stat430 a4 q2

2023-11-23

```
x1 = rep( c(-1,1), times=8)
x2 = rep( c(-1,1), times=4, each=2)
x3 = rep( c(-1,1), times=2, each=4)
y = c(5,15,19,11,4,26,-3,21,5,15,21,13,12,17,2,17)
datq2 = data.frame(y,x1,x2,x3)
datq2
```

```
##     y x1 x2 x3
## 1   5 -1 -1 -1
## 2  15  1 -1 -1
## 3  19 -1  1 -1
## 4  11  1  1 -1
## 5   4 -1 -1  1
## 6  26  1 -1  1
## 7  -3 -1  1  1
## 8  21  1  1  1
## 9   5 -1 -1 -1
## 10 15  1 -1 -1
## 11 21 -1  1 -1
## 12 13  1  1 -1
## 13 12 -1 -1  1
## 14 17  1 -1  1
## 15  2 -1  1  1
## 16 17  1  1  1
```

(a)

```
# main effect of factor A
y_A_high <- mean(datq2$y[datq2$x1==1])
y_A_low <- mean(datq2$y[datq2$x1==-1])

ME_A <- y_A_high - y_A_low
ME_A
```

```
## [1] 8.75
```

(b)

```
# Interaction Effect between A and B
x12 <- x1 * x2
data <- data.frame(y,x12)

y_AB_high <- mean(data$y[data$x12==1])
y_AB_low <- mean(data$y[data$x12==-1])
```

```r
IE_AB <- y_AB_high - y_AB_low
IE_AB
```

```
## [1] -3
```

(c)

```r
# Interaction Effect between A,B,C
x123 <- x12 * x3
data <- data.frame(y,x123)

y_ABC_high <- mean(data$y[data$x123==1])
y_ABC_low <- mean(data$y[data$x123==-1])

IE_ABC <- y_ABC_high - y_ABC_low
IE_ABC
```

```
## [1] 6
```

(d)

```r
# simplest model
mod1 <- lm(y ~ x1 + x2 + x3, data = datq2)
summary(mod1)
```

```
##
## Call:
## lm(formula = y ~ x1 + x2 + x3, data = datq2)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -10.75  -3.75  -2.25   4.50  12.25
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   12.500      1.839   6.796 1.91e-05 ***
## x1             4.375      1.839   2.379   0.0348 *
## x2             0.125      1.839   0.068   0.9469
## x3            -0.500      1.839  -0.272   0.7904
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.357 on 12 degrees of freedom
## Multiple R-squared:  0.3234, Adjusted R-squared:  0.1543
## F-statistic: 1.912 on 3 and 12 DF,  p-value: 0.1815
```

```r
beta1 <- summary(mod1)$coefficients[2]
beta1
```

```
## [1] 4.375
```

```r
# check work from (a)
all.equal(2 * beta1, ME_A)
```

```
## [1] TRUE
```

Comments: We found that 2 times estimate of beta1 equals to the estimate of main effect A, we verified our work.

(e)

```
# full model
mod2 <- lm(y ~ x1 + x2 + x3 + x1:x2 + x1:x2:x3)
summary(mod2)
```

```
##
## Call:
## lm(formula = y ~ x1 + x2 + x3 + x1:x2 + x1:x2:x3)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.250 -4.438  0.500  3.000 11.250
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   12.500      1.713   7.297 2.61e-05 ***
## x1             4.375      1.713   2.554   0.0287 *
## x2             0.125      1.713   0.073   0.9433
## x3            -0.500      1.713  -0.292   0.7763
## x1:x2         -1.500      1.713  -0.876   0.4018
## x1:x2:x3       3.000      1.713   1.751   0.1104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.852 on 10 degrees of freedom
## Multiple R-squared:  0.5109, Adjusted R-squared:  0.2664
## F-statistic: 2.089 on 5 and 10 DF,  p-value: 0.1504
```

```
# check work from (b)
beta4 <- summary(mod2)$coefficients[5]
beta4
```

```
## [1] -1.5
```

```
all.equal(2 * beta4, IE_AB)
```

```
## [1] TRUE
```

```
# check work from (c)
beta5 <- summary(mod2)$coefficients[6]
beta5
```

```
## [1] 3
```

```
all.equal(2 * beta5, IE_ABC)
```

```
## [1] TRUE
```

Comments: We found that 2 times estimates of beta4 and beta5 equals to the estimates of Interaction Effect between A&B, A&B&C separately, we verified our work.

(f)

```
X <- cbind(x1,x2,x3) # design matrix
XTX <- t(X) %*% X
XTX_inverse <- solve(XTX)
XTX_inverse
```

```
##        x1     x2     x3
## x1 0.0625 0.0000 0.0000
## x2 0.0000 0.0625 0.0000
## x3 0.0000 0.0000 0.0625
```

Comments: Obviously, the matrix is disgonal.