HW2

JLaw

2022-08-21

```
library(rethinking)
library(tidyverse)
```

Q1

The weights listed below were recorded in the !Kung census, but heights were not recorded for these individuals. Provided predicted heights and 89% compatibility intervals for each of these individuals. That is fill in the table below, using model based predictions:

Individual	weight	expected height	89% Interval
1	45		
2	40		
3	65		
4	31		

```
data("Howell1")
d <- Howell1
d2 <- d %>% filter(age >= 18)

#Define Average Weight
xbar = mean(d2$weight)

#Fit Model
m4.3 <- quap(
  alist(
    height ~ dnorm(mu, sigma),
    mu <- a + b*(weight - xbar),
    a ~ dnorm(178, 20),
    b ~ dlnorm(0, 1),
    sigma ~ dunif(0, 50)
  ),
  data = d2
)

##Get Parameter samples


weights = c(45, 40, 65, 31)
sims <- sim(m4.3, data = data.frame(weight = weights))
```

```r
mu.mean <- apply(sims, 2, mean)
mu.ci <- t(apply(sims, 2, PI, prob = .89))

knitr::kable(
  tibble(
    Individual = c(1, 2, 3, 4),
    weight = c(45, 40, 65, 31),
    `expected height`=  mu.mean,
    `89% Interval` = paste("(", round(mu.ci[, 1], 2),  ", ", round(mu.ci[, 2], 2),")")
  )
)
```

Individual	weight	expected height	89% Interval
1	45	154.4959	(146.25 , 162.61)
2	40	150.2219	(142.34 , 158.11)
3	65	172.6512	(163.84 , 181.1)
4	31	141.8506	(133.51 , 149.78)

Q2

Modle the relationship between height (cm) and the natural logarithm of weight (log-kg); log(weight). Use the entire Howell1 data frame. All rows, adults and non-adults. Use any model type from Chapter 4 that you think useful. An ordinary linear regression, a polynomial, or a spline. I recommend a plan linear regression thought. Plot the posterior predictions against the raw data.

```r
xbar = log(mean(d$weight))

q2_model <- quap(
  alist(
    height ~ dnorm(mu, sigma),
    mu <- a + b*(log(weight) - xbar),
    a ~ dnorm(178, 20),
    b ~ dlnorm(0, 1),
    sigma ~ dunif(0, 50)
  ),
  data = d
)



weight_seq <- 1:60
mu_vals <- sim(q2_model, data = data.frame(weight = weight_seq))

mu_mean <- apply(mu_vals, 2, mean)
mu_ci <- apply(mu_vals, 2, PI, .89)

plot(d$weight, d$height, col = col.alpha(rangi2, .7))
lines(weight_seq, mu_mean)
shade(mu_ci, weight_seq)
```
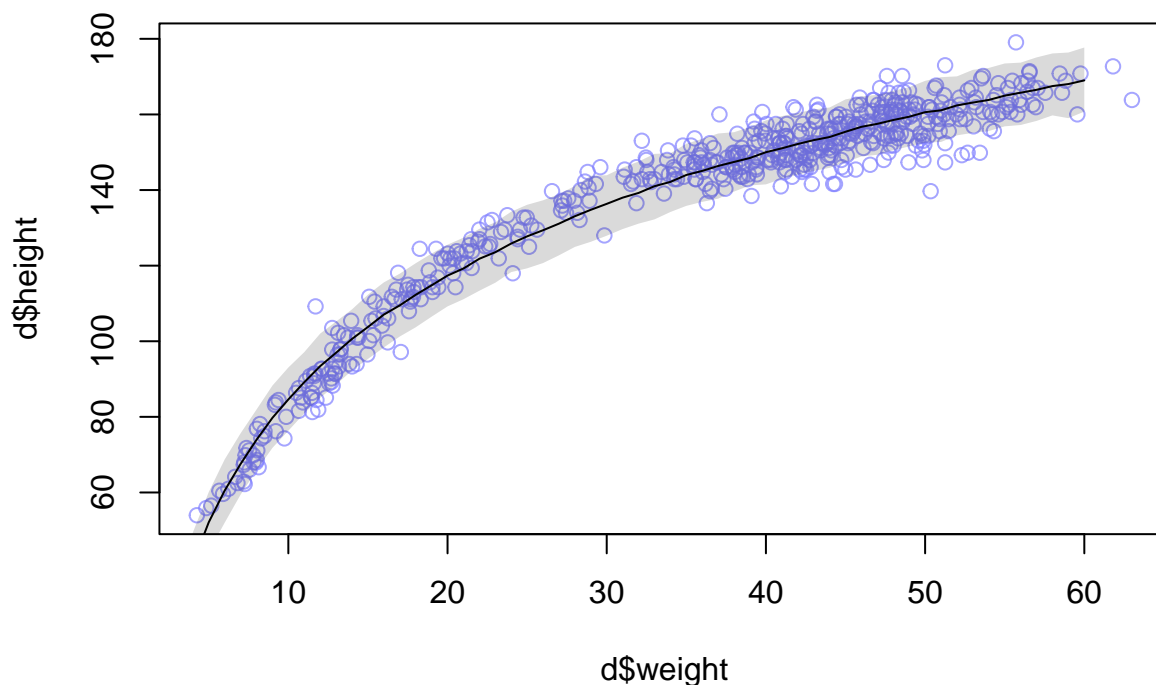
Q3

Plot the prior predictive distribution fo rthe polynomial model in Chapter 4. You can modify the code that plots the linear regression prior predictive distribution. 20 or 30 parabolas from the prior should suffice to show where the prior probability resides. Can you modify the distributions of alpha, Beta1, Beta2 so that the prior predictions stay within the biologically reasonable outcome space. That is to say, do not try to fit the data by hand. But do try to keep the curves consistent with what you know about height and weight before seeing the exact data.

```
d$weight_s <- (d$weight - mean(d$weight))/sd(d$weight)
d$weight_s2 <- d$weight_s^2

m4.5 <- quap(
  alist(
    height ~ dnorm(mu, sigma),
    mu <- a + b1 * weight_s + b2*weight_s2,
    a ~ dnorm(178, 20),
    b1 ~ dlnorm(0, 1),
    b2 ~ dnorm(0, 1),
    sigma ~ dunif(0, 50)
  ),
  data = d
)

set.seed(45)
```

```
priors <- extract.prior(m4.5)

w_seq <- seq(min(d$weight_s), max(d$weight_s), length.out = 50)
w2_seq <- w_seq^2

mu_sim <- link(m4.5, post = priors,
               data = list(weight_s = w_seq,
                           weight_s2 = w2_seq))

plot(NULL, xlim = range(w_seq), ylim = c(55, 270))
for(i in 1:50) lines(w_seq, mu_sim[i, ], col = col.alpha("black", .5))
```