STAT409 Homework2
2021320322 윤민서

1.
(a)

$S(\beta) = (y-X\beta)^T(y-X\beta)$

$$\frac{\partial S(\beta)}{\partial\beta} = \frac{\partial}{\partial\beta}\{(y-X\beta)^T(y-X\beta)\}$$
$$= \frac{\partial}{\partial\beta}(y^Ty - \beta^TX^Ty - y^TX\beta + \beta^TX^TX\beta)$$
$$= \frac{\partial}{\partial\beta}(-2y^TX\beta + \beta^TX^TX\beta)$$
$$= -2X^Ty + 2X^TX\beta \overset{set}{=} 0$$

$X^TX\beta = X^Ty$

$\Rightarrow (X^TX)^{-1}(X^TX)\beta = (X^TX)^{-1}X^Ty$

$\therefore \hat{\beta} = (X^TX)^{-1}X^Ty$

(b)

$$\|y-\bar{y}\|^2 = \|y-\hat{y}+\hat{y}-\bar{y}\|^2$$
$$= \|y-\hat{y}\|^2 + \|\hat{y}-\bar{y}\|^2 + 2(y-\hat{y})^T(\hat{y}-\bar{y})$$

$$(y-\hat{y})^T(\hat{y}-\bar{y}) = (y-\hat{y})^T(X\hat{\beta}-\bar{y})$$
$$= (y-\hat{y})^TX\hat{\beta} - (y-\hat{y})^T\bar{y} \quad \to \text{vector}$$
$$= (y-X(X^TX)^{-1}X^Ty)^TX\hat{\beta} - \underset{\text{scalar}}{\bar{y}}(y-\hat{y})^T\underset{\text{vector}}{1}$$

$$\vec{\bar{y}} = \bar{y}\vec{1}$$

$$= y^T(I-X(X^TX)^{-1}X^T)^TX\hat{\beta} - \bar{y}r^T1$$
$$= y^T(X^T - X^TX(X^TX)^{-1}X^T)^T\hat{\beta} - \bar{y}\sum_{i=1}^{n} r_i$$
$$= y^T(X^T - X^T)^T\hat{\beta}$$
$$= 0$$

$\therefore \|y-\bar{y}\|^2 = \|y-\hat{y}\|^2 + \|\hat{y}-\bar{y}\|^2$

2.

(a)

$$E(\hat{\beta}_1) = \frac{1}{S_x^2} E\left[\Sigma(x_i - \bar{x}_n) y_i - \cancelto{0}{\Sigma(x_i - \bar{x}_n)\bar{y}}\right]$$

$$= \frac{1}{S_x^2} E\left[\Sigma(x_i - \bar{x}_n)(\beta_0 + \beta_1 x_i + \varepsilon_i)\right]$$

$$= \frac{1}{S_x^2}\left[\cancelto{0}{\Sigma(x_i - \bar{x}_n)\beta_0} + \beta_1 \Sigma(x_i - \bar{x}_n) x_i + \cancelto{0}{\Sigma(x_i - \bar{x}_n) E(\varepsilon_i)}\right]$$

$$= \frac{\beta_1}{S_x^2} \Sigma(x_i - \bar{x}_n)(x_i - \bar{x}_n + \bar{x}_n)$$

$$= \frac{\beta_1}{S_x^2}\left[\Sigma(x_i - \bar{x}_n)^2 + \cancelto{0}{\bar{x}_n \Sigma(x_i - \bar{x}_n)}\right]$$

$$= \beta_1$$

$$Var(\hat{\beta}_1) = \frac{1}{(S_x^2)^2} Var\left[\Sigma(x_i - \bar{x}_n) y_i - \cancelto{0}{\Sigma(x_i - \bar{x}_n)\bar{y}}\right]$$

$$= \frac{1}{(S_x^2)^2} Var\left[\Sigma(x_i - \bar{x}_n)(\beta_0 + \beta_1 x_i + \varepsilon_i)\right]$$

$$= \frac{1}{(S_x^2)^2} Var\left[\Sigma(x_i - \bar{x}_n)\varepsilon_i\right]$$

independent

$$= \frac{1}{(S_x^2)^2} \Sigma(x_i - \bar{x}_n)^2 Var(\varepsilon_i)$$

$$= \frac{\sigma^2}{S_x^2} = \frac{\sigma^2}{\Sigma(x_i - \bar{x}_n)^2}$$

(b)

$$\hat{\beta}_1 = \frac{\Sigma(x_i - \bar{x}_n)\varepsilon_i}{S_x^2} + \frac{\Sigma(x_i - \bar{x}_n)(\beta_0 + \beta_1 x_i) - \Sigma(x_i - \bar{x}_n)\bar{y}_n}{S_x^2}$$ 으로 정리할 수 있고

ε_i에 대한 선형 결합으로 볼 수 있다.

$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$이고 정규분포의 선형 결합은 정규분포이므로 $\hat{\beta}_1$은 정규분포를 따른다.

3.

$$\frac{dA_n(c)}{dc} = \frac{1}{n}\sum_{i=1}^{n} sign(y_i - c) \overset{set}{=\!=} 0$$

$\frac{dA_n(c)}{dc}$가 0이 되려면 $sign(y_i-c)=1$인 y_i와 $sign(y_i-c)=-1$인 y_i의 개수가 같아야 한다.

$y_1, \cdots, y_n$을 정렬한 것을 $s_1 \le s_2 \le \cdots \le s_n$이라 하면

i) n이 홀수일 때, $sign(y_i-c)=1$인 y_i와 $sign(y_i-c)=-1$인 y_i의 개수가 같으므로

$\sum_{i=1}^{\frac{n-1}{2}} \{sign(s_i - c) + sign(s_{n+1-i} - c)\} = 0$이고

나머지 남은 수 $s_{\frac{n+1}{2}}$에 대하여 $sign(s_{\frac{n+1}{2}} - c) = 0$이므로 $c = s_{\frac{n+1}{2}} \Rightarrow$ median이다.

ii) n이 짝수일 때, $sign(y_i-c)=1$인 y_i와 $sign(y_i-c)=-1$인 y_i의 개수가 같으므로

$\sum_{i=1}^{\frac{n}{2}-1} \{sign(s_i - c) + sign(s_{n+1-i} - c)\} = 0$이고

나머지 남은 수 $s_{\frac{n}{2}}, s_{\frac{n}{2}+1}$에 대하여 $sign\left(\frac{s_{\frac{n}{2}} + s_{\frac{n}{2}+1}}{2} - c\right) = 0$이므로 $c = \frac{s_{\frac{n}{2}} + s_{\frac{n}{2}+1}}{2} \Rightarrow$ median이다.

4.

(a)

$$L(\theta) = \binom{n_1}{x_1}\theta^{x_1}(1-\theta)^{n_1-x_1} \times \binom{n_2}{x_2}\theta^{x_2}(1-\theta)^{n_2-x_2}$$

$$\begin{aligned} l(\theta) &= \log L(\theta) \\ &= \log\binom{n_1}{x_1} + \log\binom{n_2}{x_2} + (x_1+x_2)\log\theta + (n_1-x_1+n_2-x_2)\log(1-\theta) \end{aligned}$$

(b)

$$\frac{\partial l(\theta)}{\partial\theta} = \frac{x_1+x_2}{\theta} - \frac{n_1-x_1+n_2-x_2}{1-\theta} \overset{set}{=\!=} 0$$

$$\Rightarrow x_1+x_2-\theta(x_1+x_2) = \theta(n_1+n_2) - \theta(x_1+x_2)$$

$$\Rightarrow x_1+x_2 = \theta(n_1+n_2)$$

$$\therefore \hat{\theta} = \frac{x_1+x_2}{n_1+n_2}$$

(c)

$$\frac{\partial^2 l(\theta)}{\partial\theta^2} = -\frac{x_1+x_2}{\theta^2} - \frac{n_1-x_1+n_2-x_2}{(1-\theta)^2}$$

$$\begin{aligned} I(\theta) &= -E\left[\frac{\partial^2 l(\theta)}{\partial\theta^2}\right] \\ &= E\left[\frac{X_1+X_2}{\theta^2} + \frac{n_1-X_1+n_2-X_2}{(1-\theta)^2}\right] \\ &= \frac{n_1\theta+n_2\theta}{\theta^2} + \frac{n_1(1-\theta)+n_2(1-\theta)}{(1-\theta)^2} \\ &= \frac{n_1+n_2}{\theta} + \frac{n_1+n_2}{1-\theta} = \frac{n_1+n_2}{\theta(1-\theta)} \end{aligned}$$