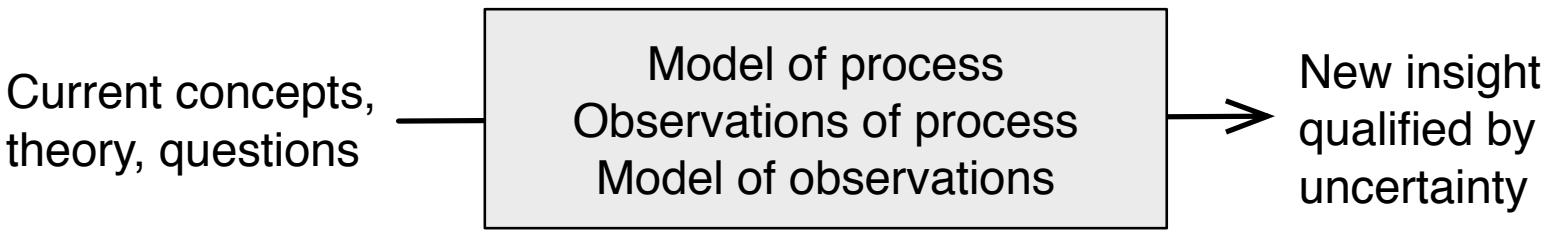
Bayesian Analysis of Designed Experiments

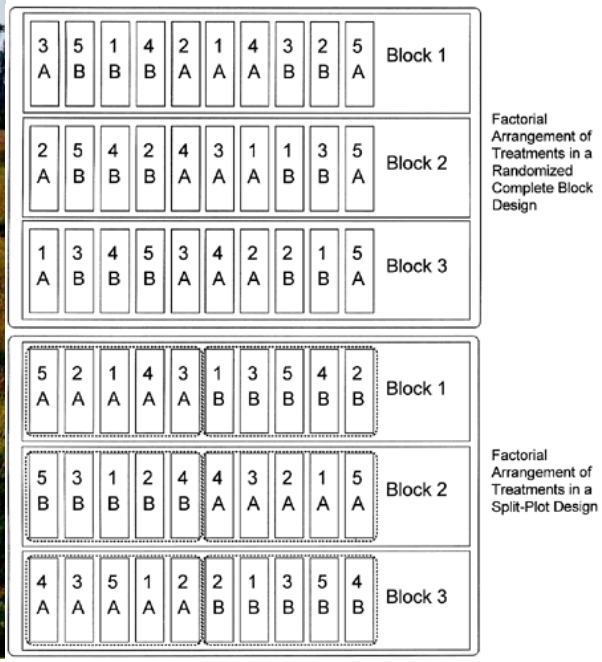
ESS 575 Models for Ecological Data

N. Thompson Hobbs

April 4, 2017


Current concepts,
theory, questions
Model of process
Observations of process
Model of observations
New insight
qualified by
uncertainty

Designed experiments



Learning outcomes

- Be able to write simple and hierarchical Bayesian models for designed experiments
- Understand
  - Design matrices
  - Effects models and means models
  - Spatial and temporal group effects (aka random effects)
- Know how to make inference on effect sizes
- Know how to make multiple comparisons of cell means

Follow-up references

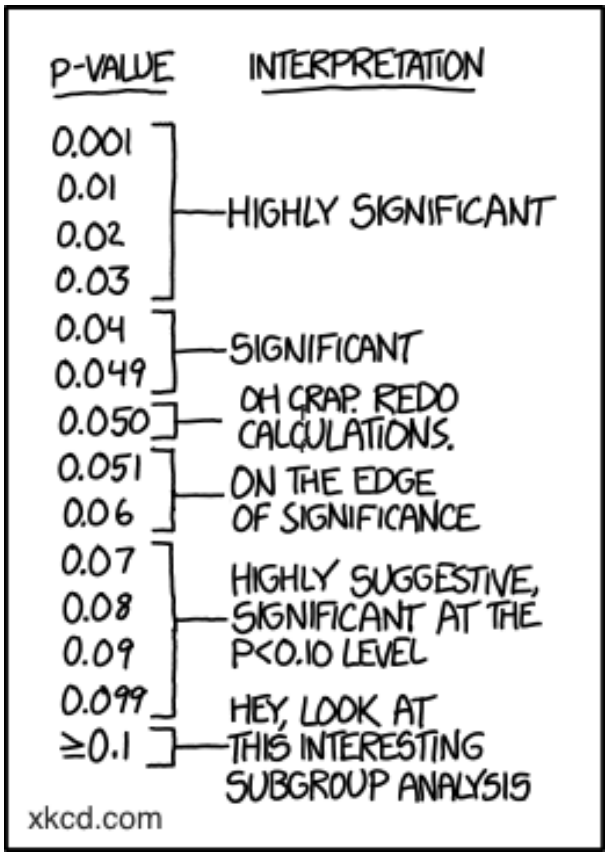
1. Hobbs and Hooten, chapters 6.2.3 and 10.2.
2. A. Gelman, J. B. Carlin, H. S. Stern, D. Dunson, A. Vehhtari, and D. B. Rubin. Bayesian data analysis. 2013 Chapman and Hall / CRC, London, UK.
3. A. Gelman, and J. Hill. 2009. Data analysis using regression and multilevel / hierarchical models. Cambridge University Press, Cambridge, UK.
4. McCarthy, M. A. 2007. Bayesian Methods for Ecology. Cambridge University Press, Cambridge, U. K.

Why Bayesian analysis for standard designs?

1. Need to account for observation and sampling uncertainty in responses (and / or predictor variables in ANCOVA) (See Hobbs and Hooten, chapters 6.2.3 and 10.2).
2. Easy to model group level effects using multi-level models
3. Easy to make inference on derived quantities of interest
4. Statements based on probability are easy to make and interpret.
5. Multiple comparisons among means are handled sensibly.

Interpreting probabilities

- The probability that at the effect of treatment exceeded zero was .87
- We can be 98.3% certain that treatment increased growth rate by more than 2 fold.
- We can be 95% certain that the control mean was $<$ 10.4 and $>$ 3.1



The new twist

$$g(\boldsymbol{\beta}, x_i) = \beta_0 + \beta_1 x_i \tag{1}$$

(2)

Up to now, x_i was some sort of quantitative observation. What if x_i denotes the presence or absence of a treatment? Or perhaps a treatment level?

Board work on reinventing "analysis of variance."

When there is no control

Before, we were able to assume that the "effect of treatment" in the control was 0. That meant the number of effects was the same as the number of cells in the experiment. But what if we have no "control?" Lets say you have J treatments but none can be considered as a reference, for example, the "treatments" might be locations or species or genotypes. Solutions that follow are given for a completely random design to keep things simple but it would be easy enough to handle a blocked design, repeated measures, etc.

When there is no control

Alternative 1: Use a means model and calculate effect sizes as departures from the grand mean.

$$[\boldsymbol{\mu}, \alpha, \varsigma^2, \sigma^2 \quad | \quad \mathbf{y}] \propto \prod_{i=1}^{n_j} \prod_{j=1}^{J} [y_{ij} | \mu_j, \sigma^2][\mu_j | \alpha, \varsigma^2][\sigma^2][\varsigma] \tag{3}$$

$$\tag{4}$$

$$\gamma_j \quad = \quad \mu_j - \alpha \tag{5}$$

When there is no control

Alternative 2: Use an effects model and assume effects are drawn from a zero-centered distribution. $\mathbf{X}$ is a $n \times J$ design matrix and $\boldsymbol{\gamma}$ is a J length vector of effects:

$$[\alpha, \varsigma^2, \sigma^2 \mid \mathbf{y}] \propto \prod_{i=1}^{n} [y_i \mid \alpha + \mathbf{X}\boldsymbol{\gamma}, \sigma^2] \tag{6}$$

$$\times \prod_{j=1}^{J} [\gamma_j \mid 0, \varsigma_\gamma^2][\sigma^2][\alpha] \tag{7}$$

When there is no control

Alternative 3: Constrain effects to sum to zero, i.e.,

$$\sum_{j=1}^{J} \gamma_j = 0 \tag{8}$$

Analysis of covariance

Analysis of covariance combines qualitative and quantitative variables in the same model. The idea is to modify the response of each cell in the experiment by measuring a predictor variable that is not affected by treatment.
These models are easily constructed in the Bayesian framework. We will see an example on Thursday. Also see McCarthy, M. A. 2007. Bayesian Methods for Ecology. Cambridge University Press, Cambridge, U. K. for some excellent, accessible examples.

Multiple comparisons of cell means

Inference on differences between cell means are calculated directly in a means model or indirectly in an effects model. They are made as the difference posterior distribution of the difference between cell means. These are analogous to single degree of freedom contrasts or Tukeys or the like, but a lot less trouble. But what about the problem of multiple comparisons?

Multiple-comparisons of cell means

Multiple comparisons are reliable if the model is hierarchical such that means or effects are drawn from a distribution. Illustrating for a means model:

$$[\boldsymbol{\mu}, \sigma^2, \alpha, \varsigma_\mu^2 \mid \quad \mathbf{y}] \propto \prod_{i=1}^{n_j} \prod_{j=1}^{J} [y_{ij} | \mu_j, \sigma^2][\mu_j | \alpha, \varsigma_\mu^2][\sigma^2][\varsigma^2] \quad [\alpha][9]$$

Subtract one cell mean from another to get posterior distribution of difference of means. Shrinkage of the distribution of means as the number of means increases assures that it becomes more difficult for the posterior distribution of a difference between to exclude 0. Neat and tidy.

This also holds for effects models where cell means are calculated from effects and the control or grand mean.

Multiple comparisons among cell means

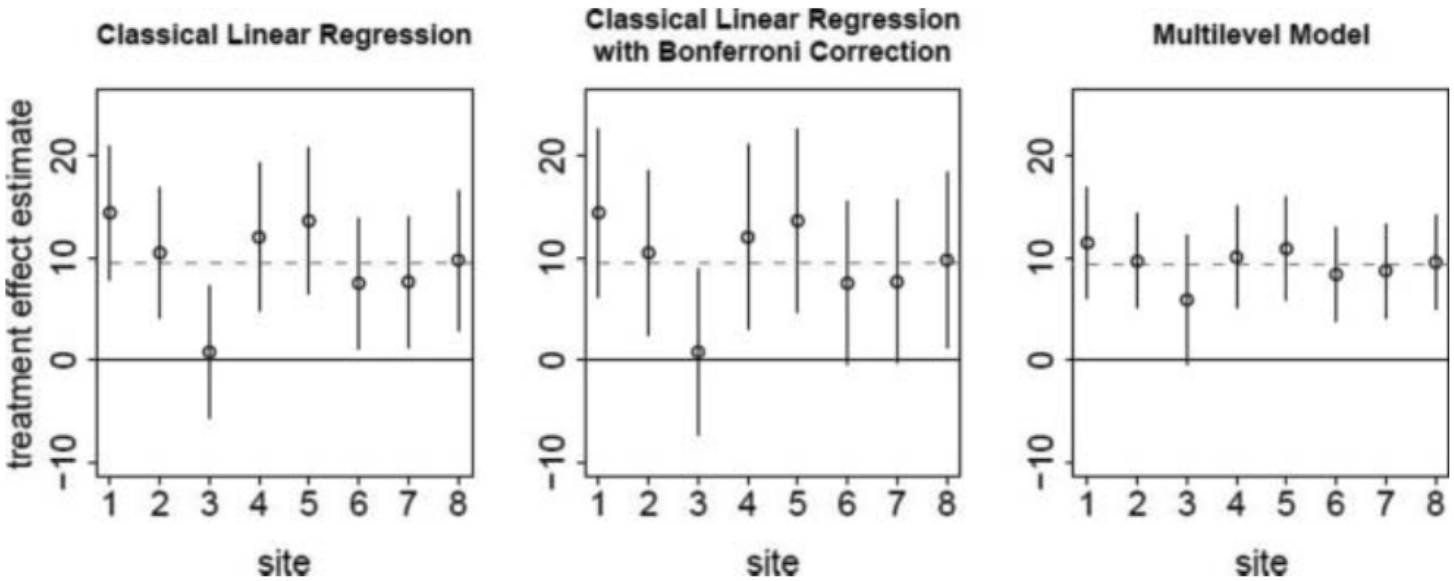


Figure 1. Treatment effect point estimates and 95% intervals across the eight Infant Health and Development Program sites. *Note.* The left panel display classical estimates from a linear regression. The middle panel displays the same point estimates as in the left panel but with confidence intervals adjusted to account for a Bonferroni correction. The right panel displays posterior means and 95% intervals for each of the eight site-specific treatment effects from a fitted multilevel model.

Gelman et al. 2012

Multiple comparisons among cell means

A. Gelman, J. Hill, and M. Yajima. Why we (usually) don't have to worry about multiple comparisons. Journal of Research on Educational Effectiveness, 5(2):189–211, 2012.

A. Gelman and E. Loken. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University, 2013.

Take home

- Analysis of designed experiments closely resembles other types of Bayesian modeling, providing enormous flexibility to the experimentalist.
- An effects model is analogous to regression except that the inputs include a design matrix. We can combine quantitative information on experimental units within the design matrix.
- A means model estimates the means of treatment cells.
- We can use moment matching and all of the hierarchical tricks we have learned to flexibly create models for analysis of designed experiments:
  - Responses and latent quantities with support 0 or 1, 0 to 1, counts, successes on trials, counts in multiple categories, real strictly non-negative, all real numbers.
  - Errors in quantitative covariates (if we have them).
  - Errors in responses
  - Group level effects (aka random effects) in space and time.