Answer 1:

library(MASS)

summary(birthwt)

#I want to look at my data to see if it is homoscedastic and if I

#can use one line to fit all or use quantile regressions

plot(birthwt$age, birthwt$bwt) # in my opinion, it is not homoscedastic



my\_linear <- lm(bwt~age, data=birthwt)

summary(my\_linear)

Call:

lm(formula = bwt ~ age, data = birthwt)

Residuals:

Min 1Q Median 3Q Max

-2294.78 -517.63 10.51 530.80 1774.92

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 2655.74 238.86 11.12 <2e-16 \*\*\*

age 12.43 10.02 1.24 0.216

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 728.2 on 187 degrees of freedom

Multiple R-squared: 0.008157, Adjusted R-squared: 0.002853

F-statistic: 1.538 on 1 and 187 DF, p-value: 0.2165

#the linear model has a coefficient of 12.43 but it's insignificant

# I will use quantile regression instead

library(quantreg)

#fitting a quantile regression for 20th quantile

my\_quant <- rq(bwt~age, data=birthwt, tau=0.2)

summary(my\_quant)

Call: rq(formula = bwt ~ age, tau = 0.2, data = birthwt)

tau: [1] 0.2

Coefficients:

coefficients lower bd upper bd

(Intercept) 2090.00000 1525.71583 2761.14299

age 10.66667 -21.42334 36.84189

#fitting a quantile regression for 50th quantile

my\_quant <- rq(bwt~age, data=birthwt, tau=0.5)

summary(my\_quant)

Call: rq(formula = bwt ~ age, tau = 0.5, data = birthwt)

tau: [1] 0.5

Coefficients:

coefficients lower bd upper bd

(Intercept) 2798.87500 1922.05497 3392.04702

age 7.12500 -12.44784 46.14190

#fitting a quantile regression for 80th quantile

my\_quant <- rq(bwt~age, data=birthwt, tau=0.8)

summary(my\_quant)

Call: rq(formula = bwt ~ age, tau = 0.8, data = birthwt)

tau: [1] 0.8

Coefficients:

coefficients lower bd upper bd

(Intercept) 3105.25000 2756.42533 3676.44756

age 23.12500 -1.71400 43.93367

**CONCLUSIONS:**

Assuming there’s causation between the mother’s age and the birth weight of the baby, I have determined that the data is not homoscedastic (based on the plot). Therefore, the linear regression cannot be applied and gives an insignificant coefficient equal to 12.

This being said, I used 3 different quantile regressions to see how the causation changes over the range of age.

The first regression for quantile 20, said that for every increase in age by 1 year, the birth weight increases by 10.6 grams.

The second regression for quantile 50, said that for every increase in age by 1 year, the birth weight increases by 7.12 grams.

The third regression for quantile 80, said that for every increase in age by 1 year, the birth weight increases by 23.1 grams.

**This means that for young mothers, when they get older by one year, the baby gets heavier by 10-7 grams, but for older mothers, the weight might increase by up to 23 grams.**