BST 760: Advanced Regression
Breheny

Project 1
Due: Thursday, April 4

The Western Collaborative Group Study carried out a population-based cohort study on men aged 39 to 59 years old regarding the onset of coronary heart disease (CHD). Data on several risk factors (smoking, cholesterol levels, blood pressure, etc.) was collected, but the primary interest of the study was to assess the relationship between CHD risk and personality. In particular, a personality test was given which categorized the subjects into two categories called "Type A" and "Type B". Type A personalities are characterized by aggressiveness and competitiveness, whereas Type B personalities are relaxed and laid-back. After recruitment, patients were monitored for the onset of CHD for about 9 years. Occurrence of CHD was determined by the diagnosis of medical experts, typically following a heart attack. The data is online `wgcs.txt` and contains the following variables:

- `Age`: Age at beginning of study
- `Height`: (inches)
- `Weight`: (pounds)
- `SBP`: Systolic blood pressure (mm Hg)
- `DBP`: Diastolic blood pressure (mm Hg)
- `Chol`: Cholesterol (mg/100 ml)
- `Ncigs`: Number of cigarettes per day
- `TypeA`: 1 if Type A, 0 if Type B
- `CHD`: 1 if CHD, 0 otherwise
- `Time`: Follow-up time (in days)
- `Arcus`: Presence of arcus senilis

Analyze this data and report its findings concerning risk factors for CHD. As noted above, be sure to address the role of personality. See `http://web.as.uky.edu/statistics/users/pbreheny/760/S13/style-guide.pdf` for details on how the report should be structured and how it will be graded.