Assignment 5—Introduction to Statistics Using R

Paul Gribble
Winter, 2017

* Due Friday Mar 24

1 Multiple Regression

You are being considered for a position as the new coach of a professional basketball team. You just finished a meeting with the owners, who are interested in replacing the existing coach because the team's performance has been going downhill. The owners of the team have attributed this downward trend to poor recruitment of new players, who have been poor performers. The owners tell you that their analysis of the data from the previous seasons indicates that one particularly weak area has been "field goals" (when a player attempts to score a basket from the 3-point line or beyond). The team's players, in particular the new recruits, tend to be poor at field goals. The problem is, as the owners tell you, that data are not available on field goal percentages, for players who are candidate recruits; so when it comes time to recruit new players there is no easy way to pick new recruits who are good field goal shooters. The owners tell you that they have been collecting data for years and years on the new recruits and their field goal percentages, and a series of other measures that are available for the new potential recruits as well. She suggests that "if you are smart enough to use the historical data to be able to predict who among the potential new recruits will be the best field goal shooters, you are hired!"

In the file called bball.csv[1] you will find historical data on 105 previous recruits. The variables are:

GAMES	# games played in previous season
PPM	average points scored per minute
MPG	average minutes played per game
HGT	height of player (centimetres)
FGP	field-goal percentage (% successful shots from 3-point line)
AGE	age of player (years)
FTP	percentage of successful penalty free throws

---

[1] http://www.gribblelab.org/stats/data/bball.csv

In the file `bball2.csv`[2] you will find data on 5 potential new recruits.

1. Use multiple regression to develop a model based on the data in `bball.csv` that will allow you to predict the field goal percentages for the 5 new recruits.

2. Which player will you suggest to the owners that they hire? Explain your decision.

3. What is the precision (in units of `FGP`) with which you can predict field goal percentage?

4. After giving them your recommendation and explaining your model, the owners ask you "does the player's age have anything to do with this?"

5. "What about the player's height? Would I be better off taking on taller players?"

6. "How many more (or fewer) field goals would a player score if he was 3 inches taller?"

---

[2]`http://www.gribblelab.org/stats/data/bball2.csv`