A shiny update to an old learning game

Robert B. Gramacy
Department of Statistics
Virginia Tech*

Abstract

A full appreciation of aspects of experimental design, modeling, and decision making in applied settings takes time, in two senses. That learning requires patience and diligence is the first, obvious, sense. The second one is that applied work, and experimentation, often play out over long time scales, during which theories are revised, model and inferential techniques are improved, and knowledge is updated. Here I present a game, borrowing liberally from one first played over forty years ago, that attempts to synergize these aspects of learning time. The goal is to reinfoce a cascade of topics in modern response surface methods, sequential design and optimization, in a stimulating, competitive and realistic environment. Interface, rules, and evaluation are described, along with a "case study" involving a crop of students at Virginia Tech.

Key words: response surface, computer experiment, experimental design, Bayesian optimization, input sensitivity, teaching game

1 The setting

In-class games are a common way to encourage learning—to interject some fun and build intition in an seemingly esoteric, or tedious technical landscape. A good game could be fundamental to retaining students in introductory statistics, say. One fine example involves using chocolate chip cookies to illustrate properties of sampling distributions (Lee, 2007). The long arc, of an "out-class" game played out over an entire semester, is attempted rather less frequently. However for some topics, like experimental design and response surface optimization, that setting is quite natural: real-life application often playes on longer temporal

* *Contact* `rbg@vt.edu` or Hutcheson Hall (MCO439), 250 Drillfield Drive, Blacksburg, VA 24061

scales, and thus in an inherently dynamic landscape. In this article I present such a game, which is actually an update of one first played over forty years ago in a setting way ahead of its time (Mead and Freeman, 1973).

The Mead and Freeman game is almost unknown, which is a shame. Although it is cited prominently in one of the canonical response surface methods texts Box and Draper (1987), which is how I found it, I could identify only eight other references from digital and print content. This is perhaps because, for many decades (70s-90s, say) the setup of the game, requiring a custom computing environment with student access, was hard to replicate. Today, with R/CRAN (R Core Team, 2017) for implementation an library support, and `Shiny` (Chang et al., 2017) for the interface via the world-wide-web, those implementation barriers are way down.

The original game involves *blackbox* evaluation of agricultural yield as a function of six nutrient levels, borrowed from Nelder (1966) and reproduced in R as follows.

```
yield <- function(N, P, K, Na, Ca, Mg)
{
  l1 <- 0.015 + 0.0005*N + 0.001*P + 1/((N+5)*(P+2)) + 0.001*K + 0.1/(K+2)
  l2 <- 0.001*((2 + K + 0.5*Na)/(Ca+1)) + 0.004*((Ca+1)/(2 + K + 0.5*Na))
  l3 <- 0.02/(Mg+1)
  return(1/(l1 + l2 + l3))
}
```

Although my updated game has kept this `yield` form in its inagural run, swapping in code for a new blackbox is a trivial matter. Section 4 suggests some attractive alternatives, some of which involve potentially demanding (i.e., real) computer simulation.

The original game involved observations of `yield` with and additive (Gaussian) block and plot-within-block effects. Such noisy `yeild` evaluations, with the ultimate goal of maximization, could be made over up to five computer sessions (simulating crop years), comprising a total of four hundred plot observations. More than one experiment could be undertaken in a single session, however information learned from those eperiments could not be used to drive experimental design during the current session, only future ones. Although the spirit of these nuances of the game has influenced my updated version, the specifics (i.e., the official rules) are quite different. These are provided in Section 2.

The many adaptations in the new game are motivated by a desire to teach a more modern statistical toolkit. Classical response surface methods, and design, emphaize low degree (first- and second-order) linear modeling. The resulting steepest ascent and ridge analysis methodology (see, e.g., Box and Draper, 1987; Myers et al., 2016) has much to recommend it, including that many of the calculations can either be performed by hand, or with a rudimentary calculator. Yet such tools represent the tip of the iceberg in modern application domains like information technology. They seem particularly crude in situations where physical experimentation is coupled with computer simulation experiments, a setup originating in physcis and engineering, but becoming increasingly commonplace the other applied sciences. Modern response surfaces methods increasingly include spatial models, like Gaussian processes, and machine learning tools like deep neural networks and regression and classification trees. Sequential design strategies like expected improvement (Jones et al., 1998) promise a more attractive approach to (so-called Bayesian) optimization via theoretical guarantees and a capacity for human-free automation.

Getting students aquainted with the strengths and drawbacks of a more modern toolkit benefits from a sandbox wherein they can play, but moreover the competitive nature of the goal—of optimizing `yield`—perhaps suggests that an element of competition may enhance the learning arena. Section 2 describes the game design, via its interface and rules, and features included to explicitrly encourage regular partiticipation. Section 3 covers student competition, timing with lecture material, and an assessment strategy designed to encourage the deployment of a range of modeling and design tools, and overall big-picture/long term thinking. Some results from a real run of the game at Virginia Tech are provided. Section 4 concludes with a discussion on lessons learned during that process, and ideas for future variations. The online supplement includes a full suite of supporting codes and teaching materials.

2 Game design

Interface

Rules

Transition to next section with evaluation on several metrics.

3 Timing, outcomes and evaluation

Homework questions

Leaderboard views

sensitivity analysis and input-dependent noise assessment

4 Discussion

Maybe point out double meanings in discussion

- shiny update in the sense of polished old furniture, but also in terms of RShiny
- update in terms of sequential design
- learning in the sense of development of both student knowledge and statistical inference.

Don't forget to discuss other `yield` functions, including maybe ATO for genuine heteroskedasticity, with reference to `hetGP`.

Acknowledgments

I am grateful to the students and colleagues at Virginia Tech who took and helped create, and promote curriculum adoption of, the class on modern response surfaces and computer experiments, for which this game is plays a central role.

References

Box, G. E. and Draper, N. R. (1987). *Empirical model-building and response surfaces.*. John Wiley & Sons.

Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. (2017). *shiny: Web Application Framework for R*. R package version 1.0.5.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). "Efficient Global Optimization of Expensive Black Box Functions." *J. of Global Optimization*, 13, 455–492.

Lee, H. K. H. (2007). "Chocolate Chip Cookies as a Teaching Aid." *The American Statistician*, 61, 4, 351–355.

Mead, R. and Freeman, K. H. (1973). "An Experiment Game." *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 22, 1, 1–6.

Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. (2016). *Response surface methodology: process and product optimization using designed experiments*. John Wiley & Sons.

Nelder, J. (1966). "Inverse polynomials, a useful group of multi-factor response functions." *Biometrics*, 22, 128–141.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.