Mean Comparisons

F-tests provide information on significance of treatment effects, but no information on what the treatment effects are. Comparisons of treatment means provide information on what the treatment effects are. Appropriate mean comparisons are based on the treatment design. When there is some structure to the treatments, the best way to compare treatment means is to use orthogonal contrasts/planned F tests which subdivide the treatment SS to isolate the variation associated with a single degree of freedom. If there is no structure to the treatments, a multiple range test can be used.

Use Orthogonal Contrasts or Factorial Comparisons when:
- treatments have a structure
- treatments are combinations of factors
- e.g. breed and sex, factorial arrangements

Use Trend Comparisons or Dose Response when:
- treatments are different levels of a factor
- e.g. fertilizer rates, seeding times

Use Multiple Range Tests when:
- no clear treatment structure
- treatments are a set of unrelated materials
- e.g. variety trials, types of insecticides

PLANNED F TESTS

Important questions can be answered regarding treatments if we plan for this in the experiment. The treatment SS can often be subdivided to isolate particular sources of variation associated with a single degree of freedom. The number of tests that can be done is equal to the degrees of freedom for treatment.

Advantages:

1. One can answer specific, important questions about treatment effects.
2. The computations are simple.
3. A useful check is provided on the treatment SS.

Appropriate Uses:

1. When treatments have an obvious structure and/or when planned single df contrasts were built into the experiment.

2. When comparing means from quantitative data such as rates of fertilizer, seeding time, plant density, etc.

3. When comparing factorial treatment combinations to estimate the main effects and interactions.

Examples:

1. A variety test with a definite structure and single df contrasts built into the design. 3 blocks.

Tall		Short	
Early	Late	Early	Late
V1	V2	V3	V4
95	84	103	90

ANOVA

Source of Var.	df
Blocks	2
Var.	(3)
Tall vs Short	1
Early vs Late	1
TS x EL	1
Error	6
Total	11

2. Quantitative data, e.g., rates of N.
N (0, 50, 100, 150 kg/ha). 3 blocks.

It is appropriate to analyze the data by regression techniques to describe the trend of the response to N. If ANOVA is done, trend comparisons should be made to determine the shape of the response curve.

ANOVA

Source of Var.	df
Blocks	2
N	(3)
Nlin	1
Nquad	1
Ncubic	1
Error	6
Total	11

3. An experiment comparing factorial treatment combinations of 2 rates of Lime and 3 rates of Phosphorus with 3 blocks. The factorial structure of the experiment allows the testing of main effects of Lime and Phosphorus, and their interaction. The appropriate analysis is an ANOVA with

single df contrasts.

ANOVA

Source of Var.	df
Blocks	2
Treatments	(5)
Lime	1
Phosphorus	2
Plin	1
Pquad	1
Error	10
Total	17

Using orthogonal contrasts:

Contrasts subdivide treatment SS to isolate sources of variation associated with a single degree of freedom

Number of tests = df for treatment
Total of tests SS = SS for treatment

Contrasts compare groups of treatments using linear combinations of treatment totals (or means)

$c_1Y_1 + c_2Y_2 + c_3Y_3 + c_4Y_4 + c_5Y_5$
Sum of coefficients $\Sigma c_i = 0$

e.g. Is the average of Y_1, Y_2 and Y_3 different than the average of Y_4 and Y_5?

$$\frac{Y_1 + Y_2 + Y_3}{3} - \frac{Y_4 + Y_5}{2} = 0?$$

$$2Y_1 + 2Y_2 + 2Y_3 - 3Y_4 - 3Y_5 = 0?$$

Contrasts are orthogonal when:

They measure independent effects
The sums of the products of the corresponding coefficients of any two contrasts = 0
Maximum number of orthogonal contrasts = df for treatment

Formula for computing SS with orthogonal comparisons:

$$SS = \frac{(\Sigma c_i Y_i)^2}{r\Sigma c_i^2}$$

SS for a contrast
c_i = comparison coefficient
Y_i = treatment total
r = number of replicates (blocks)

Sum of orthogonal contrast SS = SS for treatment

Constructing Sets of Orthogonal Contrasts

1. To compare two groups of equal size, assign coefficients of -1 to members of one group and +1 to members of the other

2. To compare two groups of unequal size, assign to the first group coefficients equal to the number of treatments in the second group, and assign to the second group coefficients of opposite sign equal to the number of members in the first group.

3. Reduce coefficients to the smallest possible integers.

4. For interactions, obtain coefficients by multiplying the corresponding coefficients of the main effects.

Example: 5 Varieties
Treatment df = 4

	Tall		Short		Control
Contrast	Early	Late	Early	Late	
	V1	V2	V3	V4	V5
New vs Cont	-1	-1	-1	-1	+4
Tall vs Short	-1	-1	+1	+1	0
Early vs Late	-1	+1	-1	+1	0
TS x EL	+1	-1	-1	+1	0

Example: 5 Herbicides
df for Treatment = 4

Contrast	Control	A	B	C	D
Control vs Herbicide	-4	+1	+1	+1	+1
AB vs CD	0	-1	-1	+1	+1
C vs D	0	0	0	-1	+1
A vs B	0	-1	+1	0	0

Example: Trend Comparisons or Dose Response
Example: 4 rates of N: 0, 50, 100 and 150 kg/ha

Contrast	N0	N50	N100	N150
Linear	-3	-1	1	3
Quadratic	1	-1	-1	1
Cubic	-1	3	-3	1

Coefficients are obtained from tables.
Contrasts describe shape of response curve: linear, quadratic, etc.
Using contrasts gives identical results to using regression!

Example: Weight gain of Lambs (lb/100 da)
Female and Male Lambs

Implants with 0 or 3 mg stilbesterol

	Block				Trt	Trt
	1	2	3	4	Total	Mean
FS0	47	52	62	51	212	53
MS0	50	54	67	57	228	57
FS3	57	53	69	57	236	59
MS3	54	65	74	59	252	63
Blk Tot	208	224	272	224	928	
Blk Mean	52	56	68	56		58

Contrast	FS0	MS0	FS3	MS3
Trt Totals	212	228	236	252
Sex: F vs M	-1	+1	-1	+1
Implant: S0 vs S3	-1	-1	+1	+1
Sex vs Implant	+1	-1	-1	+1

Analysis of Variance

Source of Var.	df	SS	MS	F	$F_{.05}$	$F_{.01}$
Total	15	854				
Blocks	3	576	192.0	24.7	3.86	6.99
Trts	(3)	208	69.3	8.9**		
Implants	1	144	144.0	18.5**	5.12	
Sex	1	64	64.0	8.23*		
IxS	1	0	0	0		
Error	9	70	7.78			

MULTIPLE RANGE TESTS

Appropriate use of MR tests:

When the treatment structure is not well understood such as for treatments that represent a set of unrelated materials, for example, varieties, herbicides, insecticides, tillage operations.

LSD-Least Significant Difference

LSD is a form of the t test.

$$t = (d - u_d)/s_d$$

Let the mean difference between 2 means ($Y_1 - Y_2 = d$) be the lower limit of the values we would expect 5% or more of the time by chance alone in drawing samples of mean differences from a population of mean differences where the mean is zero ($u_d = 0$). For the t test, replace d with LSD and u_d with 0, the formula becomes $t = LSD/s_d$. In the following, s_1^2 and s_2^2 estimate the variance of plots for treatments t_1 and t_2, and r_1 and r_2 are the number of experimental units

having t_1 and t_2.

$$t = \frac{\overline{Y_1} - \overline{Y_2}}{s_{\bar{d}}} \text{ where } s_{\bar{d}} = \sqrt{\frac{s_1^2}{r_1} + \frac{s_2^2}{r_2}}$$

$$LSD = \overline{Y_1} - \overline{Y_2} = t_{0.05}(s_{\bar{d}}$$

$$\text{For an ANOVA } LSD = t_{0.05}\sqrt{\frac{2(MSE}{r}}$$

r is the number of observations per mean

df for t is df for Error

In the ANOVA, the MSE = s^2 is used for the variance.

Points to remember in the use of the LSD.

1. The LSD should be used only if the F test is significant. This is called the "protected LSD".

2. The LSD is a fixed range test because one range is used for all treatment comparisons. It is assumed that the means being compared are adjacent in an array. Therefore, the use of the LSD to make all possible comparisons among 10 means, for example, will tend to show more mean differences to be significant than are actually significant at the 5% level. These mean differences are significant at some unknown probability, i.e., 8%, 10% etc.

There is a danger of committing a Type I error, the rejection of Ho when it is true when this is done. LSD can be used with little danger of committing a Type I error for comparing up to about 5 means.

Referring to the example of the lambs treated with stilbesterol:
From the ANOVA MSE = 7.78 with 9 df

$$LSD_{.05} = t_{.05,\ 9}\sqrt{\frac{2(MSE}{r}}$$

$$LSD_{.05} = 2.262\sqrt{\frac{2(7.78}{4}} = 4.46 \text{ lb/100 da}$$

Arrange means from low to high and compare adjacent means:

	MS3	FS3	MS0	FS0
Means	63	59	57	53
Differences	4	2	4	

No Significant differences although F test indicates significant treatment effects.

Multiple Range Tests

Provide multiple ranges for comparing means

Means do not need to be adjacent

Ranges required for significance increase as means are further apart.

Duncan's Multiple Range Test (DMR)

Starts with LSD
Adjusts depending on how far apart means are.

D = Shortest significant difference
D = R (LSD)

R = Significant studentized factor obtained from tables.
R depends on
- desired level of significance
- df for error, n
- range of means, p (p counts means to be compared and all means between them)

For lambs

$$LSD = t_{.05,\ 9}\sqrt{\frac{2(MSE}{r}} = 4.46 \text{ lb/da}$$

Calculate D's

Range of means, p	2	3	4
R, 5%, 9 df	1.00	1.04	1.07
D = R (LSD)	4.46	4.64	4.77

Order means and compare

A	B	C	D
53^{a}	57^{ab}	59^{bc}	63^{c}

A to D: difference = 10 > 4.77
A to C: difference = 6 > 4.64
A to B: difference = 4 < 4.46 NS
B to D: difference = 6 > 4.64
B to C: difference = 2 < 4.46 NS
C to D: difference = 4 < 4.46 NS

Means that are not different are:
- connected by the same line
- followed by the same letter

The results of several multiple range tests are summarized below for the lamb data.

Trt	MS3	FS3	MS0	FS0
Mean	63	59	57	53
DMR & LSD	———	———	———	
HSD	———	———		
Waller-Duncan		———		

HSD is Tukey's test - assumes all means are the maximum distance apart in the array.