

Factorial Designs II

Design of Experiments - Montgomery
Sections 5-1 - 5-3

15

Diagnostics

- Assumptions
 - 1 Errors independent, normally distributed
 - 2 Constant variance
 - 3 No interaction (if $n = 1$)
 - Tukey's Test of Nonadditivity

Normality

Histogram or QQplot of residuals

Variance and Unusual Observations

Residuals vs factor A

Residuals vs factor B

Residuals vs \hat{y}_{ij} .

15-1

Multiple Comparisons

- If no interaction
 - Compare means of each factor (average over other factor)
 - Similar procedures as CRD
 - an or bn observations used in calculations
 - Degrees of freedom error is $ab(n - 1)$
 - Use proc means or lsmeans in SAS
- If interaction significant
 - Construct interaction plot for visual description
 - Can compare means of A at fixed level of B and vice versa
 - n observations used in calculations
 - Degrees of freedom error is $ab(n - 1)$
 - means statement only does main effect comparisons
 - lsmeans statement has slice option
 - Can also do comparisons by hand

15-2

Using SAS

```
options nocenter ls=78;
goptions colors=(none);

data new;
infile 'glycogen.dat';
input trt $ resp;

if trt="a" then horm=1; if trt="a" then level=1;
if trt="A" then horm=1; if trt="A" then level=2;
if trt="b" then horm=2; if trt="b" then level=1;
if trt="B" then horm=2; if trt="B" then level=2;

proc glm;
class trt;
model resp = trt;
contrast 'level' trt 1 1 -1 -1;
contrast 'hormone' trt 1 -1 1 -1;
contrast 'interaction' trt 1 -1 -1 1;

proc glm;
class level horm;
model resp = level horm level*horm;
means level|horm / tukey lines;
lsmeans level horm level*horm / adjust=tukey tdiff;
output out=new1 r=res p=pred;

proc univariate noprint;
hist res / normal (L=1 mu=0 sigma=est) kernel (L=2);

symbol1 v=circle;
proc gplot;
plot res*pred res*level res*horm;

run;
```

15-3

Output

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	6026.83333	2008.94444	6.97	0.0022
Error	20	5767.00000	288.35000		
Corrected Total	23	11793.83333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	3	6026.83333	2008.94444	6.97	0.0022

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
level	1	5162.66667	5162.66667	17.90	0.0004
hormone	1	864.00000	864.00000	3.00	0.0989
interaction	1	0.16667	0.16667	0.00	0.9811

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	6026.83333	2008.94444	6.97	0.0022
Error	20	5767.00000	288.35000		
Corrected Total	23	11793.83333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
level	1	5162.66667	5162.66667	17.90	0.0004
horm	1	864.00000	864.00000	3.00	0.0989
level*horm	1	0.16667	0.16667	0.00	0.9811

15-4

Tukey's Studentized Range (HSD) Test for resp

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	288.35
Critical Value of Studentized Range	2.95000
Minimum Significant Difference	14.461

	Mean	N	level
A	101.083	12	2
B	71.750	12	1

	Mean	N	horm
A	92.417	12	1
A	80.417	12	2

Level of level	Level of horm	N	Mean	Std Dev
1	1	6	77.833333	24.8468643
1	2	6	65.666667	11.9443152
2	1	6	107.000000	16.5650234
2	2	6	95.166667	10.9071842

15-5

Least Squares Means

Adjustment for Multiple Comparisons: Tukey

level	resp LSMEAN	t Value	Pr > t
1	71.750000	-4.23	0.0004
2	101.083333		

horm	resp LSMEAN	t Value	Pr > t
1	92.416667	1.73	0.0989
2	80.416667		

level	horm	resp LSMEAN	LSMEAN Number
1	1	77.833333	1
1	2	65.666667	2
2	1	107.000000	3
2	2	95.166667	4

Least Squares Means for Effect level*horm
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: resp				
i/j	1	2	3	4
1		1.241001	-2.975	-1.768
		0.6091	0.0347	0.3171
2	-1.241		-4.216	-3.009
	0.6091		0.0022	0.0323
3	2.975003	4.216005		1.207001
	0.0347	0.0022		0.6297
4	1.768002	3.009003	-1.207	
	0.3171	0.0323	0.6297	

15-6

What if Interaction Present?

A grape grower is interested in maximizing the number of bushels/acre on her winery. She limits her study to the combinations of 3 varieties and 4 pesticides (12 combinations). For each combination two replicates will be obtained.

$$y_{ijk} = \mu + v_i + p_j + (vp)_{ij} + \epsilon_{ijk}$$

If there is an interaction, the grower wants to compare the pesticides for each fixed variety of grape. In other words, the grower wants to compare $\bar{y}_{i1.}$, $\bar{y}_{i2.}$, $\bar{y}_{i3.}$, and $\bar{y}_{i4.}$. Must compute appropriate standard error.

$$\begin{aligned} \bar{y}_{ij.} &= \mu + v_i + p_j + (vp)_{ij} + \bar{\epsilon}_{ij.} \\ \bar{y}_{i1.} - \bar{y}_{i2.} &= p_1 - p_2 + (vp)_{i1} - (vp)_{i2} + \bar{\epsilon}_{i1.} - \bar{\epsilon}_{i2.} \\ \text{Var}(\bar{y}_{i1.} - \bar{y}_{i2.}) &= \sigma^2/n + \sigma^2/n = 2\sigma^2/n \end{aligned}$$

Could do calculations by hand, use lsmean, or lsmean slice option.

15-7

```

options nocenter ls=78;
options colors=(none);

data new;
input varty pest resp @@;
cards;
1 1 49 1 1 39 1 2 50 1 2 55 1 3 43 1 3 38 1 4 85 1 4 73
2 1 55 2 1 41 2 2 67 2 2 58 2 3 53 2 3 42 2 4 53 2 4 48
3 1 66 3 1 68 3 2 85 3 2 92 3 3 69 3 3 62 3 4 85 3 4 99
;

proc sort; by pest varty;
proc means noprint;
var resp; by pest varty;
output out=new1 mean=mn;

symbol1 v=circle i=join; symbol2 v=square i=join; symbol3 v=triangle i=join;
proc gplot;
plot mn*pest=varty;

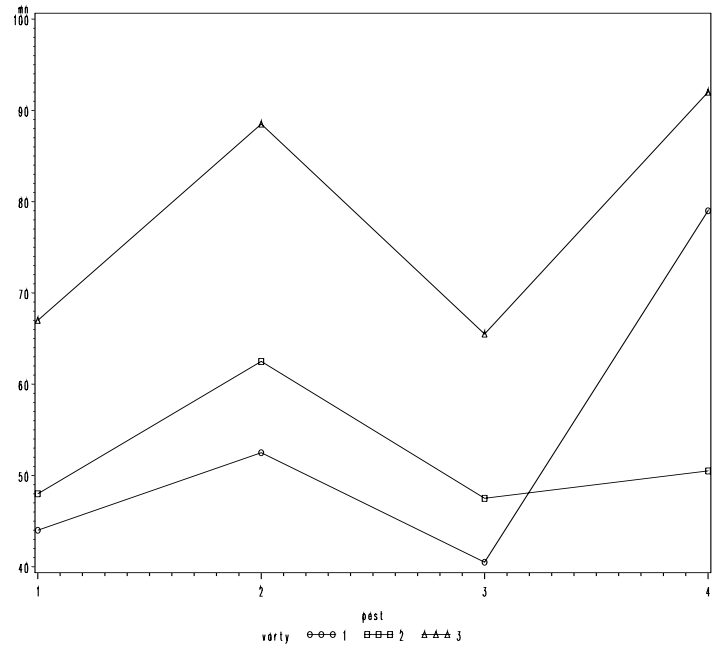
proc glm data=new;
class varty pest;
model resp = varty|pest;
lsmeans varty|pest / slice=varty adjust=tukey stderr tdiff;
output out=new1 r=res p=pred;

symbol1 v=circle i=none;
proc gplot;
plot res*pred;
run;

```

15-8

Interaction Plot



15-9

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	6680.458333	607.314394	14.36	<.0001
Error	12	507.500000	42.291667		
Corrected Total	23	7187.958333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
varty	2	3397.583333	1698.791667	40.17	<.0001
pest	3	2227.458333	742.486111	17.56	0.0001
varty*pest	6	1055.416667	175.902778	4.16	0.0171

Least Squares Means
Adjustment for Multiple Comparisons: Tukey

varty*pest Effect Sliced by varty for resp

varty	DF	Sum of Squares	Mean Square	F Value	Pr > F
1	3	1819.000000	606.333333	14.34	0.0003
2	3	297.375000	99.125000	2.34	0.1245
3	3	1166.500000	388.833333	9.19	0.0020

15-10

Least Squares Means

Adjustment for Multiple Comparisons: Tukey

varty	pest	resp LSMEAN	Standard Error	Pr > t	LSMEAN Number
1	1	44.0000000	4.5984599	<.0001	1
1	2	52.5000000	4.5984599	<.0001	2
1	3	40.5000000	4.5984599	<.0001	3
1	4	79.0000000	4.5984599	<.0001	4
2	1	48.0000000	4.5984599	<.0001	5
2	2	62.5000000	4.5984599	<.0001	6
2	3	47.5000000	4.5984599	<.0001	7
2	4	50.5000000	4.5984599	<.0001	8
3	1	67.0000000	4.5984599	<.0001	9
3	2	88.5000000	4.5984599	<.0001	10
3	3	65.5000000	4.5984599	<.0001	11
3	4	92.0000000	4.5984599	<.0001	12

Least Squares Means for Effect varty*pest
t for H0: LSMean(i)=LSMean(j) / Pr > |t|
Dependent Variable: resp

i/j	1	2	3	4	5	6
1		-1.30705 0.9623	0.538196 1.0000	-5.38196 0.0055	-0.61508 0.9999	-2.84475 0.2695
2	1.307048 0.9623		1.845244 0.7699	-4.07491 0.0422	0.691967 0.9998	-1.5377 0.9024
3	-0.5382 1.0000	-1.84524 0.7699		-5.92016 0.0024	-1.15328 0.9838	-3.38295 0.1242
4	5.381962 0.0055	4.074914 0.0422	5.920158 0.0024		4.766881 0.0142	2.537211 0.3992
5	0.615081 0.9999	-0.69197 0.9998	1.153278 0.9838	-4.76688 0.0142		-2.22967 0.5595
6	2.844751 0.2695	1.537703 0.9024	3.382948 0.1242	-2.53721 0.3992	2.22967 0.5595	

15-11

Using SAS functions

In the previous example, SAS, when making the Tukey correction, assumed we would compare all 66 pairs of means instead of the $6 \times 3 = 18$ of interest. Here is a way to better adjust for this in SAS. It involves using built in SAS functions.

1. There are three sets of comparisons (i.e., contrasts). Each set is orthogonal to another set. Think of this experiment as 12 treatments. Each set of contrasts involves a different set of four treatments.
2. For each set, we're looking at all pairwise comparisons.

Because the sets are orthogonal, consider breaking up the overall α level among the three groups using orthogonality principles and then compute Tukey's minimum significant difference for each pairwise comparison within a group. In this case, the critical value of the test statistic is 3.56411 instead of 3.97016.

```
data test;
  input FWE alev blev df mse;
/* First determine alpha level for orthogonal contrast groups */
  cut1 = probmc('maxmod',.,1-FWE,df,alev);
  group_alpha = 2*(1-probt(cut1,df));
/* Now compute Tukey's minimum significant difference */
  qval = probmc('range',.,1-group_alpha,df,blev)/sqrt(2);
  sig_dist = qval*sqrt(2*mse/(df/(alev*blev)+1));
  output;
datalines;
0.05 3 4 12 42.29166718
;
proc print;
run;
```

15-12

Sample Size/Power

- Assume both factors are fixed
- Have three different Φ^2
 - Factor A: $\Phi^2 = nbD_A^2/2a\sigma^2$
 - Factor B: $\Phi^2 = naD_B^2/2b\sigma^2$
 - Interaction : $\Phi^2 = nD_{AB}^2/2((a-1)(b-1)+1)\sigma^2$
- $df_E = ab(n-1)$
- df_{Num} depends on factor
- Compute n for each factor and choose largest n
- If $D_A = D_B = D_{AB}$, will be interaction Φ

15-13

Example: Hormone Study

Want to find any difference larger than 25 with 80% probability ($\alpha = .05$ and $\hat{\sigma}^2 = 288.35$)

$$\Phi_A^2 = 1.084n \quad \Phi_B^2 = 1.084n \quad \Phi_{AB}^2 = 0.542n$$

$$df_E = 4(n-1) \text{ and } df_{Num}=1$$

For $df_{Num}=1 \rightarrow n = 4 < .80$ and $n = 5 \approx 85\%$.

Example: Fruit Yield

Want to find any difference larger than 5 bushels with 80% probability ($\alpha = .05$ and $\hat{\sigma}^2 = 42.29$)

$$\Phi_A^2 = 0.388n \quad \Phi_B^2 = 0.222n \quad \Phi_{AB}^2 = 0.042n$$

$$df_E = 12(n-1) \text{ and } df_{Num}=2, 3 \text{ or } 6$$

For $df_{Num}=2 \rightarrow n = 8 \approx 80\%$.

For $df_{Num} = 3 \rightarrow n = 8 \approx 55\%$, $n = 13 \approx 80\%$

For $df_{Num} = 6 \rightarrow n = 45 \approx 80\%$

15-14

No Interaction Assumption

- Only one replicate, error and interaction confounded
- Experimenter may feel no interaction exists
- Allows tests of main effects
- Similar to additivity assumption - RCB design
- Check residual plot for violations - curvilinear
- Tukey's Test for Non-Additivity

15-15