1930026143 薛劭杰

(a) $x_i = 0, 10, 20, 30, 40, 50, 60$ (Independent variable)

$y_i = 16, 27, 28, 39, 39, 48, 51$ (Dependent variable)

$\bar{x} = 30$, $\bar{y} \approx 35.43$

$S_{xx} = \sum_{i=1}^{7} (x_i - \bar{x})^2 = 2800$

$S_{xy} = \sum_{i=1}^{7} (x_i - \bar{x})(y_i - \bar{y}) = 1580$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \approx 0.5643$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 35.43 - (30 \times 0.5643) = 18.5014$

So $\hat{y} = 0.5643\hat{x} + 18.5014$

(b) $b_0 = 18.5014$, $b_1 = 0.5643$

b_0: when leaving home Zero minute after 7 am, it take 18.5014 minutes to travel to school

b_1: After 7 am, One minute away from home, it will take more 0.5643 minutes to travel to school

(c) $\hat{y} = 18.5014 + 0.5643\hat{x}$,

t = the time leaving home after 7am + the time travel to school

$= y + x = 18.5014 + 1.5643x < 90$

we get that $x < 45.7064$ which means that the lastest time base on the model is 45.7064 minutes.

Exercise 2.

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

(a) y is dependent variable, x is independent variable.
β_0, β_1 are the regression cofficients which is unknown.
and β_0 can be considered to intercept, β_1 can be considered to slople

(b) 1. Assume that the mean of the random error value equal to zero

2. Assume that the variance of the random error value is not depend on the value of x.

3. Assume that the distriution of the random potential error has normal distriution.

4. Assume that the value of the ramdom potential error is not depend on each other error values.

5. Assume the the D of the $\beta_0, \beta_1 > 0$ and the two partial derivatives of β_0 are always bigger than 0

c. $\hat{y} = \beta_0 + \beta_1 \hat{x} + \varepsilon$, we should make the value of ε be minimum.

$$Q_2 = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i^2)$$

and we can get the

$$\frac{\partial Q}{\partial \beta_0} = -2(n\bar{y} - n\beta_0 - n\beta_1\bar{x}) = 0$$

$$\frac{\partial Q}{\partial \beta_1} = -2(n\bar{y} - n\beta_0 - \sum \beta_1 x_i^2) = 0.$$

$$\frac{\partial Q}{\partial \beta_0} \cdot \bar{x} - \frac{\partial Q}{\partial \beta_1} \Rightarrow n\bar{x}\bar{y} - n\beta_1\bar{x}^2 - \sum_{i=1}^{n} x_i y_i + \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \beta_1 (\sum x_i^2 - n\bar{x}^2) = \sum x_i y_i - n\bar{x}\bar{y}$$

$$\beta_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum x_i^2 - 2n\bar{x}^2 + \sum \bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$ should be minimum.

$\Downarrow$

$\beta_0 = \bar{y} - b\bar{x}$, so $(\bar{x}, \bar{y})$ on the regression line.

Exercise 3.

a. The term "regression" was proposed by Francis Galton.
The regression phenomenon was that the heights of descendants of tall ancestors tend to regress down towards a normal average.

b. For least square estimate, we should find a function that minimize the square of the sum of all potient errors
$\hat{y} = \beta_0 + \beta_1 \hat{x} + \varepsilon$, ε should be minimum and we can find the β_0, β_1 to make it minimum ($\sum_{i=1}^{n} e_i^2 = \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$)
For L_1-norm estimate, we should find a function that minimize the absolute of the sume of all potient errors ($\sum_{i=1}^{n} |e_i| = \sum_{i=1}^{n} |y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i|$), we can also find parameter β_0, β_1 to make it minimum.
For Robust estimate, just as its name implies, it can resist model deviations and resist abnormal observation distrbances. and we can find a parameter K, when the absolute value of errors (e_i) smaller than K, we will ~~th~~ take the square of e_i; otherwise, we will take the square of K.

```r
# Input the data of Xi and Yi
x <- c(0, 10, 20, 30, 40, 50, 60)
y <- c(16, 27, 28, 39, 39, 48, 51)

# Calculate the value
avg_x <- sum(x)/ length(x)
avg_y <- sum(y)/ length(y)

Sxx <- sum((x-avg_x)^2)
Syy <- sum((y-avg_y)^2)
Sxy <- sum((x-avg_x)*(y-avg_y))

# Get the b0, b1
b1 <- Sxy / Sxx
b0 <- avg_y - b1*avg_x

# Output the result
result <- matrix(nrow = 1, ncol = 7, dimnames = list(c("Value"),
          c("avg_x", "avg_y", "SSxx", "SSyy", "SSxy", "b0", "b1")))
result[1,1] <- avg_x
result[1,2] <- avg_y
result[1,3] <- Sxx
result[1,4] <- Syy
result[1,5] <- Sxy
result[1,6] <- b0
result[1,7] <- b1
result

# Plot the graphic, y = b0+b1*x
curve(b0+b1*x,0,60)
plot(x,y, type = "p", main = "The time spent going to school and time away from home after 7 am", col="blue",
     xlab = "Time away from home after 7 am", ylab = "The time spent going to school")
abline(a=b0, b=b1, col="red", lwd=2)
```

```
> result[1,3] <- Sxx
> result[1,4] <- Syy
> result[1,5] <- Sxy
> result[1,6] <- b0
> result[1,7] <- b1
> result
      avg_x    avg_y SSxx     SSyy SSxy   b0        b1
Value    30 35.42857 2800 929.7143 1580 18.5 0.5642857
>
> # Plot the graphic, y = b0+b1*x
> curve(b0+b1*x,0,60)
> plot(x,y, type = "p", main = "The time spent going to school and time away from home after 7 am", col="blue",
+      xlab = "Time away from home after 7 am", ylab = "The time spent going to school")
> abline(a=b0, b=b1, col="red", lwd=2)
>
>
>
```