Statistical Methods AEMA-610

Simple Cross-Over Design

R. I. Cue

Department of Animal Science
Faculty of Agricultural and Environmental Sciences
McGill University

Roger.Cue@McGill.ca

Fall 2003

Contents

1 Cross Over Design

1.1 General comments

In experiments where the animal/person/experimental unit remains on the treatment from the start of the experiment until the end we can call this a continuous trial. Completely Randomised Design (One-Way ANOVA), Two-Way ANOVA (Randomised Complete Block Designs) and Factorial Models are all examples of continuous trials. In a cross-over (also called a change-over trial), however, each animal will receive consecutively two or more experimental treatments during the course of the experiment; this has similarities with the Latin Square design. The period of comparison (C.P.) is therefore divided into a number of sub-periods, which are sometimes refered to as $C.P._1$, $C.P._2$, etc. We could think of the cross-over design as being a 2-by-2 Latin Square replicated several times contemporaneously.

In a continuous trial, particularly with animals, it is common to place animals on a standard diet/treatment, prior to their random allocation to the experimental treatments. For example, one might have a standardisation period (S.P.) prior to the experiment; this might be the preceeding lactation if one was carrying out a whole (complete) lactation study with dairy cattle, or it might be the weight gain in the month preceeding the start of the trial in a feeding trial. We take account of, or exploit, the high repeatability of lactation milk yield from one lactation to another, or the relatively high corrrelation between successive weights on a growth trial; all these with the objective of reducing the experimental error, by covariance adjustment for the measures taken during the standardisation period. Since, in the change-over design, two or more treatments are contrasted on the same experimental unit (e.g. animal, cow) the between-experimental unit (between cow) variation does not enter into the experimental error. Thus, the covariance feature is not needed, and the standardisation period (S.P.) plays a minor role, if any. However, in view of the value of standardising experimental conditions it would seem eminently desirable to routinely employ a short standardisation period, although such data will not (and cannot) be used in the analysis. The basic cross-over design and analysis presented here assumes that there are no carry-over effects, or equivalently, that they are removed by any 'washout' period between the treatment periods, or that the length of time on

the treatments is sufficient to remove such residual effects. For a more advanced consideration of cross-over designs (which include this simple two-factor crossover as well as Latin squares) where carry-over effects may be present see Ratkowsky et al; Cross-over experiments, Design, Analysis and Application.

1.2 Description

The basic cross-over or simple reversal trial can be defined as one in which two treatments (A and B) are studied, and each animal (cow, experimental unit) receives both treatments in either of the sequences A, B or B, A. Thus, the basic pattern of the design is simply:

Basic Pattern

Comparison	Sequence Group	
period	1	2
1	A	B
2	B	A

where the letters in the table represent the treatments. The two periods should be the same length (of time). The experimental units (animals, cow, people) available for the experiment should be allocated to the two sequence groups at random. Usually the same number of animals should be allocated to both groups, since this provides the maximum information per experimental unit, and equivalently the smallest sampling variances. If an odd number of experimental units (animals) are available, however, the numer of animals allocated to one sequence can exceed by one the number allocated to the other sequence. There is no need to discard animals (experimental units) just to obtain equal numbers in the two sequence groups. Higher precision will be obtained than by leaving out the odd animal, although it should be recognised that the information (in the statistical sense) per unit is not quite maximum.

The cross-over design exploits the fact that in each time period we have both treatments; hence comparisons between treatments are free of period effects. We effectively remove the period effect from the comparison of treatments. Likewise, each animal receives both treatments, so the comparison of treatments is within animal, thereby removing between-animal variation from the treatment differences.

1.3 Linear Model

Linear model for dairy cow example

$Y_{ijk} = \mu + seq_i + cow_{ij} + per_k + trt_h + e_{ijk}$

where Y_{ijk} = the performance during the k^{th} period of the
j^{th} cow in the i^{th} group (i = 1,2; j = 1, 2, .., n_i; k = 1,2)
μ = the overall mean effect
seq_i = the effect of the i^{th} sequence group (i = 1,2)
cow_{ij} = the effect of the j^{th} cow on the i^{th} sequence
(j = 1, 2, .. , n_i), $cow_{ij} \sim N(0, \sigma^2_{cow})$
per_k = the effect of the k^{th} period (k = 1,2), $per_k \sim N(0, \sigma^2_{period})$
trt_h = the effect of the h^{th} treatment (h = 1,2; being a function of i and k)
e_{ijk} = the random error, $e_{ijk} \sim N(0, \sigma^2_e)$

1.4 Parameters of the model

Parameters of the model are the mean (μ), the effect of the sequence group (seq_i), the variance amongst animals (experimental units) (σ^2_{cow}), the variance amongst periods (σ^2_{per}), the effect of the treatment (trt_h), and the random residual variation (σ^2_e). We are considering that periods are a random effect and not a consistent difference that we wish to estimate. It is

possible to consider periods as a fixed effect if they correspond to repeatable time periods (perhaps months of the year, etc).

1.5 Matrix Equations

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ . \\ . \\ Y_{1n_12} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ . \\ . \\ Y_{2n_21} \\ Y_{2n_22} \end{bmatrix} = \begin{bmatrix} \mu & seq_1 & seq_2 & a_{11} & . & a_{1n_1} & a_{21} & a_{22} & . & a_{2n_2} & p_1 & p_2 & trt_1 & trt_2 \\ 1 & 1 & 0 & 1 & . & 0 & 0 & 0 & . & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & . & 0 & 0 & 0 & . & 0 & 0 & 1 & 0 & 1 \\ . & & & & & & & & & & & & & \\ 1 & 1 & 0 & 0 & . & 1 & 0 & 0 & . & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & . & 1 & 0 & 0 & . & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & . & 0 & 1 & 0 & . & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & . & 0 & 1 & 0 & . & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & . & 0 & 0 & 1 & . & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & . & 0 & 0 & 1 & . & 0 & 0 & 1 & 0 & 1 \\ . & & & & & & & & & & & & & \\ 1 & 0 & 1 & 0 & . & 0 & 0 & 0 & . & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & . & 0 & 0 & 0 & . & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ seq_1 \\ seq_2 \\ animal_{11} \\ animal_{12} \\ . \\ animal_{1n_1} \\ animal_{21} \\ animal_{22} \\ . \\ animal_{2n_2} \\ period_1 \\ period_2 \\ trt_1 \\ trt_2 \end{bmatrix} + \begin{bmatrix} \\ e_{111} \\ e_{112} \\ e_{121} \\ . \\ . \\ e_{1n_12} \\ e_{211} \\ e_{212} \\ e_{221} \\ . \\ . \\ e_{2n_2} \end{bmatrix}$$

$$Y = Xb + e$$

$$X'X\tilde{b} = X'Y$$

$$\tilde{b} = (X'X)^{-}X'Y$$

solution vector, $\tilde{b}$

$$= \begin{pmatrix} \mu \\ seq_1 \\ seq_2 \\ animal_{11} \\ animal_{12} \\ . \\ animal_{1n_1} \\ animal_{21} \\ animal_{22} \\ . \\ animal_{2n_2} \\ period_1 \\ period_2 \\ trt_1 \\ trt_2 \end{pmatrix}$$

1.6 Example data set

Period	Trt		Data			
			Sequence	group 1		
		Cow 1	Cow 2	Cow 3	Cow 4	
1	1	29.9	54.0	41.6	28.5	
2	2	27.8	49.7	38.4	26.5	
			Sequence	group 2		
		Cow 5	Cow 6	Cow 7	Cow 8	Cow 9
1	2	22.2	55.5	43.5	33.2	18.2
2	1	21.4	49.1	41.3	34.3	17.1

1.7 Derivation of CONTRASTS

Treatments: Consider the fitted values

$$\begin{array}{cc} & \hat{Y}_{111} = \tilde{\mu} + s\tilde{e}q_1 + c\tilde{o}w_{11} + p\tilde{e}r_1 + t\tilde{r}t_1 \\ - & \hat{Y}_{112} = \tilde{\mu} + s\tilde{e}q_1 + c\tilde{o}w_{11} + p\tilde{e}r_2 + t\tilde{r}t_2 \\ \hline \hat{Y}_{111} - \hat{Y}_{112} & = (p\tilde{e}r_1 - p\tilde{e}r_2) + (t\tilde{r}t_1 - t\tilde{r}t_2) \end{array}$$

$$\begin{array}{cc} & \hat{Y}_{211} = \tilde{\mu} + s\tilde{e}q_2 + c\tilde{o}w_{21} + p\tilde{e}r_1 + t\tilde{r}t_2 \\ - & \hat{Y}_{212} = \tilde{\mu} + s\tilde{e}q_2 + c\tilde{o}w_{21} + p\tilde{e}r_2 + t\tilde{r}t_1 \\ \hline \hat{Y}_{211} - \hat{Y}_{212} & = (p\tilde{e}r_1 - p\tilde{e}r_2) + (t\tilde{r}t_2 - t\tilde{r}t_1) \end{array}$$

$$\begin{array}{cc} \text{Then} & (\hat{Y}_{111} - \hat{Y}_{112} - (\hat{Y}_{211} - \hat{Y}_{212} \\ = & (p\tilde{e}r_1 - p\tilde{e}r_2) + (t\tilde{r}t_2 - t\tilde{r}t_1) \\ - & (p\tilde{e}r_1 - p\tilde{e}r_2) + (t\tilde{r}t_2 - t\tilde{r}t_1) \\ \hline = & 2(t\tilde{r}t_1 - t\tilde{r}t_2) \end{array}$$

Thus we can see that $\frac{1}{2}[(\hat{Y}_{111} - \hat{Y}_{112} - (\hat{Y}_{211} - \hat{Y}_{212}]$ provides us with a CONTRAST between the two treatments free of BOTH period effects and animal effects.

1.8 Analysis using SAS/MIXED

```
data cross;
input per trt seq cow my;
cards;
1 1 1 1 29.9
2 2 1 1 27.8
1 1 1 2 54.0
2 2 1 2 49.7
1 1 1 3 41.6
```

```
2 2 1 3 38.4
1 1 1 4 28.5
2 2 1 4 26.5
1 2 2 5 22.2
2 1 2 5 21.4
1 2 2 6 55.5
2 1 2 6 49.1
1 2 2 7 43.5
2 1 2 7 41.3
1 2 2 8 33.2
2 1 2 8 34.3
1 2 2 9 18.2
2 1 2 9 17.1
;

proc mixed;
classes per trt seq cow;
model my = seq trt;
random cow(seq) per;
lsmeans trt;
estimate 'trt 1-2' trt 1 -1;
lsmeans seq;
run;
```

1.9 Parameter Estimates And Significance

Covariance parameters	
Cow(Seq)	171.1866
period	2.5773
Residual	2.4777

Model Fitting Information	
Observations	18
Res. Log Likelihood	-49.4708
Akaike's Information Criterion	-52.4708
Schwarz's Bayesian Criterion	-53.5329
-2 Res. Log Likelihood	98.9416

Tests of Fixed Effects				
Source	NDF	DDF	Type III F	Pr > F
Sequence	1	7	0.16	0.7054
Trt	1	7	0.42	0.5372

trt_A - trt_B	0.4841	± 0.7462
lsmeans		
Trt A	35.557	± 4.564
Trt B	35.073	± 4.564
Sequence 1	37.050	± 6.663
Sequence 2	33.580	± 5.981

Note, that since sequence 2 has one more experimental unit (cow) than sequence 1 it arrives at having a smaller sampling variance and standard error for the Least squares mean. The standard errors for the two treatments are equal, due to the balance of the design. If these data had been analysed using SAS PROC/GLM we would have obtained essentially the same estimates of the Least squares means for the treatments, but the standard errors of these Least squares means would have been a factor of 10 times too small!