate the two means being compared.

  Note that when $n_1=n_2=r$, the formula can be written:

  $$SED = \sqrt{\frac{2 \times EMS}{r}}$$

  where r is the number of replications.

- Find the relevant t , using df_{error} with required level of significance (***two-tailed***).
- Calculate $LSD = t \times SED$
- Compare this LSD with the differences between the pairs of means and make a decision as to which pairs ARE significantly different.

Example: *A Trial with 5 treatments was replicated 4 times. Given the following results, find which pairs of means are significantly different.*

Treatment	A	B	C	D	E
Mean	8.0	6.5	3.0	4.0	5.4

With $MS_{error} = 4.53 \; and \; df_{error} = 15$

Calculations:

- $$SED = \sqrt{\left(\frac{2 \times 4.53}{4}\right)} = 1.505$$
- t = 2.132 at 5% level and t = 2.947 at 1% level

  => 5% LSD = 2.132 x 1.505 = 3.21

- compare pairs of means to see if difference is larger than the LSD:

Results:

Means being compared	Difference	> 3.21	signif @ 5% ?
A and B	1.5	No	No
A and C	5	Yes	Yes
A and D	4	Yes	Yes
A and E	2.6	No	No
B and C	3.5	Yes	Yes
B and D	2.5	No	No
B and E	1.0	No	No
C and D	1	No	No
C and E	2.5	No	No
D and E	1.5	No	No

1% LSD = 2.947 x 1.505 = 4.435

Results:

Comparing	Difference	>4.435	signif @ 1% ?
A and C	5	Yes	Yes
A and D	4	No	No
B and C	3.5	No	No

That is, at the 5% level, mean of A is significantly different to C and D; mean of B is significantly different to C; no other means are significantly different.
At the 1% level, mean of A is significantly different to C; no other means are significantly different.

There are various ways of showing the above, including graphic methods. The way you display your findings is your own choice but you will probably find it much easier to interpret the above if you carry out the comparisons as follows:

- First arrange the means in order:

Treatment	A	B	E	D	C
Mean	8.0	6.5	5.4	4.0	3.0

- Now subtract the LSD from the largest mean and draw a line under all the means that are larger than this value. (These means are not sig. different from each other because the difference is less than the LSD).

- Repeat for the second largest mean, drawing a line as above.

- Continue until you have compared the smallest and second smallest means with each other.

Results:

Treatment	A	B	E	D	C
Mean	8.0	6.5	5.4	4.0	3.0

Note: A disadvantage of this method is that as the number of comparisons increases, so does the probability of Type 1 error (i.e., probability of saying means are different when they are not.)

.

Practise

1. The following is data from an experiment to investigate the crop yield from five different brands of seed. Each brand was tested in 4 different locations.

	Location					
Brand	A	B	C	D	Total	$(\text{Total})^2$
1	12	11	20	17	60	3600
2	2	0	18	2	22	484
3	4	10	12	8	34	1156
4	5	15	16	7	43	1849
5	1	4	14	11	30	900
Total	24	40	80	45	189	7989
$(\text{Total})^2$	576	1600	6400	2025	10601	

a) Construct the two-way analysis of variance table, and test the location effect for significance at the 1% level.
b) Establish which field(s) if any, are significantly different (use 1% LSD).
c) Establish which brand(s) if any, are significantly different.

2. Osage Orange, or hedge apple (*Maclura pomifer)*¸is a tree found commonly in t he Great Plains of the USA. It is used often for fence posts and rails. The following analysis was from an experiment to test the rot-resistance of the heartwood following different "accelerated aging" treatments (Yoshimura, *Faculty Agr. Mie Uni. Bull. 27:225).*

   Two species of fungus were used, white rot (*polyporus versicolor*) and brown-rot (*Poria monticola*).

   Two incubation periods were used, 90 days and 120 days, and there were three accelerated aging treatments **and** a control. The response variate measured was the percentage weight loss of heartwood wafers after the incubation periods. For the analysis, Log_{10}(% wt loss) was used to improve the residual plots.

Analysis of Variance for Log_{10}(%wt loss)					
Source	DF	SS	MS	F	P
Fungus	1	2.75102	2.7510	143.45	0.000
incubation	1	0.6041	0.6041	31.46	0.000
aging	3	11.1771	3.7257	194.27	0.000
F * Inc	1	0.1836	0.1863	9.703	0.003
F*aging	3	5.2324	1.7441	90.95	0.000
I*aging	3	0.6275	0.2092	10.91	0.000
F*I*A	3	0.1400	0.04667	2.43	$0.05 < p < 0.1$
Error	80	1.5342	0.0192		
Total	95	22.2526			

		Means	
Fungus	N		
1	48	0.27452	
2	48	0.61308	
incubation	N		
1	48	0.36447	
2	48	0.52313	
aging	N		
Control	24	0.08705	
Ether	24	0.13519	
Methanol	24	0.67197	
Eth-Meth	24	0.64835	
Fungus	**incubation**	N	Means
1	1	24	0.15113
1	2	24	0.39791
2	1	24	0.57782
2	2	24	0.64835
Fungus	**aging**	N	Means
1	Control	12	0.1370
1	Ether	12	0.1793
1	Methanol	12	0.4106
1	Eth-Meth	12	0.3712
2	Control	12	0.0371
2	Ether	12	0.0911
2	Methanol	12	0.9334
2	Eth-Meth	12	1.3908
Incubation	**aging**	N	Means
1	Control	12	0.0931
1	Ether	12	0.1313
1	Methanol	12	0.5233
1	Eth-Meth	12	0.7102
2	Control	12	0.0810
2	Ether	12	0.1391
2	Methanol	12	0.8206
2	Eth-Meth	12	1.0518

(e) Calculate the 5% LSD for incubation x aging interaction means and determine which means are significantly different

Solutions

1.

Source	df	SS	MS	f
Brand	4	211.2	52.8	3.7795
Location	3	334.15	111.38	7.9728
Error	12	167.6	13.97	
Total	19	712.95		

At 1%, $t_{0.005,12)}$ = 3.055 (two tailed remember)

$$LSD_{Location} = 3.055 \times \sqrt{\frac{13.97 \times 2}{5}} = 7.22$$

From the earlier data,

Location				
	A	B	C	D
Totals	24	40	80	45
Means	4.8	8	16	9

Differences: A - B = 3.2; A – C = 13.2*; A – D = 4.2
B – C = 8*; B – D = 1
C – D = 7

Differences marked * are significantly different.

A faster way of recognising the above result is to put the means in order:

Location				
	C	D	B	A
Totals	80	45	40	24
Means	16	9	8	4.8

Now examine the differences, stopping when the difference is no longer significant.

ie 16 – 4.8 sig; 16 – 8 sig; 16– 9 NS
9 – 4.8 NS (stop)
8 – 4.8 NS (stop)

$$LSD_{Brand} = 3.055 \times \sqrt{\frac{13.97 \times 2}{4}} = 8.074$$

Brand	Total	Mean
1	60	12
2	22	5.5
3	34	8.4
4	43	10.75
5	31	7.75

Put in order:

Brand	Total	Mean
1	60	12
4	43	10.75
3	34	8.4
5	31	7.75
2	22	5.5

12 – 7.75 NS
Since highest minus lowest is NS, none of the others will be significant.

No significant differences between brands.

2. e. $LSD = t \times \sqrt{\frac{EMS \times 2}{r}} = 1.99 \times \sqrt{\frac{0.0192 \times 2}{12}} = 0.11257$

Analyse the differences by assigning the same letter to all values whose difference is less than the LSD (a quicker method)

- Order the means for incubation x aging, and test against LSD:
- For ease of explaining this, each combination mean is assigned a letter
- Assign a letter (a) to the smallest mean
- Calculate the difference between this mean and following means. If the difference is less than the LSD, assign the same letter to these means. When the difference becomes greater than the LSD, given this mean (E) a new letter (b)

A	2	control	0.0810	a
B	1	control	0.0931	a
C	1	ether	0.1313	a
D	2	ether	0.1391	a
E	1	methanol	0.5233	b
F	1	eth-meth	0.7102	
G	2	methanol	0.8206	
H	2	eth-meth	1.0518	

- Now compare the following means with this mean and repeat the process:

A	2	control	0.0810	a
B	1	control	0.0931	a
C	1	ether	0.1313	a
D	2	ether	0.1391	a
E	1	methanol	0.5233	b
F	1	eth-meth	0.7102	c
G	2	methanol	0.8206	
H	2	eth-meth	1.0518	

- continue until all means have been compared:

A	2	control	0.0810	a
B	1	control	0.0931	a
C	1	ether	0.1313	a
D	2	ether	0.1391	a
E	1	methanol	0.5233	b
F	1	eth-meth	0.7102	c
G	2	methanol	0.8206	c
H	2	eth-meth	1.0518	d

Means with same letter are not significantly different.
That is, A, B, C and D are significantly different to E, F, G and H
E is significantly different to F, G and H
F and G are significantly different to H
H is significantly different to all.

CONTRASTS **To calculate Sums of Squares:**

You need to identify

- the particular treatment being investigated
- the number of levels for this treatment (this dictates which orthogonal polynomial coefficients to use)
- the totals for each of these treatments (T)
- the number of values used to calculate these totals (r)
- the type of contrast – linear, quadratic…
- the coefficients for this contrast from the orthogonal table provided (l)

The formula: $$Sum\ of\ \ Squares = \frac{(\Sigma lT)^2}{\Sigma(l)^2 \times r}$$

The process:

Numerator

- Multiply each treatment total by its respective coefficient $(l \times T)$
- Add these products ΣlT
- Square the result $(\Sigma lT)^2$

Denominator-

- Square each coefficient l^2
- Add these squared coefficients $\Sigma(l^2)$
- Multiply the result by the number of values making up the treatment totals ($\Sigma(l^2) \times r$)

Divide numerator by denominator $\frac{(\Sigma lT)^2}{\Sigma(l)^2 \times r}$

If required:

- Calculate Linear MS
- Compare $F = \frac{MS(linear)}{MS(error)}$ with F table as per usual.

Example

The lab test was based on the paper by Carlo *et al*. who studied the plasma levels over time of people taking aspirin in two forms. There were 12 subjects; there were two bases (water and Alka-Seltzer) and plasma were sampled at 20, 45, 120, 300 minutes. Here is the ANOVA table of their results.

Analysis of Variance for plasma

Source	**DF**	**SS**	**MS**	**F**	**P**
subject	11	1426.52	129.68	4.65	0.000
solution	1	234.38	234.38	8.40	0.005
time	3	4438.65	1479.55	53.00	0.000
solution*time	3	1490.81	496.94	17.80	0.000
Error	77	2149.48	27.92		
Total	95	9739.83			

Means:

solution	N	plasma	time	N	plasma
1	48	29.854	20	24	28.562
2	48	32.979	45	24	38.792
			120	24	36.646
			300	24	21.667

A table of interaction **totals** was given as:

Rows: time Columns: solution

	1	2	All
20	254.0	431.5	685.5
45	430.0	501.0	931.0
120	463.5	416.0	879.5
300	285.5	234.5	520.0
All	1433.0	1583.0	3016.0

It was noted that if the logarithms of time were taken, you would get 1.30, 1.65, 2.08 and 2.48, for 20, 45, 120 and 300 respectively. These are approximately equally spaced so that standard coefficients for orthogonal polynomials can be used.

Calculate the **quadratic** sums of squares for the **time effect.** You will need to select the correct coefficients from the table. Note, from the ANOVA table, $df = 95 \Rightarrow n = 96$. Since there are **4** levels of time, each level total consists of 24 values.

Answer:

Table of totals for time with $r = 24$ are,

Time	20	45	120	300
total, T	685.5	931.0	879.5	520.0

Number of treatments is 4, so quadratic coefficients are:

coefficient, l	1	–1	–1	1

$$SS = \frac{\left((1\times 685.5)+(-1\times 931)+(-1\times 879.5)+(1\times 520)\right)}{\left[(1^2)+(-1)^2+(-1)^2+1^2\right]\times 24} = \frac{(-605)^2}{96} = 3812.76$$

By comparison for a ***linear*** sum of squares,

coefficient: l	-3	-1	1	3

$$SS = \frac{\left((-3\times 685.5)+(-1\times 931)+(1\times 879.5)+(3\times 520)\right)}{\left[(-3^2)+(-1)^2+(1)^2+3^2\right]\times 24} = \frac{(-548)^2}{480} = 12513$$

Orthogonal polynomials

$n=3$		$n=4$			$n=5$			
X_1	X_2	X_1	X_2	X_3	X_1	X_2	X_3	X_4
-1	+1	-3	+1	-1	-2	+2	-1	+1
0	-2	-1	-1	+3	-1	-1	+2	-4
+1	+1	+1	-1	-3	0	-2	0	+6
		+3	+1	+1	+1	-1	-2	-4
					+2	+2	+1	+1
↑ *linear*	↑ *quad*	↑ *linear*	↑ *quad*	↑ *cubic*	↑ *linear*	↑ *quad*	↑ *cubic*	↑ *quart*

Example

A study on the growth of cultured mammalian liver cells examined the effect of various levels of vitamins on growth. (Broad et al 1980). The experiment was laid out as a RCBD, with the treatments comprising a $2 \times 3 \times 4$ *factorial with 2 levels of Vitamin B1, 3 levels of vitamin A, and 4 levels of riboflavin. There were five blocks.*
*The following is an analysis of variance table with some items deleted (indicated by ***), and below that, tables of various treatment combinations.*

Source of variance	DF	Sum of Squares	Mean Square	F ratio
Block	4	0.909000		
A	*	*******	*******	13.280
B_1	1	0.30000	0.300000	19.714
$A \times B_1$	2	0.037500	0.018750	1.232
Riboflavin	3	0.441667	0.147222	9.675
Interaction				
$A \times$ Riboflavin	6	0.195833	0.032639	2.145
$B_1 \times$ Riboflavin	*	*******	*******	****
$A \times B_1 \times$ Riboflavin	6	0.200970	0.033465	2.201
Residual	92	1.4000000	0.015217	
Total	119	4.405804		

	A			
	1	2	3	Total
B_1-0	3	5	6.5	14.5
B_1-+	1.5	3.5	3.5	8.5
Total	4.5	8.5	10	23

	Riboflavin				
	1	2	3	4	Total
B_1-0	2	2.5	5.5	4.5	14.5
B_1-+	1	4	2.5	1	8.5
Total	3	6.5	8	5.5	23

	A			
riboflavin	1	2	3	Total
1	1	1	1	3
2	1.5	2	3	6.5
3	1.5	3	3.5	8
4	0.5	2.5	2.5	5.5
Total	4.5	8.5	10	23

Calculate the Linear and Quadratic Sums of Squares for Vitamin A, and Riboflavin

Linear SS for Vitamin A:

- Number of levels = 3 Treatment Sums = 4.5, 8.5, 10
- Orthogonal coefficients are -1, 0, 1
- Hence $\boldsymbol{LinearSS} = \dfrac{\{-1(4.5)+0(8.5)+1(10)\}^2}{\left[(-1)^2+(0)^2+(1)^2\right]\times 40} = 0.378$ $\Rightarrow \boldsymbol{MS} = \dfrac{0.378}{3} = 0.126$

$$\Rightarrow \boldsymbol{F} = \frac{0.126}{0.015217} = 8.283$$

$$QuadraticSS = \frac{\{1(4.5)+-2(8.5)+1(10)\}^2}{\{(1)^2+(-2)^2+(1)^2\}\times 40} = 0.026$$

Linear SS for Vitamin B cannot be done since only 2 treatments.

Linear SS for Riboflavin:

- Number of treatments = 4 Treatment Sums = 3, 6.5, 8, 5.5
- Orthogonal coefficients are -3, -1, 1, 3

$$\boldsymbol{LinearSS} = \frac{\{-3(3)+-1(6.5)+1(8)+3(5.5)\}^2}{\{(-1)^2+(0)^2+(1)^2+(3)^2\}\times 30} = 0.135 \qquad \Rightarrow \boldsymbol{MS} = \frac{0.135}{4} = 0.03375$$

$$\Rightarrow \boldsymbol{F} = \frac{0.03375}{0.015217} = 2.218$$

Quadratic SS for Riboflavin:

$$\boldsymbol{QuadraticSS} = \frac{\{1(3)+-1(6.5)+-1(8)+1(5.5)\}^2}{\{(1)^2+(-1)^2+(-1)^2+(1)^2\}\times 30} = 0.3 \quad \Rightarrow \boldsymbol{MS} = \frac{0.3}{4} = 0.075$$

$$\Rightarrow \boldsymbol{F} = \frac{0.075}{0.015217} = 4.9287$$

Using the information about Osage Orange given in Practice question 2.

(i) Calculate the sum of squares for the contrast Ether *versus* [Methanol and Eth- Meth]. (Note that *means* are given, not the totals.)

(ii) Determine if this contrast is significant. (Use 5% level of significance.)

Solution

First calculate totals, and then multiply these by the appropriate coefficients.

Control	0.08705	2.0892	0	0.0000
Ether	0.13519	3.2446	2	6.4891
Methanol	0.67197	16.1273	-1	-16.1273
Eth_Meth	0.88100	21.1440	-1	-21.1440

Sum the last column: -30.7822, then calculate $30.7822^2 \div (24\times6) = 6.580$.
Because this has 1 DF, the SS is also equal to the MS. Divide this by the EMS from the ANOVA table. We get 342.7, a very high value so highly significant.

An experiment to measure the effect of fertilizers (Nitrogen and Sulphur) on the yield of wheat (kg per plot) was laid out in a randomized complete block design. There were three replicated blocks with treatmens consisting of a 3 by 4 factorial of nitrogen at 0, 180, 230, and sulphur at four levels, 0, 10, 20 and 40.

Here is a partial output. You are required to complete the missing parts of the table.

Analysis of variance for Yield

Source of variance	DF	Sum of Squares	Mean Square	F ratio	p
Block	2	30.191	15.095	3.34	0.054
Nitrogen	*	*******	*******	******	
Sulphur	3	107.051	35.684	7.91	0.001
Nitrogen*Sulphur	6	75.874	12.646	2.80	0.035
Residual	22	99.296	4.513		
Total	**	*******			

S = 2.12449 R-sq = 86.81% R – sq (adj) = 79.02%

Means:

Nitrogen	N	Yield
0	12	6.025
180	12	12.767
230	12	13.975

Sulphur	N	Yield
0	9	8.067
10	9	11.278
20	9	11.678
40	9	12.667

Nitrogen	Sulphur	N	Yield
0	0	3	5.600
0	10	3	7.733
0	20	3	5.233
0	40	3	5.533
180	0	3	8.267
180	10	3	12.067
180	20	3	15.267
180	40	3	15.467
230	0	3	10.333
230	10	3	14.033
230	20	3	14.533
230	40	3	17.000

(a) Calculate the SS, MS, DF, F- ratio and significance (*p*) for Nitrogen
Note that the values listed are means not totals.

(b) Calculate the 5% LSD for the interaction means.

(c) Calculate the 1% LSD for the nitrogen means.

EITHER

(d) Calculate the SS due the linear response of yield on sulphur levels, and test for significance.

OR (e) Explain the cause of the significant interaction, using graphs if necessary.

Solutions

(a) With 3 blocks, 3 levels of nitrogen and 4 levels of sulphur, $r = 3\times 3\times 4 = 36$.
Using the means given to calculate the treatment totals and the overall total:

$$SS_{nitrogen} = \frac{(12\times 6.025)^2 + (12\times 12.767)^2 + (12\times 13.975)^2}{12} - \frac{(12\times 32.767)^2}{36}$$

$$= 440.465$$

From this value, $df_{nitrogen} = 2,\ \ MS_{nitrogen} = \frac{440.465}{2} = 220.233,$

$$F_{nitrogen} = \frac{220.233}{4.513} = 48.000,\ \ and\ \ p = 0.000$$

$$df_{total} = 35 \quad SS_{total} = (30.191 + 440.465 + 107.051 + 75.874 + 99.296)$$

$$= 752.877$$

(b) $t_{(0.025,22)} = 2.074 \Rightarrow$

$$5\%LSD_{interaction} = 2.074\times\sqrt{\frac{4.513\times 2}{3}} = 3.5975$$

(c) $t_{(0.005,22)} = 2.819 \Rightarrow$

$$1\%LSD = 2.819\times\sqrt{\frac{4.513\times 2}{12}} = 2.445$$

(d) Since there are 4 levels of sulphur, coefficients for Linear SS are -3, -1, 1, and 3.
Each mean must by multiplied by 9 to get the totals.

$$Linear\ \ SS_{sulphur} = \frac{({}^-3\times 9\times 8.067 + {}^-1\times 9\times 11.278 + 1\times 9\times 11.678 + 3\times 9\times 12.667)^2}{\left[({}^-3)^2 + ({}^-1)^2 + 1^2 + 3^2\right]\times 9}$$

$$= 90.738$$