Question 2

Joseph Froelicher

11/5/2020

# Data  
hw7 <- data.frame(  
 id = 1:7,  
 gender = c(0,1,0,1,0,1,0),  
 chol = c(254,402,288,354,220,451,405),  
 wtkg = c(57,79,63,84,30,76,65),  
 age = c(23,57,28,46,34,57,52)  
)

# Linear regression cholesterol and weight  
fit\_lm <- lm(chol ~ hw7$wtkg, data = hw7)  
  
# Info for table  
sum <- summary(fit\_lm)  
sum

##   
## Call:  
## lm(formula = chol ~ hw7$wtkg, data = hw7)  
##   
## Residuals:  
## 1 2 3 4 5 6 7   
## -55.86 10.14 -44.22 -56.50 10.78 70.32 65.32   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 97.395 89.782 1.085 0.3275   
## hw7$wtkg 3.727 1.340 2.781 0.0388 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 59.5 on 5 degrees of freedom  
## Multiple R-squared: 0.6074, Adjusted R-squared: 0.5289   
## F-statistic: 7.735 on 1 and 5 DF, p-value: 0.03884

ci\_lm <- confint(fit\_lm)  
ci\_lm

## 2.5 % 97.5 %  
## (Intercept) -133.3957971 328.186461  
## hw7$wtkg 0.2823573 7.172412

# Part A

table <- cbind(sum$coefficients, ci\_lm)  
table

## Estimate Std. Error t value Pr(>|t|) 2.5 % 97.5 %  
## (Intercept) 97.395332 89.781670 1.084802 0.32752222 -133.3957971 328.186461  
## hw7$wtkg 3.727385 1.340174 2.781269 0.03884374 0.2823573 7.172412

# Part B

y = 97.395332 + 3.7273847

# Part C

Intercept: 97.395332  
Average cholesterol if weight is zero kilograms is 97.395332 (probably not interpretable).

# Part D

95% confidence interval: [-133.3957971, 328.1864611]  
We can say with 95% confidence that at weight zero, average Cholesterol is between -133.3957971 and 328.1864611.

# Part E

Fail to reject the hypothesis that the true intercept is 0, based on our confidence interval [-133.3957971, 328.1864611] contains 0.

# Part F

Estimated slope: 3.7273847 For a unit increase in weight, for persons with weight between 30 and 84 kilograms, average cholesterol increases by 3.7273847 mg/100ml.

# Part G

95% confidence interval: [0.2823573, 7.1724122] We can say with 95% confidence that the average increase in Cholesterol is between 0.2823573 and 7.1724122 for a unit increase in weight.

# Part H

The 95% confidence interval for the slope of the regression line of weight and cholesterol does not contain 0. Therefore, we reject that hypothesis that the true slope of the regression line of weight and cholesterol is equal to 0.

# Part I

There is a significant increase in plasma levels of total cholesterol (mg/100mL) for increasing weight (kg) in individuals (p = 0.388). On average, the plasma levels of total cholesterol increases by 3.7273847 mg/100mL (95% CI: 0.2823573 to 7.1724122 mg/100mL) for every 1 kg increase in weight.

# Part J

pred <- predict(fit\_lm, newdata = data.frame(hw7\_new=50), interval='prediction')

## Warning: 'newdata' had 1 row but variables found have 7 rows

hw7\_new <- cbind(hw7, pred)  
  
ggplot(data = hw7\_new, aes(x = wtkg, y = chol)) +  
 geom\_point() +  
 geom\_line(aes(y = lwr), color = "purple", linetype = "twodash")+  
 geom\_line(aes(y = upr), color = "purple", linetype = "twodash")+  
 geom\_smooth(method = lm, se = TRUE, color = "orange") +  
 labs(title = 'Simple Linear Regression: Weight vs. Cholesterol', x = 'Weight (kg)', y = 'Cholesterol (mg/100mL)')

## `geom\_smooth()` using formula 'y ~ x'

