Permutation Tests with SAS

```
/* testmultest1.sas */
options linesize=79 noovp formdlim='_';
title 'Permutation test example from lecture: One-sided p = 0.10';
data scramble;
      input group Y;
      datalines;
      1 1.3
      1 6.0
      1 3.0
      2 5.6
      2 6.5
      2 7.1
      ;

proc glm;
      title2 'F and t-tests with proc glm';
      class group;
      model Y = group;
      estimate 'groupdiff' group 1 -1;

proc multtest permutation pvals; /* seed = here if you want */
      title2 'Permutation test, Monte Carlo p-value';
      class group;
      test mean(Y);
      /* Or test mean(Y / lowertailed); */
      contrast 'multtest groupdiff' 1 -1;


/* Commented out

Multiple comparison with permutation testing: 9.1 manual page 2957 says

The permutation method "... creates pseudo-data sets by sampling
observations without replacement from each within-stratum pool of
observations. An entire data set is thus created, and p-values for all
tests are computed on this pseudo-data set. A counter records whether the
minimum p-value from the pseudo-data set is less than or equal to the
actual p-value for each base test. (If there are R tests, then there are R
such counters.) This process is repeated a large number of times, and the
proportion of resampled data sets where the minimum pseudo-p-value is less
than or equal to an actual p-value is the adjusted p-value reported by PROC
MULTTEST. The algorithms are described by Westfall and Young (1993)."
```

                              The GLM Procedure

                          Class Level Information

                    Class   Levels   Values

                    group   2   1 2

                 Number of Observations Read   6
                 Number of Observations Used   6

                              The GLM Procedure

Dependent Variable: Y

                                     Sum of
Source   DF   Squares   Mean Square   F Value   Pr > F

Model   1   13.20166667   13.20166667   4.24   0.1087

Error   4   12.46666667   3.11666667

Corrected Total   5   25.66833333

            R-Square   Coeff Var   Root MSE   Y Mean

            0.514317   35.90661   1.765408   4.916667

Source   DF   Type I SS   Mean Square   F Value   Pr > F

group   1   13.20166667   13.20166667   4.24   0.1087

Source   DF   Type III SS   Mean Square   F Value   Pr > F

group   1   13.20166667   13.20166667   4.24   0.1087

                                      Standard
   Parameter   Estimate   Error   t Value   Pr > |t|

   groupdiff   -2.96666667   1.44144989   -2.06   0.1087

                    Permutation test, Monte Carlo p-value

                        The Multtest Procedure

                        Model Information

Test for continuous variables   Mean t-test
Degrees of Freedom Method   Pooled
Tails for continuous tests   Two-tailed
Strata weights   None
P-value adjustment   Permutation
Center continuous variables   No
Number of resamples   20000
Seed   220552383

                    Contrast Coefficients

                                            group

 Contrast   1   2

 multtest groupdiff   Centered   1   -1

                Continuous Variable Tabulations

                                                   Standard
        Variable   group   NumObs   Mean   Deviation

        Y   1   3   3.4333   2.3798
        Y   2   3   6.4000   0.7550

                            p-Values

        Variable   Contrast   Raw   Permutation

        Y   multtest groupdiff   0.1087   0.2030

Ran it again and got p = 0.1974, 0.2021, 0.1956, etc.

```
/* 305mpotato.sas */
options linesize=79 noovp formdlim='_';
title 'Rotten potatoes';
title2 'Illustrate randomization tests: Just bacteria type';

data spud;
     infile 'potato2.data' firstobs=2; /* Skip the first line that R uses */
     input id bact temp rot;

proc means;
     class bact;
     var rot;

proc glm;
     title3 'Use proc glm for comparison';
     class bact;
     model rot=bact;
     /* Contrast 1 versus average of 2 and 3 */
     contrast '1vs23' bact 2 -1 -1;
     /* Pairwise Multiple Comparisons */
     lsmeans bact / pdiff adjust=bon;
     lsmeans bact / pdiff adjust=tukey;

proc multtest permutation pvals;
     title3 'Permutation test of a single contrast';
     class bact;
     test mean(rot);
     contrast '1vs23' 2 -1 -1;

proc multtest permutation pvals;
     title3 'Multiple comparison permutation tests: All pairwise ';
     class bact;
     test mean(rot);
     contrast '1vs2' 1 -1  0;
     contrast '1vs3' 1  0 -1;
     contrast '2vs3' 0  1 -1;
```

                          The MEANS Procedure

                          Analysis Variable : rot

              N
     bact Obs   N   Mean   Std Dev   Minimum   Maximum
     ----
        1   18   18   5.2777778   4.1981166   0   10.0000000

        2   18   18   9.1666667   6.6177124   0   23.0000000

        3   18   18   13.7777778   7.7121413   2.0000000   26.0000000

                          Rotten potatoes   2
        Illustrate randomization tests: Just bacteria type
                   Use proc glm for comparison

                        The GLM Procedure

                     Class Level Information

                  Class   Levels   Values

                  bact   3   1 2 3

             Number of Observations Read   54
             Number of Observations Used   54

_____

                          Rotten potatoes   3
        Illustrate randomization tests: Just bacteria type
                   Use proc glm for comparison

                        The GLM Procedure

Dependent Variable: rot

| | | Sum of | | | |
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	2	651.814815	325.907407	8.09	0.0009
Error	51	2055.222222	40.298475		
Corrected Total	53	2707.037037			

R-Square	Coeff Var	Root MSE	rot Mean
0.240785	67.47989	6.348108	9.407407

Source	DF	Type I SS	Mean Square	F Value	Pr > F
bact	2	651.8148148	325.9074074	8.09	0.0009

Source	DF	Type III SS	Mean Square	F Value	Pr > F
bact	2	651.8148148	325.9074074	8.09	0.0009

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
1vs23	1	460.4537037	460.4537037	11.43	0.0014

                         Rotten potatoes   4
        Illustrate randomization tests: Just bacteria type
                  Use proc glm for comparison

```
                       The GLM Procedure
                    Least Squares Means
        Adjustment for Multiple Comparisons: Bonferroni

                                      LSMEAN
             bact        rot LSMEAN   Number

              1          5.2777778       1
              2          9.1666667       2
              3         13.7777778       3


        Least Squares Means for effect bact
        Pr > |t| for H0: LSMean(i)=LSMean(j)


              Dependent Variable: rot

         i/j           1           2           3

          1                     0.2158      0.0006
          2         0.2158                  0.1019
          3         0.0006      0.1019
```

_____

                         Rotten potatoes   5
        Illustrate randomization tests: Just bacteria type
                  Use proc glm for comparison

```
                       The GLM Procedure
                    Least Squares Means
        Adjustment for Multiple Comparisons: Tukey

                                      LSMEAN
             bact        rot LSMEAN   Number

              1          5.2777778       1
              2          9.1666667       2
              3         13.7777778       3


        Least Squares Means for effect bact
        Pr > |t| for H0: LSMean(i)=LSMean(j)


              Dependent Variable: rot

         i/j           1           2           3

          1                     0.1677      0.0006
          2         0.1677                  0.0845
          3         0.0006      0.0845
```

For later comparison,

```
                       p-Values

        Variable    Contrast          Raw    Permutation

        rot         1vs2           0.0719         0.1695
        rot         1vs3           0.0002         0.0004
        rot         2vs3           0.0340         0.0845
```

Illustrate randomization tests: Just bacteria type
Permutation test of a single contrast

The Multtest Procedure

Model Information

Test for continuous variables   Mean t-test
Degrees of Freedom Method   Pooled
Tails for continuous tests   Two-tailed
Strata weights   None
P-value adjustment   Permutation
Center continuous variables   No
Number of resamples   20000
Seed   790560353

Contrast Coefficients

                                          bact

Contrast   1   2   3

1vs23   Centered   2   -1   -1

Continuous Variable Tabulations

                                               Standard
        Variable   bact   NumObs   Mean   Deviation

        rot   1   18   5.2778   4.1981
        rot   2   18   9.1667   6.6177
        rot   3   18   13.7778   7.7121

p-Values

        Variable   Contrast   Raw   Permutation

        rot   1vs23   0.0014   0.0016

Illustrate randomization tests: Just bacteria type
Multiple comparison permutation tests: All pairwise

The Multtest Procedure

Model Information

Test for continuous variables	Mean t-test
Degrees of Freedom Method	Pooled
Tails for continuous tests	Two-tailed
Strata weights	None
P-value adjustment	Permutation
Center continuous variables	No
Number of resamples	20000
Seed	790815719

Contrast Coefficients

bact

Contrast		1	2	3
1vs2	Centered	1	-1	0
1vs3	Centered	1	0	-1
2vs3	Centered	0	1	-1

Continuous Variable Tabulations

Variable	bact	NumObs	Mean	Standard Deviation
rot	1	18	5.2778	4.1981
rot	2	18	9.1667	6.6177
rot	3	18	13.7778	7.7121

p-Values

Variable	Contrast	Raw	Permutation
rot	1vs2	0.0719	0.1695
rot	1vs3	0.0002	0.0004
rot	2vs3	0.0340	0.0845