Ch 21 – Randomized Complete Block Designs (RCBDs)

Recall, a Completely Randomized Design (CRD) is one in which factors, factor levels and sample sizes are determined by the experimenter, then treatments are assigned to experimental units completely at random.

If the experimenter can partition the experimental units into rather homogeneous sets or groups, then it may be advantageous to restrict randomization in view of this. For example, plots of ground in the same field may be more similar than plots of ground in different fields, products produced from raw materials from the same batch may be more similar than those produced from raw materials from different batches, observations collected on the same day may be more similar than those collected on different days, observations collected on the sample person may be more similar than those collected on different people, and so on.

RCBD: Consider an experiment to compare r treatments. The experiment is conducted in a *randomized complete block design* if $r \times n_b$ experimental are partitioned into n_b *blocks* (or groups) of size r, each of the r treatments is observed once in each block, and treatment are assigned to experimental units at random for each block.

Generally speaking, a blocking factor is one that the experimenter is not directly interested in studying (so it's not a treatment factor), but it's one that the experimenter anticipates will be a nonnegligible source of variability.

Example: Risk Premium example, p895

Fundamentally, blocking is a trade-off. When one blocks, the blocking factor is modeled, reducing the number of error degrees of freedom—a bad thing. However, if one blocks well so that experimental units within blocks are relatively homogeneous, then the error variance σ^2 is smaller than it would be for a CRD—a good thing. For a RCBD σ^2 represents variability of experimental units within blocks, whereas for a CRD σ^2 represents variability of all experimental units used in the experiment.

How to randomize: Treatments are randomly assigned to experimental units within each block.

Example: ch20eg.txt

Standard model:

$$Y_{ij} = \mu + \rho_i + \tau_j + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2) \text{ and independent}$$

This is analogous to a single-replicate two-factor experiment, and the analysis is carried out largely the same way. However, one is generally not interested in analysis of the block effects (unless this may be of interest for future experiments), and there is no randomization test for block effects.

Data plot: Plotting Y*Treatmnt=Block or Y*Block=Treatment is useful for seeing if there are clear treatment (or block) effects.

Example: ch21eg.txt

Data analysis: One is generally only interested in analysis of the treatment effects—not the block effects.

Treatment contrasts: Treatment contrasts are those of the form $L_T = \sum_j c_j \tau_j$, $(\sum_j c_j = 0)$. Because of the design structure, the least squares estimate is $\hat{L}_T = \sum_j c_j \bar{Y}_{\cdot j}$, with $S^2(\hat{L}_T) = (\hat{\sigma}^2/n_b)(\sum_j c_j^2)$.

Individual or simultaneous inferences may be used to compare treatment effects.

ANOVA: Same as for single replicate two-factor experiment, but the test for block effects is not justified by randomization—there is no randomization test for block effects. If one treatment factor, testing $H_0 : \tau_1 = \cdots = \tau_r$. See Table 21.2, p899.

Example: ch21eg.txt

Comments:

- The model assumes additivity. Nonetheless, even if interactions exist between treatments and blocks, if block effects are substantial, blocking is probably prudent and beneficial.
- If block effects are random, the analysis is unchanged
- **Nonparametric methods** can be obtained by ranking the observations separately within each block, (see pp1138–1139). The F-statistics F_R^* computed from the ranks has asymptotically an $F(r-1,(r-1)(n_b-1))$ distribution. Asymptotically, the minimum significant difference for the Bonferroni method for pairwise differences based on treatment mean ranks is $B[r(r+1)/6n_b]^{1/2}$, where $B = z(1-\alpha/2g)$ for $g = \binom{r}{2} = r(r-1)/2$.
- **21.6 Use of more than one blocking variable:** The blocks may represent one blocking factor, or they may correspond to combinations of levels to two or more blocking factors. If for example there are two blocking factors, one can still "block in one direction" as in a RCBD, or one can "block in two directions", as one does in a row-column design, in which case the model typically involves row and column effects.
- **21.7 More than one replication per block:** If circumstances allow, one could for example have each block consist of two complete replicates. Then treatment-block interactions can be modeled, while still having error degrees of freedom available. If one of the treatments is a control, one could simply replicate the control the same number of times in each block, ideally about $\sqrt{r}$ times.
- **21.8 Factorial treatments:** The treatments in a RCBD can be factorial treatments, with every combination of treatment factors occuring once in each block. The model used may still assume additivity between treatment and block effects, or it may include lower order interactions between block and treatment effects. In the ANOVA, variability associated with block effects is still modeled and extracted.
- **Planning sample sizes:** This is the same for a RCBD as for an equireplicate 2-factor experiment, except one is only concerning with the number of blocks (i.e. replicates) needed to achieve the desired power for testing treatment effects, or to achieve confidence intervals of specified width.
- **Incomplete block designs:** One can also have incomplete blocks, as is necessary if the block size is less than the number of treatments. Such designs tend to be less efficient, since the least squares estimates of treatment contrasts must be adjusted for block effects.
- **Lost observations:** When one or more observations are lost, unless only complete blocks are lost, least squares estimates are more complicated. Then one can still do an exact analysis of the data, but in ANOVA testing treatment effects adjusted for block effects, and perhaps avoiding use of Tukey's method since it is not known to be appropriate—i.e. exact or conservative. The old fashion way would be to estimate missing values and adjust the analysis accordingly, at least if only one or a small number of observations were lost.
- **Advantages and disadvantages of a RCBD:** see pp894–895.

Residual analysis and remedial methods: Same as before, but plots of residuals versus treatments and blocks may be revealing for unequal variances. If variances only change from block to block, this will have little or no effect on treatment contrasts. If variances only change from treatment to treatment, then one could work around this for pairwise comparisons by pairing data and using the Bonferroni method.

Homework: For the data of problem 21.5, do residual analysis, plot the data, test for treatment effects, and apply Tukey's method. Discuss all aspects. Do the same for the data of problem 21.13, but use a model that includes treatment-block interactions.