

Biostatistics - MED131/MBG211 - Homework III

Dec 15, 2022

Note: Deadline for submission is December 29th, 2022, 16:00

Please provide a pdf file answering the questions

Once again, you'll be working on the "prostate cancer" dataset under `data/prostate_cancer.csv`. You may read the data directly from the GitHub repository:

```
URL <- "https://raw.githubusercontent.com/egeulgen/MED131_22_23/main/data/prostate_cancer.csv"
prca_df <- read.csv(URL)
```

The main aim of collecting this data set was to inspect the association between prostate-specific antigen (PSA) and prognostic clinical measurements in men with advanced prostate cancer. Data were collected on 97 men who were about to undergo radical prostatectomy.

For the second question, you'll be working with the "Birthweight" dataset under `data/Birthweight_dataset.csv`. You may read the data directly from the GitHub repository:

```
URL <- "https://raw.githubusercontent.com/egeulgen/MED131_22_23/main/data/Birthweight_dataset.csv"
bw_df <- read.csv(URL)
```

This dataset contains information on newborn babies and their parents.

Please answer the following questions using R.

1. [20 pts] Fit the following simple linear regression model and interpret the results. Transform the dependent variable if necessary

$$PSA_i = \beta_0 + \beta_2 age_i$$

2. [40 pts] Fit the following multiple linear regression model and interpret the results. Transform the dependent variable if necessary

$$PSA_i = \beta_0 + \beta_2 age_i + \beta_3 (seminal\ vesical\ invasion)_i + \beta_4 I(Gleason\ score_i = 7) + \beta_5 I(Gleason\ score_i = 8) \\ + \beta_6 I(Gleason\ score_i = 7) age_i + \beta_7 I(Gleason\ score_i = 8) age_i$$

3. [40 pts] Defining low birth weight (naming it `low`) as `Birthweight < 2.5`, fit the following logistic regression model and interpret the results.

variable	description
mage	Maternal age (years)
Gestation	Gestation (the period of time between conception and birth) (weeks)
mppwt	Mother's pre-pregnancy weight (kg)

$$\text{logit}\left(\frac{P(\text{low} = TRUE)}{P(\text{low} = FALSE)}\right) = \beta_0 + \beta_1 \text{mage} + \beta_2 \text{Gestation} + \beta_3 \text{mppwt}$$