

Intro to R for Epidemiologists

Lab 8 (3/5/15)

Data

This lab will use the combined male and female diabetes dataset (from Homework 2). The full dataset can be found on the website under lab 8 as “diabetes.csv”. This dataset also contains diabetes status as computed in homework 2.

Part 1. Multiple logistic regression

1. Read in the dataset "diabetes.csv"
2. Use multiple logistic regression (one regression model) to estimate the associations between a set of variables (total cholesterol, hdl cholesterol, age, height, and weight) and diabetes status.
3. Create a data frame as below of the variable name, odds ratio from multiple logistic regression, corresponding p-value from the regression, and lower and upper confidence bounds for a 95% confidence interval.

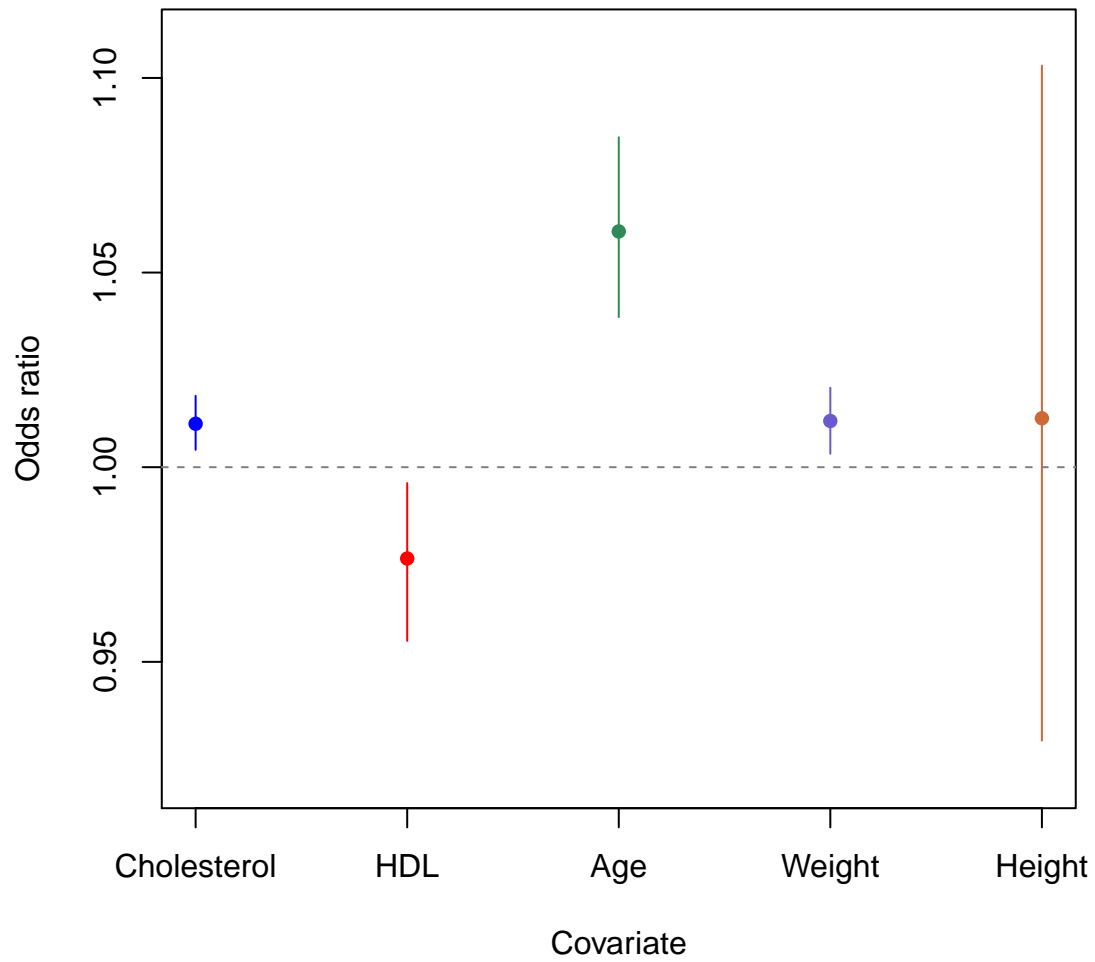
##	Variable	OR	p-value	LB	UB
## 1	chol	1.0111629	1.422292e-03	1.0044689	1.0183326
## 2	hdl	0.9765422	2.421062e-02	0.9553714	0.9959181
## 3	age	1.0605720	1.032982e-07	1.0385658	1.0847577
## 4	weight	1.0118751	5.496561e-03	1.0034372	1.0204154
## 5	height	1.0125684	7.739832e-01	0.9297465	1.1031758

Part 2. Plotting regression results

Create the plot displaying the odds ratios from Part 1 and corresponding 95% confidence intervals as shown on the following page. The colors for this plot can be specified using: `cols <- c("blue", "red", "seagreen4", "slateblue", "sienna3")`.

1. Plot the odds ratios. Be sure to label your axes.
 - Specify `ylim = c(0.92, 1.11)` and `axes = F`.
 - Your x values will be 1:5 and your y values will be the odds ratios.
2. Add the correct axes using the R function `axis`
3. Add a box around the plot using `box()`.
4. Add confidence intervals to each point using `segments`.
5. Add a horizontal dashed grey line (Hint: specify `lty = 2`)

Associations between covariates and diabetes



Part 3. Kaplan-Meier Curves

For this part, we will use the `kidney` dataset in the `survival` package, which gives recurrence times to infection for kidney patients.

1. Create a survival object for the data using `Surv` for time followed until recurrence.
2. Plot the Kaplan-Meier curves for time until recurrence by disease type.

Survival Estimates for Kidney patients

