Quiz2

jbellier

12 décembre 2016

Question 1

Consider the following data :

```
x <- c(0.61, 0.93, 0.83, 0.35, 0.54, 0.16, 0.91, 0.62, 0.62)
y <- c(0.67, 0.84, 0.6, 0.18, 0.85, 0.47, 1.1, 0.65, 0.36)
```

Give a p-value for the two-sided hypothesis test of wether β_1 from a linear regression model is 0 or not.

- 2.325
- 0.391
- 0.05296
- 0.025

```
fit <- lm(y~x)
summary(fit)$coefficients
```

```
##               Estimate Std. Error   t value   Pr(>|t|)
## (Intercept) 0.1884572  0.2061290 0.9142681 0.39098029
## x           0.7224211  0.3106531 2.3254912 0.05296439
```

The p-value that we are looking for is then : ***0.05296***.

Question 2

Consider the previous problem, give the estimate the residual standard deviation.

- 0.4358
- 0.05296
- 0.223
- 0.3552

```
summary(fit)$sigma
```

```
## [1] 0.2229981
```

So the answer is ***0.223***.

Question 3

In the mtcars data set, fit a linear regression model of weight (predictor) on mpg (outcome). Get a 95% confidence interval for the expected mpg at the average weight. What is the lower endpoint ?

- 21.190
- -6.486
- 18.991
- -4.00

```
fit2 <- lm(mpg~wt,data=mtcars)
coeffs <- summary(fit2)$coefficients
newwt= data.frame(wt=c(mean(mtcars$wt)))
newwt
```

```
##        wt
## 1 3.21725
```

```
p1 = data.frame(predict(fit2, newdata=newwt,interval="confidence"))
p1
```

```
##        fit      lwr      upr
## 1 20.09062 18.99098 21.19027
```

Answer : : The lower boud of con fidence interval is then ***18.991***

Question 4

Refer to the previous question. Read the help file for mtcars. What is the weight coefficient interpreted as ?

- The estimated expected change in mpg per 1,000 lb increase in weight
- The estimated expected 1,000 lb change in weight per mpg increase
- It can't be interpreted without further information
- The estimated expected change in mpg per 1 lb increase in weight

Answer : the correct choice is the first one. According to the help file of mtcars, the weight is measured in 1,000 lbs. As the model determines mpg as the output with the weight as predictor, it is the estimated expected change in mpg per 1,000 lb increase in weight.

Question 5

Consider again mtcars data set and a linear regression model with mpg as predicted by weight (1,000 lbs). A new car is coming weighting 3,000 pounds. Construct a 95% prediction interval for its mpg. What is the upper endpoint ?

- -5.77
- 27.57
- 14.93
- 21.25

```
fit3 <- lm(mpg~wt,data=mtcars)
newwt= data.frame(wt=c(3))
p2 = data.frame(predict(fit2, newdata=newwt,interval="predict"))
p2
```

```
##        fit      lwr      upr
## 1 21.25171 14.92987 27.57355
```

Answer :   the upper point is *27.57*.

Question 6

Consider again mtcars data set and a linear regression model with mpg as predicted by weight (1,000 lbs). A
"short" ton is defined as 2,000 lbs. Construct a 95% prediction interval for the expected change in mpg per 1
short ton increase in weight. Give the lower endpoint.

- -6.486
- -9.000
- -12.973
- 4.2026

Answer :   As the weight unit is twice the original one, this means that all the predictor values need to be
divived by two. So the coefficients will be multiplied by 2.

```
fit3 <- lm(mpg~I(wt*0.5),data=mtcars)
confint(fit3)[2,]
```

```
##      2.5 %    97.5 %
## -12.97262  -8.40527
```

The lower endpoint of the interval is *-12.973*.

Question 7

If my X from a linear regression is measured in centimeters and I convert it to meters what would happen to
the slope coefficient ?

- It would get divided by 100.
- It would get multiplied by 100.
- It would get divided by 10.
- It would get multiplied by 10.

Answer :   the correct answer is : *It would get multiplied by 100*.

Question 8

I have an outcome Y and a predictor X and fit a linear regression model with $Y = \beta_0 + \beta_1 * X + \epsilon$ to obtain $\hat{\beta}_0$ and $\hat{\beta}_1$. What would be the consequence to the subsequent slope and intercept if I were to refit the model with a new regressor $X + c$ for some constant c ?

- The new intercept would be $\hat{\beta}_0 + c\hat{\beta}_1$.
- The new intercept would be $\hat{\beta}_0 - c\hat{\beta}_1$.
- The new slope would be $c\hat{\beta}_1$.
- The new slope would be $\hat{\beta}_1 - c$.

Answer : If we add a constant to the predictor, it should be substracted to keep the consistency of the equations. So ***The new intercept would be*** $\hat{\beta}_0 - c\hat{\beta}_1$.

Question 9

Refer back to the mtcars dataset with mpg as an outcome and weight (wt) as the predictor. About what is the ratio of the sum of the squared errors, $\Sigma_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ when comparing a model with just an intercept (denominator) to the model with intercept an slope (numerator) ?

- 0.75
- 0.25
- 0.50
- 4.00

```
fit4 <- lm(mpg~wt,data=mtcars)
fit5 <- lm(mpg~1,data=mtcars)

num <- sum((predict(fit4)-mtcars$mpg)^2)
den <- sum((predict(fit5)-mtcars$mpg)^2)
num/den
```

```
## [1] 0.2471672
```

Answer : The correct answer is then ***0.25***.

Question 10

Do the residuals always sum to zero in linear regression ?

- The residuals always sum to 0
- If an intercept is included, the residuals most likely won't sum to 0
- If an intercept is included, then they will sum to 0
- The residuals never sum to 0

Answer : ***The residuals always sum to 0 if an intercept is defined.*** Otherwise, they may not.