Test 1 Info Sheet

Stat 341 – Spring 2017

Logistics

- In-class test: Friday, March 3
- Take-home test: Due Monday, March 6
- Both portions cover through chapter 5 of *Statistical Rethinking*

Topics

This list isn't meant to be exhaustive, but I hope it is useful in preparing for the test.

Techniques

1. Grid Method

2. Quadratic approximation using `map()`

3. Posterior sampling

   a. manually (for example, when using the grid method)

   b. using `extract.samples()`

4. Using posterior samples to investigate

   a. model paramters

   b. model predictions (for averages and for individuals)

   - Creating counterfactual data (for example, for use with `link()` and `sim()`)

5. Centering

6. A non-exhaustive list of R commands you should know

   - Plotting commands (`lattice`, `ggplot2`, or `ggformula` versions) for the types of plots we have been making.

   - from `rethinking`: `map()`, `alist()`, `precis()`, `extract.samples()`, `link()`, `sim()`, `chainmode()`

   - `apply()`

   - working with data frames: `data_frame()`, `expand.grid()`, `mutate()`, `filter()`

   - distributions: `dnorm()`, `dbinom()`, `dunif()`, `rnorm()`, `rbinom()`, `runif()`

   - miscelaneous: `zscore()`, `str()`, `head()`, `sample()`

Concepts

1. Conditional probability and connections to Bayesian inference

   - definition of conditional probability

   - problems like the pandas problem

2. posterior $\propto$ prior $\cdot$ likelihood

3. What grid apprid approximation is, how it works, and why it is limited to models with only a few parameters.

4. What posterior sampling is and what can be done with posterior samples.

5. Why quadratic approximation is called that and roughly how it works.

   - potential problems with quadratic approximation (and why we will need a third method eventually)

6. Creating models

   - selecting priors that are reasonable
   - creating a model relationship (equation with variables and parameters)
   - (additive) linear models
   - handling categorical variables

7. Interpreting models

   - interpreting model coefficients (and their uncertainty).
   - interpreting model predictions (and their uncertainty).
   - interpreting σ in models where the resopnse is modeled with a normal distribution
   - plots for models (especially those with multiple predictors)
   - residuals
   - masking, spurious relationships and other issues that can be revealed by fitting models with multiple predictors.

8. Why and how to center variables used in models