MATH 3080 Lab Project 11

Aini Liang

March 29, 2018

Remember: I expect to see commentary either in the text, in the code with comments created using `#`*, or (preferably) both!* ***Failing to do so may result in lost points!***

Problem 1

Analyze the relationship between wrist size ('wrist') and circumference of the abdomen ('abdomen') by using the data contained with the `fat` (*UsingR*) dataset. (Load the package and examine the dataset first.)

1. Is a linear model appropriate?

```
# Your code here

library(UsingR)
```

```
## Loading required package: MASS
```

```
## Loading required package: HistData
```

```
## Loading required package: Hmisc
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

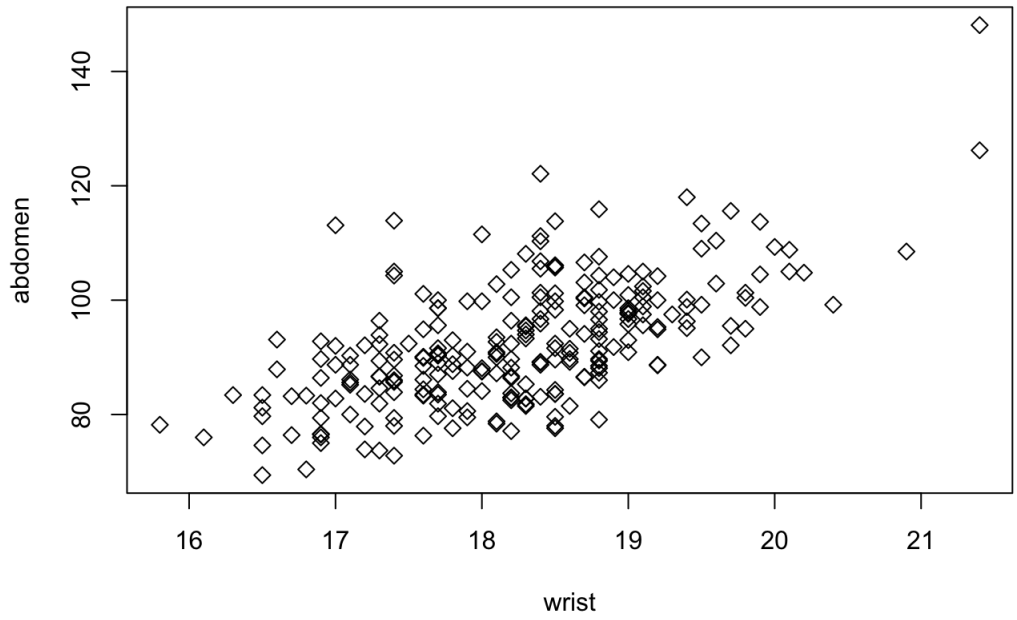
```
##
## Attaching package: 'UsingR'
```

```
## The following object is masked from 'package:survival':
##
##     cancer
```

```
attach(fat)
cor(wrist, abdomen)
```

```
## [1] 0.6198324
```

```
plot(wrist, abdomen, pch = 5)
```




```
fit <- lm(abdomen ~ wrist)
```

2. What is the equation of the least squares line? Are the assumptions met?

```
# Your code here

attach(fat)
```

```
## The following objects are masked from fat (pos = 3):
##
##     abdomen, age, ankle, bicep, BMI, body.fat, body.fat.siri,
##     case, chest, density, ffweight, forearm, height, hip, knee,
##     neck, thigh, weight, wrist
```

```
fit <- lm(abdomen ~ wrist)
summary(fit)
```

```
##
## Call:
## lm(formula = abdomen ~ wrist)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.538  -5.590  -0.647   4.338  32.848
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -37.9542    10.4638  -3.627 0.000347 ***
## wrist         7.1592     0.5732  12.489  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.479 on 250 degrees of freedom
## Multiple R-squared:  0.3842, Adjusted R-squared:  0.3817
## F-statistic:   156 on 1 and 250 DF,  p-value: < 2.2e-16
```

3. What proportion of the observed variation in 'abdomen' can be explained through 'wrist'?

```
# Your code here 38.42% of the observed variation in abdomen can be
# explained through wrist
```

4. In the summary of your linear model what are the p-values for the t statistic associated with 'wrist' and the F statistic for the model?

Why are they so small? Why are they useful?

```
# Your code here
summary(fit)$coefficients[, 4]
```

```
##  (Intercept)        wrist
## 3.472679e-04 3.874236e-28
```

```
summary(fit)$r.squared
```

```
## [1] 0.3841922
```

5. Is there strong evidence that the true slope is greater than 3? Use the estimate for the slope and its standard error to devise a (one sided) test.

```
# Your code here Yes.
```

6. Construct a confidence interval for the true, average abdomen circumference of individuals with wrist size of 17cm.

```
# Your code here
library(UsingR)
attach(fat)
```

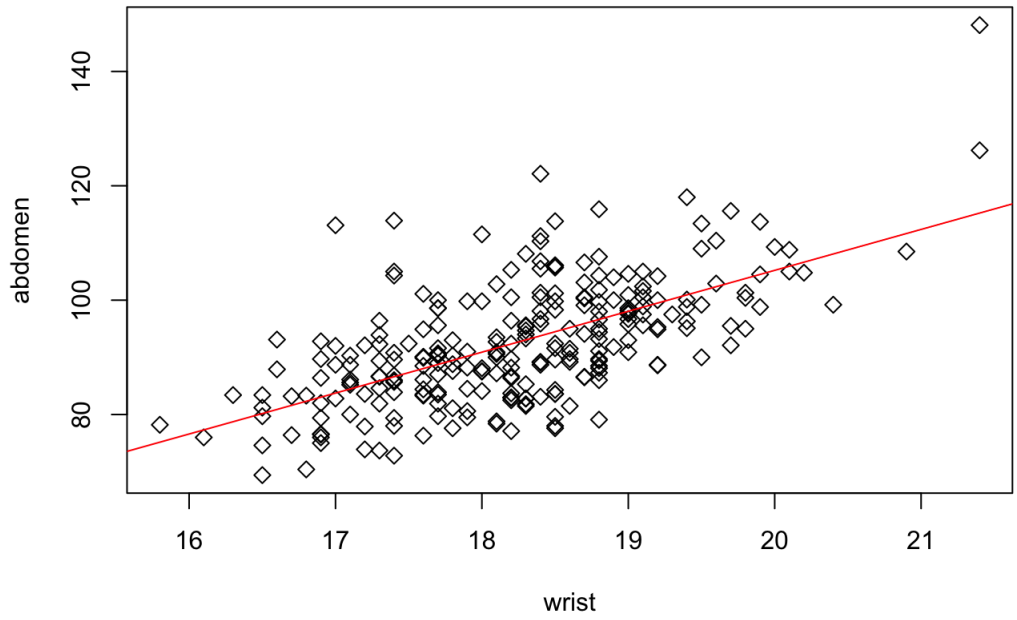
```
## The following objects are masked from fat (pos = 3):
##
##     abdomen, age, ankle, bicep, BMI, body.fat, body.fat.siri,
##     case, chest, density, ffweight, forearm, height, hip, knee,
##     neck, thigh, weight, wrist
```

```
## The following objects are masked from fat (pos = 4):
##
##     abdomen, age, ankle, bicep, BMI, body.fat, body.fat.siri,
##     case, chest, density, ffweight, forearm, height, hip, knee,
##     neck, thigh, weight, wrist
```

```
fit = lm(abdomen ~ wrist)
summary(fit)
```

```
##
## Call:
## lm(formula = abdomen ~ wrist)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.538  -5.590  -0.647   4.338  32.848
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -37.9542    10.4638  -3.627 0.000347 ***
## wrist         7.1592     0.5732  12.489  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.479 on 250 degrees of freedom
## Multiple R-squared:  0.3842, Adjusted R-squared:  0.3817
## F-statistic:   156 on 1 and 250 DF,  p-value: < 2.2e-16
```

```
plot(wrist, abdomen, pch = 5)
abline(fit, col = "red")
```




```
confint(fit, level = 0.99)
```

```
##                  0.5 %     99.5 %
## (Intercept) -65.114361 -10.793955
## wrist         5.671238   8.647118
```

```
c.int = predict(fit, data.frame(wrist = c(17)), int = "c")
```

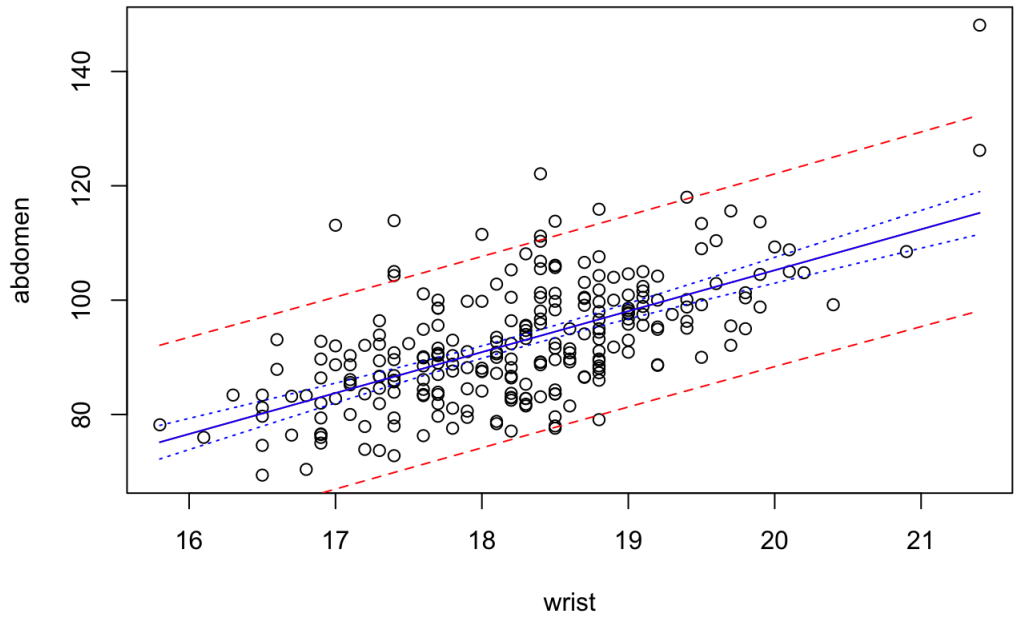
7. Construct a prediction interval for the abdomen circumference of the next individual with wrist size of 17cm.

```
# Your code here
p.int = predict(fit, data.frame(wrist = c(17)), int = "p")
```

8. Plot your confidence bands and prediction bands.

```
# Your code here
x.val = seq(min(wrist), max(wrist), length = 1000)
pb = predict(fit, data.frame(wrist = x.val), int = "p")
cb = predict(fit, data.frame(wrist = x.val), int = "c")

# plotting prediction and confidence bands #
plot(wrist, abdomen)
matlines(x.val, pb, lty = c(1, 2, 2), col = "red")
matlines(x.val, cb, lty = c(1, 3, 3), col = "blue")
```


abdomen
80
100
120
140
16
17
18
19
20
21
wrist