Potatoes with Contrasts

```
/* potato4.sas */
options linesize=79 noovp formdlim='_' nodate;
title 'Rotten potatoes';

data spud;
     infile 'potato2.data' firstobs=2; /* Skip the first line that R uses */
     input id bact temp rot;
     combo = 10*temp+bact; /* First digit is Temp, second is Bact */

proc means mean stddev;
     class bact temp;
     var rot;
/* Get better looking output from proc tabulate below */

proc glm;
     title3 'Standard 2-way (2-factor) ANOVA with proc glm';
     class bact temp;
     model rot=temp|bact;
     means temp|bact;

proc tabulate;
     class bact temp;
     var rot;
     table (temp all),(bact all) * (mean*rot);

/* Now generate the tests for main effects and interaction, and also 2 tests
   for bacteria type, once just for low temp and once just for high. Use
   contrasts in proc glm.

               BACTERIA TYPE
TEMP      1          2         3
  1     mu11       mu12      mu13
  2     mu21       mu22      mu23

   Definition: A contrast is a linear combination whose coefficients
               add to zero. We can test whether collections of contrasts
               of cell means are all to equal zero.   */

proc glm;
     title3 'Test contrasts with proc glm';
     class combo;
     model rot=combo;
     means combo / tukey scheffe;
     contrast 'Main Effect for Temperature'
        combo  1  1  1  -1 -1 -1;
     contrast 'Main Effect for Bacteria'
        combo  1 -1  0   1 -1  0,
        combo  0  1 -1   0  1 -1;
     contrast 'Temperature by Bacteria Interaction'
        combo  1 -1  0  -1  1  0,
        combo  0  1 -1   0 -1  1;
     contrast 'Bact Marginal 1 vs 2'
        combo  1 -1  0   1 -1  0;
     contrast 'Bact Marginal 1 vs 3'
        combo  1  0 -1   1  0 -1;
```

```
     contrast 'Bact Marginal 2 vs 3'
        combo  0  1 -1   0  1 -1;
     contrast 'Bacteria Just for Low Temp'
        combo  1 -1  0   0  0  0,
        combo  0  1 -1   0  0  0;
     contrast 'Bacteria Just for High Temp'
        combo  0  0  0   1 -1  0,
        combo  0  0  0   0  1 -1;
     /* Components of the Interaction */
     contrast 'Temp Effect: Bact 1 vs 2'  /* (mu11-mu21) - (mu12-mu22) */
        combo  1 -1  0  -1  1  0;
     contrast 'Temp Effect: Bact 1 vs 3'  /* (mu11-mu21) - (mu13-mu23) */
        combo  1  0 -1  -1  0  1;
     contrast 'Temp Effect: Bact 2 vs 3'  /* (mu12-mu22) - (mu13-mu23) */
        combo  0  1 -1   0 -1  1;
     /* One could protect the last 3 tests with a Bonferroni correction as
        follow-ups to the significant interaction. Use 0.05/3 = 0.01667 */

/* Another strategy is to look at all the tests above as Scheffe followups
   to the initial significant F test for diffeence among the 6 means.
   Need Scheffe critical values. */

proc iml;
     title3 'Table of critical values for all possible Scheffe tests';
     numdf = 5;   /* Numerator degrees of freedom for initial test */
     dendf = 48;  /* Denominator degrees of freedom for initial test */
     alpha = 0.05;
     critval = finv(1-alpha,numdf,dendf);
     zero = {0 0}; S_table = repeat(zero,numdf,1);  /* Make empty matrix */
     /* Label the columns */
     namz = {"Number of Contrasts in followup test"
             "    Scheffe Critical Value"}; mattrib S_table colname=namz;
     do i = 1 to numdf;
        s_table(|i,1|) = i;
        s_table(|i,2|) = numdf/i *  critval;
     end;
     reset noname; /* Makes output look nicer in this case */
     print "Initial test has"  numdf " and " dendf "degrees of freedom."
           "Using significance level alpha = " alpha;
     print s_table;

/* Another, less conservative approach is protect pairwise comparisons of
   marginal means not as followups to the overall test that SAS reports
   first, but as followups to a significant main effect. This is a less
   conservative procedure. SAS makes it convenient, but only do it this
   way ONLY if sample sizes are equal or proportional. For unbalanced
   designs, specify your own contrasts, because what SAS will do is likely
   not what you want.

   One more thing: The proc glm below shows how to save and examine
   residuals. */
```

```
proc glm;
     title3 'Two-factor ANOVA with followups to main effect (save residuals)';
     class bact temp;
     model rot=temp|bact;
     means bact / scheffe;
     output residual=rotres; /* Creates a new SAS data set with an
                                additional variable called rotres.
                                Residuals are deviations from the
                                cell sample means. Are they normal?
                                Equal variance? */

proc univariate normal plot;
     title3 'Examine residuals';
     var rotres;
```

Rotten potatoes 1

The MEANS Procedure

Analysis Variable : rot

bact	temp	N Obs	Mean	Std Dev
1	1	9	3.5555556	4.2752518
	2	9	7.0000000	3.5355339
2	1	9	4.7777778	3.1135903
	2	9	13.5555556	6.3267510
3	1	9	8.0000000	4.5552168
	2	9	19.5555556	5.5251948

Rotten potatoes 2

Standard 2-way (2-factor) ANOVA with proc glm

The GLM Procedure

Class Level Information

Class	Levels	Values
bact	3	1 2 3
temp	2	1 2

Number of Observations Read	54
Number of Observations Used	54

Standard 2-way (2-factor) ANOVA with proc glm

The GLM Procedure

Dependent Variable: rot

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1652.814815	330.562963	15.05	<.0001
Error	48	1054.222222	21.962963		
Corrected Total	53	2707.037037			

R-Square	Coeff Var	Root MSE	rot Mean
0.610562	49.81676	4.686466	9.407407

Source	DF	Type I SS	Mean Square	F Value	Pr > F
temp	1	848.0740741	848.0740741	38.61	<.0001
bact	2	651.8148148	325.9074074	14.84	<.0001
bact*temp	2	152.9259259	76.4629630	3.48	0.0387

Source	DF	Type III SS	Mean Square	F Value	Pr > F
temp	1	848.0740741	848.0740741	38.61	<.0001
bact	2	651.8148148	325.9074074	14.84	<.0001
bact*temp	2	152.9259259	76.4629630	3.48	0.0387

Standard 2-way (2-factor) ANOVA with proc glm

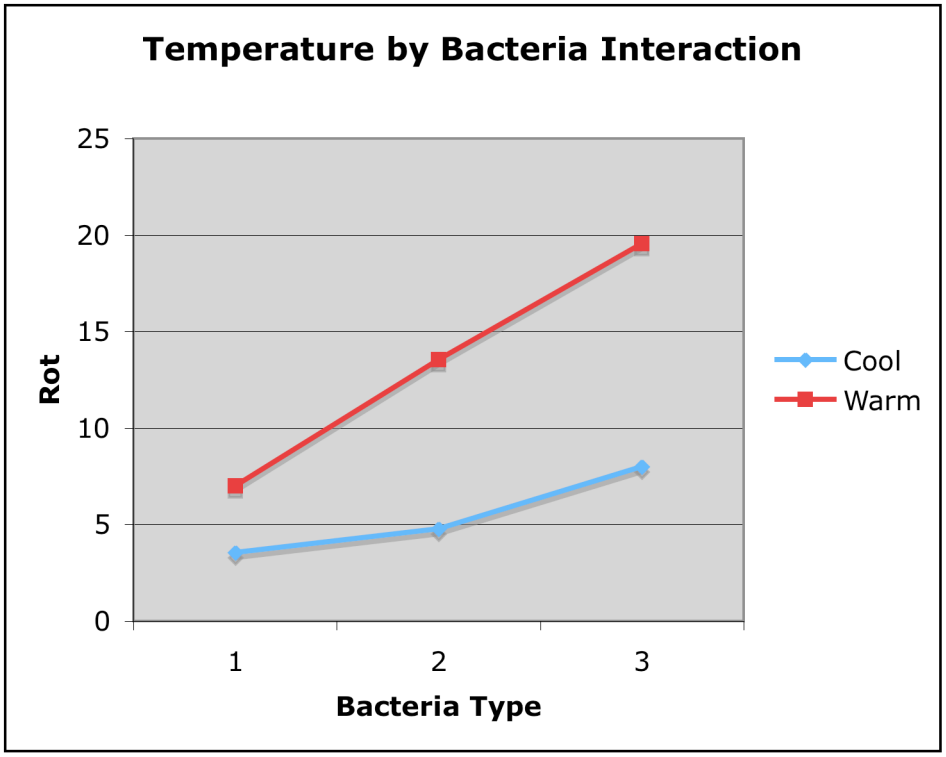
The GLM Procedure

Level of temp	N	rot Mean	rot Std Dev
1	27	5.4444444	4.31752541
2	27	13.3703704	7.27031979

Level of bact	N	rot Mean	rot Std Dev
1	18	5.2777778	4.19811660
2	18	9.1666667	6.61771242
3	18	13.7777778	7.71214135

Level of bact	Level of temp	N	rot Mean	rot Std Dev
1	1	9	3.5555556	4.27525178
1	2	9	7.0000000	3.53553391
2	1	9	4.7777778	3.11359028
2	2	9	13.5555556	6.32675097
3	1	9	8.0000000	4.55521679
3	2	9	19.5555556	5.52519482

	bact			
	1	2	3	All
	Mean	Mean	Mean	Mean
	rot	rot	rot	rot
temp				
1	3.56	4.78	8.00	5.44
2	7.00	13.56	19.56	13.37
All	5.28	9.17	13.78	9.41



Test contrasts with proc glm

The GLM Procedure

Class Level Information

Class	Levels	Values
combo	6	11 12 13 21 22 23

Number of Observations Read 54
Number of Observations Used 54

Test contrasts with proc glm

The GLM Procedure

Dependent Variable: rot

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1652.814815	330.562963	15.05	<.0001
Error	48	1054.222222	21.962963		
Corrected Total	53	2707.037037			

R-Square	Coeff Var	Root MSE	rot Mean
0.610562	49.81676	4.686466	9.407407

Source	DF	Type I SS	Mean Square	F Value	Pr > F
combo	5	1652.814815	330.562963	15.05	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
combo	5	1652.814815	330.562963	15.05	<.0001

The GLM Procedure

Tukey's Studentized Range (HSD) Test for rot

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	48
Error Mean Square	21.96296
Critical Value of Studentized Range	4.19724
Minimum Significant Difference	6.5567

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	combo
A	19.556	9	23
A			
B A	13.556	9	22
B			
B C	8.000	9	13
B C			
B C	7.000	9	21
C			
C	4.778	9	12
C			
C	3.556	9	11

Scheffe's Test for rot

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

Means with the same letter are not significantly different.

Scheffe Grouping	Mean	N	combo
A	19.556	9	23
A			
B A	13.556	9	22
B			
B C	8.000	9	13
B C			
B C	7.000	9	21
C			
C	4.778	9	12
C			
C	3.556	9	11

Test contrasts with proc glm

The GLM Procedure

Dependent Variable: rot

Contrast	DF	Contrast SS	Mean Square
Main Effect for Temperature	1	848.0740741	848.0740741
Main Effect for Bacteria	2	651.8148148	325.9074074
Temperature by Bacteria Interaction	2	152.9259259	76.4629630
Bact Marginal 1 vs 2	1	136.1111111	136.1111111
Bact Marginal 1 vs 3	1	650.2500000	650.2500000
Bact Marginal 2 vs 3	1	191.3611111	191.3611111
Bacteria Just for Low Temp	2	94.8888889	47.4444444
Bacteria Just for High Temp	2	709.8518519	354.9259259
Temp Effect: Bact 1 vs 2	1	64.0000000	64.0000000
Temp Effect: Bact 1 vs 3	1	148.0277778	148.0277778
Temp Effect: Bact 2 vs 3	1	17.3611111	17.3611111

Contrast	F Value	Pr > F
Main Effect for Temperature	38.61	<.0001
Main Effect for Bacteria	14.84	<.0001
Temperature by Bacteria Interaction	3.48	0.0387
Bact Marginal 1 vs 2	6.20	0.0163
Bact Marginal 1 vs 3	29.61	<.0001
Bact Marginal 2 vs 3	8.71	0.0049
Bacteria Just for Low Temp	2.16	0.1264
Bacteria Just for High Temp	16.16	<.0001
Temp Effect: Bact 1 vs 2	2.91	0.0943
Temp Effect: Bact 1 vs 3	6.74	0.0125
Temp Effect: Bact 2 vs 3	0.79	0.3784

Table of critical values for all possible Scheffe tests

Initial test has 5 and 48 degrees of freedom.
Using significance level alpha = 0.05

Number of Contrasts in followup test	Scheffe Critical Value
1	12.042571
2	6.0212853
3	4.0141902
4	3.0106426
5	2.4085141

The GLM Procedure

Class Level Information

Class	Levels	Values
bact	3	1 2 3
temp	2	1 2

Number of Observations Read	54
Number of Observations Used	54

The GLM Procedure

Dependent Variable: rot

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1652.814815	330.562963	15.05	<.0001
Error	48	1054.222222	21.962963		
Corrected Total	53	2707.037037			

R-Square	Coeff Var	Root MSE	rot Mean
0.610562	49.81676	4.686466	9.407407

Source	DF	Type I SS	Mean Square	F Value	Pr > F
temp	1	848.0740741	848.0740741	38.61	<.0001
bact	2	651.8148148	325.9074074	14.84	<.0001
bact*temp	2	152.9259259	76.4629630	3.48	0.0387

Source	DF	Type III SS	Mean Square	F Value	Pr > F
temp	1	848.0740741	848.0740741	38.61	<.0001
bact	2	651.8148148	325.9074074	14.84	<.0001
bact*temp	2	152.9259259	76.4629630	3.48	0.0387

Table of critical values for all possible Scheffe tests
Two-factor ANOVA with followups to main effect (save residuals)

The GLM Procedure

Scheffe's Test for rot

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

Alpha	0.05
Error Degrees of Freedom	48
Error Mean Square	21.96296
Critical Value of F	3.19073
Minimum Significant Difference	3.9462

Means with the same letter are not significantly different.

Scheffe Grouping	Mean	N	bact
A	13.778	18	3
B	9.167	18	2
B			
B	5.278	18	1

Table of critical values for all possible Scheffe tests
Examine residuals

The UNIVARIATE Procedure
Variable: rotres

Moments

N	54	Sum Weights	54
Mean	0	Sum Observations	0
Std Deviation	4.45993109	Variance	19.8909853
Skewness	-0.3768585	Kurtosis	-0.1194702
Uncorrected SS	1054.22222	Corrected SS	1054.22222
Coeff Variation	.	Std Error Mean	0.60691975

Basic Statistical Measures

Location		Variability	
Mean	0.00000	Std Deviation	4.45993
Median	0.22222	Variance	19.89099
Mode	-3.55556	Range	21.00000
		Interquartile Range	7.00000

NOTE: The mode displayed is the smallest of 2 modes with a count of 5.

Tests for Location: Mu0=0

Test	-Statistic-		-----p Value------	
Student's t	t	0	Pr > |t|	1.0000
Sign	M	1	Pr >= |M|	0.8919
Signed Rank	S	7	Pr >= |S|	0.9526

Tests for Normality

Test	--Statistic---		-----p Value------	
Shapiro-Wilk	W	0.977486	Pr < W	0.4002
Kolmogorov-Smirnov	D	0.119787	Pr > D	0.0514
Cramer-von Mises	W-Sq	0.072424	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.448962	Pr > A-Sq	>0.2500

Quantiles (Definition 5)

Quantile	Estimate
100% Max	9.444444
99%	9.444444
95%	6.444444
90%	5.222222
75% Q3	3.444444
50% Median	0.222222
25% Q1	-3.555556
10%	-5.555556

Table of critical values for all possible Scheffe tests
Examine residuals

The UNIVARIATE Procedure
Variable: rotres

Quantiles (Definition 5)

Quantile	Estimate
5%	-7.000000
1%	-11.555556
0% Min	-11.555556

Extreme Observations

Lowest	Obs	Highest	Obs
-11.55556	54	5.44444	3
-10.55556	31	5.44444	7
-7.00000	17	6.44444	46
-6.55556	33	7.00000	43
-6.00000	44	9.44444	32

Stem	Leaf	#	Boxplot
9	4	1	
8			
7	0	1	
6	4	1	
5	0244	4	
4	2444	4	
3	000004444	9	+-----+
2	04	2	| |
1	04	2	| |
0	2244	4	*--+--*
-0	8886	4	| |
-1	6000	4	| |
-2	80	2	| |
-3	666660	6	+-----+
-4	860	3	
-5	60	2	
-6	60	2	
-7	0	1	
-8			
-9			
-10	6	1	
-11	6	1	

Table of critical values for all possible Scheffe tests
Examine residuals

The UNIVARIATE Procedure
Variable: rotres

Normal Probability Plot