Multilevel-modeling group meeting

Jan 17th, 2017

Matt Espe

Problem

The CA rice breeding program has claimed annual yield increases over time through the release of improved cultivars. This conclusion is supported by data that shows each new cultivar out-performing previous ones by an average of 5%. However, statewide yields in CA rice production have been plateaued for the last 20 years, suggesting no appreciable increase in yield potential.

There are two possible explanations: (1) annual improvements in yield potential are being canceled out by an effect in the opposite direction, or (2) the comparison between new and old cultivars is flawed. It is unknown if it is possible to separate these two effects in a linear model, since they are related to one another.

Data

I have access to 30,000 observations from managed variety trials over the last 21 years. These trials have side-by-side plantings between old and new cultivars on farmers' fields, managed similar to a commercial operation. Over the 21 years, there have been roughly 1,500 unique cultivars planted. I additionally have estimates of the contributions of environmental factors to yield variability.

Model structure

The following effects are reasonable:

- Technology improvement: expected to be positive, breeding/cultivar independent.

- Cultivar improvement: expected to be positive, expected to become "fixed" at year of release (i.e., the cultivar does not change once entered) - Big question: can it be separated from technology improvements?

- Cultivar loss-of-fitness: expected to be negative, beginning at release.

A simplified model is:

$$y = \beta_{tech.improvement} + \beta_{cult.improvement} + \beta loss - of - fitness + \text{other effects} + \epsilon$$