

Modeling HW3 Part 2

Kerui Cao

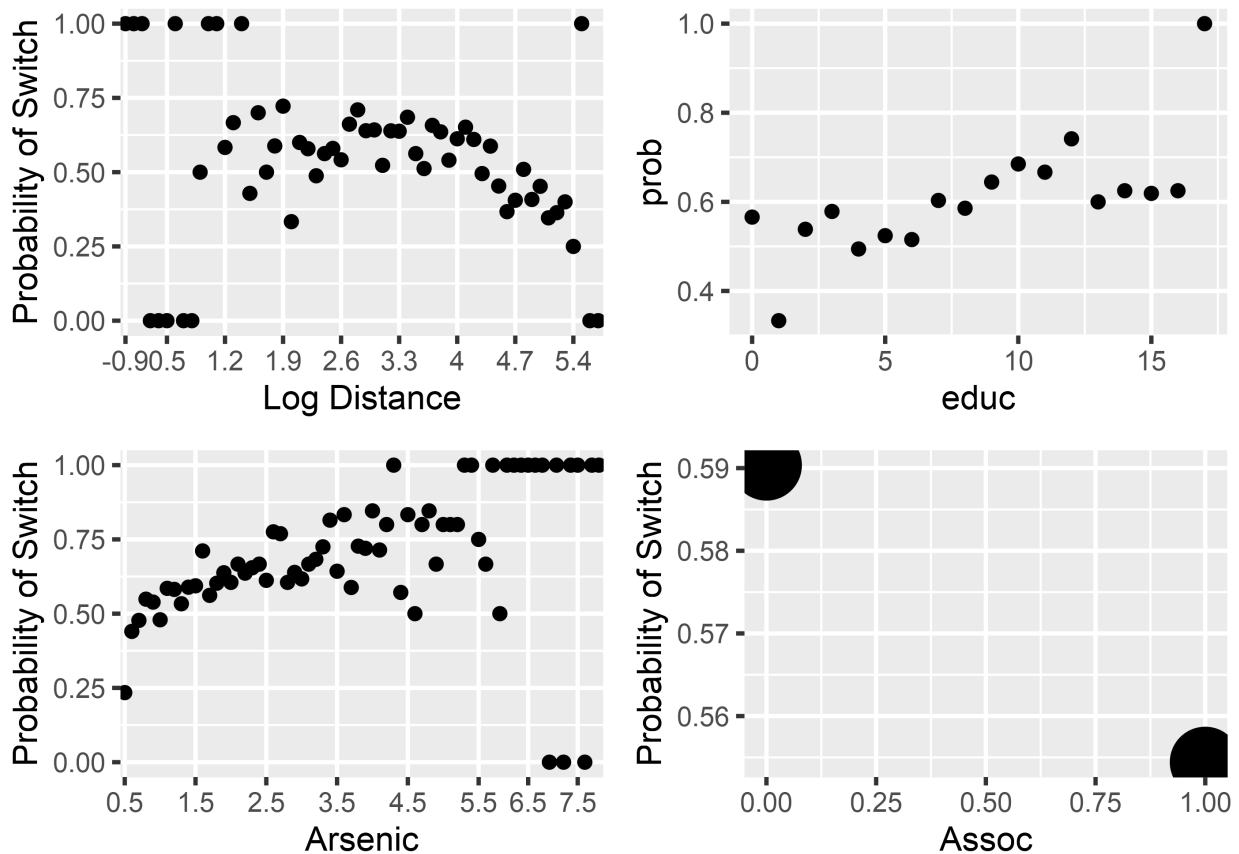
9/29/2019

Graphing logistic regressions:

the well-switching data described in Section 5.4 of the Gelman and Hill are in the folder **arsenic**.

1. Fit a logistic regression for the probability of switching using `log` (distance to nearest safe well) as a predictor.

Since we know nothing about the real meaning of these variables, we can only try to use information criteria as model selection standard.



2. Make a graph similar to Figure 5.9 of the Gelman and Hill displaying $\Pr(\text{switch})$ as a function of distance to nearest safe well, along with the data.
3. Make a residual plot and binned residual plot as in Figure 5.13.
4. Compute the error rate of the fitted model and compare to the error rate of the null model.
5. Create indicator variables corresponding to `dist < 100`, `100 <= dist < 200`, and `dist > 200`. Fit a logistic regression for $\Pr(\text{switch})$ using these indicators. With this new model, repeat the computations and graphs for part (1) of this exercise.