19-12

Nikhil Gopal

4/1/2022

19.12

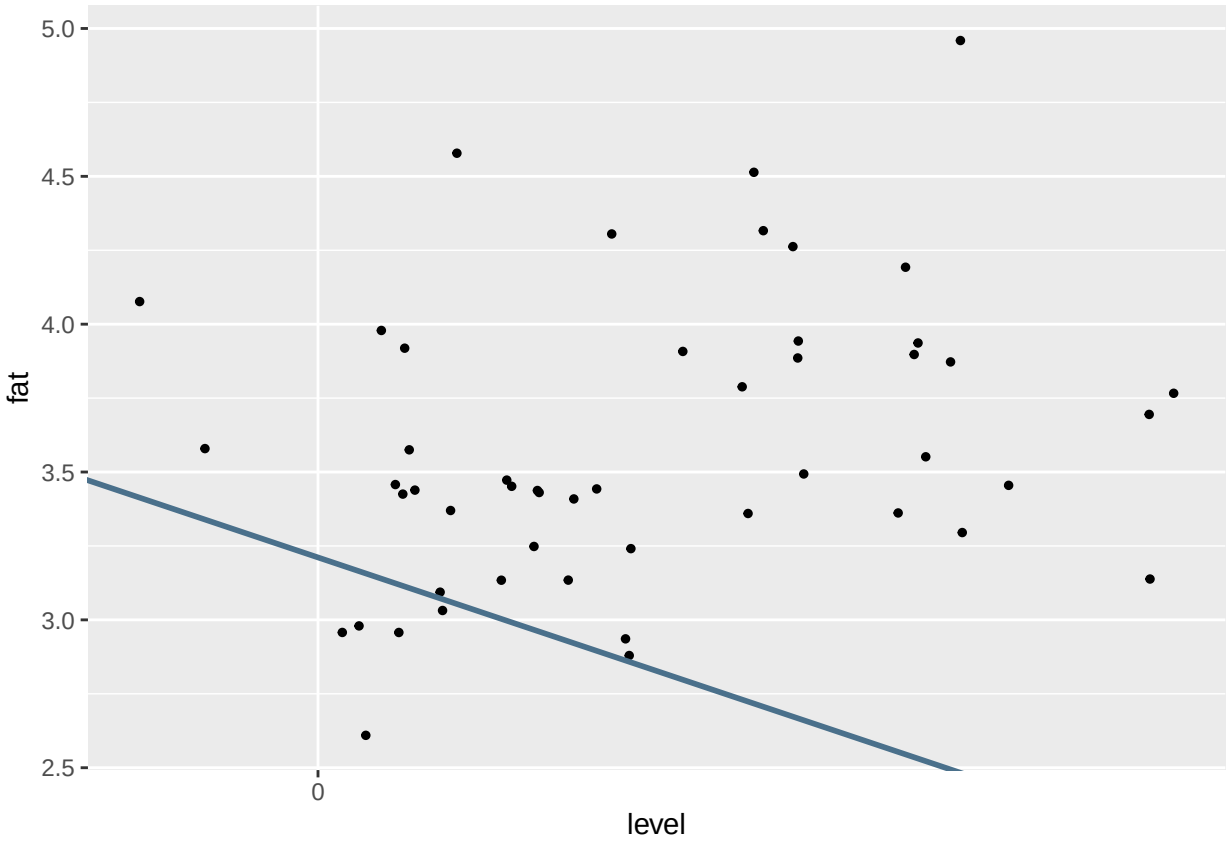
Working through your own example: Continuing the example from the final exercises of the earlier chapters, estimate the causal effect you defined in Exercise 18.18 using a regression of y on the treatment indicator z, at least one pre-treatment predictor x, and their interaction. Plot the data and the fitted regression lines for treatment and control, and discuss the assumptions that are required for this to be a good estimate of the causal effect of interest.

```
data <- read.csv("cows.csv")

fit_3 <- stan_glm(fat~level*initial.weight, data = data, refresh = 0)

print(fit_3)
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      fat ~ level * initial.weight
##  observations: 50
##  predictors:   4
## ------
##                       Median MAD_SD
## (Intercept)            3.2    0.7
## level                 -2.5    3.5
## initial.weight         0.0    0.0
## level:initial.weight   0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.4    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

I could not easily find data relating to what I wrote for 18.18, so I decided to use the cows dataset from the previous homework. We regressed the level of treatment onto the amount of milk fat produced, and also included the interaction of the cow's initial weight with the treatment level. We are seeking to determine if the treatment level has an effect on the amount of milk produced, and we assume that all cows were assigned to groups randomly. Since we are including an interaction term, we are assuming that some of the relationship between the treatment and the outcome variables depends on the value of the interaction term.

```
base <- ggplot(data, aes(x = level, y = fat)) +
  geom_point(size = 1, position = position_jitter(height = 0.05, width = 0.1)) +
  scale_x_continuous(breaks = c(0,1))

base + geom_abline(intercept = coef(fit_3)[1], slope = coef(fit_3)[2],
                   color = "skyblue4", size = 1)
```



This model appears to fit the data poorly. The treatment effect was -2.5, and the both the initial weight and the interaction term had coefficients of 0.