

MODEL BUILDING

Model building refers to the process of deciding what model to use for the context. So far we have just talked about one-way models, cell-means models, two-way main-effect models, and two-way complete models. We will be discussing more (and more complex) models as the semester progresses, but we have already seen in the reaction time data that the two-way and main effects models give different test statistics.

Sometimes existing well-supported theory or knowledge guides the choice of model, but sometimes the choice needs to be made empirically -- on the basis of data. For example, if we use a complete two-way model, test for interaction, and find no evidence of interaction, we might decide that, in the future, it makes sense to use a main effects model for that situation (but it might be better to make such a decision only on the basis of two experiments, each with enough power to detect meaningful interaction, if both give no evidence of interaction).

Here are some considerations to be taken into account in empirical model building:

1. *Model building and inference to answer the original questions should be done using different data.*

Example: If we decided to use the two-way complete model for the reaction time data, tested for interaction, and found no evidence of interaction, it would *not* be legitimate to switch to the main-effects model for further hypothesis testing and confidence intervals with the same data. This would be changing the model on the basis of data; the significance levels and confidence levels produced by the new model in this instance would not be correct.

Instead, we would need to either continue with the complete two-way model to address our original questions, or else collect new data and use the main effects model with the new data.

However, it would be legitimate to use the finding that the data provide no evidence of interaction to consider using a main-effects model for *future* experiments in the same context.

Sometimes data for use in model building and data for use in addressing the original questions are collected at the same time. To really do it right, one needs one data set for model building, one for model verification, and then after a model is accepted, another data set for inference addressing the original question. However, this is not always possible because of constraints, so one may need to do the best one can with the data available (e.g., address the original question based on the full model, and then treat model building as "data snooping"; point out limitations and uncertainty in the write-up). Using a complete model and sticking with it through the analysis is one alternative.

2. Model building based on hypothesis tests raises legitimate questions of multiple comparisons. An alternate approach, sometimes called a "prediction based approach," is

based on Mallows's C statistic, discussed in M 384G. Other model-evaluation techniques (Akaike's criterion, Bayesian information criterion, ...) are also used.

3. Although there is some difference of opinion, there seems to be fairly general agreement that if an interaction term is included in a model, then each factor involved in the interaction should also be included in the model. In other words, if the interaction of factors A and B is included in the model, then both A and B should be included. Models following this rule are called *hierarchical*.