Question 1

Joseph Froelicher

December 3, 2020

Part A

```
fit1 = lm(iq ~ expose, data = lead)
summary(fit1)
```

```
##
## Call:
## lm(formula = iq ~ expose, data = lead)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -51.705 -10.127   1.295  10.295  46.295
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  102.705      1.767  58.121  < 2e-16 ***
## expose        -7.770      2.901  -2.678  0.00842 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.61 on 122 degrees of freedom
## Multiple R-squared:  0.05553,    Adjusted R-squared:  0.04779
## F-statistic: 7.173 on 1 and 122 DF,  p-value: 0.008421
```

```
a = fit1$coefficients[2]
```

On average, those who are exposed to lead have a -7.7703456 point lower IQ score than those who were not exposed.

Part B

```
fit2 = lm(iq ~ expose + sex, data = lead)
summary(fit2)
```

```
##
## Call:
## lm(formula = iq ~ expose + sex, data = lead)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -52.664  -9.935   0.921   9.839  45.336
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  106.002      4.439  23.878   <2e-16 ***
## expose        -7.916      2.911  -2.719   0.0075 **
## sex           -2.338      2.887  -0.810   0.4196
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.63 on 121 degrees of freedom
## Multiple R-squared:  0.06062,    Adjusted R-squared:  0.04509
## F-statistic: 3.904 on 2 and 121 DF,  p-value: 0.02274
```

```
b = fit2$coefficients[2]
```

In the presence of sex, those who are expose to lead have a -7.9163116 point lower IQ score than those who were not exposed.

Part C

```
fit3 = lm(sex ~ expose, data = lead)
summary(fit3)
```

```
##
## Call:
## lm(formula = sex ~ expose, data = lead)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4103 -0.4103 -0.3478  0.5897  0.6522
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.41026    0.05550  25.412   <2e-16 ***
## expose      -0.06243    0.09111  -0.685    0.495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4901 on 122 degrees of freedom
## Multiple R-squared:  0.003833,   Adjusted R-squared:  -0.004332
## F-statistic: 0.4695 on 1 and 122 DF,  p-value: 0.4945
```

```
c1 = (fit1$coefficients[2] - fit2$coefficients[2]) / fit1$coefficients[2] * 100
c2 = summary(fit2)$coefficients[3, 4]
```

Sex is not a confounder of the association between IQ and lead exposure. We should report the result from part A, because it is not accounting for sex in the model. We have established that sex (-1.8785011% change, p = 0.4196026) is not significant, and therefore, there is no reason to include it in the model.