

A SAS® Macro Program to Perform Permutation Testing for Group*Time Interaction in an Unbalanced Repeated-Measures Analysis of Variance.

Mark S. Litaker and Bernard Gutin

Office of Biostatistics and Department of Pediatrics, Medical College of Georgia, Augusta, GA.

ABSTRACT

In a study to evaluate the effects of physical training on cardiovascular risk factors, 79 obese 7-11 year old children were randomly assigned to participate in an exercise program during the first or second 4 months of the study period. Measurements were made at months 0, 4, and 8 with the goal of identifying differences in patterns of changes in risk factor levels between the two groups across the study period. SAS PROC GLM was used to calculate analyses of variance (ANOVA) with SUBJECT as a random effect and GROUP and TIME as fixed effects. The statistical test which is of interest is the F-test for GROUP*TIME interaction.

Distributions of residuals for some variables indicated non-normality, often due to a few relatively large residuals. Permutation testing provides a method for evaluating the statistical significance of the observed F-tests, while avoiding problematic distributional assumptions. A SAS macro program is presented which samples the permutation distribution of the F-statistic under the null hypothesis, and prints percentile values for the empirical distribution and a p-value for the permutation test.

INTRODUCTION

The defining characteristic of a permutation or randomization test is that the null distribution of a statistic of interest is generated by re-ordering the data. Under the null hypothesis, the observed test statistic is the result of randomly choosing one of these orderings. The significance level of the test is defined as the proportion of the possible orderings which yield a test statistic at least as extreme as that from the observed sample. Typically, the permutation distribution is characterized by randomly re-ordering the data and calculating the test statistic

for a relatively large sample of these orderings, rather than by complete enumeration.

Permutation tests provide a method to evaluate samples which do not meet the distributional assumptions of conventional statistical tests, without requiring the loss of the magnitude of the observed values, as is the case with rank-based nonparametric methods.

STUDY DESCRIPTION

79 apparently healthy obese 7-11 year-old children were randomly assigned to receive physical training in either the first or second four-month period of the study. A variety of measurements were made at baseline, after 4 months of physical training, and after 4 months of no physical training. Measurements included anthropometrics, body composition by dual energy x-ray absorptiometry and by magnetic resonance imaging, ultrasonographic evaluation of left-ventricular and hemodynamic characteristics, evaluation of serum lipids, insulin, leptin, and hemostatic variables (fibrinogen, PAI-1, d-dimer), and physical activity and diet recalls.

STATISTICAL ANALYSIS

The primary analysis technique used in the evaluation of measurements from this study was mixed-model analysis of variance, with SUBJECT as a random effect and GROUP and TIME as fixed effects. The test which is of interest is that for GROUP*TIME interaction, which reflects a difference in response profiles between the groups across the study period. SAS PROC GLM (SAS Institute, Inc., 1996) was used to perform the analyses. This approach allows the maximum available number of subjects to be used in the analysis, since missing values at some time points do not cause deletion of a subject's non-missing observations.

Distributions of residuals were examined using SAS PROC UNIVARIATE, which revealed distinctly non-normal distributions for many of the variables considered. In most cases, transformation of the data did not adequately normalize the residuals. For these variables, permutation testing was used to evaluate the significance of the interaction effect, and, secondarily, to provide an indication of the robustness of the normal-based analysis.

A PERMUTATION TEST

The null hypothesis for this study scenario is that the observed sample is randomly chosen from the set of all possible assignments of group and time categories to this set of measurements. Thus, the permutation distribution of the test statistic under the null hypothesis is the distribution resulting from calculating the GROUP*TIME F-statistic for all possible assignments of the observed data to the group and time categories, preserving the original sample sizes for GROUP and SUBJECT. Since complete enumeration of the permutation distribution involves an impractically large number of arrangements, the null distribution is approximated by taking a sample of random permutations of the data. The application of this approach involves randomization at two levels: subjects are randomly assigned to group categories, preserving the original sample sizes; and observations within subjects are randomly assigned to time points. This approach retains the blocking on subjects, so that subject differences do not become part of the error term under the null. Missing values are allowed, and are randomized to a time point as are actual observations. The macro reads a data set which has one record per subject, including the three observations of each dependent variable, and re-orders it to have a single observation per record and thus three records per subject. The usual ANOVA is performed on the observed data, and prints the dependent variable name along with the F-statistic and p-value for the GROUP*TIME interaction term. The permutation distribution of the F-statistic is then

sampled by randomly assigning subjects to the two groups, maintaining the original sample sizes, and randomly re-ordering each subject's observations. The ANOVA is calculated for the re-ordered data, and the F-statistic for the interaction term is written to a text file. The procedure is repeated to provide sufficient observations to adequately characterize the distribution. The p-value for the permutation test is calculated as the proportion of analyses yielding F-statistics at least as large as that from the observed data. PROC MEANS and an indicator variable are used to calculate this proportion. The program prints the p-value and the 80th through 99th percentiles of the estimated permutation distribution.

EXAMPLE OUTPUT

F-test for group*time,
using the observed data

NAME	_SOURCE_	F	PROB
INSULIN	GROUP*MONTH	3.40541	0.036379

permutation test group*time p-value

dependent variable: insulin

Analysis Variable : P (p-value)

N	Mean
2000	0.0350000

percentiles of the permutation
distribution of F (group*time)

dependent variable: insulin

P80	P81	P82	P83	P84	P85	P86
1.74	1.80	1.86	1.92	1.99	2.03	2.10
P87	P88	P89	P90	P91	P92	P93
2.15	2.22	2.31	2.40	2.48	2.59	2.75
P94	P95	P96	P97	P98	P99	
2.90	3.08	3.29	3.63	4.05	4.81	

RESULTS

The results of the permutation tests are nearly identical to those of the normal-based F-tests for all of the dependent variables considered. The analysis of left ventricular mass provides an example of the performance of the permutation test when the normality assumption is not rejected. The normal-based analysis yields a p-value of .2170 for the GROUP*TIME F-test, and the permutation test p-value is .2190. The distribution of residuals for the ratio of low frequency to high frequency power at rest, a heart-period variability measure, is non-normal but relatively symmetric. The interaction p-values are .4848 for the F-test and .5025 for the permutation test. The distribution of residuals for root mean square of successive differences in heart period shows marked skewness due to a relatively few extreme values. The p-values for the interaction term are .8260 and .8545 for the F- and permutation tests, respectively.

DISCUSSION

The similarity of results between the permutation and normal-based tests is expected in the case of well-behaved data, and is not surprising for dependent variables with non-normal but symmetrically-distributed residuals. Manly (1991, p79) states that permutation and normal-based tests typically agree closely unless the data are "quite clearly extremely non-normal." The close correspondence of results of the two methods when applied to data with moderately skewed distributions of residuals provides reassurance of the robustness of the usual analysis of variance procedure.

Apparently, none of the dependent variables considered here demonstrated sufficient skewness to cause large discrepancies between the results of the permutation and normal-based test procedures.

CONCLUSION

Permutation testing provides a means to analyze data for which distributional assumptions are questionable. Since the results of permutation tests are typically very similar to those of

normal-based tests for well-behaved data, the use of a randomization test is not necessary if a valid conventional test is available. In the case of data which markedly departs from satisfying the assumptions of the usual analysis of variance, permutation testing provides a means of evaluating the significance of a test statistic without concern that the result may be an artifact of distributional vagaries.

A PERMUTATION TEST MACRO

```
libname permute ' ';
*-----*;
* provide a name for the      *;
* dependent variable          *;
*-----*;
%let dep = insulin;

*-----*;
* identify 3 measurements of the *;
* dependent variable           *;
*-----*;
%let dep1 = pinsul; * baseline;
%let dep2 = m4insul; * month 4;
%let dep3 = m8insul; * month 8;

*-----*;
*Set data with one record per subject
* three observations per record *;
*-----*;

data hrv;
set permute.hrv;
run;

*-----*;
* analysis of observed data    *;
*-----*;
* rearrange to have a single
observation per record;
data hrv2;
set hrv;
do i = 0 to 8 by 4;
    subject = subject;
    group = group;
    month = i;
```

```

if month eq 0 then &dep = &dep1;
else if month eq 4 then &dep = &dep2;
else if month eq 8 then &dep = &dep3;
output;
end;
keep subject group month &dep;
run;

```

```

* analysis of variance;

```

```

proc glm data=hrv2 noprint
outstat=fstat;
class subject group month;
model &dep = subject(group) group
month group*month /ss3;
random subject(group);

```

```

* save the observed value of the
F-statistic ;

```

```

data fstat;
set fstat;
if _source_ = 'GROUP*MONTH';
call symput('obs_f',f);
run;
proc print data=fstat;
title 'F-test for group*time, using
the observed data';
var _name_ _source_ f prob;
run;

```

```

*****;
* Macro to randomly sample          *;
* the permutation distribution      *;
* of the group*time f-statistic    *;
* given the observed sample        *;
*****;

```

```

%macro permute(reps);
  %do i = 1 %to &reps;
    *-----*;
    reassign subjects randomly to groups
    (by re-ordering group assignments
     randomly to the records)
    *-----*;
  data group;
  set hrv;

```

```

grp=group;
ran1=ranuni(-1);
keep grp ran1;
run;
proc sort data=group;
by ran1;
run;
data reorder;
merge hrv group;
group=grp;
drop grp ran1;
run;
proc sort data=reorder;
by group;

```

```

data rework;
set reorder;
do i = 0 to 8 by 4;
  subject = subject;
  group = group;
  month = i;
  if month eq 0 then &dep = &dep1;
  else if month eq 4 then &dep = &dep2;
  else if month eq 8 then &dep = &dep3;
  output;
end;
keep subject group month &dep;
run;

```

```

*-----*;
* randomly re-order observations    *;
* within subjects                  *;
*-----*;

```

```

data subject;
set rework;
keep subject group month;
proc sort;
by subject month;
run;
data randobs;
set rework;
subj=subject;
ran=ranuni(-1);
keep &dep subj ran;
proc sort data=randobs;
by subj ran;

```

```
run;
data work;
merge subject randobs;
drop subj ran;
run;
*-----*;
*  analysis of variance for the  *;
*  randomized data                *;
*-----*;
proc glm data=work noprint
outstat=stats;
class subject group month;
model &dep = subject(group) group
month group*month /ss3;
random subject(group);

* Write the group*month F to a text
file;

data stats;
set stats;
if _source_ = 'GROUP*MONTH';
file "&dep .txt" mod;
put f;
run;
    %end;
%mend permute;

%permute(2000)
run;
*-----*;
Calculate p-value for the permutation
test, and print percentiles of the
permutation distribution
*-----*;
data fstats;
infile "&dep .txt";
input f;
p = 0;
if f ge &obs_f then p = 1;
label p='(p-value)';
run;
proc means data=fstats n mean;
title1 'permutation test group*time
p-value';
title2 "dependent variable: &dep";
var p;
```

```
run;
proc univariate data=fstats noprint;
var f;
output out=ptiles pctlpts=80 81 82 83
84 85 86 87 88 89 90 91 92 93
94 95 96 97 98 99 pctlpre=p;
proc print data=ptiles;
title1 'percentiles of the permutation
distribution of F (group*time)';
title2 "dependent variable: &dep";
run;
```

REFERENCES

SAS Institute, Inc. (1996), *SAS/STAT® Software: Changes and Enhancements through Release 6.11*, Cary, NC: SAS Institute Inc.

Manly BFG (1991), *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, London.

SAS and SAS/STAT are registered trademarks or trademarks of SAS Institute, Inc. in the USA and other countries. ® indicates USA registration.

AUTHOR

Mark S. Litaker
Office of Biostatistics
Medical College of Georgia
Augusta, GA 30912-4900
(706) 721-4846
mlitaker@mail.mcgc.edu