Calculating Standard Errors for Rodent Experiments in which Treatments Are Allocated by Cage

Stephen Senn, Competence Centre for Methodology and Statistics, Luxembourg Institute of Health

Introduction

In many rodent experiments, whole cages of rodents are allocated the same treatment. This means that naive, but standardly applied, formulae are not correct. This note explains what needs to be done to deal with this.

The basic variance component model

Variances

It is assumed that the random variation (that is to say any variability in excess of the average effect) can be explained by two additive variance terms: a between-cage term, γ^2 and a within -cage (between animal) term ϕ^2. If the animals within cages can be assumed independent, that is to say if it may be assumed that animals in a first given cage are no more similar to each other in terms of their response than they are to animals in a second cage given the same treatment, then the first variance component is zero, so we have $\gamma^2 = 0$. More generally, however this is not the case. Given the model above, the total variation for an animal chosen at random is described by a variance $\sigma^2 = \gamma^2 + \phi^2$.

Covariances

The above model implies that there are the following covariances. First, between animals in the same cage, the covariance is γ^2. Second, between animals in different cages, the covariance is 0.

Correlations

The correlation between animals in the same cage is $\rho = {\gamma^2}/{\sigma^2} = {\gamma^2}/{(\gamma^2 + \phi^2)}$ and between animals in different cages is 0.

Alternative formulation

An alternative formulation proceeds directly from the total variance and the correlation coefficient. Thus we have the total variance σ^2 and the covariance between animals in the same cage of $\rho\sigma^2$. This form is slightly more general. It is equivalent to the other form if $\rho \geq 0$. However, it also allows for the possibility that $\rho < 0$. Such a negative correlation can arise, for example, if animals in a cage compete for food. In that case the variation in weight from cage to cage can actually be less than would be predicted by treating animals as independent.

For the rest of this note, however, the additive variance formulation will be assumed.

Variances of treatment means and contrasts

Assume that each treatment is assigned to m cages of k animals, with $n = mk$ animals per treatment. In that case the variance of the treatment mean is

$$\frac{\gamma^2}{m} + \frac{\phi^2}{mk}.$$

The variance of any pairwise contrast is simply twice this figure.

Note that a naive analysis would assume a single variance, say ψ^2 , estimated largely from within-cage variation (and entirely so if there is only one cage per treatment) and $n = mk$ independent observations and hence a variance of

$$\frac{\psi^2}{mk} \text{ or } \frac{\psi^2}{n}.$$

This latter figure will only be correct if $\gamma^2 = 0$.

Analysis

For the balanced case, with equal number of animals per cage, as is well known, a fully efficient analysis using the summary measures approach can be performed(Senn, S. J. et al., 2000). All that is necessary is to reduce the figures to a mean per cage. These cage means can be put into standard software for the two sample t-test, for a case with two treatments, or using one-way analysis of variance, for a case with more than two treatments.

More generally, analysis can be carried out using standard software capable of analysing mixed models. It will be necessary to nominate 'cage' as a random effect. For the balanced case, this will yield exactly the same result as the summary measures approach. For the unbalanced case it will be more efficient, since the summary measures approach will weight cages with different numbers of animals equally, whereas the mixed model will give more weight according to cages with more animals.

The situation is analogous to that of cluster randomised clinical trials with cages, equivalent to centres, and animals equivalent to patients. This means that the literature on such trials is a useful source for planning and analysis(Campbell, M. J. & Walters, S. J., 2014; Donner, A. & Klar, J., 2000).

Note that conventional linear models cannot be used to analyse such designs: if the cage effects are fitted as fixed, they will be confounded with treatments, and no estimate of treatment effects will be possible. If they are ignored, the standard errors will be calculated incorrectly.

Complications

Unfortunately, optimal weighting in the unbalanced case requires knowledge of the *true* ratio γ^2/φ^2. However, this ratio is unknown and has to be estimated from the data(Senn, S., 2015). In general, it is imperfectly estimated and this has two consequences 1) a naive analysis will not maintain correct significance

levels and in general the P-values will be lower than they ought to be so that 'significance' will be given too easily. 2) The weights will be suboptimal.

The first difficulty can be dealt with using the Kenward and Roger correction, which is incorporated in many software packages(Kenward, M. G. & Roger, J. H., 1997). There is no ideal solution to the second. The necessary ratio is estimated from the data and the precision with which it is estimated depends on the degrees of freedom involved(Senn, S., 2017).

Design considerations

Approximate sample sizes can be calculated using the accompanying Excel spreadsheet[1]. However, nQuery® has a module for cluster randomised trials that can be used for such animal experiments also.

In general, it is necessary to have an adequate number of cages, per treatment because, first, this is what drives the element γ^2/m, second, the degrees of freedom for the t-test reflect this, third, in the unbalanced case the weights depend on the estimated variances and the precision of these, in turn, depends on degrees of freedom, the variance with the fewer degrees of freedom being γ^2.

If the design is balanced, or at least approximately balanced, some of the issues of weighting are finessed and in any case, this design is the most efficient for most obvious purposes.

References

Campbell, M. J., & Walters, S. J. (2014). *How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research*. Chichester: Wiley.

Donner, A., & Klar, J. (2000). *Design and Analysis of Cluster Randomisation Trials in Health Research: In Health Research*. London: Hodder & Stoughton Educational.

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics,* **53**(3), 983-997

Senn, S. (2015). Various varying variances: The challenge of nuisance parameters to the practising biostatistician. *Statistical Methods in Medical Research,* **24**(4), 403-419

Senn, S. (2017). Sample size considerations for n-of-1 trials. *Statistical Methods in Medical Research*

Senn, S. J., Stevens, L., & Chaturvedi, N. (2000). Repeated measures in clinical trials: simple strategies for analysis using summary measures. *Statistics in Medicine,* **19**(6), 861-877

[1] *Cluster Improved*. Available as an Excel worksheet in Stats Notes