Week 15 Live Session

w203 Instructional Team

Announcements

No class next week (or ever!)

Lab 4 Debrief

Congratulations on submitting your final lab!

Output variables:

Most teams settled on birthweight, a few included specifications with a second metric for comparison.

It's unclear how well birthweight represents health, since there is a range of weights that would be considered healthy.

One option would be to model the low birthweight flag. If you plug this in as your outcome variable, you have what is called a linear probability model. Your model will typically output a number between 0 and 1, which is interpreted as the probability of low birthweight. This is perfectly valid for this class, but a more accepted technique is logistic regression, which is guaranteed to always output a number between 0 and 1.

You could also experiment with modeling the number of pounds underweight.

Transformations:

Several teams used the square of mother's age successfully in their models. Interestingly, it's hard to see curvature in the scatterplots. Some of you plugged in the squared term, then observed how much curvature there was in the best-fitting parabola. Others saw that adding the square term improved AIC and the significance of the original mother's age variable, or just that the new term was statistically significant. These are great examples of data-driven model building.

How does including mother's age squared help? It explains more of the variation in the outcome, so helps us keep our standard errors low. It makes it a bit harder to understand the effects of age, but that isn't the coefficient we really care about.

There also seems to be a non-linearity evident in scatterplots of prenatal visits versus birthweight. This is central, and we would have wanted to see teams try to model this.

1. log of visits
2. include a square term
3. include an indicator for visits below a threshold
4. include a separate slope for visits below a threshold

The square term is the most flexible, but harder to present to many audiences. The two slopes would let you tell a clear story: each visit has X effect, but the first 10 visits have an extra benefit of Y...

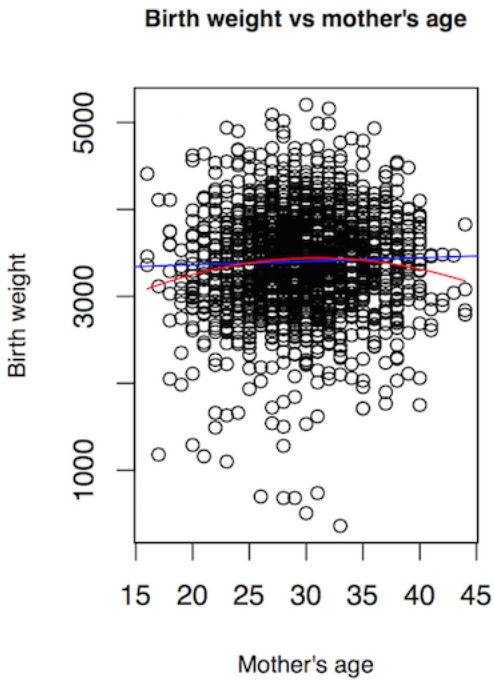

Figure 1: Effect of mother's age - Chris, Victoria, Frederic

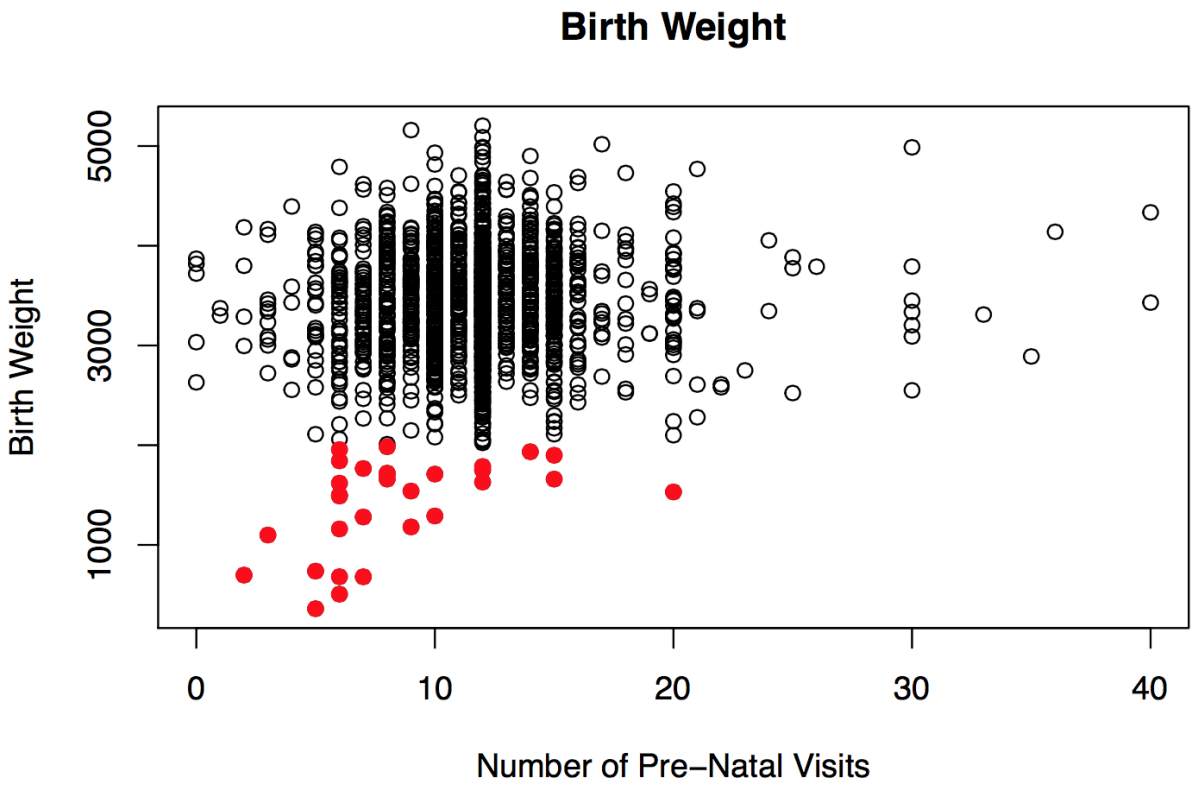

Figure 2: Highlighting low birthweight - Todd

Input versus outcome variables:

Make sure you have a clear idea of what variables are an outcome of prenatal care. These should not go on the right hand side.

For example, low birthweight should only appear on the left, as should interaction terms derived from low birthweight. Low birthweight is, of course, a great predictor of birthweight, but this doesn't tell us anything and reduces all other coefficients. APGAR scores should only appear on the left. These variables are clearly outcomes, and you should not include them as predictors in any specification, even your most expansive one.

Many teams correctly noted that drinks and cigarettes are problematic variables, because they are partly outcome variables. There could be causal pathways from prenatal care to these variables. For this reason, these variables should be held back from you main specification, but included in an alternate specification as a robustness check.

The model building process

In real situations, we often work in cycles.

Examine plots -> fit specification -> test assumptions -> alter specification -> examine plots / run F-tests ->...

We want a model with plausible assumptions, that is realistic, parsimonious, highlights features we care about, allows test we want to run. Balancing these aims is part science and part art.

Reproducability Exercise

In this exercise we will be looking at the dataset called pill_test.RData . It contains the results of an IQ test administered immediately after a subject has taken a pill containing an experimental nootropic compound. Each compound was experimentally coded as a different "color". A control group also was given a pill that did not contain an experimental compound. These nootropic compounds are extremely expensive to manfacture, so our budget will only allow ten pills of each type to be used for th purposes of scientific experimentation.

Test 1: One Pill vs. Control

First test the hypothesis of the blue pill having an effect on IQ test. Discuss your results

Test 2: Two Pills vs. Control

Now test the blue pill and the brown pill against the control. Discuss your results

Test 3: All pills vs. Control

Lastly test all pills against the control. Discuss your results. Would you have made different conclusions if you had done this test first?

More about Reproducibility

Q1. In light of this week's reproducibility discussion, how would you interpret a study based on publically-available country-by-country indicators? As a simple example, say you read an article that claims that government regulation is associated with greater internet penetration.

Q2. In your industry or field of main interest, do you expect most hypothesized relationships to be true or false? In what way does reproducibility affect the way you interpret results?

Course Evaluation

Log in at: https://course-evaluations.berkeley.edu

Thank you!