

Use contrasts to analyze data with empty cells

Because the factorial effects are ambiguous, it makes no sense to test hypotheses about the factorial model parameters. For example, are the column effects above zero or nonzero? What does make sense is to look at simple effects and to set up contrasts that make factorial-like comparisons where possible. For example, levels 1 and 2 of factor A are complete, so we can compare those two levels with a contrast. Note that the difference of row means is 72.5, and $\alpha_2 - \alpha_1$ is 72.5 in both decompositions. We might also want to compare level 1 of factor B with level 2 of factor B for the two levels of factor A that are complete. There are many potential ways to choose interesting contrasts for designs with empty cells.

10.2 Multiple Comparisons

F-tests in factorial ANOVA not usually adjusted for multiple comparisons

The perceptive reader may have noticed that we can do a lot of F-tests in the analysis of a factorial, but we haven't been talking about multiple comparisons adjustments. Why this resounding silence, when we were so careful to describe and account for multiple testing for pairwise comparisons? I have no good answer; common statistical practice seems inconsistent in this regard. What common practice does is treat each main effect and interaction as a separate "family" of hypotheses and make multiple comparisons adjustments within a family (Section 9.1) but not between families.

Be wary of isolated significant interactions

We sometimes use an informal multiple comparisons correction when building hierarchical models. Suppose that we have a three-way factorial, and only the three-way interaction is significant, with a p -value of .04; the main-effects and two-factor interactions are not near significance. I would probably conclude that the low p -value for the three-way interaction is due to chance rather than interaction effects. I conclude this because I usually expect main effects to be bigger than two-factor interactions, and two-factor interactions to be bigger than three-factor interactions. I thus interpret an isolated, marginally significant three-way interaction as a null result. I know that isolated three-way interaction can occur, but it seems less likely to me than chance occurrence of a moderately low p -value.

We could also adopt a predictive approach to model selection (as in Section 5.4.9) and choose that hierarchical model that has lowest Mallows' C_p . Models chosen by predictive criteria can include more terms than those chosen via tests, because the C_p criterion corresponds to including terms with F-tests greater than 2.