Three Problems

John Poulsen

November 13, 2018

See the three applied statistics problems. Try to finish two of them in class, working in groups.

Problem 1

To assess the effect of large doses (grams) of Vitamin C on incidence of colds, a large number of volunteers (>2000) were randomly given tablets to take daily, without knowing the dose of Vitamin C. The data include: vitamin c dose (`Dose`), total number of volunteers (`Number`), and number of volunteers witout the illness (`WithoutIllness`).

Is there evidence that the chances of having a cold are associated with the dose of Vitamin C?

To download the data, load the `Sleuth3` package and the file `ex2113`.

```
require(Sleuth3)
dat <- ex2113
```

What does the intercept represent in the model? How does dose of Vitamin C affect the odds of getting a cold?

Problem 2

In Arahazar upazila, Bangladesh, researchers labeled wells with their level of arsenic and an indication of whether the well was "safe" or "unsafe." People using unsafe wells were encouraged to switch. After several years, the researchers returned to determine whether each household using an unsafe well had changed to a safe well.

`switch`: whether or not the household switched from an unsafe well to a safe `well`: no or yes `arsenic`: the level of arsenic contamination in the household's original well, in hundreds of micrograms per liter `dist`: distance in meters to the closest known safe well `educ`: education in years of the head of the household.

Download the data from Sakai:

```
wells <- read.table ("wells.txt")
```

1. First, model the relationship between `switch` and `dist`. Note that the effect of distance is very small. Rescale the `dist` to make more sense.
2. Graph the effect of distance on the probability of switching wells.
3. Fit the model, `switch ~ dist * arsenic + educ`. What is the minimum adequate model? What is the goodness of fit?

Problem 3

In a study on horseshoe crabs, each female horseshoe crab had a male crab attached to her in her nest. The study investigated factors that influence whether the female crab had any other males, called satellites, residing near her. Explanatory variables investigated include:

- `Sa`: the number of satellite males
- `C`: female crab's color
- `S`: spine condition
- `Wt`: weight (of the female)
- `W`: carapace width

Model the effects of carapace width on number of satellites. What is the effect of `W`? Does the model fit the data well?

Add `C` to the model. Does it make the model fit better?

Plot the effect of `W` on the number of satellites, `Sa` at two different levels of color.