

Some notes on what I will put in a Chapter on Experimental Design

My first comment is that almost all textbooks give information on *analysing* experiments under the heading 'Experimental Design', which is wrong. Certainly part of the design part is making sure that you can analyse the data, but the actual analysis has nothing to do with design. I am sure that you were expecting analysis in such a Chapter, so we should call the Chapter 'Design and Analysis of Experiments'. And I would begin with Design, and then move on to Analysis.

The book somewhere must have an emphasis on determining an appropriate sample size. I would prefer this to occur throughout the book (I hate examples where we are given a sample of (say) 16 observations, but there's no justification of why there are 16 in the sample). This could occur when we look at single samples, and then when we consider two independent samples, perhaps on the basis of getting a confidence interval of a target length, or perhaps in determining the power of a test. I'd prefer this to arise in situ, so to speak, because if it gets left until a Chapter of Design & Analysis, the rot has already set in.

In the actual Chapter, there'd be extensive discussion on replication and randomisation, and then on blocking. I'd emphasise that we want an experiment that will give us data so that, if we find a statistically significant difference between two treatments, there are only two valid logical conclusions: either a big fluke has occurred, or there's a difference between the effects of the treatments. This means that we have to eliminate alternative causes that might arise through confounding.

I'd talk about a Completely Randomised Design (CRD), which addresses the issues of randomisation and replication. There'd be discussion of the possible need for a control treatment (to avoid the problems of the previous paragraph). From there I would go to the Randomised Complete Block Design (RCBD), and thence to the Latin Square (LS) design (because everyone expects that to be there). But I'd discuss the disadvantages of Latin Square designs as well as the advantages. The development of the RCBD and LS design would be pitched as one extension of the CRD, based on the blocking structure of the design. Then I'd discuss the other extension, where we extend the treatment structure through a 2-factor factorial structure, and possibly a 3-factor factorial structure as well.

When I get on to analysis of the data, I would include discussion of residual analysis to ensure that the assumptions of the analysis are met (at least approximately).

I'd have a final section that would emphasise that this is just a small collection of the designs in use, and that more designs are being devised all the time. I'd want to say that we haven't considered random effects, nested treatment structures, incomplete block designs, and the design and analysis for Generalized Linear Models.