Computer-intensive tests with SAS `proc multtest`

```
/* sleep.sas */
options nodate linesize=79 noovp formdlim=' ';
title 'Randomization and Bootstrap tests on Student''s Sleep data';

data bedtime;
     infile 'sleep.data' firstobs=2; /* Skip the first line that R uses */
     input id extra group;

proc multtest permutation nsample=5000 seed=12345;
     title2 'Permutation Test';
     class group;
     test mean(extra);

proc multtest bootstrap nsample=5000 seed=54321;
     title2 'Bootstrap Test';
     class group;
     test mean(extra);
```

```
          Randomization and Bootstrap tests on Student's Sleep data          1
                              Permutation Test

                          The Multtest Procedure

                            Model Information

       Test for continuous variables:              Mean t-test
       Tails for continuous tests:                 Two-tailed
       Strata weights:                             None
       P-value adjustment:                         Permutation
       Center continuous variables?                No
       Number of resamples:                        5000
       Seed:                                       12345


                          Contrast Coefficients

                                          group

                Contrast                  1                 2

                Trend                     0                 1


                     Continuous Variable Tabulations

                                                                    Standard
          Variable    group    NumObs             Mean           Deviation

          extra       1            10          0.7500              1.7890
          extra       2            10          2.3300              2.0022
```

p-Values

Variable	Contrast	Raw	Permutation
extra	Trend	0.0792	0.0916

Bootstrap Test

The Multtest Procedure

Model Information

Test for continuous variables:	Mean t-test
Tails for continuous tests:	Two-tailed
Strata weights:	None
P-value adjustment:	Bootstrap
Center continuous variables?	Yes
Number of resamples:	5000
Seed:	54321

Contrast Coefficients

	group	
Contrast	1	2
Trend	0	1

Continuous Variable Tabulations

Variable	group	NumObs	Mean	Standard Deviation
extra	1	10	0.7500	1.7890
extra	2	10	2.3300	2.0022

p-Values

Variable	Contrast	Raw	Bootstrap
extra	Trend	0.0792	0.0774

Rotten Potatoes Again

```
/* potatorand.sas */
options linesize=79 noovp formdlim='_' nodate;
title 'Randomization tests on 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 tabulate;
     class bact temp;
     var rot;
     table (temp all),(bact all) * (mean*rot);

proc glm;
     title2 'Standard 2-way ANOVA with proc glm';
     class bact temp;
     model rot=temp|bact;

proc multtest permutation nsample=5000 seed=12345;
     title2 'Permutation Test: A Bad Example';
     class combo;
     test mean(rot);

proc multtest permutation nsample=5000 seed=12345;
     title2 'Permutation Test for Temperature';
     class combo;
     test mean(rot);
     contrast 'Main Effect of Temperature'
               1  1  1  -1 -1 -1;  /* Note difference from proc glm syntax */

proc multtest permutation nsample=5000 seed=12345;
     title2 'Permutation Test for Bacteria';
     class combo;
     test mean(rot);
     contrast  'Bacteria 1'  1 -1  0   1 -1  0;
     contrast  'Bacteria 2'  0  1 -1   0  1 -1;
     /* They will be tested simultaneously */

proc multtest permutation nsample=5000 seed=12345;
     title2 'Permutation Test for Temperature by Bacteria';
     class combo;
     test mean(rot);
     contrast 'Temp*Bact 1'  1 -1  0  -1  1  0;
     contrast 'Temp*Bact 2'  0  1 -1   0 -1  1;

proc glm;
     title3 'Component with proc glm';
     class combo;
     model rot=combo;
     contrast  'Bacteria 1'  combo  1 -1  0   1 -1  0;
     contrast  'Bacteria 2'  combo  0  1 -1   0  1 -1;
     contrast 'Temp*Bact 1'  combo  1 -1  0  -1  1  0;
     contrast 'Temp*Bact 2'  combo  0  1 -1   0 -1  1;
```

```
/* But the interaction F had p = 0.0387, and the main effect for bacteria had
   p < 0.0001. The lesson here is that (unfortunately) WHICH CONTRASTS
   YOU SPECIFY MAKES A DIFFERENCE. The good news is that you can specify
   redundant contrasts and they'll all be tested simultaneously.
   But you do have to think of them in advance.  */

proc multtest permutation nsample=5000 seed=12345;
     title2 'Permutation Test for Bacteria: All pairwise differences';
     title3 'Between Marginal Means';
     class combo;
     test mean(rot);
     contrast  'Bacteria 1vs2'  1 -1  0   1 -1  0;
     contrast  'Bacteria 1vs3'  1  0 -1   1  0 -1;
     contrast  'Bacteria 2vs3'  0  1 -1   0  1 -1;

/* Two final comments:

   1. Specifying a list of variables in the test statement will make it
      multivariate. All contrasts for all variables will be simultaneously
      protected against Type I error. And of course it's non-parametric
      as well as completely valid for small samples.

   2. Proc multtest has several special and very attractive methods for
      multiple binary dependent variables, and these can be used in
      conjunction with "means," so you can analyze a bunch of quantitative
      dependent variables and a bunch of binary dependent variables all
      at the same time, without paying the price of a Bonferroni correction.
```

---

	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

The GLM Procedure

Class Level Information

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

Number of observations 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

Permutation Test: A Bad Example

The Multtest Procedure

Model Information

Test for continuous variables:	Mean t-test
Tails for continuous tests:	Two-tailed
Strata weights:	None
P-value adjustment:	Permutation
Center continuous variables?	No
Number of resamples:	5000
Seed:	12345

Contrast Coefficients

	combo				
Contrast	11	12	13	21	22
Trend	0	1	2	3	4

Contrast Coefficients

	combo
Contrast	23
Trend	5

Continuous Variable Tabulations

Variable	combo	NumObs	Mean	Standard Deviation
rot	11	9	3.5556	4.2753
rot	12	9	4.7778	3.1136
rot	13	9	8.0000	4.5552
rot	21	9	7.0000	3.5355
rot	22	9	13.5556	6.3268
rot	23	9	19.5556	5.5252

p-Values

Variable	Contrast	Raw	Permutation
rot	Trend	<.0001	<.0001

Permutation Test for Temperature

The Multtest Procedure

Model Information

Test for continuous variables:	Mean t-test
Tails for continuous tests:	Two-tailed
Strata weights:	None
P-value adjustment:	Permutation
Center continuous variables?	No
Number of resamples:	5000
Seed:	12345

Contrast Coefficients

	combo			
Contrast	11	12	13	21
Main Effect of Temper	1	1	1	-1

Contrast Coefficients

	combo	
Contrast	22	23
Main Effect of Temper	-1	-1

Continuous Variable Tabulations

Variable	combo	NumObs	Mean	Standard Deviation
rot	11	9	3.5556	4.2753
rot	12	9	4.7778	3.1136
rot	13	9	8.0000	4.5552
rot	21	9	7.0000	3.5355
rot	22	9	13.5556	6.3268
rot	23	9	19.5556	5.5252

p-Values

Variable	Contrast	Raw	Permutation
rot	Main Effect of Temper	<.0001	<.0001

Permutation Test for Bacteria

The Multtest Procedure

Model Information

Test for continuous variables:	Mean t-test
Tails for continuous tests:	Two-tailed
Strata weights:	None
P-value adjustment:	Permutation
Center continuous variables?	No
Number of resamples:	5000
Seed:	12345

Contrast Coefficients

	combo			
Contrast	11	12	13	21
Bacteria 1	1	-1	0	1
Bacteria 2	0	1	-1	0

Contrast Coefficients

	combo	
Contrast	22	23
Bacteria 1	-1	0
Bacteria 2	1	-1

Continuous Variable Tabulations

Variable	combo	NumObs	Mean	Standard Deviation
rot	11	9	3.5556	4.2753
rot	12	9	4.7778	3.1136
rot	13	9	8.0000	4.5552
rot	21	9	7.0000	3.5355
rot	22	9	13.5556	6.3268
rot	23	9	19.5556	5.5252

p-Values

Variable	Contrast	Raw	Permutation
rot	Bacteria 1	0.0163	0.0298
rot	Bacteria 2	0.0049	0.0088

Permutation Test for Temperature by Bacteria

The Multtest Procedure

Model Information

Test for continuous variables:	Mean t-test
Tails for continuous tests:	Two-tailed
Strata weights:	None
P-value adjustment:	Permutation
Center continuous variables?	No
Number of resamples:	5000
Seed:	12345

Contrast Coefficients

	combo			
Contrast	11	12	13	21
Temp*Bact 1	1	-1	0	-1
Temp*Bact 2	0	1	-1	0

Contrast Coefficients

	combo	
Contrast	22	23
Temp*Bact 1	1	0
Temp*Bact 2	-1	1

Continuous Variable Tabulations

Variable	combo	NumObs	Mean	Standard Deviation
rot	11	9	3.5556	4.2753
rot	12	9	4.7778	3.1136
rot	13	9	8.0000	4.5552
rot	21	9	7.0000	3.5355
rot	22	9	13.5556	6.3268
rot	23	9	19.5556	5.5252

p-Values

Variable	Contrast	Raw	Permutation
rot	Temp*Bact 1	0.0943	0.1604
rot	Temp*Bact 2	0.3784	0.5756

Component with proc glm

The GLM Procedure

Class Level Information

Class	Levels	Values
combo	6	11 12 13 21 22 23

(Skipping the usual output ...)

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Bacteria 1	1	136.1111111	136.1111111	6.20	0.0163
Bacteria 2	1	191.3611111	191.3611111	8.71	0.0049
Temp*Bact 1	1	64.0000000	64.0000000	2.91	0.0943
Temp*Bact 2	1	17.3611111	17.3611111	0.79	0.3784

Permutation Test for Bacteria: All pairwise differences
Between Marginal Means

The Multtest Procedure

Model Information

Test for continuous variables:	Mean t-test
Tails for continuous tests:	Two-tailed
Strata weights:	None
P-value adjustment:	Permutation
Center continuous variables?	No
Number of resamples:	5000
Seed:	12345

Contrast Coefficients

	combo			
Contrast	11	12	13	21
Bacteria 1vs2	1	-1	0	1
Bacteria 1vs3	1	0	-1	1
Bacteria 2vs3	0	1	-1	0

Contrast Coefficients

	combo	
Contrast	22	23
Bacteria 1vs2	-1	0
Bacteria 1vs3	0	-1
Bacteria 2vs3	1	-1

Continuous Variable Tabulations

Variable	combo	NumObs	Mean	Standard Deviation
rot	11	9	3.5556	4.2753
rot	12	9	4.7778	3.1136
rot	13	9	8.0000	4.5552
rot	21	9	7.0000	3.5355
rot	22	9	13.5556	6.3268
rot	23	9	19.5556	5.5252

p-Values

Variable	Contrast	Raw	Permutation
rot	Bacteria 1vs2	0.0163	0.0416
rot	Bacteria 1vs3	<.0001	<.0001
rot	Bacteria 2vs3	0.0049	0.0126

1. Specifying a list of variables in the test statement will make it multivariate. All contrasts for all variables will be simultaneously protected against Type I error. And of course it's non-parametric as well as completely valid for small samples.

2. Proc multtest has several special and very attractive methods for multiple binary dependent variables, and these can be used in conjunction with "means," so you can analyze a bunch of quantitative dependent variables and a bunch of binary dependent variables all at the same time, without paying the price of a Bonferroni correction.