
   a) Use the `load` command to read in the saved dataframe `hills.RData` from the course website (i.e., www.stat.washington.edu/handcock/506/Data).

   b) Fit a linear model taking as the dependent variable the record time in minutes (`time`) and as independent variables the distance in miles (`distance`) and the climb in feet (`climb`).

   c) Use the R function `summary` to obtain the summary of the fitted model. What are the coefficients of the fitted model?

   d) Look at the fit graphically. To do this use the `plot` function. In considering the resulting plots, what do you conclude about the quality of the fit?

   e) Try to fit a generalized additive model. To do this the following commands will be useful:

   ```r
   > library(mgcv)
   > hills.gam <- gam(time ~ s(distance) + s(climb), data=hills)
   > summary(hills.gam)
   > plot(hills.gam)
   ```

   Try to understand this commands and play around with their arguments. Note: If the `mgcv` package does not exist on your installation, it can be downloaded from the r-project.org website.

   Does the time appear to be linear in distance? Does the time appear to be linear in climb? Do you think the linear model is appropriate?

2) This question involves nonparametric regression with the Great Barrier Reef data. In a survey of the fauna on the sea bed in an area lying between the coast of northern Queensland and the Great Barrier Reef, data were collected at a number of locations. Data and the needed functions are in the R library `sm`, discussed in Bowman and Azzalini (1997). This which can be downloaded from the r-project.org website.

   Try the following commands:

   ```r
   > #
   > # Load the SM package
   > #
   ```
> library(sm)
> #
> # Read information about the SM package
> #
> help(sm)
> #
> # Load in the "trawl" data from the SM package
> #
> provide.data(trawl)
> par(mfrow=c(2,2))
> plot(Longitude, Latitude, type="n")
> points(Longitude[Zone == 1], Latitude[Zone == 1])
> text(Longitude[Zone == 0], Latitude[Zone == 0], "o")
> Zone93 <- (Year == 1 & Zone == 1)
> Position <- cbind(Longitude - 143, Latitude)
> sm.regression(Latitude[Zone93], Score1[Zone93], h=0.1)
> sm.regression(Position[Zone93,], Score1[Zone93],
> h=c(0.1, 0.1), eye=c(8,-6, 5), xlab="Longitude - 143")
> sm.regression(Longitude[Zone93], Score1[Zone93], h=0.1)
> par(mfrow=c(1,1))

The top left panel displays the sampling points. The dots are taken from a closed region, where commercial fishing is not allowed to protect the nature. The circles refer to areas outside, which were surveyed to allow comparisons. Due to large numbers and types of species captured in the survey, the response variable is expressed as a score, on a log weighted scale, which combines information across species. The two right panels of the plot show nonparametric estimates of the relationship between the catch score within the closed zone and the spatial coordinates, latitude and longitude, for the 1993 survey.

First, here is the output:

> #
> # Load the SM package
> #
> library(sm)
> Library 'sm', version 2; Copyright (C) 1997, 2000 A.W.Bowman & A.Azzalini
> type help(sm) for summary information
> #
> # Read information about the SM package
> #
> help(sm)
> #
> # Load in the "trawl" data from the SM package
> #
> provide.data(trawl)
> Data file being loaded
> par(mfrow=c(2,2))
> plot(Longitude, Latitude, type="n")
> points(Longitude[Zone == 1], Latitude[Zone == 1])
> text(Longitude[Zone == 0], Latitude[Zone == 0], "o")
> Zone93 <- (Year == 1 & Zone == 1)
> Position <- cbind(Longitude - 143, Latitude)
> sm.regression(Latitude[Zone93], Score1[Zone93], h=0.1)
> sm.regression(Position[Zone93,], Score1[Zone93],
+ c(0.1, 0.1), eye=c(8,-6, 5), xlab="Longitude - 143")
> sm.regression(Longitude[Zone93], Score1[Zone93], h=0.1)
> par(mfrow=c(1,1))

a) Can we observe a change in the catch score with latitude?

**Solution** (10): First, here is the plot from the output:

![Plot of catch score against latitude](image1)

b) Can we observe a change in the catch score with longitude?