Contents

9

Split-plot and nested designs

In this chapter we look at ***split-plot*** designs. The terminology comes from agriculture. One treatment factor might be applied to large plots of land. A second treatment treatment factor is then applied to smaller areas, nested within the plots. The original plots are then split up for the second treatment. The more general term is ***split-unit*** design because the experimental unit doesn't have to be a plot of land. For instance in steel production one factor might be applied to 350 tons of steel while it is being produced and a second factor might be applied to ingots weight 10 kilograms from one production run (called a "heat").

There is a clear economic advantage to split-plot experiments. For instance, it would be extremely expensive to increase the number of heats of steel being made and quite inexpensive to look at a large number of ingots from each heat. Cox (1958, Chapter 7.4) mentions another issue. We might have a limit on the natural size of blocks that stops us from putting all AB combinations into one block. Split-plot designs can fit into the block size more easily.

We will also look at two closely related topics. These are nested ANOVAs and cluster randomized trials.

9.1 Split-plot experiments

In an agriculural setting, one might drive a tractor for one kilometer while applying factor A (e.g., fertilizer) to plots of land. Then factor B (e.g., seed type) could be applied to smaller units within that kilometer, called sub-plots.

We do this when factor A is expensive to change in time or money, while factor B is inexpensive to change. The plots for factor A serve almost exactly like blocks for factor B. They're somewhat different from the usual blocks because

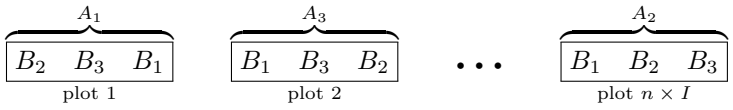
we want have purposely made them differ and want to study them in their own right.

When we do computer experiments or Monte Carlo simulations, it often makes sense to analyze them as designed experiments. If one factor A is set at the beginning of an hour long computation and then a second factor B can be set when there are just two minutes left in the computation, then a split-plot design makes sense. We choose A, compute for 58 minutes, save the internal state of the computation, and then vary factor B several times.

A split-plot design will ordinarily give us better comparisons for the inner factor B because it is blocked by the outer factor and the outer factor is not blocked. This is perhaps not surprising. If it is so much cheaper to vary factor B then it is expected that we can study it with more precision.

To begin with, we will suppose that both A and B are fixed effects. We consider random effects and mixed effects later.

Let's vary factor A at $I \geqslant 2$ levels. We will have n plots at each of those levels for a total of $n \times I$ plots. We depict them as follows



Here, factor A varies at the level of whole plots. Each level i appears n times. Factor B varies at the level of sub-plots: $j = 1, \ldots, J$, with $J = 3$ in the diagram. Each level j appears nI times

The n appearances of each level of A could be from n replicates or from a completely randomized design where I treatments were applied n times each in random order to $n \times I$ plots. When replicates are used it is usual to include an additive shift for them in the model.

We can compare levels of B using 'within-plot' differences such as $Y_{ijk} - Y_{ij'k}$ for levels $j \neq j'$ of factor B. We can compare levels of A using 'between-plot' differences such as $\bar{Y}_{i\bullet k} - \bar{Y}_{i'\bullet k}$. We expect between-plot differences to be less informative when the plots vary a lot.

The AB interaction is estimated with 'within-plot' differences. More precisely, it uses between plot differences of within plot differences that are also within plot differences of between plot differences:

$$\begin{aligned}
\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet j\bullet} + \bar{Y}_{\bullet\bullet\bullet} &= \bigl(\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet}\bigr) - \bigl(\bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet}\bigr) \\
\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} &= \frac{1}{J}\sum_{j'=1}^{J} \bar{Y}_{ij\bullet} - \bar{Y}_{ij'\bullet} \qquad \text{within} \\
\bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet} &= \frac{1}{I}\sum_{i=1}^{I} \bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} \qquad \text{also within.}
\end{aligned}$$

We see from the last line that the interaction is an average of within-plot differences and so it gets within-plot accuracy.

The ANOVA for Factor A is based on $\bar{Y}_{i\bullet k}$ for $i = 1, \ldots, I$ and $k = 1, \ldots, n$. If we the plots are in a randomized block design then our analysis uses those $I \times n$ numbers in a table like the following:

Source	df
Replications	$n-1$
A	$I-1$
Whole plot error	$(I-1)(n-1)$
Total	$In-1$

If the plots are not in n blocks of I units then we use

Source	df
A	$I-1$
Whole plot error	$I(n-1)$
Total	$In-1$

Let's prefer the blocked analysis. Then the sub-plot ANOVA table is

Source	df	
B	$J-1$	
AB	$(I-1)(J-1)$	
"Sub-plot error"	$I(J-1)(n-1)$	by subtraction
Total	$I(J-1)n$	by subtraction

The subtraction to get the subplot degrees of freedom is

$$(IJn-1) - \underbrace{In-1}_{\text{whole plots}} - (J-1) - (I-1)(J-1) = I(J-1)(n-1).$$

It is the same df as for I replicates of a $J \times n$ experiment. The subtraction to get the total degrees of freedom is

$$(IJn-1) - (In-1) = I(J-1)n.$$

We still have to figure out the sums of squares that go in these tables. We could do a full $I \times J \times n$ ANOVA with replicates $k = 1, \ldots, n$ treated as a third factor C crossed with factors A and B. Then the subplot error sum of squares is $\mathrm{SS}_{ABC} + \mathrm{SS}_{BC}$. The whole plot error sum of squares is SS_{AC}. The replicates sum of squares in the whole plot analysis is SS_C. The sums of squares for A, B and AB are, unsurprisingly, SS_A, SS_B and SS_{AB}.

There is a different analysis in Montgomery (1997, Chapter 12-4). In our notation, his model is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + c_k + (\beta c)_{jk} + (\alpha\beta c)_{ijk} + \varepsilon_{ijk},$$

where γ_k is the variable for block $k = 1, \ldots, n$. That analysis treats blocks as random effects and also allows them to have interactions with the fixed effects. Most other authors choose models in which blocks simply do not interact with

other effects, for better or worse. There would be no way to estimate $\mathbb{E}(\varepsilon_{ijk}^2)$ separately from SS_{ABC} in that model, without another level of replication. We will stay with the analysis based on the two tables described above. However, if you are in a setting where you suspect that there could be meaningful interactions between blocks and treatments, then Montgomery's approach provides a way forward.

Yandell (1997, Chapter 23) is a whole chapter on split-plot designs mostly motivated by agriculture.

9.2 More about split-plots

Jones and Nachtsheim (2009) focus on split-plot designs in industry. They consider unsuspected split-plots and quote Cuthbert Daniel as saying that most or all industrial experiments are split-plots. They raise an issue of unrecognized split-plot experiments. Suppose that an experiment varies oven temperature A, position in the oven B and recipe C. We think of an $A \times B \times C$ experiment.

Suppose that temperature is at three levels and is done on a random order schedule

350 400 375 375 350 400 $\cdots$

The question that arises is whether after run three they turned the oven off and back on again or just kept it going at 375 degrees. If the other temperature changes involved resetting the oven temperature but this one didn't, then the experiment may really be partly of split plot type, with the two consecutive runs at 375 degrees being a double-size whole plot. It is also possible that the operators might undo the randomization into:

350 350 400 400 375 375 $\cdots$

to save time and expense which would be a genuine split-plot. Jones and Nachtsheim (2009) advocate taking care to make sure that the analysis matches how the experiment was done.

In a ***split-split-plot*** experiment, the sub-plots are split further into sub-sub-plots for a third treatment factor C. The analysis involves three tables, one at the plot level, one at the sub-plot level and a new third one at the sub-sub-plot level.

In a ***strip-plot*** experiment, there are two whole plot factors that when crossed define the whole plots. The name comes from agriculture. A tractor might drive North to South placing 8 different kinds of fertilizer on the ground. Later that year a crop-dusting plain might fly East to West trajectories over the farm land, spraying 12 different pesticide treatments. That crossed structure generates 96 different whole plots. Each of those whole plots can then be divided into 4 subplots for four kinds of brocolli. Yandell (1997, Chapter 24.2) discusses strip-plot experiments.

9.3 Nested ANOVAs

Sometimes the levels of B are only defined and make sense with respect to a specific setting of factor A. Kirk (2013) describes an experiment where rats were exposed to ionized air. There were four animals per cage. There were eight cages. Each cage and the animals in it got exposed to either positive or negative ionization. They then studied a measure of the animals' activity level. We can sketch the setup as follows:

Cage	1	2	3	4	5	6	7	8
+ve:	••••	••••	••••	••••				
−ve:					••••	••••	••••	••••

We could also depict it this way:

Cage	1	2	3	4
+ve:	••••	••••	••••	••••
−ve:	••••	••••	••••	••••

However, in this diagram there is no meaningful connection between cage 1 at positive ionization and cage 1 at negative ionization. If we treat 'cage' as a factor then its meaning is dependent on the ionization level. It would not make sense to make comparisons between cage numbers 1, 2, 3 and 4 in general.

In the setting above we say that the factor 'cage' is nested within the factor 'ionization'. If we were studying schools we might label first grade classrooms with numbers 1 through 4 but classroom would ordinarily be nested within school. Schools can be nested within school boards and those can be nested within counties within states. Nesting is a hierarchical relationship often drawn using branching tree diagrams instead of a grid of boxes formed by horizontal and vertical lines as we have for crossed effects.

When B is nested within A, we write $B(A)$. This is a little odd because we just put A inside the parentheses. But since we read left to right, "B nested within A" gets the label $B(A)$.

Whether something is nested or crossed can become subtle. Ingots might ordinarily be nested within heats. For instance, we might randomly select three ingots from each heat to study. Then they are clearly nested. Then again somebody might always take the first and last and middle ones from a subsequent step in the production line. In that case the first one out of one heat does have a meaningful connection with the first one out of another heat and we have a crossed structure.

Suppose we get Y_{ijk} for animal k in cage j that gets treatment i. Then for $1 \leqslant i \leqslant I$, $1 \leqslant j \leqslant J$ and $1 \leqslant k \leqslant n$ our model has

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk} \qquad \text{or } \varepsilon_{k(ij)}.$$

For identifiability we set $\sum_{i=1}^{I} \alpha_i = 0$ and $\sum_{j=1}^{J} \beta_{j(i)} = 0$ for each $i = 1, \ldots, I$. We do not require $\beta_{j(i)}$ to sum to zero over i for any j because that would be a

sum of values that had no meaningful connection. The ANOVA decomposition for this setting is

$$\begin{aligned}
\mathrm{SS}_{E(A,B)} &= \sum_{ijk}(Y_{ijk} - \bar{Y}_{ij\bullet})^2 \\
\mathrm{SS}_{B(A)} &= \sum_{ijk}(\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet})^2 = n\sum_{ij}(\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet})^2 \\
\mathrm{SS}_{A} &= \sum_{ijk}(\bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet})^2 = Jn\sum_{i}(\bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet})^2, \quad \text{and} \\
\mathrm{SST} &= \sum_{ijk}(Y_{ijk} - \bar{Y}_{\bullet\bullet\bullet})^2 = \mathrm{SS}_{E(A,B)} + \mathrm{SS}_{B(A)} + \mathrm{SS}_{A}.
\end{aligned}$$

The new quantity is

$$\mathrm{SS}_{B(A)} = n\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet})^2.$$

For each level i, it has $J-1$ degrees of freedom and so it has $I(J-1)$ degrees of freedom in total.

Recall that the AB interaction has $(I-1)(J-1)$ df. This $B(A)$ sum of squares gets $I(J-1)-(I-1)(J-1) = J-1$ more df. They are the df for the B main effect. The B main effect is meaningless when $j=1$ has no persistent meaning as i varies. As a result we lump the B main effect in with the prior AB interaction to get $\mathrm{SS}_{B(A)} = \mathrm{SS}_B + \mathrm{SS}_{AB}$.

9.4 Expected mean squares and random effects

Now let's consider a model with a random effect B nested within another random effect A. The model has

$$Y_{ijk} = \mu + a_i + b_{j(i)} + \varepsilon_{ijk}$$

where $a_i \overset{\text{iid}}{\sim} (0, \sigma_A^2)$, independently of $b_{j(i)} \overset{\text{iid}}{\sim} (0, \sigma_B^2)$, and $\varepsilon_{ijk} \overset{\text{iid}}{\sim} (0, \sigma^2)$. When we say something has distribution (μ, σ^2) it is like $\mathcal{N}(\mu, \sigma^2)$ but without assuming normality. Derivations of F distributions require a normal distribution but expected mean squares do not. The expected mean squares in this setting are

$$\begin{aligned}
\mathbb{E}(\mathrm{MS}_A) &= \sigma^2 + n\sigma^2_{B(A)} + nJ\sigma^2_A \\
\mathbb{E}(\mathrm{MS}_{B(A)}) &= \sigma^2 + n\sigma^2_{B(A)}, \quad \text{and} \\
\mathbb{E}(\mathrm{MS}_E) &= \sigma^2.
\end{aligned}$$

Let's derive the first one. We start with $\mathrm{SS}_A = nJ\sum_{i=1}^{I}(\bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet})^2$. So it is just nJ times a sample variance among

$$\bar{Y}_{i\bullet\bullet} = \mu + a_i + \bar{b}_{\bullet(i)} + \bar{\varepsilon}_{i\bullet\bullet} \sim \Big(\mu, \sigma_A^2 + \frac{\sigma^2_{B(A)}}{J} + \frac{\sigma^2}{nJ}\Big).$$

We know from formulas for a sample variance that

$$\mathbb{E}(\mathrm{SS}_A) = nJ(I-1)\Big(\sigma_A^2 + \frac{\sigma^2_{B(A)}}{J} + \frac{\sigma^2}{nJ}\Big)$$

and so

$$\mathbb{E}(\mathrm{MS}_A) = nJ\Big(\sigma_A^2 + \frac{\sigma^2_{B(A)}}{J} + \frac{\sigma^2}{nJ}\Big) = \sigma^2 + n\sigma^2_{B(A)} + nJ\sigma_A^2.$$

The others are similar.

If B is a random effect nested in a fixed effect A, then

$$\begin{aligned}
\mathbb{E}(\mathrm{MS}_A) &= \sigma^2 + n\sigma^2_{B(A)} + \frac{nJ\sum_{i=1}^I \alpha_i^2}{I-1}\\
\mathbb{E}(\mathrm{MS}_{B(A)}) &= \sigma^2 + n\sigma^2_{B(A)}\\
\mathbb{E}(\mathrm{MS}_E) &= \sigma^2
\end{aligned}$$

and we see that σ_A^2 is replaced by a sample variance among the α_i. If both A and B are fixed, then

$$\begin{aligned}
\mathbb{E}(\mathrm{MS}_A) &= \sigma^2 + \frac{nJ\sum_{i=1}^I \alpha_i^2}{I-1}\\
\mathbb{E}(\mathrm{MS}_{B(A)}) &= \sigma^2 + n\frac{\sum_{i=1}^I\sum_{j=1}^J \beta^2_{j(i)}}{I(J-1)}, \quad \text{and}\\
\mathbb{E}(\mathrm{MS}_E) &= \sigma^2.
\end{aligned}$$

9.5 Additional models

There are more nesting patterns. We might have A, $B(A)$ and C nested within B within A, i.e., $C(B(A))$. Or we could have B and C and $B \times C$ nested within A. Or we could have A crossed with B while C is nested within A. For instance hospital A crossed with drug B and ward C nested within A. These designs can be completely randomized or arranged in randomized block structures. There is a comprehensive treatment of these situations in Kirk (2013, Chapter 11). The online version has 71 pages with worked examples and formulas for expected mean squares.

It is striking how complicated an analysis can become based on combinations of a small number of choices based on how things are nested or crossed and the way the blocks are structured. We are faced with a small but combinatorial explosion of cases. It would be very useful to have a tool such as a probabiliistic programming language that lets a user describe how the data were organized and then sets up the analysis.

9.6 Cluster randomized trials

In cluster randomized trials, we might apply a treatment at random to a whole village or a school or a sports team or a county or a marketing region, like the Bay Area versus Pheonix or Chicago. The experimental unit is a cluster of one of those types. We might also be able to get data on individuals within the cluster. Perhaps people or, in a marketing context, stores.

Let the data be Y_{ij} for $i = 1, \dots, n$ and $j = 1, \dots, n_i$. Suppose that cluster i got treatment $\mathrm{trt}(i) \in 1, \dots, I$ where ordinarily I is much less than n. We can model the individual data via

$$Y_{ij} = \mu + \alpha_{\mathrm{trt}(i)} + \varepsilon_{ij},$$

and we can model the cluster data via

$$\bar{Y}_{i\bullet} = \mu + \alpha_{\mathrm{trt}(i)} + \bar{\varepsilon}_{i\bullet}, \quad i = 1, \dots, n.$$

The most straitforward analysis is to model the cluster level data using either a randomization (permutation) approach or a one way ANOVA. It seems like a shame to greatly reduce the sample size from $N = \sum_{i=1}^{n} n_i$ individuals to just $n \ll N$ clusters. It is then tempting to analyze the data on the individual level. It would however be quite wrong to analyze the individual data as if they were N independent measurements. There are instead inter-cluster correlations. An individual level analysis can be quite unreliable if it considers the individuals to be independent when they are in fact correlated. Suppose for instance that ε_{ij} have variance σ^2, correlations ρ within a cluster and are independent between clusters. Then

$$\mathrm{var}(\bar{Y}_{i\bullet}) = \mathrm{var}(\bar{\varepsilon}_{i\bullet}) = \frac{\sigma^2}{n_i}\big(1 + (n_i - 1)\rho\big).$$

It is larger by a factor $1 + (n_i - 1)\rho$ than it would be under independence. This factor is called the ***design effect***.

If we knew ρ then we could consider an individual level analysis. For two treatments we could work out

$$\mathrm{var}\big(\mathrm{avg}(Y_{ij} \,|\, A) - \mathrm{avg}(Y_{ij} \,|\, B)\big)$$

in terms of ρ and σ^2 and all the n_i. Murray (1998) is an entire book on cluster randomized trials also known as ***group randomized trials***.

Bibliography

Cox, D. R. (1958). *Planning of experiments*. Wiley.

Jones, B. and Nachtsheim, C. J. (2009). Split-plot designs: What, why, and how. *Journal of quality technology*, 41(4):340–361.

Kirk, R. E. (2013). *Experimental design: Procedures for the Behavioral Sciences*. SAGE Publications, Inc., Thousand Oaks, CA.

Montgomery, D. C. (1997). *Design and analysis of experiments*. John wiley & sons, 4 edition.

Murray, D. M. (1998). *Design and analysis of group-randomized trials*, volume 29. Oxford University Press, USA.

Yandell, B. S. (1997). *Practical data analysis for designed experiments*. CRC Press.