ANALYSIS OF 2 X 2 CROSS-OVER TRIALS USING AN INTEGRATED MACRO PACKAGE

Harvie Chang and Tamara Strang
Marion Merrell Dow Inc., Cincinnati, OH 45215

Abstract

Two by two cross-over designs are frequently used in clinical trials. Jones and Kenward (1989) have outlined a framework for analyzing 2 x 2 cross-over trials. The framework includes techniques which can be used to quantitatively and visually assess the association between the treatments and the response of interest. We have developed a macro package, using SAS® procedures, to analyze data from a 2 x 2 cross-over design according to the framework.

Most of the techniques implemented in our macro package are commonly used to analyze cross-over data. These techniques are subject profiles plots, a sequences-by-periods plot, an analysis of variance table, residuals plots, and a confidence interval. The package also includes a less commonly used, but very powerful, plot which allows one to visually assess the carry-over and treatment effects.

Introduction

In the development of a treatment for a disease, the effectiveness of the clinical candidate is assessed by performing clinical trials. A 2 x 2 cross-over design is one of the study designs often used for such trials.

In a 2 x 2 cross-over trial, each subject is randomly assigned to one of two treatment sequences; treatment A in period one followed by treatment B in period two or treatment B in period one followed by treatment A in period two. Since the treatment effect can be tested by the typically smaller within-subject variation, a smaller sample size is required to detect a treatment effect. Thus, the advantage of a 2 x 2 cross-over design is its efficient use of resources. As the availability of subjects and the cost of studies are always two of the realistic concerns for clinical trials, a 2 x 2 cross-over design has its indispensable place.

Jones and Kenward have proposed a framework for analyzing 2 x 2 cross-over designs in the book Design and Analysis of Cross-Over Trials. We have written a macro package using SAS to implement portions of the framework. For illustration, we will follow one of the examples used in Chapter 2 of their book. The data were originally presented by Patel (1983) in a study of bronchial asthma.

Example:

Objective: To compare the efficacy of treatments A and B with respect to the Forced Expired Volume in one second (FEV1) in liters.

Variables:

y_1 = FEV1 measurement taken in the first treatment period.

y_2 = FEV1 measurement taken in the second treatment period.

Data:

SEQUENCE 1 (AB)			SEQUENCE 2 (BA)		
Subject	Period 1 y_1	Period 2 y_2	Subject	Period 1 y_1	Period 2 y_2
1	1.28	1.33	1	3.06	1.38
2	1.60	2.21	2	2.68	2.10
3	2.46	2.43	3	2.60	2.32
4	1.41	1.81	4	1.48	1.30
5	1.40	0.85	5	2.08	2.34
6	1.12	1.20	6	2.72	2.48
7	0.90	0.90	7	1.94	1.11
8	2.41	2.79	8	3.35	3.23
			9	1.16	1.25

Linear model:

$$y_{ijk} = \mu + s_{ik} + \pi_j + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + e_{ijk}$$

where,

y_{ijk} is the response of subject k in period j under sequence i, i=1,2; j=1,2; k=1,...,n_i,

μ is the overall mean,

s_{ik} is the between-subject random error assumed to be i.i.d. $N(0,\sigma^2_s)$,

π_j is the effect of period j, $\pi_1 + \pi_2 = 0$,

$d_{(i,j)}$ is the treatment assigned in period j under sequence i,

τ_l is the direct treatment effect, l=1,2; $\tau_1 + \tau_2 = 0$,

λ_l is the carry-over effect of the treatments, $\lambda_{d(1,0)} = 0$, $\lambda_1 + \lambda_2 = 0$, and

e_{ijk} is the within-subject random error assumed to be i.i.d $N(0,\sigma^2)$.

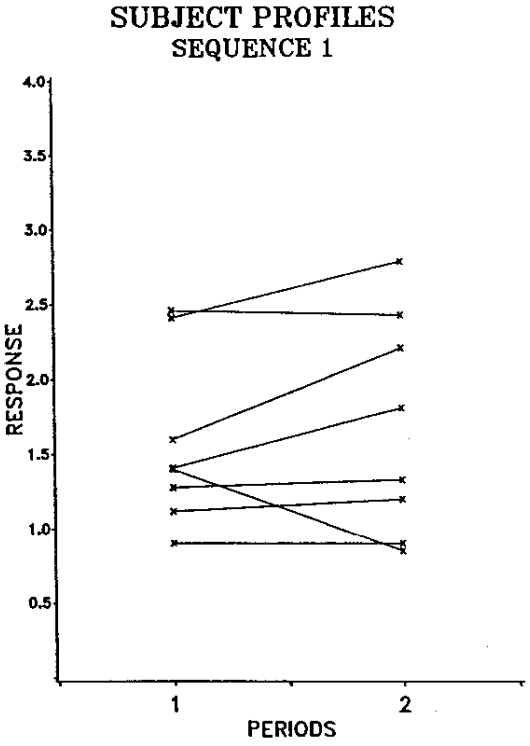
Analysis Techniques

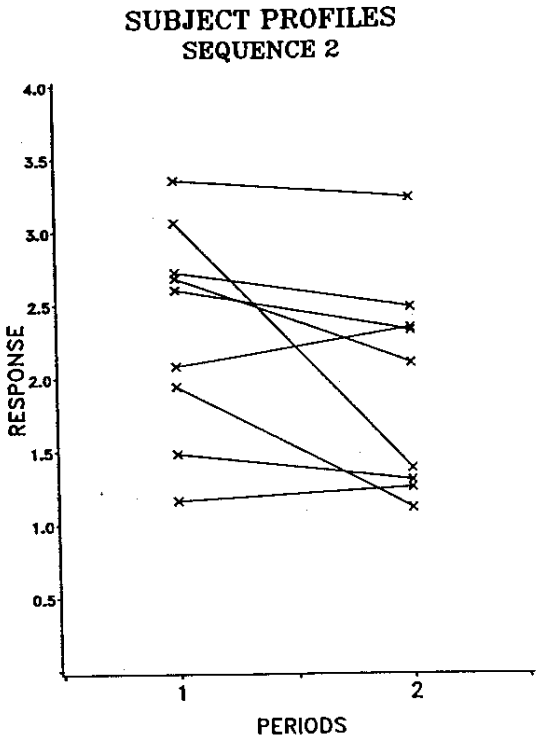
A. Preliminary Plots of the Data

Preliminary plots can provide a general idea of the characteristics of the data.

1. Plot of the Subject Profiles

A subject profiles plot presents the treatment responses for each subject. The plot is used to assess the strength as well as the consistency of the treatment effect for all subjects. The profiles are plotted by sequence since the subjects in different sequences receive the treatments in a different order. The plots obtained using the example data are shown below.

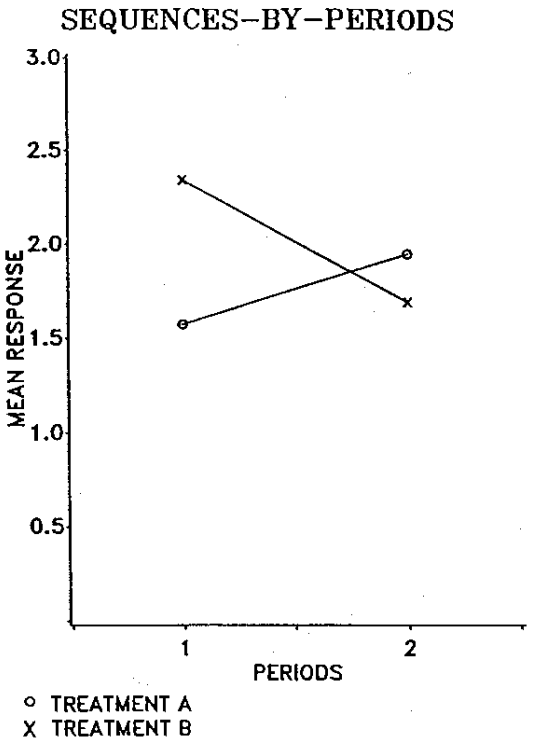




The plots show that the FEV1 values tend to be higher for treatment B than for treatment A. Note, for subject 1 in sequence 2 (BA), the difference between the response to A and to B is large, relative to the other subjects.

2. Plot of the Sequences-By-Periods Means

A sequences-by-periods plot presents the relation between the mean response to treatments A and B in both periods. For each period, the mean responses are plotted by treatment. The means which represent the same treatment are connected by a line. The example plot is shown below.



The plot shows that, in period 1, the treatment with the higher mean response is treatment B, while, in period 2, it is treatment A.

B. Analysis of 2 x 2 Cross-Over Data Using t-tests

The expectations of the response y_{ijk} in each sequence and period are as follows:

Period	Sequence 1 (AB)	Sequence 2 (BA)
1	$\mu + \pi + \tau$	$\mu + \pi - \tau$
2	$\mu - \pi - \tau + \lambda$	$\mu - \pi + \tau - \lambda$

Where $\pi = \pi_1 = -\pi_2$ since $\pi_1 + \pi_2 = 0$, and likewise $\tau = \tau_1 = -\tau_2$, $\lambda = \lambda_1 = -\lambda_2$.

With the above table we can perform hypothesis tests:

1) Test of carry-over effects (H_0:$\lambda = 0$)

The "subject totals"

$$t_{1k} = y_{11k} + y_{12k} \text{ and } t_{2k} = y_{21k} + y_{22k}$$

have expectations

$$E(t_{1k}) = 2\mu + \lambda \text{ and } E(t_{2k}) = 2\mu - \lambda,$$

respectively.

The null hypothesis that $\lambda = 0$ can be tested by using a two sample t-test to compare the mean subject totals for the two sequences. Grizzle (1965) suggested using 0.10 as the significance level for this preliminary test.

If the hypothesis $\lambda = 0$ is rejected, only period 1 data should be used for comparison. If it is not rejected, we can assume $\lambda = 0$, and proceed with the following tests.

2) Test of treatment effects (H_0: $\tau = 0$)

The "subject differences"

$$d_{1k} = y_{11k} - y_{12k} \text{ and } d_{2k} = y_{21k} - y_{22k}$$

have expectations

$$E(d_{1k}) = 2\pi + 2\tau \text{ and } E(d_{2k}) = 2\pi - 2\tau,$$

respectively.

The null hypothesis that $\tau = 0$ can be tested by using a two sample t-test to compare the mean subject differences for the two sequences.

3) Test of period effects (H_0: $\pi = 0$)

The subject cross-over differences

$$c_{1k} = y_{11k} - y_{12k} \text{ and } c_{2k} = y_{22k} - y_{21k}$$

have expectations

$$E(c_{1k}) = 2\pi + 2\tau \text{ and } E(c_{2k}) = -2\pi + 2\tau,$$

respectively.

The null hypothesis that $\pi = 0$ can be tested by using a two sample t-test to compare the mean cross-over differences for the two sequences.

C. Plot of the Convex Hulls

The convex hulls plot allows one to visualize the strength of the carry-over and treatment effects and to assess the variability within and between treatment sequences in terms of subject totals and subject differences.

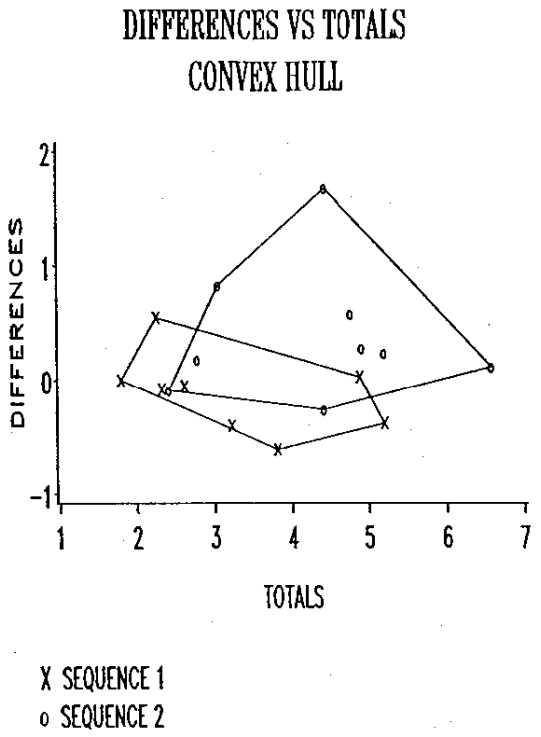
A convex hull is simply a frame of plotted points. The frame is formed by connecting the outermost points so that all points are either on or within the frame. To aid in the comparison of the two sequences, the convex hulls are plotted for both sequences.

As mentioned above, the carry-over effects were tested by comparing the subject totals for the two sequences. Assuming there were no carry-over effects, the treatment effects were tested by comparing the subject differences for the two sequences.

The degree of separation between the convex hulls, along the horizontal axis, suggests the magnitude of the carry-over effects. Likewise, the degree of separation between the convex hulls, along the vertical axis, suggests the magnitude of the treatment effects.

The area of the convex hull represents the variability within the sequence; the larger the area, the greater the variability. The positions of the two hulls reflects the variability between the sequences; the greater the overlap of the hulls, the lesser the between-sequence variability.

The differences versus totals plot obtained using the example data is shown below.



The overlap of the hulls, with respect to the horizontal axis, indicates that there is little evidence of carry-over effects. Given this, the lack of overlap, with respect to the vertical axis, suggests a treatment difference.

Based on the areas of the two hulls, the variability within sequence 2 (BA) appears larger than the variability within sequence 1 (AB). However, if the previously identified outlier is deleted (here represented by the point total=4.44, difference=1.68), the areas of the two hulls will be relatively comparable. With justification, the data can be reanalyzed without this outlier.

D. Analysis of 2 x 2 Cross-Over Data Using Analysis of Variance (ANOVA)

As an alternative to two sample t-tests, analysis of variance (ANOVA) can be used to perform hypothesis tests for a 2 x 2 cross-over study.

1. ANOVA Table Using Actual Data

The ANOVA table obtained using the example data is shown below.

SOURCE	DF	SS	MS	F	PROB
Carry-Over	1	2.2212	2.22124	2.63565	0.12531
B-S Residual	15	12.6415	0.84277		
Direct Treatments	1	0.5574	0.55742	4.67585	0.04716
Periods	1	0.1637	0.16373	1.37339	0.25951
W-S Residual	15	1.7882	0.11921		
Total	33	17.4102			

Note that the F test for carry-over effects uses the between-subject mean square as the denominator.

The table shows, at a significance level of 0.10, as suggested by Grizzle (1965), that the carry-over effects are not significant. Thus, it is appropriate to test whether the period and treatment effects are significant. The table gives evidence, at a significance level of 0.05, that the time of treatment (periods) did not influence the response while the type of treatment did. A 95%

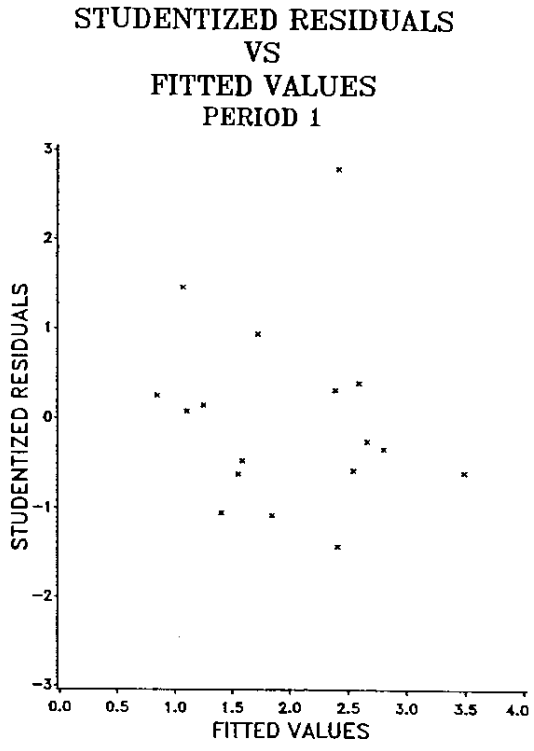
confidence interval for the treatment difference can be obtained from our macro package.

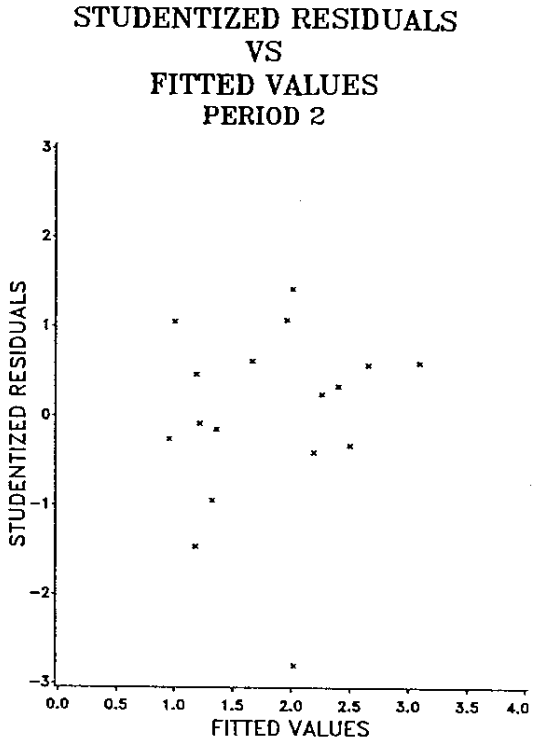
2. Analysis of Residuals

The residuals resulting from the ANOVA should always be analyzed to determine whether model assumptions have been met.

a. Plot of the studentized residuals versus the fitted values

A plot of the studentized residuals versus the fitted values allows one to identify outliers and to assess whether the assumptions of the model are correct with respect to the within-subject random error. If the assumptions of the model are correct, the points should be randomly scattered around the horizontal line defined by studentized residual equal to zero. The example plot is shown below.



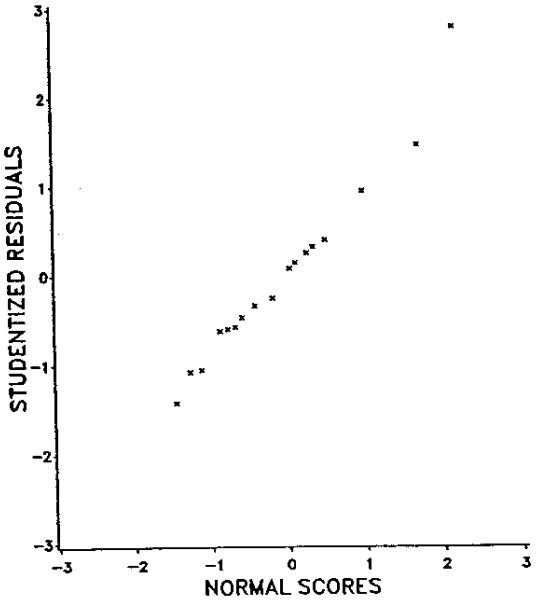


The plotted points appear to fall in a random fashion about the horizontal line zero and thus support model assumptions regarding the within-subject random error. Note that the subject who had the exceptionally large drop in the subject profiles plot has a studentized residual of 2.79 and appears to be an outlier.

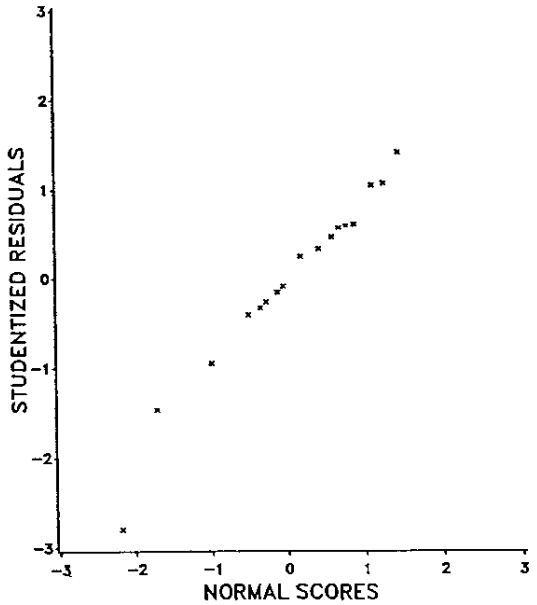
b. Plot of the Studentized Residuals Versus the Expected Normal Scores

A plot of the studentized residuals versus their corresponding expected normal order statistics will appear fairly linear if the residuals are normally distributed. Using Montgomery's (1984) formula for normal order statistics, the example plot is shown below.

STUDENTIZED RESIDUALS
VS
NORMAL SCORE
PERIOD 1



STUDENTIZED RESIDUALS
VS
NORMAL SCORE
PERIOD 2



The plot, with the exception of the outlier previously mentioned, appears to be fairly linear and thus supports the model assumption that the within-subject random error is normally distributed. The Shapiro-Wilk test of normality, although not shown here, is included in our macro package.

Acknowledgements

The authors would like to thank Dr. Eric M. Chi for advice and Mrs. Nikki Griffin for administrative support.

References

Grizzle, J.E. (1965), "The Two-Period Change-Over Design and its Use in Clinical Trials". Biometrics, 21, 467-80.

Jones, Byron and Kenward, Michael G. (1989), Design and Analysis of Cross-Over Trials, New York: Chapman and Hall.

Montgomery, Douglas C. (1984), Design and Analysis of Experiments, 2nd Edition. New York: John Wiley & Sons.

Patel, H.I. (1983), "Use of Baseline Measurements in the Two-Period Cross-Over Design". Communications in Statistics - Theory and Methods, 12, 693 - 712.

Information

A copy of the macro package can be obtained from the authors upon request:

Harvie Chang
Marion Merrell Dow Incorporated
Department of Biostatistics
Cincinnati, Ohio 45215
(513) 948-6001

Tamara Strang
Marion Merrell Dow Incorporated
Department of Biostatistics
Cincinnati, Ohio 45215
(513) 948-7383

Notes

SAS is a registered trademark of SAS Institute INC., Cary, NC, USA.