Homework #2

Yunting Chiu

2021-01-31

1. (**1.2**) The members of a health spa pay annual membership dues of $$$300 plus a charge of $$$2 for each visit to the spa. Let Y denote the dollar cost for the year for a member and X the number of visits by the member during the year. Express the relation between X and Y mathematically. Is it a functional relation or a statistical relation (that is, is the relation deterministic or stochastic)?

- deterministic: the output of the model is entirely determined by the values of the parameters and the initial conditions; stochastic: random, unpredictable.
- The association between X and Y is: Y = 300 + 2X (dollars).
  This is a functional relationship because a effect in the value of the X will cause the corresponding change in the value of the Y. X and Y will not have a uncertain effect.

2. (**1.6**) Suppose the regression parameters are $\beta 0 = 200$ and $\beta 1 = 5.0$.

(a) Plot the regression equation.

(b) Predict the response for X = 10, 20, and 40.

(b)

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

$Y = 200 + 5X_i + \varepsilon_i$ $\quad \boxed{\varepsilon_i \sim N(0, 1)}$

$E\{Y_i\} = 200 + 5X_i + E\{\varepsilon_i\}$ → 0

$E\{Y_i\} = 200 + 5 \times 10 = 250$ # (if X=10)

$E\{Y_i\} = 200 + 5 \times 20 = 300$ # (if X=20)

$E\{Y_i\} = 200 + 5 \times 40 = 400$ # (if X=40)

(c) Explain the meaning of parameters $\beta 0$ and $\beta 1$.

- $\beta 0$ = intercept
- $\beta 1$: one unit change in X, generates a $\beta 1$ unit change in Y.

3. (**1.10**) An analyst in a large corporation studied the relation between current annual salary (Y) and age (X) for the 46 computer programmers presently employed in the company. The analyst concluded that the relation is curvilinear, reaching a maximum at 47 years. Does this imply that the salary for a programmer increases until age 47 and then decreases? Explain.

4. The time it takes to transmit a file always depends on the file size. Suppose you transmitted 30 files, with the average size of 126 Kbytes and the standard deviation of 35 Kbytes. The average transmittance time was 0.04 seconds with the standard deviation of 0.01 seconds. The correlation coefficient between the time and the size was 0.86. Based on this data, fit a linear regression model and predict the time it will take to transmit a 400 Kbyte file.

```
library(tidyverse)
```

```
## -- Attaching packages ---- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts ------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
aaa <- read_tsv("./data/CH01PR19.txt")
```

```
## Parsed with column specification:
## cols(
##   `3.897    21` = col_character()
## )
```

```
aaa
```

```
## # A tibble: 119 x 1
##    `3.897    21`
##    <chr>
##  1 3.885    14
##  2 3.778    28
##  3 2.540    22
##  4 3.028    21
##  5 3.865    31
##  6 2.962    32
##  7 3.961    27
##  8 0.500    29
##  9 3.178    26
## 10 3.310    24
## # ... with 109 more rows
```