Mean Comparisons

ANOVA tests the null hypothesis:

$$\mu_1 = \mu_2 = \ldots = \mu_I$$

What we really want to know is which means differ and how.

$$\mu_1 = \mu_2$$
$$\mu_1 = \mu_I$$
$$\mu_2 = \mu_I$$

Mean Comparisons

Approaches

- Multiple Comparison Procedures
- Contrasts
- Curve Fitting

Mean Comparisons
Considerations

Qualitative – classification variables
- Multiple Comparison Procedures
- Contrasts

Quantitative – numerical variables
- Orthogonal Polynomial Contrasts
- Curve Fitting

Least Significant Difference

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{d}}{s_{\bar{d}}} = \frac{LSD}{s_{\bar{d}}}$$

$$LSD = t_P \times s_{\bar{d}}$$

Where t_P is the critical t value at a given P level.

Least Significant Difference

$$LSD = t_{.05} \times S_{\bar{d}} = t_{.05} \times \sqrt{\frac{2MSE}{r}}$$

$$LSD = t_{.05} \times \sqrt{2} \times S_{\bar{x}}$$

$$LSD = t_{.05} \times \sqrt{2} \times \frac{S}{\sqrt{r}}$$

One-Factor ANOVA

Switchgrass Example

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Cultivar	9	355.5271	39.50301	4.57	0.0002
Error	**50**	432.5709	**8.651419**		
Total	59	788.098			

Least Significant Difference

Switchgrass Example

$$S_{\bar{x}} = \sqrt{\frac{MS_E}{r}} = \sqrt{\frac{8.651}{6}} = 1.20$$

$$S_{\bar{d}} = \sqrt{\frac{2MS_E}{r}} = \sqrt{\frac{2(8.651)}{6}} = 1.70$$

$$LSD = t_{.05} \times S_{\bar{d}} = 2.008 \times 1.70 = 3.41$$

Least Significant Difference

Switchgrass Example

$$LSD = 3.41$$

Cultivar	Mean Yield (Mg/ha)			
C	14.2	a		
G	13.9	a		
D	12.7	a	b	
A	12.6	a	b	
H	9.6		b	c
E	9.4		b	c
I	9.4		b	c
J	9.1			c
B	8.3			c
F	6.7			c

LSD Approximations

$$LSD_{.05} = 3 \times S_{\bar{x}} = 3(1.2) = 3.6$$

$$LSD_{.10} = 2.5 \times S_{\bar{x}} = 2.5(1.2) = 3.0$$

$$LSD_{.01} = 3.5 \times S_{\bar{x}} = 3.5(1.2) = 4.2$$

Useful Formulae

$$S_{\bar{x}} = SEM = \sqrt{\frac{MS_E}{r}} = \frac{RMSE}{\sqrt{r}}$$

$$95\%\ CI = \pm t_{.05} \times SEM$$

$$S_{\bar{d}} = SED = \sqrt{\frac{2MS_E}{r}} = \sqrt{2} \times \frac{RMSE}{\sqrt{r}}$$

$$LSD = t_{.05} \times SED$$

Example Calculations

Switchgrass Example

$$SEM = \frac{RMSE}{\sqrt{r}} = \frac{2.491}{\sqrt{6}} = 1.20$$

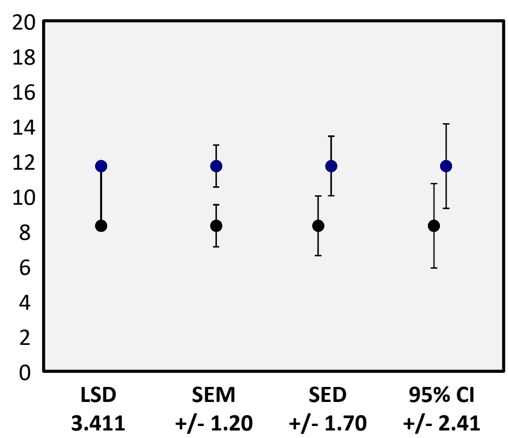
$$95\%\ CI = \pm t_{.05} \times SEM = \pm 2.008 \times 1.20 = \pm 2.41$$

$$SED = \sqrt{2} \times SEM = 1.414 \times 1.2 = 1.70$$

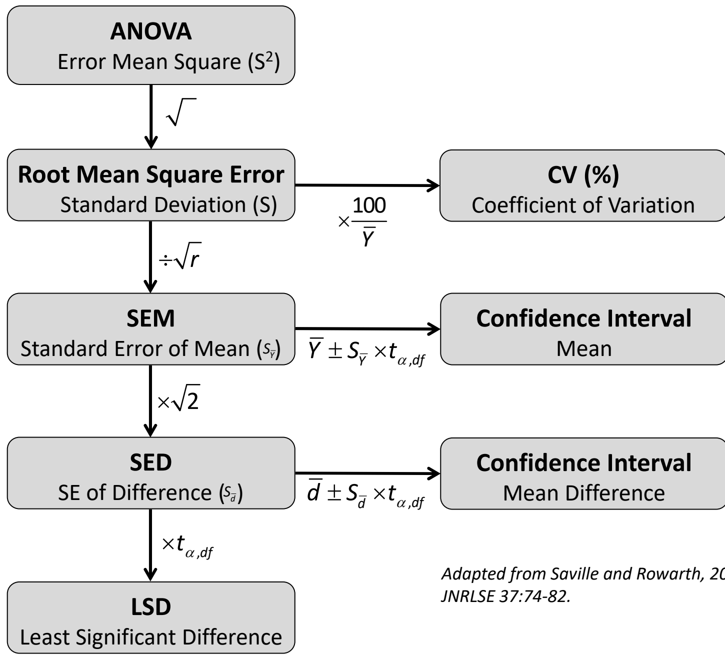
$$LSD = t_{.05} \times SED = 2.008 \times 1.70 = 3.41$$

Comparison of Error Statistics

Switchgrass Variety Trial



Relationships Among Error Statistics



Adapted from Saville and Rowarth, 2008 JNRLSE 37:74-82.

Hypothesis Testing

Error Rates

<u>Comparisonwise Error Rate:</u>

CER = P(Reject H_0 | H_0 is True)

$$\alpha_C = \frac{number\ Type\ I\ errors}{number\ \ nonsignificant\ comparisons\ tested}$$

Hypothesis Testing

Error Rates

Experimentwise Error Rate:

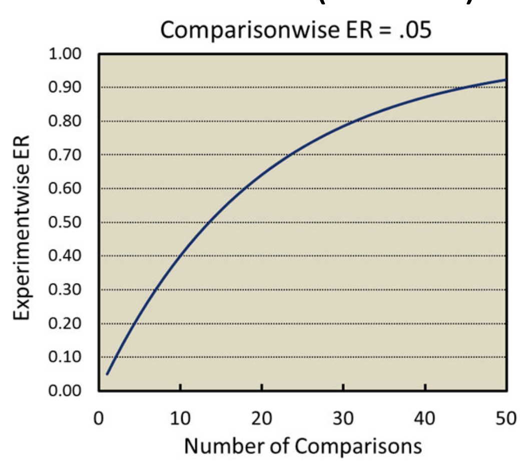
EER = P(Reject at least one H_{01} ... H_{0n} | all H_0 are True)

$$\alpha_E = \frac{\text{number Type I errors}}{\text{number experiments with at least one true null hypothesis}}$$

Possible Comparisons = t(t – 1)/2

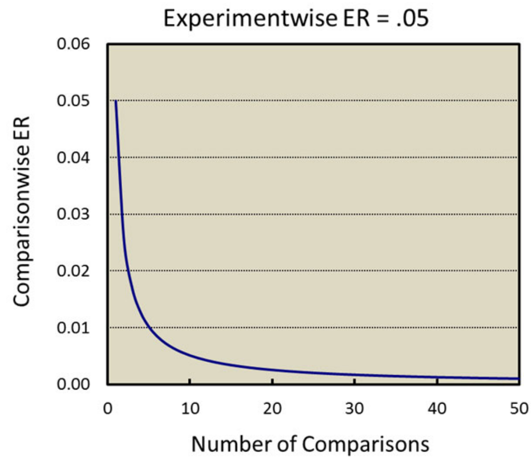
Relationship Between CER and EER

$$EER = 1 - (1 - CER)^k$$



Relationship Between CER and EER

$$CER = 1 - (1 - EER)^{1/k}$$



Mean Comparison Procedures

Other Common Approaches

- Tukey's Significant Difference
- Bonferroni Adjustment
- Duncan's Multiple Range Test (DMRT)
- Dunnett's Test
- Bayes Least Significant Difference (BLSD)
- Student-Newman-Kuels

Tukey's Honest Significant Difference

Switchgrass Example

$$HSD = q_{.05} \times S_{\bar{x}} = q_{.05} \times \sqrt{\frac{MS_E}{r}}$$

$$HSD = 4.68 \times 1.20 = 5.62$$

Tukey's Honest Significant Difference

Switchgrass Example

Cultivar	Mean Yield (Mg/ha)	LSD			HSD		
C	14.2	a			a		
G	13.9	a			a	b	
D	12.7	a	b		a	b	
A	12.6	a	b		a	b	
H	9.6		b	c	a	b	c
E	9.4		b	c	a	b	c
I	9.4		b	c	a	b	c
J	9.1			c	a	b	c
B	8.3			c		b	c
F	6.7			c			c

Bonferroni Adjustment
Definition

The Bonferroni p value is calculated by multiplying the nominal α-level by the number of planned comparisons:

$$\alpha_{adj} = \alpha_c \times n$$

- The more comparisons made, the higher the adjusted alpha
- e. g. if you plan to make 10 comparisons at α = 0.05 then the adjusted α is 0.5
- When the adjusted value > 1 it is set to 1

Bonferroni Adjustment
Reducing the Number of Comparisons

To use the Bonferroni Adjustment to control experimentwise error rate:

- Calculate the unadjusted α-level that results in the desired adjusted α-level for the family of comparisons:

$$\alpha_c = \frac{\alpha_{adj}}{n}$$

- The calculated α is used for the planned pairwise comparisons using a *t* test.

Bonferroni Adjustment

Switchgrass Example

Cultivar	Mean Yield (Mg/ha)	LSD			HSD			Bon *	
C	14.2	a			a			a	
G	13.9	a			a	b		a	
D	12.7	a	b		a	b		a	
A	12.6	a	b		a	b		a	
H	9.6		b	c	a	b	c	a	b
E	9.4		b	c	a	b	c	a	b
I	9.4		b	c	a	b	c	a	b
J	9.1			c	a	b	c	a	b
B	8.3			c		b	c	a	b
F	6.7			c			c		b

* Note that when used for all possible pairwise comparisons the Bonferroni method is more conservative than Tukey.

Duncan's Multiple Range Test

Background

- Multiple range tests compare sets of means based on the studentized range statistic q_r.
- DMRT is just one of several (e. g. SNK) that use this general approach.
- The idea is to reduce the number of comparisons made by comparing groups containing multiple means.
- This reduces the experimentwise error rate.
- MRT are more conservative than an LSD but less so than the HSD.

Duncan's Multiple Range Test
Procedure

- Calculate the Least Significant Range of each size group of means:

$$DMRT_i = q_i S_{\overline{X}}$$

$DMRT_2$ = 2.8406(1.2008) = 3.411
$DMRT_3$ = 2.9872(1.2008) = 3.587
$DMRT_4$ = 3.0846(1.2008) = 3.704
$DMRT_5$ = 3.1546(1.2008) = 3.788
$DMRT_6$ = 3.2079(1.2008) = 3.852
$DMRT_7$ = 3.2512(1.2008) = 3.904
$DMRT_8$ = 3.2862(1.2008) = 3.946
$DMRT_9$ = 3.3153(1.2008) = 3.981
$DMRT_{10}$ = 3.3403(1.2008) = 4.011

LSD = 3.41
HSD = 5.62

Duncan's Multiple Range Test
Procedure

- Sort treatment means
- Compare largest mean with smallest using the LSR for spanning 10 means:

  $\delta_{C\text{-}F} = 7.52 > 4.011 \therefore * @ \alpha = 0.05$
- Continue comparing largest mean with next smallest using the appropriate LSR until a NS result occurs

  $\delta_{C\text{-}B} = 5.87 > 3.981 \therefore *$
  $\delta_{C\text{-}J} = 5.07 > 3.946 \therefore *$
  $\delta_{C\text{-}E} = 4.80 > 3.904 \therefore *$
  $\delta_{C\text{-}I} = 4.79 > 3.852 \therefore *$
  $\delta_{C\text{-}H} = 4.54 > 3.788 \therefore *$
  $\delta_{C\text{-}A} = 1.60 < 3.704 \therefore$ NS
  $\delta_{C\text{-}D} = 1.43 < 3.587 \therefore$ NS
  $\delta_{C\text{-}G} = 0.26 < 3.411 \therefore$ NS

Duncan’s Multiple Range Test
Procedure

- **Exception Rule:** a mean difference cannot be considered * if the two means fall within a subset of means already determined to be NS.
- Repeat the process starting with the second largest mean
  - δ_{G-F} 7.26 > 3.981 ∴ *
  - δ_{G-B} 5.61 > 3.946 ∴ *
  - δ_{G-J} 8.85 > 3.904 ∴ *
  - δ_{G-E} 4.54 > 3.852 ∴ *
  - δ_{G-I} 4.53 > 3.788 ∴ *
  - δ_{G-H} 4.28 > 3.704 ∴ *
  - δ_{G-A} 1.16 < 3.587 ∴ NS
  - δ_{G-D} 0.26 < 3.411 ∴ NS
- Repeat the process starting with the third largest mean, and so on until the last comparison
- There are potentially n(n – 1)/2 comparisons
- Use lines or letters to indicate means that are not different

Duncan’s Multiple Range Test
Switchgrass Example

Cultivar	Mean Yield (Mg/ha)	LSD			DMRT			HSD			Bon *	
C	14.2	a			a			a			a	
G	13.9	a			a			a	b		a	
D	12.7	a	b		a	b		a	b		a	
A	12.6	a	b		a	b		a	b		a	
H	9.6		b	c		b	c	a	b	c	a	b
E	9.4		b	c		b	c	a	b	c	a	b
I	9.4		b	c		b	c	a	b	c	a	b
J	9.1			c		b	c	a	b	c	a	b
B	8.3			c			c		b	c	a	b
F	6.7			c			c			c		b

Mean Comparisons

PROC ANOVA / GLM

```
proc anova;
    class trt;
    model yield = trt;
    means trt / lsd;
    means trt / duncan;
    means trt / tukey;
    means variety / bon;
    means variety / duncan;
run;
```

Mean Comparisons

Recommendations

- Avoid using MCP for making all possible pairwise comparisons
- Use MCP for preplanned comparisons
- Limit the number of comparisons
- For most agronomic data using an unprotected LSD provides a good balance between controlling Type I and II errors
- Use Tukey only in situations when the consequences of committing a Type I error are extreme but recognize that by doing so you increase the probability of committing a Type II error
- Other MCPs are not recommended