A Two-Phase Wheat Variety Trial

(This example is based on a problem described to us by Kathy Haskard, BiometricsSA.) In the field phase of this experiment 49 lines of wheat are investigated using a randomized complete block design with four blocks. The produce of each plot is to be analysed using a gas chromatograph in which just seven samples can be processed in a single run. A 7×7 balanced lattice square design with four replicates is used to assign the blocks, plots and lines to four intervals in each of which there are seven runs at which samples are processed at seven consecutive times.

The sets for this experiment are analyses, plots and lines and the tiers are $\mathcal{F}_{\text{analyses}} = \{\text{Intervals}, \text{Runs}, \text{Times}\}$, $\mathcal{F}_{\text{plots}} = \{\text{Blocks}, \text{Plots}\}$, and $\mathcal{F}_{\text{lines}} = \{\text{Lines}\}$. There are two randomizations: lines to plots and plots to analyses.

There are two aspects of this experiment that prevent the use of two randomizations that are composed: (i) lines are randomized to plots within blocks in the first randomization; (ii) plots within blocks are to be randomized in the second randomization to the levels of, not a single analyses factor, but two different analyses factors, and their combinations. These obstacles could be overcome, and composed randomizations employed, by assigning the 49 plots within a block completely at random to the 49 run-time combinations within an interval. However, when it comes to the analysis of variance, it would not be feasible to isolate Runs within Intervals and Times within Intervals effects, Lines being hopelessly confounded with these terms. So we need to choose different partitions of Plots within Blocks to randomize to different sets of analyses factors. A partition into 7 parts of Plots within Blocks will be randomized to Runs within Intervals and another partition into 7 parts to Times within Intervals so that all 49 Plots within a block are randomized to the 49 Runs by Times combinations within an interval. A pseudofactor with 7 levels and nested within blocks, say P_1, is set up; it indexes which plots are randomized to which runs. Similary, a second pseudofactor, say P_2, is set up for plots randomized to times. The selection of partitions of Plots within Blocks needs to be done so that, in the analysis of variance, Lines are balanced with the three terms involving Runs and Times. Clearly, to choose these partitions it is necessary to know which Lines are associated with which Plots within Blocks — that is, to know the outcome of the first randomization. In this example, a balanced lattice square design to specify that the plots with certain lines should be processed at the same run in a particular interval and at the same time in a particular interval. This divison of the lines into groups for randomizing in a particular interval, simultaneously results in a division into groups of the plots in the block assigned to that interval. Again a pair of Lines pseudofactors for each interval can be set up, one to indicate those lines that occur in the same run in that interval and the other to indicate those that occur at the same time in that interval. Suppose the pairs are labelled: (L_1, L_2), (L_3, L_4), (L_5, L_6) and (L_7, L_8).

The randomization diagram for this experiment is given in Figure 1. It has two new features. The dashed oval shows the pseudotier created by the first randomization. One diamond, with lines going towards it, indicates that four of the Lines pseudofactors and the correspondng Plots pseudofactor are taken CJB24 in combination, but that not all combinations of the Lines pseudofactors with the Plots pseudofactor occurs: those that do are randomized to Runs or Times within Intervals, as appropriate. The diamond shows combinations of levels of factors from different tiers, in contrast to a black circle that shows combintions from the same tier.

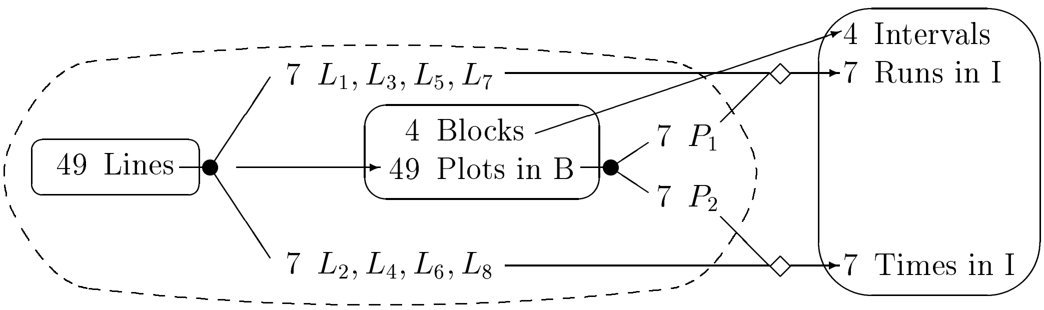

Figure 1: Randomized-inclusive randomizations in the wheat experiment

The permutation group in the first randomization is $S_{49} \wr S_4$ on plots. In the second randomization it is $(S_7 \times S_7) \wr S_4$ on analyses. Thus the effect of the first randomization is not to constrain the second randomization but to constrain the choice of initial systematic plan allocating the diamond combinations to analyses.

Note that, like the sensory and the corn-seed germination experiment, this is a two-phase experiment and involves different units in its field and laboratory phases. However, it differs from the other examples in that it involves r-inclusive, rather than composed, multiple randomizations. That is, factors from both tiers of the first randomization, $\mathcal{F}_{\text{plots}}$ and $\mathcal{F}_{\text{lines}}$, are explicitly included in the randomized factors for the second randomization.

The sets for this experiment are analyses, plots and lines. Because pseudofactors were introduced for lines and plots, the sets of factors on which the structure for the analysis will be based are $\mathcal{F}_{\text{analyses}} = \{\text{Intervals, Runs, Times}\}$, $\mathcal{F}_{\text{plots}} \cup \mathcal{H}_{\text{plots}} = \{\text{Blocks, Plots, P}_1, \text{P}_2\}$, and $\mathcal{F}_{\text{lines}} \cup \mathcal{H}_{\text{lines}} = \{\text{Lines, L}_1, \text{L}_2, \text{L}_3, \text{L}_4, \text{L}_5, \text{L}_6, \text{L}_7, \text{L}_8\}$.

The structure formulae derived from these sets are

$$\begin{gathered} 4 \text{ Intervals} / (7 \text{ Runs} * 7 \text{ Times}) \\ 4 \text{ Blocks} / (49 \text{ Plots} // (7 \text{ P}_1 + 7 \text{ P}_2)) \\ 49 \text{ Lines} // (7 \text{ L}_1 + 7 \text{ L}_2 + 7 \text{ L}_3 + 7 \text{ L}_4 + 7 \text{ L}_5 + 7 \text{ L}_6 + 7 \text{ L}_7 + 7 \text{ L}_8). \end{gathered}$$

and the Hasse diagrams for the structures derived from them are given in Figure 2. It is with respect to the structures summarized in these figures that the design is structurally balanced. The non-zero efficiency factors are

$$\begin{aligned} \lambda_{\text{Intervals}, \text{Blocks}} &= 1 \\ \lambda_{\text{Runs[I]}, \text{P}_1\text{[B]}} = \lambda_{\text{Times[I]}, \text{P}_2\text{[B]}} &= 1 \\ \lambda_{\text{P}_1\text{[B]}, \text{L}_i} = \lambda_{\text{P}_2\text{[B]}, \text{L}_j} &= 0.25 \qquad i = 1, 3, 5, 7; j = 2, 4, 6, 8 \\ \lambda_{\text{Plots[B]}, \text{L}_k} &= 0.75 \qquad k = 1 \ldots 8 \end{aligned}$$

The analysis of variance table for this example is given in Table 1.

Source		DF		Efficiency factor
Intervals	3			
Blocks		3		
Runs [Intervals]	24			
Plots [Blocks]		24		
Lines			24	1/4
Times [Intervals]	24			
Plots [Blocks]		24		
Lines			24	1/4
Runs # Times [Intervals]	144			
Plots [Blocks]		144		
Lines			48	3/4
Residual			96	
Total	195			

Table 1: Analysis variance table for the two-phase wheat experiment

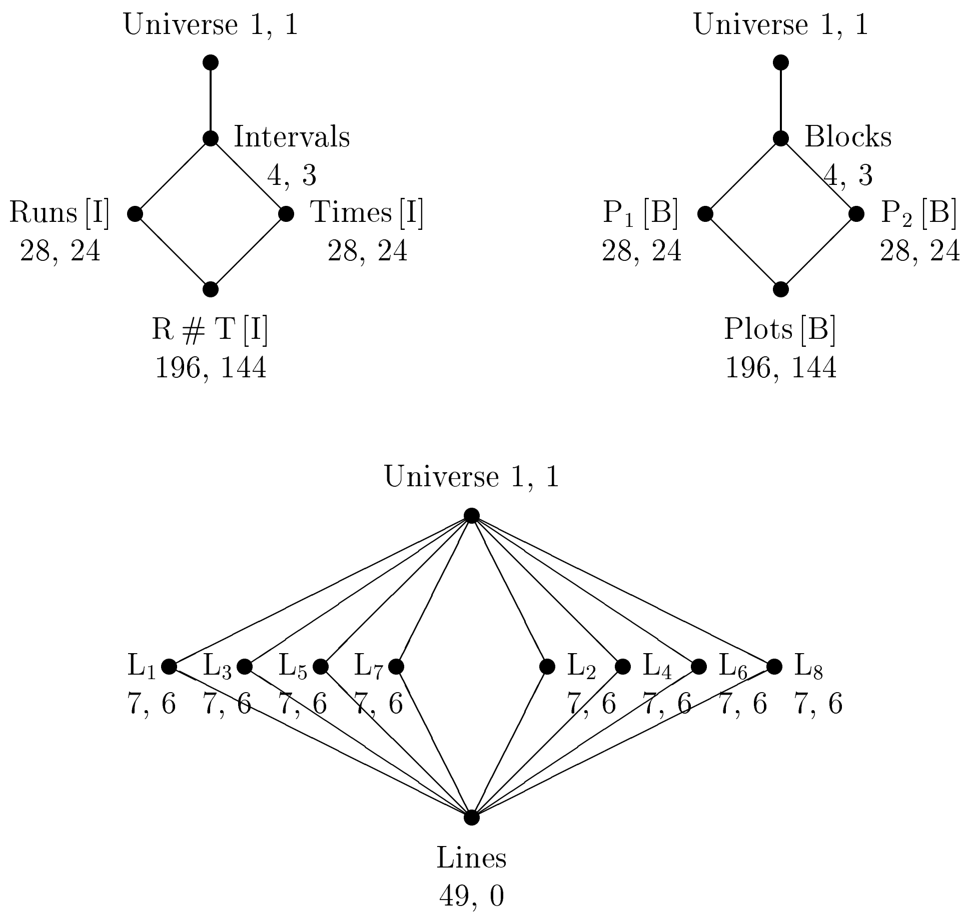

Figure 2: Hasse diagrams for analyses, plots and lines from the two-phase wheat experiment