

Homework #6 Solutions

(1) $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, where $\varepsilon_1, \dots, \varepsilon_n \sim \text{iid } N(0, \sigma^2)$ and $\sum_{i=1}^n X_i = 0$

(a) $\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon}$, where $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2 \underline{I})$

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \underline{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

(b) (i) $\underline{b} = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y} = \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum X_i^2} \end{bmatrix} \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$. So $\boxed{b_0 = \bar{Y}}$
 $\boxed{b_1 = \frac{\sum X_i Y_i}{\sum X_i^2}}$

(ii) $\underline{Cov}(\underline{b}) = \sigma^2 (\underline{X}'\underline{X})^{-1} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{\sum X_i^2} \end{bmatrix}$. So $\boxed{\text{Var}(b_0) = \frac{\sigma^2}{n}, \text{Var}(b_1) = \frac{\sigma^2}{\sum X_i^2}}$
 $\boxed{\text{Cov}(b_0, b_1) = 0}$

(iii) $\hat{Y}_h = b_0 + b_1 X_h = \underline{X}_h' \underline{b}$, where $\underline{X}_h' = [1 \ X_h]$

$\underline{\text{Var}}(\hat{Y}_h) = \underline{X}_h' \underline{Cov}(\underline{b}) \underline{X}_h = [1 \ X_h] \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{\sum X_i^2} \end{bmatrix} \begin{bmatrix} 1 \\ X_h \end{bmatrix} = \underline{\underline{\sigma^2 \left(\frac{1}{n} + \frac{X_h^2}{\sum X_i^2} \right)}}$

(2) β_ℓ is the difference in mean response from a 1 unit

(a) increase in X_ℓ (the ℓ^{th} input), with all other input levels held fixed.

(b) $\underline{b} = \begin{bmatrix} 158.5 \\ -1.14 \\ -0.44 \\ -13.5 \end{bmatrix}$ (c) Each of the input variables (age, severity, anxiety) has a negative effect on patient satisfaction.

(d) see attached for $\hat{\underline{Cov}}(\underline{b})$.

HW 6 Computing

Data from Exercise 6.15

Our goal is to study the relationship between patient satisfaction (y) and patient's age (x1), severity of illness (x2), and anxiety level (x3).

:

```
hw6.data = read.table(  
  'http://users.stat.ufl.edu/~rrandles/sta4210/Rclassnotes/data/textdatasets/Ku  
tnerData/Chapter%20%206%20Data%20Sets/CH06PR15.txt'  
)  
colnames(hw6.data)=c("satisfaction","age","illness","anxiety")  
str(hw6.data)
```

```
## 'data.frame': 46 obs. of 4 variables:  
## $ satisfaction: int 48 57 66 70 89 36 46 54 26 77 ...  
## $ age : int 50 36 40 41 28 49 42 45 52 29 ...  
## $ illness : int 51 46 48 44 43 54 50 48 62 50 ...  
## $ anxiety : num 2.3 2.3 2.2 1.8 1.8 2.9 2.2 2.4 2.9 2.1 ...
```

```
hw6.mod = lm(satisfaction ~ age + illness + anxiety, data = hw6.data)
```

```
summary(hw6.mod)
```

```
##  
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 158.4913 18.1259 8.744 5.26e-11 ***  
## age -1.1416 0.2148 -5.315 3.81e-06 ***  
## illness -0.4420 0.4920 -0.898 0.3741  
## anxiety -13.4702 7.0997 -1.897 0.0647 .  
## ---
```

```
vcov(hw6.mod)
```

```
## (Intercept) age illness anxiety  
## (Intercept) 328.5478428 0.93283693 -6.87207388 -6.8081417  
## age 0.9328369 0.04613853 -0.03223004 -0.4716488  
## illness -6.8720739 -0.03223004 0.24203030 -1.7916031  
## anxiety -6.8081417 -0.47164876 -1.79160306 50.4051837
```