Quiz 3

Quiz, 7 questions

7/7 points (100%)

Congratulations! You passed!

Next Item

1/1 points

1. Consider the **`mtcars`** data set. Fit a model with mpg as the outcome that includes number of cylinders as a factor variable and weight as confounder. Give the adjusted estimate for the expected change in mpg comparing 8 cylinders to 4.

- ○ 33.991
- ○ -3.206
- ◉ -6.071

Correct

```
fit <- lm(mpg ~ factor(cyl) + wt, data = mtcars)
summary(fit)$coef
```

```
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept)    33.991     1.8878  18.006 6.257e-17
## factor(cyl)6   -4.256     1.3861  -3.070 4.718e-03
## factor(cyl)8   -6.071     1.6523  -3.674 9.992e-04
## wt             -3.206     0.7539  -4.252 2.130e-04
```

- ○ -4.256

1/1 points

2. Consider the **`mtcars`** data set. Fit a model with mpg as the outcome that includes number of cylinders as a factor variable and weight as a possible confounding variable. Compare the effect of 8 versus 4 cylinders on mpg for the adjusted and unadjusted by weight models. Here, adjusted means including the weight variable as a term in the regression model and unadjusted means the model without weight included. What can be said about the effect comparing 8 and 4 cylinders after looking at models with and without weight included?.

- ○ Including or excluding weight does not appear to change anything regarding the estimated impact of number of cylinders on mpg.
- ○ Within a given weight, 8 cylinder vehicles have an expected 12 mpg drop in fuel efficiency.
- ◉ Holding weight constant, cylinder appears to have less of an impact on mpg than if weight is disregarded.

Correct

It is both true and sensible that including weight would attenuate the effect of number of cylinders on mpg.

- ○ Holding weight constant, cylinder appears to have more of an impact on mpg than if weight is disregarded.

1/1 points

3. Consider the **`mtcars`** data set. Fit a model with mpg as the outcome that considers number of cylinders as a factor variable and weight as confounder. Now fit a second model with mpg as the outcome model that considers the interaction between number of cylinders (as a factor variable) and weight. Give the P-value for the likelihood ratio test comparing the two models and suggest a model using 0.05 as a type I error rate significance benchmark.

- ○ The P-value is small (less than 0.05). So, according to our criterion, we reject, which suggests that the interaction term is not necessary.
- ◉ The P-value is larger than 0.05. So, according to our criterion, we would fail to reject, which suggests that the interaction terms may not be necessary.

Correct

```
fit1 <- lm(mpg ~ factor(cyl) + wt, data = mtcars)
fit2 <- lm(mpg ~ factor(cyl) * wt, data = mtcars)
summary(fit1)$coef
```

```
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept)    33.991     1.8878  18.006 6.257e-17
## factor(cyl)6   -4.256     1.3861  -3.070 4.718e-03
## factor(cyl)8   -6.071     1.6523  -3.674 9.992e-04
## wt             -3.206     0.7539  -4.252 2.130e-04
```

```
summary(fit2)$coef
```

```
##                 Estimate Std. Error t value  Pr(>|t|)
## (Intercept)       39.571      3.194 12.3895 2.058e-12
## factor(cyl)6     -11.162      9.355 -1.1932 2.436e-01
## factor(cyl)8     -15.703      4.839 -3.2448 3.223e-03
## wt                -5.647      1.359 -4.1538 3.128e-04
## factor(cyl)6:wt    2.867      3.117  0.9197 3.662e-01
## factor(cyl)8:wt    3.455      1.627  2.1229 4.344e-02
```

```
anova(fit1, fit2)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ factor(cyl) + wt
## Model 2: mpg ~ factor(cyl) * wt
##   Res.Df RSS Df Sum of Sq    F Pr(>F)
## 1     28 183
## 2     26 156  2      27.2 2.27   0.12
```

- ○ The P-value is small (less than 0.05). Thus it is surely true that there is an interaction term in the true model.
- ○ The P-value is larger than 0.05. So, according to our criterion, we would fail to reject, which suggests that the interaction terms is necessary.
- ○ The P-value is small (less than 0.05). So, according to our criterion, we reject, which suggests that the interaction term is necessary
- ○ The P-value is small (less than 0.05). Thus it is surely true that there is no interaction term in the true model.

1/1 points

4. Consider the **`mtcars`** data set. Fit a model with mpg as the outcome that includes number of cylinders as a factor variable and weight inlcuded in the model as

```
lm(mpg ~ I(wt * 0.5) + factor(cyl), data = mtcars)
```

How is the wt coefficient interpreted?

- ○ The estimated expected change in MPG per half ton increase in weight.
- ○ The estimated expected change in MPG per half ton increase in weight for for a specific number of cylinders (4, 6, 8).
- ◉ The estimated expected change in MPG per one ton increase in weight for a specific number of cylinders (4, 6, 8).

Correct

- ○ The estimated expected change in MPG per half ton increase in weight for the average number of cylinders.
- ○ The estimated expected change in MPG per one ton increase in weight.

1/1 points

5. Consider the following data set

```
x <- c(0.586, 0.166, -0.042, -0.614, 11.72)
y <- c(0.549, -0.026, -0.127, -0.751, 1.344)
```

Give the hat diagonal for the most influential point

- ○ 0.2287
- ○ 0.2025
- ○ 0.2804
- ◉ 0.9946

Correct

```
influence(lm(y ~ x))$hat
```

```
##      1      2      3      4      5
## 0.2287 0.2438 0.2525 0.2804 0.9946
```

```
## showing how it's actually calculated
xm <- cbind(1, x)
diag(xm %*% solve(t(xm) %*% xm) %*% t(xm))
```

```
## [1] 0.2287 0.2438 0.2525 0.2804 0.9946
```

1/1 points

6. Consider the following data set

```
x <- c(0.586, 0.166, -0.042, -0.614, 11.72)
y <- c(0.549, -0.026, -0.127, -0.751, 1.344)
```

Give the slope dfbeta for the point with the highest hat value.

- ○ -.00134
- ○ -0.378
- ○ 0.673
- ◉ -134

Correct

```
influence.measures(lm(y ~ x))
```

```
## Influence measures of
##   lm(formula = y ~ x) :
##
##      dfb.1_     dfb.x     dffit cov.r   cook.d   hat inf
## 1  1.0621 -3.78e-01    1.0679 0.341 2.93e-01 0.229    *
## 2  0.0675 -2.86e-02    0.0675 2.934 3.39e-03 0.244
## 3 -0.0174  7.92e-03   -0.0174 3.007 2.26e-04 0.253    *
## 4 -1.2496  6.73e-01   -1.2557 0.342 3.91e-01 0.280    *
## 5  0.2043 -1.34e+02 -149.7204 0.107 2.70e+02 0.995    *
```

1/1 points

7. Consider a regression relationship between Y and X with and without adjustment for a third variable Z. Which of the following is true about comparing the regression coefficient between Y and X with and without adjustment for Z.

- ◉ It is possible for the coefficient to reverse sign after adjustment. For example, it can be strongly significant and positive before adjustment and strongly significant and negative after adjustment.

Correct

See lecture 02_03 for various examples.

- ○ Adjusting for another variable can only attenuate the coefficient toward zero. It can't materially change sign.
- ○ The coefficient can't change sign after adjustment, except for slight numerical pathological cases.
- ○ For the coefficient to change sign, there must be a significant interaction term.