Chemical trts: A, B, C, D
liquid
powder

STAT 5200 Handout #7: Multiple Comparison Procedures (Ch. 5)

Example 1: [Beet Lice, same data as Handout #4] Here, look at adjusting for multiple comparisons (all pairwise treatment comparisons). See Handout #5 for more options.

```
/* Enter data */
data lice;
     input Chemical $ Numlice @@;
     datalines;
A 12  B 10  C 23  D 32  A 13  B 21  C 14  D 26  A 26  B 34
C 14  D 24  A 13  B 15  C 20  D 16  A 17  B  5  C 27  D 32
A 24  B 22  C 25  D 18  A 14  B 12  C 17  D 33  A 10  B 25
C 18  D 16  A  6  B 18  C 29  D 34  A  4  B 12  C 14  D 18
A  2  B  2  C 31  D  9  A 10  B  2  C  5  D 19  A  8  B 10
C 13  D 29  A  6  B 22  C 18  D 30  A  7  B 17  C 23  D 18
A 13  B 20  C 16  D 25  A 18  B 19  C 13  D 20  A 10  B 20
C 23  D 21  A 18  B 12  C  4  D 27  A  3  B 11  C 16  D 31
A  4  B 16  C 17  D 25  A 18  B  5  C  9  D 33  A 13  B 11
C 28  D 24  A 10  B 17  C 23  D 16  A 21  B 16  C 19  D 24
; run;
```

```
/* Look at a contrast: Suppose A, B, and C are applied
   in liquid form, but chemical D is applied in powder
   form (and is a current "gold standard").
   Test null "liquid mean = powder mean"
 */
proc glm data=lice;
 class Chemical;
 model Numlice = Chemical;
 contrast 'Ave of A,B,C versus D'
          Chemical 0.33333 0.33333 0.33333 -1;
 title1 'Contrast Example';
run;
```

CONTRAST 'name' FACTOR w_1 w_2 ... w_g;
numeric coeffs in Contrast ψ
need to be in same order as CLASS variable table levels (alphanumeric order)

Contrast Example

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Ave of A,B,C versus D	1	1482.963333	1482.963333	31.00	<.0001

name
# of contrasts simultaneously
p-value for $H_0: \psi = 0$

```
/* Look at pairwise comparisons */
proc glm data=lice;
  class Chemical;
  model Numlice = Chemical ;
  /* request raw p-values for pairwise comparisons: */
  lsmeans Chemical / pdiff lines;
  /* REGWQ MCP: */
  means Chemical / REGWQ lines;
  /* Tukey MCP: */
  lsmeans Chemical / pdiff adjust=Tukey lines;
  /* Bonferroni MCP: */
  lsmeans Chemical / pdiff adjust=Bonferroni lines;
  /* Dunnett MCP, with chemical 'D' as control: */
  means Chemical / Dunnett('D') lines;
  title1 'Example pairwise comparisons';
  title2 '(In practice, you''d just do one of these)';
run;
```

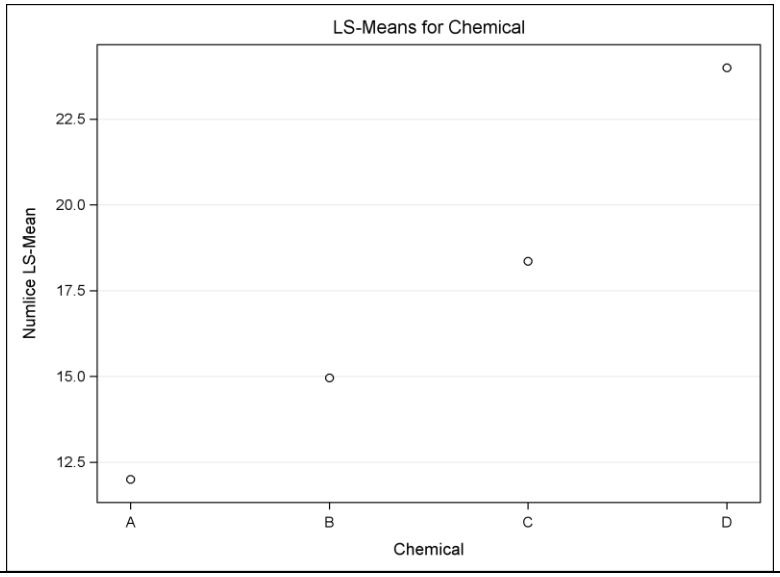
no adjustment for multiple testing

*

Example pairwise comparisons

Least Squares Means

Chemical	Numlice LSMEAN	LSMEAN Number
A	12.0000000	1
B	14.9600000	2
C	18.3600000	3
D	24.0000000	4



Least Squares Means for effect Chemical Pr > |t| for H0: LSMean(i)=LSMean(j) Dependent Variable: Numlice				
i/j	**1**	**2**	**3**	**4**
1		0.1336	0.0016	<.0001
2	0.1336		0.0854	<.0001
3	0.0016	0.0854		0.0049
4	<.0001	<.0001	0.0049	

pdiff (but need to adjust for mult. testing to control the Type I error rate)

T Comparison Lines for Least Squares Means of Chemical				
LS-means with the same letter are not significantly different.				
		Numlice LSMEAN	**Chemical**	**LSMEAN Number**
	A	24.00	**D**	4
	B	18.36	**C**	3
	B			
C	B	14.96	**B**	2
C				
C		12.00	**A**	1

LINES

Note: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.

Ryan-Einot-Gabriel-Welsch Multiple Range Test for Numlice

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	96
Error Mean Square	47.84083

Number of Means	**2**	**3**	**4**
Critical Range	4.444667	4.6572764	5.1150605

Means with the same letter are not significantly different.

REGWQ Grouping		Mean	N	Chemical
	A	24.000	25	D
	B	18.360	25	C
	B			
C	B	14.960	25	B
C				
C		12.000	25	A

Adjustment for Multiple Comparisons: Tukey

Least Squares Means for effect Chemical Pr > |t| for H0: LSMean(i)=LSMean(j) Dependent Variable: Numlice				
i/j	**1**	**2**	**3**	**4**
1		0.4338	0.0085	<.0001
2	0.4338		0.3100	<.0001
3	0.0085	0.3100		0.0247
4	<.0001	<.0001	0.0247	

Tukey Comparison Lines for Least Squares Means of Chemical				
LS-means with the same letter are not significantly different.				
		Numlice LSMEAN	**Chemical**	**LSMEAN Number**
	A	24.00	**D**	4
	B	18.36	**C**	3
	B			
C	B	14.96	**B**	2
C				
C		12.00	**A**	1

Adjustment for Multiple Comparisons: Bonferroni

Least Squares Means for effect Chemical Pr > |t| for H0: LSMean(i)=LSMean(j) Dependent Variable: Numlice				
i/j	**1**	**2**	**3**	**4**
1		0.8013	0.0095	<.0001
2	0.8013		0.5126	<.0001
3	0.0095	0.5126		0.0292
4	<.0001	<.0001	0.0292	

Bonferroni Comparison Lines for Least Squares Means of Chemical				
LS-means with the same letter are not significantly different.				
		Numlice LSMEAN	**Chemical**	**LSMEAN Number**
	A	24.00	**D**	4
	B	18.36	**C**	3
	B			
C	B	14.96	**B**	2
C				
C		12.00	**A**	1

Dunnett's t Tests for Numlice

Note: This test controls the Type I experimentwise error for comparisons of all treatments against a control.

Alpha	0.05
Error Degrees of Freedom	96
Error Mean Square	47.84083
Critical Value of Dunnett's t	2.38676
Minimum Significant Difference	4.6693

Comparisons significant at the 0.05 level are indicated by *.**				
Chemical Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
C - D	-5.640	-10.309	-0.971	***
B - D	-9.040	-13.709	-4.371	***
A - D	-12.000	-16.669	-7.331	***

Example 2: ["Cleanliness" paper by Schnall et al. in Psychological Science]
Researchers hypothesized that cleanliness would reduce the severity of moral judgments. Forty subjects were asked to perform a "priming" task: from sets of scrambled words, underline three words to form a sentence. Subjects were randomly assigned (in a balanced design) to either "neutral" sets of words or "cleanliness" sets of words. (In the cleanliness set, half of the sets each subject saw included words like pure, washed, clean, etc.) After this priming task, each subject was asked to rate how wrong six moral dilemmas were, on a scale of 0 (perfectly okay) to 9 (extremely wrong).

```
/* cleanliness data */
data clean; input dilemma $ meanC sdC meanN sdN; cards;
 dog      5.70 2.39 6.55 2.52
 trolley  1.85 1.50 2.75 2.38
 wallet   4.95 2.35 5.45 2.86
 plane    6.05 2.39 6.45 2.56
 resume   4.65 2.28 5.40 2.26
 kitten   6.70 2.49 8.25 1.48
 ;
data clean; set clean;
   sp = ( (19*sdC**2 + 19*sdN**2) / (19+19)  )**0.5;
   tstat = (meanC-meanN)/(sp*(1/19+1/19)**0.5);
   raw_p = cdf('t', tstat, 19+19);
   /* "raw_p" must be variable name for PROC MULTTEST */

proc print data=clean;
   var dilemma meanC meanN raw_p;
   title1 'Results of Cleanliness Questions';

proc multtest pdata=clean bonferroni;
   title1 'Bonferroni-Adjusted P-values';
run;
```

do p-val adjustment on raw-p

Note that we use one-sided p-values here because of direction of research hypothesis.

Results of Cleanliness Questions

Obs	dilemma	meanC	meanN	tstat	raw_p
1	dog	5.70	6.55	-1.06679	0.14640
2	trolley	1.85	2.75	-1.39447	0.08564
3	wallet	4.95	5.45	-0.58878	0.27975
4	plane	6.05	6.45	-0.49784	0.31073
5	resume	4.65	5.40	-1.01834	0.15748
6	kitten	6.70	8.25	-2.33246	0.01254

Bonferroni-Adjusted P-values

p-Values		
Test	Raw	Bonferroni
1	0.1464	0.8784
2	0.0856	0.5138
3	0.2797	1.0000
4	0.3107	1.0000
5	0.1575	0.9449
6	0.0125	0.0752

× 6 (but truncate at 1.0)

P(at least one Type I error) < .05

Compare to .05 threshold, will control the overall Type I error rate at .05 (FWER)