

Stat 539 Spring 2006  
Homework 8 due 6 April

A study of workers in the U.S. cotton industry was commissioned to discover factors that relate to contaction of the lung disease byssinosis. In the byss.dat data set the columns in order are

Yes:	Number suffering from byssinosis
No:	Number not suffering from byssinosis
Dust:	Dustiness of workplace, 1 = high, 2 = medium, 3 = low
Race:	1 = white, 2 = other
Sex:	1 = male, 2 = female
Smoker:	1 = yes, 2 = no
Emp.Length:	Length of employment, 1 = "<10 years", 2 = "10-20 years", 3 = ">20 years"

These are cross-sectional data.

- Use chi-squared (or binomial or Fisher's Exact) tests to explore the relationships of the various predictors to byssinosis.
- Compute observed odds ratios for probability of developing byssinosis from your analyses above. In all cases use the first category as the reference category.
- Do any of the predictor variables seem highly related? Which ones? How could that affect conclusions?
- Write a summary of your analyses using these "classical" methods

We will use logistic regression with this data set soon. For now restrict analysis to tests and estimates as in part a.

[Click here for the byss.dat data set.](#)