

Assignment # 4, Stat 8053, Fall 2011

Reading

The reading for this week is the vignette for the `quantreg` package and Chapter 13 of Faraway.

No Class

Early warning: A midterm exam will be handed out on Wednesday October 26, and will be due on Monday, October 31. There will be no class on Friday, November 4.

Problems

The following problems are due on Friday, October 21, 2010, in class.

1. A paper <http://www.nber.org/papers/w10522> by Persico, Postlewaite and Silverman (1994) claims to show that, at least for white males in England, taller men make more money (or specifically, men who were taller at age 16). This problem works with part of the data they used, based on the NDCS, a study of all children born in England during the week beginning March 3, 1958. Our sample consist of 1599 white males with complete records up to age 33. The response variable is `lngwge33`, the natural log of the wage rate at age 33. The predictors are `height16`, height at age 16 `htchange`, change in height from age 16 to age 23; `yrsschl`, years of schooling, `evmar33` equal to one if ever married by age 33, zero otherwise; `divsep33`, one if divorced or separated at age 33, zero otherwise; `siblings` number of siblings, and `momskldi` equal to one if the person's mother held a skilled occupation, zero otherwise.

Use quantile regression and tree methods to explore the question of interest: do taller men earn more? How much more?

To load the data, use:

```
> loc <- "http://www.stat.umn.edu/~sandy/courses/8053/Data/ndcs.txt"
> ndcs <- read.table(url(loc), header=TRUE)
```

2. Use the data in the file `Mroz` in the `car` package. The response is `lpf`. See the help file `?Mroz` for a description of the predictors. Compare using a single tree via `rpart` and bagged-trees using `randomForest`. Also, compare results to other methods we have covered.