A PERMUTATION TEST FOR A REPEATED MEASURES DESIGN

James J. Higgins

William Noble

Recommended Citation

A PERMUTATION TEST FOR A REPEATED MEASURES DESIGN

By James J. Higgins and William Noble

Kansas State University

Abstract

Multivariate permutation tests have advantages over conventional methods in analyzing repeated measures designs. The tests are exact for all sample sizes regardless of the underlying population distribution from which the observations are selected. More importantly the tests do not require a priori assumptions about the form of the correlation structure, obviating the need to check Huynh-Feldt conditions. An example is given of how a multivariate permutation test may be conducted in a context frequently encountered in agricultural research. The SAS program corresponding to this example is also given.

1   Introduction

Permutation tests have long been part of the statistician's tool kit. Earliest examples date back to Fisher's famous tea-tasting experiment [2]. However, the volume of computation one must do in order to carry out such tests even for relatively modest-size data sets have discouraged their widespread adoption. Now with high speed computing power at the researcher's desk top, these tests are receiving new attention.

In this article we consider the classical repeated measures situation in which different treatments are applied to two groups of subjects, and each subject is then measured at several times. We first consider the classical split-plot analysis of variance approach to analyzing such repeated measures data and describe some of the correlation structure assumptions which validate the analysis. We then describe a permutation test which can be used in place of the split-plot analysis. This permutation test makes no demands upon the correlation structure and also does not require asymptotic approximations.

We also provide an almost entirely automated SAS program for performing permutation tests in situations similar to the one described in this article. This program

only requires that the user enter the raw data, the number of subjects corresponding to each treatment, and the number of iterations to be used in generating an approximation to the permutation distribution. The program then returns the results of the permutation test and also provides a profile plot which is useful in interpreting the results.

2 ANOVA Or Permutation Test?

A hazard which one always faces when using a traditional analysis of variance procedure is that the theoretical assumptions which validate the procedures (normality, equal variances, independence, etc.) may not be satisified. In practice, the assumptions are never exactly satisfied and the results obtained are always approximations. The danger inherent in the procedures is that the assumptions may be violated to such an extent that the approximations are useless.

Many tests have been developed to check whether the traditional analysis of variance assumptions are satisfied. For example, one can use normal plots to check the normality assumption, Hartley's F-max to test for equal variances, and runs tests to check on independence. (See Devore [1], Ott [7], and Rawlings [8], respectively.)

Permutation tests have three advantages over traditional ANOVA procedures. First, permutation tests have no distributional assumptions. They only rely upon the random assignment of treatments to experimental units. This randomization is not an assumption which may or may not be satisfied to one degree or another, but is instead a physical procedure which one carries out while performing the experiment. Once this procedure has been carried out, the validity of the corresponding permutation test is guaranteed. A second advantage of a permutation test is that it is exact. Traditional ANOVA methods use approximations which decrease in accuracy as the number of subjects decreases. Stated another way, the p-value (or values) which one obtains at the conclusion of a traditional ANOVA is an approximation to an exact p-value. The accuracy of the approximation is a function of the number of replications, with complete accuracy obtained only if one has infinitely many replications. On the other hand, the p-value which one obtained at the conclusion of a permutation test is *exact* - its accuracy does not depend on the number of replications. A third advantage of permutation tests is their simplicity. Since the tests rely only on the randomization procedure, they are conceptually simple and it is easy to convince researchers of their validity.

There are two traditional arguments against the use of permutation tests. The first is that it is too hard to obtain the permutation distribution. To some extent this argument is still valid, because many of the major computing packages do not yet

have permutation tests built in. This is one of the tradeoffs which must be considered when one is deciding whether or not to use a permutation test. However, a researcher with a little programming skill and with a PC on his or her desktop will find few problems in programming permutation tests such as the one described in this article. The second traditional argument against permutation tests has been that have less statistical power compared to the usual analysis of variance procedures. This will certainly be the case - *in those cases where the usual ANOVA assumptions are valid.* However, when the ANOVA traditional assumptions are not valid, permutation tests can have considerably more power. Moreover, it is frequently possible to further increase the power of a permutation test by tailoring the test statistic to fit the conditions under which the experiment was performed.

3   The Traditional Split-Plot ANOVA

In this section we give a brief outline of some of the traditonal analysis of variance methods of analyzing split-plot/repeated measures experiments. For a much more complete description, see Milliken and Johnson [6], especially Chapters 5, 24, 26, and 27.

As an illustration of a typical repeated measure design, consider the following.

Beef Carcass Example

The goal of this experiment was to compare two methods of treating beef carcasses in terms of their effect on pH measurements taken on the carcasses over time. To make the comparisons, twelve beef carcasses were assigned randomly, six to each of the two treatments. The pH measurements were made on each of the twelve carcasses at each of the six times. The experiment resulted in the following data.

Treatment 1: Control

Carcass	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6
1	6.81	6.16	5.92	5.86	5.80	5.39
2	6.68	6.30	6.12	5.71	6.09	5.28
3	6.34	6.22	5.90	5.38	5.20	5.46
4	6.68	6.24	5.83	5.49	5.37	5.43
5	6.79	6.28	6.23	5.85	5.56	5.38
6	6.85	6.51	5.95	6.06	6.31	5.39

Treatment 2: Electrical Stimulation and Hot Boning

Carcass	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6
7	6.64	5.91	5.59	5.41	5.24	5.23
8	6.57	5.89	5.32	5.41	5.32	5.30
9	6.84	6.01	5.34	5.31	5.38	5.45
10	6.71	5.60	5.29	5.37	5.26	5.41
11	6.58	5.63	5.38	5.44	5.17	5.62
12	6.68	6.04	5.62	5.31	5.41	5.44

Split-Plot or Repeated Measure?

This example is typical of the situation one encounters in split-plot and repeated measures designs. There is a larger experimental unit (the *wholeplot*) and a smaller experimental unit (the *subplot*) obtained by splitting the larger experimental unit into parts. The wholeplots in this example are the twelve carcasses and the subplots are the carcasses at the six different times. There are also two treatments: the *wholeplot treatment*, which in this example is identified as the two ways of treating the whole carcass, and the *subplot treatment*, which in this example is time.

What distinguishes this experiment as being a repeated measure, rather than a split-plot, design is the fact that the subplot treatment (time) cannot be randomly assigned to the subplot experimental units. The consequence of being unable carry out the randomization is that measurements on subplot units within a given wholeplot may be correlated.

Split-Plot Analysis

The traditional method of analyzing a repeated measures design has been to treat it as if it were a split-plot design and hope that possible correlations between observations do not greatly affect the results.

Regarding treatment as the wholeplot and time as the subplot, a split-plot analysis of the data yields the following ANOVA table.

```
Source                    DF      SS       MS       F      Pr > F
METHOD                     1    1.5901   1.5901    20.37   0.0001
METHOD*CARCASS(METHOD)    10    0.7806   0.0781
TIME                       5   13.9372   2.7874   106.64   0.0001
METHOD*TIME                5    0.8673   0.1735     6.64   0.0001
ERROR                     50    1.3069   0.0261
CORRECTED TOTAL           71   18.4821
```

Since there is a significant `time` × `treatment` interaction, it is still appropriate to make multiple comparisons of the treatments at the different times. The appropriate least significant difference (LSD) is

$$LSD_\alpha = t^*_{\alpha/2} \sqrt{\frac{2 \times \widehat{\sigma_\epsilon^2 + \sigma_e^2}}{r}},$$

where

$$\widehat{\sigma_\epsilon^2 + \sigma_e^2} = \frac{ERROR(TRT) + (b-1)ERROR(TIME)}{B},$$

and

$$t^*_{\alpha/2} = \frac{t_{\alpha/2,(a-1)(r-1)} ERROR(TRT) + t_{\alpha/2,a(r-1)(b-1)} (b-1)ERROR(TIME)}{ERROR(TRT) + (b-1)ERROR(TIME)}.$$

The development of this split-plot analysis can be found, for example, in Section 24.2 Milliken and Johnson [6]. In particular, see equations (24.2.6) and (24.2.7) contained therein. For this example, LSD(0.05) = 0.23 was obtained, which showed treatment differences at times 2 through 5, but not times 1 or 6.

Discussion of Split-Plot Analysis

Split-plot analyses on repeated measures designs have led to theoretical investigations of the conditions under which such analyses are valid. Huynh and Feldt [3] found that the split-plot analysis procedure remains valid when, in addition to the usual assumptions of multivariate normality, the variance-covariance matrices satisfy certain conditions which have come to be known as the *Huynh-Feldt* conditions. A special case of the Huynh-Feldt conditions occurs when one has *sphericity*, i.e., constant correlation within wholeplots. Thus, the usual split-plot analysis will be valid under the usual multivariate normality assumptions if, in addition, the measurements within wholeplots have constant correlation.

What do these theoretical results mean practically? In SAS there is a RE-PEATED option within PROC GLM which performs a chi-square test for sphericity (constant correlation). When this test was applied to the beef carcass data, it resulted in a chi-square value of 14.9 for 14 degrees of freedom and a p-value of 0.0015. Although it is still possible that the more general Huynh-Feldt conditions will be satisfied, further doubt has been cast upon the traditional split-plot analysis.

If Not Split-Plot, Then What?

There are at least two options at this stage. One option is to use Box's adjustment of the split-plot analysis (see Milliken and Johnson [6]). Another option is to

move to a multivariate analysis of the data. As it turned out, both the adjusted split-plot analysis and the multivariate analysis yielded the same conclusion as the previous split-plot analysis. Moreover, the mean separation of treatments at each time obtained by CONTRAST statements in SAS gives the same mean separation as the previous analysis. (This is nothing more than ordinary t-tests for comparing treatments at each time.)

However, one should keep in mind that both the adjusted split-plot and multivariate analyses carry with them their own set of assumptions, violation of which will affect the accuracy of the results to an unknown degree. Moreover, when statistical analyses begin to reach this level of complexity, a statistician performing such analyses should begin to consider whether the researcher will be able to comprehend and explain to his or her colleagues the analysis being performed. This is another one of the tradeoffs which should be considered when deciding whether to use the traditional analysis or the conceptually simple permutation test described in the following section.

4   Permutation Test Applied to Beef Carcass Data

Unlike the traditional analysis of variance methods described in the previous section, the multivariate permutation test which we will use makes no demands upon either the covariance structure or the distribution of the observations. Moreover, the logic behind the procedure is conceptually simple, relying only upon the random assignment of (wholeplot) treatments to experimental units.

The Permutation Test Algorithm

The permutation test algorithm can be applied to the beef carcass data as follows.

1. Consider all possible ways of assigning the 12 carcasses, 6 to a treatment. The number of such possibilities is

$$\binom{12}{6} = \frac{12!}{6!6!} = 924.$$

2. For each of the 924 treatment assignments, compute an "appropriate" statistic for testing for differences between treatments.

3. These 924 numbers form the **permutation distribution**, i.e., the reference distribution against which comparisons will be made. If the statistic from the actual data is in the extreme 5% of the 924 possible statistics, conclude that treatments are different at 5% level of significance.

In deciding what to permute, one should use the following rule:

The permutations should correspond to the randomization.

In the beef carcass example, the randomization consisted of assigning wholeplot treatments to wholeplot experimental units. Thus, the appropriate permutation test involves permuting among wholeplot experimental units. Since there was no random assignment of subplot treatments to subplot experimental units, then there is no corresponding permutation among subplot experimental units.

Maximum Absolute t Statistic

For the example under consideration, the statistic used was the **maximum absolute t statistic**, denoted by $|t|_{max}$. This statistic is obtained by first computing the univariate t statistic for differences between methods at each of the six times. (The reason for using t values, rather than means, is to obtain a standardized scale of measurements for plotting values across time.) Denoting these six statistics as $t_1, t_2, \ldots, t_6$, $|t|_{max}$ is then defined to be $\max_i |t_i|$, i.e., the maximum of the absolute values of the six t statistics. The SAS program in Appendix A was used to calculate the absolute t statistics at the individual times and also the maximum absolute t statistic. The following output was produced.

```
RESULTS FOR ORIGINAL DATA:
TIME =          1       ABSOLUTE t = 0.2517093
TIME =          2       ABSOLUTE t = 4.7970662
TIME =          3       ABSOLUTE t = 6.6607153
TIME =          4       ABSOLUTE t = 3.3086666
TIME =          5       ABSOLUTE t = 2.3872679
TIME =          6       ABSOLUTE t =   0.33064
MAX ABS t FOR ORIGINAL DATA = 6.6607153
```

An Approximation to the Permutation Distribution

As outlined above, the next step is to obtain the permutation distribution by computing $|t|_{max}$ for each of the 924 rearrangements of the data. If $|t|_{max}$ for the actual data exceeds the upper 5% point of the permutation distribution, then conclude that there is a difference between the two treatments.

Exact calculation of the permutation distribution involves the enumeration of all possible permutations. For the beef carcass data there are only 924 permutations and this is computationally feasible. However, if there had been, say, 15 animals

carcasses allocated to each treatment, then the number of permutations would have been

$$\binom{30}{15} = \frac{30!}{15!15!} = 155,117,520.$$

This is *not* computationally feasible. Fortunately, there is an alternative to enumeration. One can obtain an approximation to the permutation distribution by taking random samples of the permutations, using sampling with replacement. This is the procedure which is used in the SAS program in the Appendix. The program uses the following algorithm to obtain an approximate permutation distribution.

1. A random number generator is used to randomly assign to assign six carcasses to one method and six to the other.

2. The maximum absolute t statistic is calculated for this random assignment.

3. Steps 1 and 2 are repeated a total of 1000 times.

4. The upper 5% point of the 1000 values is used as the critical value.

It should be noted that although the distribution generated in this way is an approximate distribution, the approximation can be made arbitrarily good simply by doing more simulations.

The SAS program found in Appendix A will generate the approximate permutation distribution and calculate the 80th, 85th, 90th, 95th, and 99th percentiles of the approximate distribution. For the beef carcass data the following output was obtained.

```
PERCENTILES OF APPROXIMATE PERMUTATION DISTRIBUTION
    P80        P85        P90        P95        P99
  2.18031    2.32313    2.52484    2.93173    3.55844
```

The normal approximation to the binomial can be used estimate the accuracy of the approximations obtained by the simulations. Recall that 95% confidence intervals resulting from this approximation have the form

$$p \pm 2\sqrt{\frac{p(1-p)}{k}},$$

where k is the number of simulations. For example, corresponding to the value of 2.93173 in the output we obtain

$$p \pm 2\sqrt{\frac{p(1-p)}{k}} = 0.95 \pm \sqrt{\frac{0.95(1-0.95)}{1000}} = 0.95 \pm 0.0138.$$

Thus, the value `2.93173` actually corresponds to a percentile somewhere between the 93.72th and 96.38th.

Multiple Comparisons

The maximum absolute t statistic also allows one to make comparisons at individual times. The procedure for making a comparison at time i is to compare $|t_i|$ to the upper 5% point of the permutation distribution. If $|t_i|$ exceeds this upper 5% point, then conclude that the treatments differ at time i.

A useful way to display the results of this permuation test is to make a **profile plot** by plotting the absolute t statistic $|t_i|$ versus time i. If one plots these values as well as the critical value of the $|t|_{max}$ statistic, then those $|t_i|$ values falling above the critical value correspond to times at which treatments differ significantly. A profile plot for the beef carcass data is given in Figure 1. From the profile plot one can see that the permutation test procedure would make the conclusion that differences exist at times 2, 3, and 4.

Error Rate

One of the advantages of using the maximum absolute t statistic is that it controls the **experiment-wise error rate** at the 5% level under the null hypothesis of no differences between treatments. This means that if there are no differences between treatments at any of the six times, then the probability of declaring a difference to exist is 5%. This is because

$$P(\text{declaring a difference to exist}|\text{no differences})$$
$$= P(\text{at least one } |t_i| > 5\% \text{ critical point})$$
$$= P(|t|_{max} > 5\% \text{ critical point}) = 0.05.$$

In fact, the permutation test procedure does even better than this, because it also controls the experiment-wise error rate at the 5% level under a number of other partial null hypotheses. For example, suppose that the true situation is that differences exist at times 2 and 3, but not at times 1, 4, 5, or 6. Then the probability of declaring a difference to exist when it doesn't is

$$P(\max\{|t_1|, |t_4|, |t_5|, |t_6|\} > 5\% \text{ critical point})$$
$$\leq P(\max\{|t_1|, |t_2|, |t_3|, |t_4|, |t_5|, |t_6|\} > 5\% \text{ critical point})$$
$$= P(|t|_{max} > 5\% \text{ critical point}) = 0.05.$$

Thus, in terms of error rate, the permutation test serves the same purpose as the Bonferroni adjustment, but yields an exact error rate.

5 Concluding Comments

In the case in which there are more than two treatments, the corresponding permutation test would involve calculation of an F-statistic, rather than a t-statistic, at each time point. Permutation tests similar to the one described above can also be applied in simpler standard agricultural designs such as completely randomized and randomized complete block designs. Such permutation tests carry with them all of the potential advantages described in Section 2, including the lack of distributional assumptions and the exactness of the permutation distribution.

Finally, it should be emphasized that it is not necessary to use the "maximum absolute t statistic" when one performs a permutation test. We used this test statistic because it seemed appropriate for the beef carcass data. However, in other situations, other test statistics may be more appropriate. In other words, the test statistic can be "tailored" to fit the special features of the problem at hand. This gives permutation tests an added flexibility not found in the traditional analysis techniques.

A SAS Program

Description of Program

In order to use the program below, the user must have a version of SAS which implements both the SAS/IML programming language (see [5]) and the SAS Macro Facility (see [4]). However, knowledge of these two options are not required in order to use the program. The program with the beef carcass data took about five minutes to run on a 486/33mhz PC.

The program may be used in the following repeated-measures scenario.

- There are two wholeplot treatments.

    - The first wholeplot treatment is applied to m (wholeplot) experimental units.

    - The second wholeplot treatment is applied to n (wholeplot) experimental units.

- Each wholeplot experimental unit is then measured at several times. (The program will automatically label these as TIME = 1, TIME = 2, etc.) The individual times correspond to the subplot experimental units. The program requires that each wholeplot experimental unit be measured the same number of times.

The program below corresponds to the beef carcass example described in the article. Comparison of the example with the program shows that rows of data (ending in commas) correspond to wholeplot experimental units, and columns within rows correspond to subplot experimental units within wholeplot experimental units.

The user is required to enter the following items:

1. The raw data. The first m rows of the data correspond to the m wholeplot experimental units assigned to the first wholeplot treatment. The second n rows of the data correspond to the n wholeplot experimental units assigned to the second wholeplot treatment. Each column in the data corresponds to a different time (subplot experimental units).

2. m and n

3. The number of iterations to be used to obtain an approximation to the permutation distribution. This quantity is labeled `itmax` in the program. (The default value, unless changed by the user, is `itmax=1000`.)

The program produces three sets of output.

1. The absolute t statistic corresponding to each measurement time.

2. The 80th, 85th, 90th, 95th, and 99th percentiles of the approximate permutation distribution.

3. A profile plot with a reference line corresponding to the 95th percentile. A different reference line may be drawn by changing the

```
plot x*time='*' p95*time='-' / overlay;
```

statement near the end of the program. For example, to produce an 80th percentile reference line, change the statement to

```
plot x*time='*' p80*time='-' / overlay;
```

To get the appropriate title, change the

```
title 'PROFILE PLOT WITH 95th PERCENTILE REFERENCE LINE';
```

statement to

```
title 'PROFILE PLOT WITH 80th PERCENTILE REFERENCE LINE';
```

Program

```
/* REPEATED MEASURES PERMUTATION TEST */
options ps=30 ls=70 nonumber;
proc iml;
file print;
```

```
/* USER MUST INITIALIZE data, m, n, AND itmax */
/* data MATRIX: ROWS=WHOLEPLOTS=SUBJECTS, COLUMNS=SUBPLOTS=TIME */
/* (TREATMENTS RANDOMLY ASSIGNED TO ROWS) */
data={6.81 6.16 5.92 5.86 5.80 5.39,
      6.68 6.30 6.12 5.71 6.09 5.28,
      6.34 6.22 5.90 5.38 5.20 5.46,
      6.68 6.24 5.83 5.49 5.37 5.43,
      6.79 6.28 6.23 5.85 5.56 5.38,
      6.85 6.51 5.95 6.06 6.31 5.39,
      6.64 5.91 5.59 5.41 5.24 5.23,
      6.57 5.89 5.32 5.41 5.32 5.30,
      6.84 6.01 5.34 5.31 5.38 5.45,
      6.71 5.60 5.29 5.37 5.26 5.41,
      6.58 5.63 5.38 5.44 5.17 5.62,
      6.68 6.04 5.62 5.31 5.41 5.44};
/* TREATMENT 1 ASSIGNED TO FIRST m ROWS */
m=6;
/* TREATMENT 2 ASSIGNED TO LAST n ROWS */
n=6;
/* itmax=NUMBER OF SIMULATIONS */
/*        FOR APPROXIMATE PERMUTATION DISTRIBUTION */
itmax=1000;

/* nr=NUMBER OF ROWS OF DATA */
nr=nrow(data);
/* nc=NUMBER OF COLUMNS OF DATA */
nc=ncol(data);
/* INITIALIZE VECTOR OF ABSOLUTE t STATISTICS */
abst=j(nc,1,0);

/* MACRO: TWO-SAMPLE t-TEST */
/* FIRST m ROWS OF VECTOR d AGAINST LAST n ROWS OF VECTOR d */
%macro ttest(d,m,n);
d1=&d[1:&m];
d2=&d[&m+1:&m+&n];
m1=sum(d1)/nrow(d1);
m2=sum(d2)/nrow(d2);
s1s=(ssq(d1)-sum(d1)**2/nrow(d1))/(nrow(d1)-1);
s2s=(ssq(d2)-sum(d2)**2/nrow(d2))/(nrow(d2)-1);
sps=((nrow(d1)-1)*s1s + (nrow(d2)-1)*s2s)/(nrow(d1)+nrow(d2)-2);
sp=sqrt(sps);
t=(m1-m2)/(sp*sqrt(1/nrow(d1) + 1/nrow(d2)));
%mend ttest;
```

```
/* MACRO: MAXIMUM ABSOLUTE t STATISTIC ACROSS COLUMNS OF DATA */
%macro maxabs(matrix);
do time=1 to nc;
vec=&matrix[,time];
%ttest(vec,m,n);
abst[time]=abs(t);
end;
maxabst=max(abst);
%mend maxabs;

/* MAXIMUM ABSOLUTE t STATISTIC FOR ORIGINAL DATA */
%maxabs(data);
put @1 'RESULTS FOR ORIGINAL DATA:';
create b var{time x};
do time=1 to nc;
 x=abst[time];
 put @1 'TIME = ' time +5 'ABSOLUTE t = ' x;
 use b;
 append;
end;
put @1 'MAX ABS t FOR ORIGINAL DATA = ' maxabst;
close b;

/* GENERATE APPROXIMATE PERMUTATION DISTRIBUTION */
create a var{maxabst};
do j=1 to itmax;
/* PERMUTE ROWS OF ORIGINAL DATA */
r=rank(uniform(j(nr,1,0)));
permdata=data;
do k=1 to nr;
 permdata[k,]=data[r[k],];
end;
/* MAXIMUM ABSOLUTE t STATISTIC FOR PERMUTED DATA */
%maxabs(permdata);
append;
end;
close a;

proc univariate data=a noprint;
 var maxabst;
 output out=pctl
        pctlpts=80 85 90 95 99
        pctlpre=p;
```

```
run;

title 'PERCENTILES OF APPROXIMATE PERMUTATION DISTRIBUTION';
proc print data=pctl noobs;
run;

title 'PROFILE PLOT WITH 95th PERCENTILE REFERENCE LINE';
data pctl;
 set pctl;
 id=1;
data b;
 set b;
 id=1;
data c;
 merge b pctl;
 by id;
 label x='ABSOLUTE t';

proc plot data=c nolegend;
 plot x*time='*' p95*time='-' / overlay;
run;

quit;
```

References

[1] Jay L. Devore. *Probability and Statistics for Engineering and the Sciences.* Brooks/Cole Publishing Company, Pacific Grove, California, 3rd edition, 1991.

[2] Ronald A. Fisher. *The Design of Experiments.* Hafner Publishing Co., New York, third edition, 1971.

[3] H. Huynh and L.S. Feldt. Conditions under which mean square ratios in repeated measures designs have exact F-distributions. *Journal of the American Statistical Association*, 65:1582–1589, 1970.

[4] SAS Institute Inc. *SAS Guide to Macro Processing, Version 6.* SAS Institute Inc., Cary, NC, first edition, 1987.

[5] SAS Institute Inc. *SAS/IML Users's Guide.* SAS Institute Inc., Cary, NC, release 6.03 edition, 1988.

[6] George A. Milliken and Dallas E. Johnson. *Analysis of Messy Data*, volume 1. Van Nostrand Reinhold, New York, 1984.

[7] R. Lyman Ott. *An Introduction to Statistical Methods and Data Analysis*. Duxbury Press, Belmont, California, 4th edition, 1993.

[8] John O. Rawlings. *Applied Regression Analysis: A Research Tool*. Wadsworth and Brooks/Cole, Pacific Grove, California, 1988.

DEPARTMENT OF STATISTICS
KANSAS STATE UNIVERSITY
MANHATTAN, KS 66506-0802

PROFILE PLOT WITH 95th PERCENTILE REFERENCE LINE

```
        |
     8 +
        |
        |
        |                            *
        |
   A 6 +
   B    |
   S    |
   O    |            *
   L    |
   U 4 +
   T    |
   E    |                                  *
        |   -         -         -         -         -         -
   t    |                                             *
     2 +
        |
        |
        |
        |  *                                              *
     0 +
        ---+---------+---------+---------+---------+---------+---------+--
           1         2         3         4         5         6
```
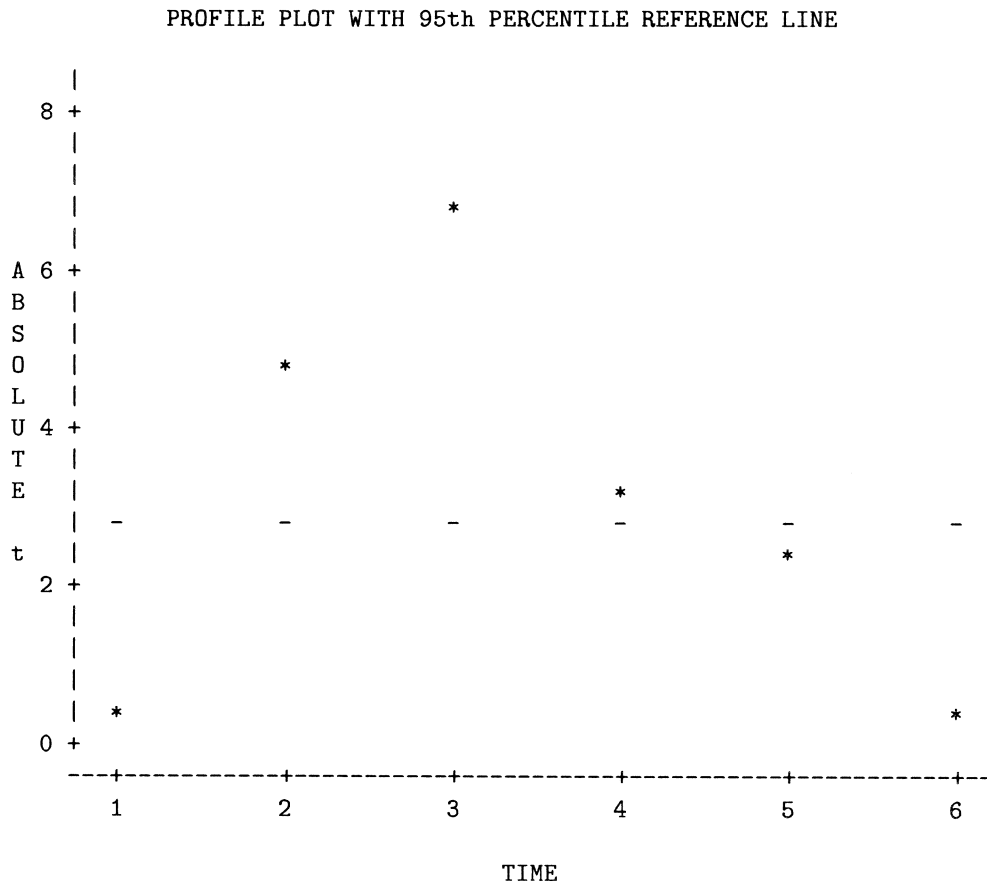
TIME

Figure 1: Profile Plot for Beef Carcass Data