Week 5: Bayesian linear regression and introduction to Stan

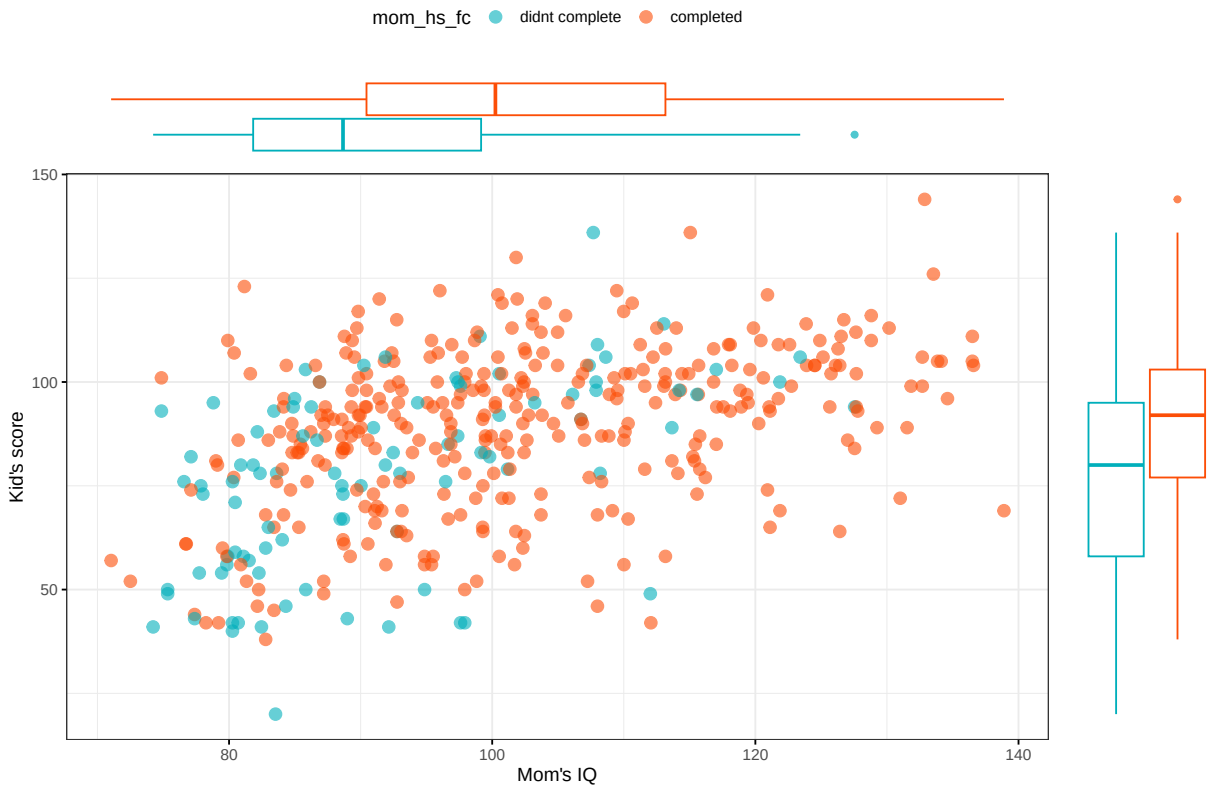
Quynh Vu

2023-02-13

Question 1: Use plots or tables to show three interesting observations about the data. Remember: Explain what your graph/ tables show & Choose a graph type that's appropriate to the data type

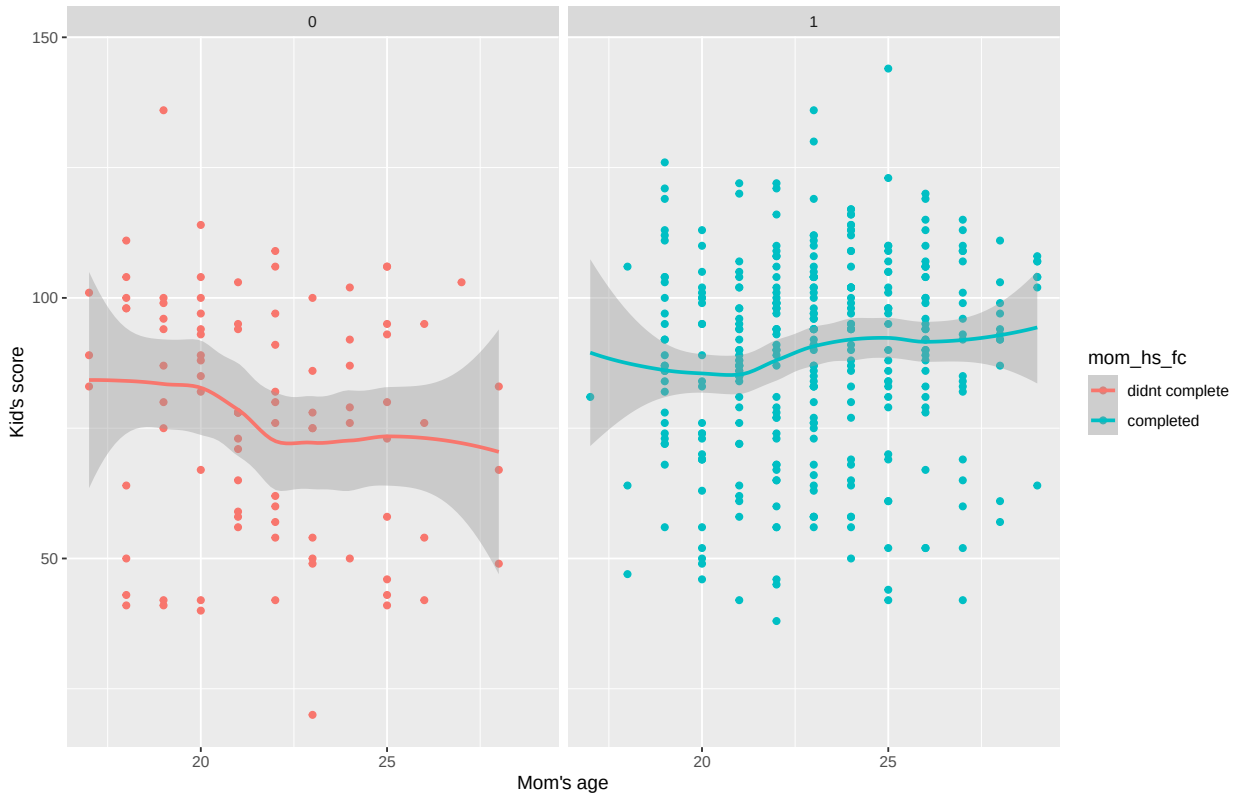
First, we noticed two interesting observations of kids' scores whose moms finished high school. In particular, one kid scored 123, above 75% of their peers, despite the mom having a below-average IQ of 81.2. In contrast, the other kid scored 69, below 25% of their peers, whereas the mom had the highest IQ of 138.89.

Plot 1: Mom's IQ and Kid's score



On the other hand, kids' scores appeared to decrease as the mom who didn't finish high school aged and seemed to increase for those whose moms finished high school, and the latter tended to perform better overall. However, one kid whose mom did not complete high school scored 136, higher than the average score of those whose mom completed high school, given that his/her mom had an above-average IQ of 108.

Plot 2: Kid's score and Mom's age



Three interesting observations about the data are

```
## # A tibble: 3 x 5
##   kid_score mom_hs mom_iq mom_age mom_hs_fc
##       <int>  <dbl>  <dbl>   <int> <fct>
## 1        98      1   89.4      25 completed
## 2        69      1  139.       20 completed
## 3       136      0  108.       19 didnt complete
```

Question 2: Change the prior to be much more informative (by changing the standard deviation to be 0.1). Rerun the model. Do the estimates change? Plot the prior and posterior densities.

The values of $\hat{R}$ are 1, suggesting the two chains have mixed well. However, the $N(80, 0.1^2)$ prior yields an estimate closer to the assigned mean μ of 80 and a slightly larger estimate for σ while requiring smaller effective sample sizes.

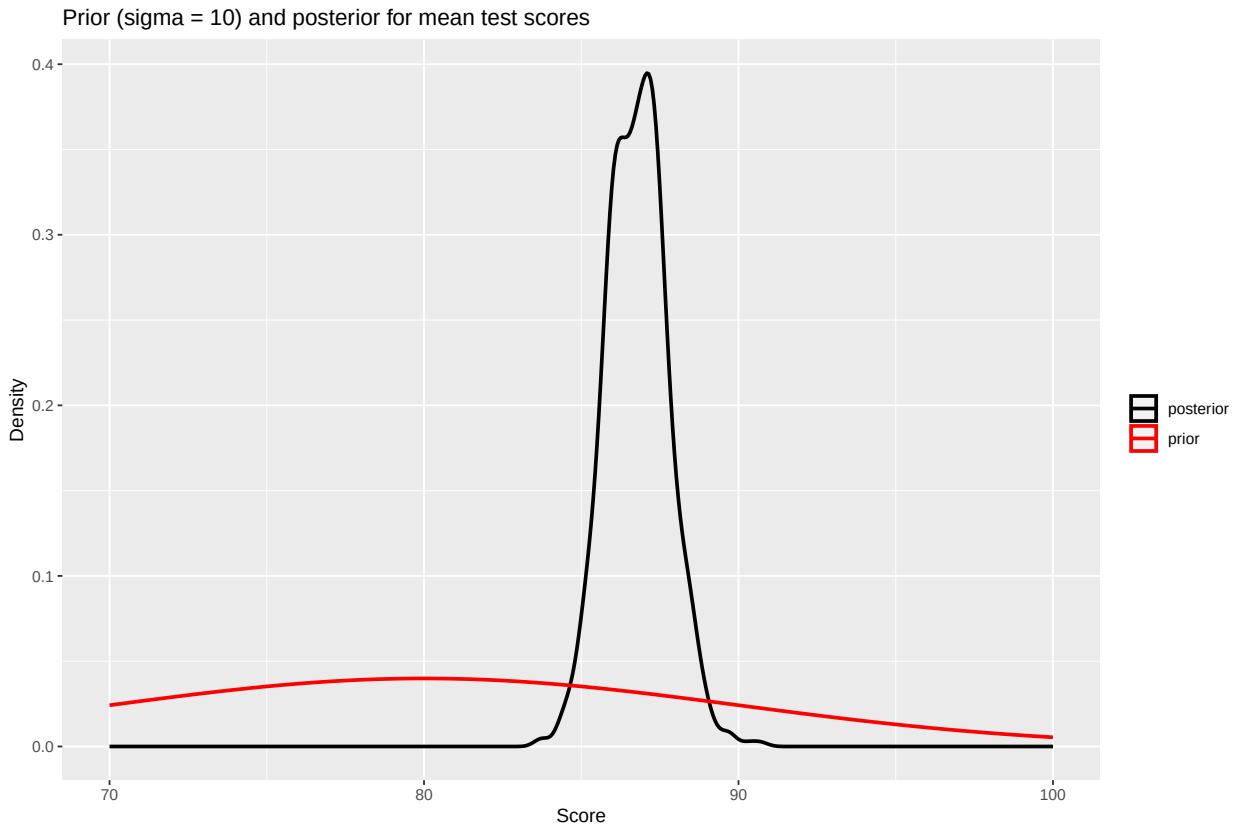
```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 5.8e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.58 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:   1 / 500 [  0%]  (Warmup)
```

```
## Chain 1: Iteration:  50 / 500 [ 10%]  (Warmup)
## Chain 1: Iteration: 100 / 500 [ 20%]  (Warmup)
## Chain 1: Iteration: 150 / 500 [ 30%]  (Warmup)
## Chain 1: Iteration: 200 / 500 [ 40%]  (Warmup)
## Chain 1: Iteration: 250 / 500 [ 50%]  (Warmup)
## Chain 1: Iteration: 251 / 500 [ 50%]  (Sampling)
## Chain 1: Iteration: 300 / 500 [ 60%]  (Sampling)
## Chain 1: Iteration: 350 / 500 [ 70%]  (Sampling)
## Chain 1: Iteration: 400 / 500 [ 80%]  (Sampling)
## Chain 1: Iteration: 450 / 500 [ 90%]  (Sampling)
## Chain 1: Iteration: 500 / 500 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.036 seconds (Warm-up)
## Chain 1:                0.013 seconds (Sampling)
## Chain 1:                0.049 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 1.2e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.12 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:   1 / 500 [  0%]  (Warmup)
## Chain 2: Iteration:  50 / 500 [ 10%]  (Warmup)
## Chain 2: Iteration: 100 / 500 [ 20%]  (Warmup)
## Chain 2: Iteration: 150 / 500 [ 30%]  (Warmup)
## Chain 2: Iteration: 200 / 500 [ 40%]  (Warmup)
## Chain 2: Iteration: 250 / 500 [ 50%]  (Warmup)
## Chain 2: Iteration: 251 / 500 [ 50%]  (Sampling)
## Chain 2: Iteration: 300 / 500 [ 60%]  (Sampling)
## Chain 2: Iteration: 350 / 500 [ 70%]  (Sampling)
## Chain 2: Iteration: 400 / 500 [ 80%]  (Sampling)
## Chain 2: Iteration: 450 / 500 [ 90%]  (Sampling)
## Chain 2: Iteration: 500 / 500 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 0.042 seconds (Warm-up)
## Chain 2:                0.016 seconds (Sampling)
## Chain 2:                0.058 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 1.4e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.14 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:   1 / 500 [  0%]  (Warmup)
## Chain 3: Iteration:  50 / 500 [ 10%]  (Warmup)
## Chain 3: Iteration: 100 / 500 [ 20%]  (Warmup)
## Chain 3: Iteration: 150 / 500 [ 30%]  (Warmup)
## Chain 3: Iteration: 200 / 500 [ 40%]  (Warmup)
```

```
## Chain 3: Iteration: 250 / 500 [ 50%]  (Warmup)
## Chain 3: Iteration: 251 / 500 [ 50%]  (Sampling)
## Chain 3: Iteration: 300 / 500 [ 60%]  (Sampling)
## Chain 3: Iteration: 350 / 500 [ 70%]  (Sampling)
## Chain 3: Iteration: 400 / 500 [ 80%]  (Sampling)
## Chain 3: Iteration: 450 / 500 [ 90%]  (Sampling)
## Chain 3: Iteration: 500 / 500 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 0.023 seconds (Warm-up)
## Chain 3:                0.012 seconds (Sampling)
## Chain 3:                0.035 seconds (Total)
## Chain 3:
```

```
## Inference for Stan model: anon_model.
## 3 chains, each with iter=500; warmup=250; thin=1;
## post-warmup draws per chain=250, total post-warmup draws=750.
##
##           mean se_mean   sd     2.5%      25%      50%      75%    97.5% n_eff
## mu       86.78    0.04 0.97    84.96    86.08    86.78    87.41    88.69   698
## sigma    20.39    0.03 0.68    19.14    19.91    20.42    20.86    21.70   392
## lp__  -1525.76    0.07 1.03 -1528.52 -1526.15 -1525.43 -1525.05 -1524.78   228
##       Rhat
## mu    1.00
## sigma 1.00
## lp__  1.02
##
## Samples were drawn using NUTS(diag_e) at Mon Feb 13 04:22:55 2023.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

```
## # A tibble: 2 x 7
##   .variable .value .lower .upper .width .point .interval
##   <chr>      <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>
## 1 mu          86.8   85.6   88.0    0.8 median qi
## 2 sigma       20.4   19.4   21.3    0.8 median qi
```



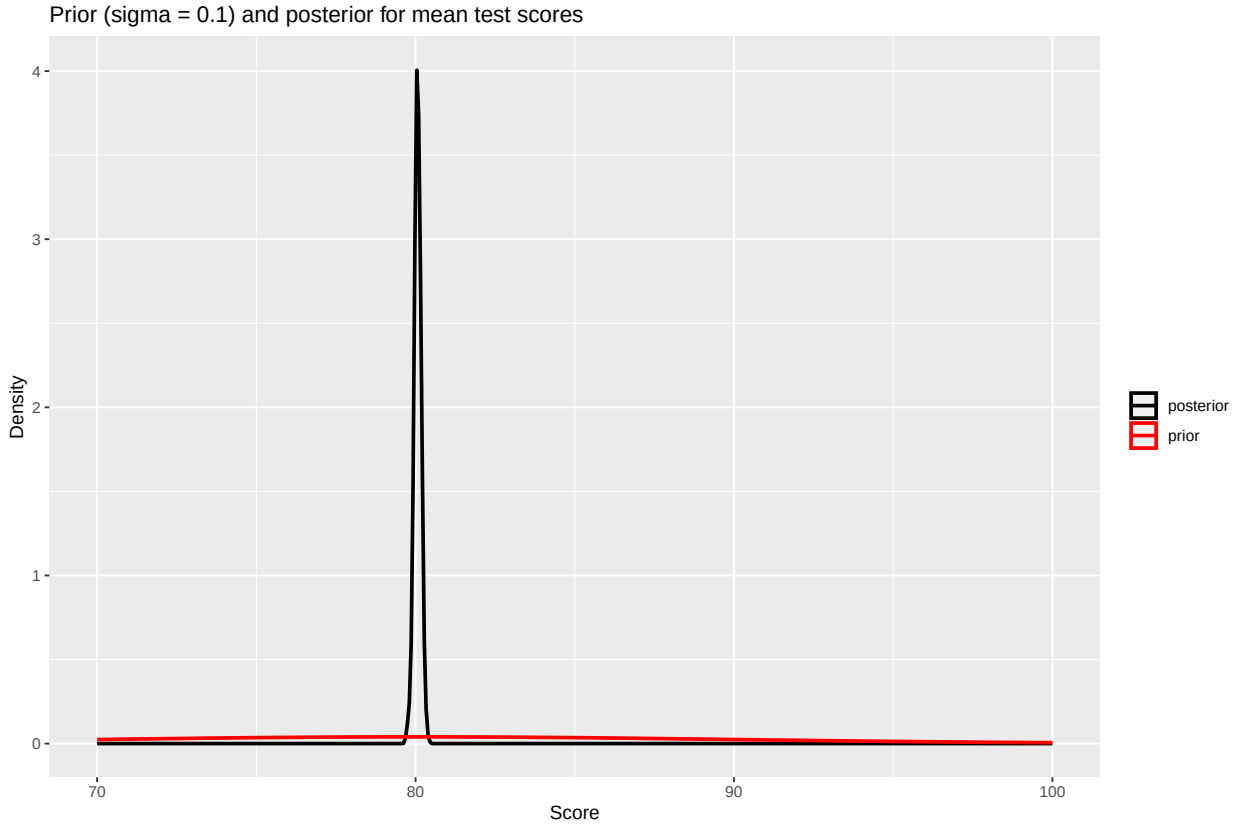

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 7e-06 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.07 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:   1 / 500 [  0%]  (Warmup)
## Chain 1: Iteration:  50 / 500 [ 10%]  (Warmup)
## Chain 1: Iteration: 100 / 500 [ 20%]  (Warmup)
## Chain 1: Iteration: 150 / 500 [ 30%]  (Warmup)
## Chain 1: Iteration: 200 / 500 [ 40%]  (Warmup)
## Chain 1: Iteration: 250 / 500 [ 50%]  (Warmup)
## Chain 1: Iteration: 251 / 500 [ 50%]  (Sampling)
## Chain 1: Iteration: 300 / 500 [ 60%]  (Sampling)
## Chain 1: Iteration: 350 / 500 [ 70%]  (Sampling)
## Chain 1: Iteration: 400 / 500 [ 80%]  (Sampling)
## Chain 1: Iteration: 450 / 500 [ 90%]  (Sampling)
## Chain 1: Iteration: 500 / 500 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.009 seconds (Warm-up)
## Chain 1:                0.009 seconds (Sampling)
## Chain 1:                0.018 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
```

```
## Chain 2:
## Chain 2: Gradient evaluation took 7e-06 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.07 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:   1 / 500 [  0%]  (Warmup)
## Chain 2: Iteration:  50 / 500 [ 10%]  (Warmup)
## Chain 2: Iteration: 100 / 500 [ 20%]  (Warmup)
## Chain 2: Iteration: 150 / 500 [ 30%]  (Warmup)
## Chain 2: Iteration: 200 / 500 [ 40%]  (Warmup)
## Chain 2: Iteration: 250 / 500 [ 50%]  (Warmup)
## Chain 2: Iteration: 251 / 500 [ 50%]  (Sampling)
## Chain 2: Iteration: 300 / 500 [ 60%]  (Sampling)
## Chain 2: Iteration: 350 / 500 [ 70%]  (Sampling)
## Chain 2: Iteration: 400 / 500 [ 80%]  (Sampling)
## Chain 2: Iteration: 450 / 500 [ 90%]  (Sampling)
## Chain 2: Iteration: 500 / 500 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 0.008 seconds (Warm-up)
## Chain 2:                0.007 seconds (Sampling)
## Chain 2:                0.015 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 7e-06 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.07 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:   1 / 500 [  0%]  (Warmup)
## Chain 3: Iteration:  50 / 500 [ 10%]  (Warmup)
## Chain 3: Iteration: 100 / 500 [ 20%]  (Warmup)
## Chain 3: Iteration: 150 / 500 [ 30%]  (Warmup)
## Chain 3: Iteration: 200 / 500 [ 40%]  (Warmup)
## Chain 3: Iteration: 250 / 500 [ 50%]  (Warmup)
## Chain 3: Iteration: 251 / 500 [ 50%]  (Sampling)
## Chain 3: Iteration: 300 / 500 [ 60%]  (Sampling)
## Chain 3: Iteration: 350 / 500 [ 70%]  (Sampling)
## Chain 3: Iteration: 400 / 500 [ 80%]  (Sampling)
## Chain 3: Iteration: 450 / 500 [ 90%]  (Sampling)
## Chain 3: Iteration: 500 / 500 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 0.011 seconds (Warm-up)
## Chain 3:                0.008 seconds (Sampling)
## Chain 3:                0.019 seconds (Total)
## Chain 3:
```

```
## Inference for Stan model: anon_model.
## 3 chains, each with iter=500; warmup=250; thin=1;
## post-warmup draws per chain=250, total post-warmup draws=750.
##
##           mean se_mean   sd     2.5%      25%      50%      75%    97.5% n_eff
```

```
## mu       80.07    0.00 0.10    79.86    80.00    80.06    80.13    80.26   595
## sigma    21.39    0.03 0.73    20.08    20.87    21.36    21.89    22.85   450
## lp__  -1548.41    0.07 1.06 -1551.45 -1548.74 -1548.07 -1547.63 -1547.40   225
##       Rhat
## mu    1.01
## sigma 1.00
## lp__  1.00
##
## Samples were drawn using NUTS(diag_e) at Mon Feb 13 04:22:58 2023.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

```
## # A tibble: 2 x 7
##   .variable .value .lower .upper .width .point .interval
##   <chr>      <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>
## 1 mu          80.1   79.9   80.2    0.8 median qi
## 2 sigma       21.4   20.5   22.3    0.8 median qi
```



Question 3 $Score = \alpha + \beta X$ where $X = 1$ if the mother finished high school and zero otherwise.

a) Confirm that the estimates of the intercept and slope are comparable to results from `lm()`

Simple linear regression:

```
##
## Call:
## lm(formula = kid_score ~ mom_iq, data = kidiq)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -56.753 -12.074   2.217  11.710  47.691
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 25.79978    5.91741    4.36 1.63e-05 ***
## mom_iq       0.60997    0.05852   10.42  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 18.27 on 432 degrees of freedom
## Multiple R-squared:  0.201,  Adjusted R-squared:  0.1991
## F-statistic: 108.6 on 1 and 432 DF,  p-value: < 2.2e-16
```

Simple Bayesian regression: $Score|\alpha, \beta, \sigma \sim N(\alpha + \beta X, \sigma^2)$

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0.000105 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 1.05 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 1: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 1: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 1: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 1: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 1: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 1: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 1: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 1: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 1: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 1: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 1: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.21 seconds (Warm-up)
## Chain 1:                0.116 seconds (Sampling)
## Chain 1:                0.326 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 1.9e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.19 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:   1 / 1000 [  0%]  (Warmup)
```

```
## Chain 2: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 2: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 2: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 2: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 2: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 2: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 2: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 2: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 2: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 2: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 2: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 0.219 seconds (Warm-up)
## Chain 2:                0.083 seconds (Sampling)
## Chain 2:                0.302 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 2.3e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.23 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 3: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 3: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 3: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 3: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 3: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 3: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 3: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 3: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 3: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 3: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 3: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 0.196 seconds (Warm-up)
## Chain 3:                0.087 seconds (Sampling)
## Chain 3:                0.283 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 4).
## Chain 4:
## Chain 4: Gradient evaluation took 1.6e-05 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.16 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 4: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 4: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 4: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 4: Iteration: 400 / 1000 [ 40%]  (Warmup)
```

```
## Chain 4: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 4: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 4: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 4: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 4: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 4: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 4: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 0.144 seconds (Warm-up)
## Chain 4:                0.087 seconds (Sampling)
## Chain 4:                0.231 seconds (Total)
## Chain 4:
```

```
##             mean    se_mean       sd      2.5%       25%      50%      75%
## alpha   78.04012 0.07178254 2.001407 74.058052 76.682423 78.02434 79.39570
## beta[1] 11.15604 0.08126388 2.253988  6.541375  9.664472 11.17659 12.70274
##            97.5%    n_eff     Rhat
## alpha   82.20658 777.3801 1.000981
## beta[1] 15.55085 769.3216 1.002132
```

The Bayesian estimates are analogous to the linear regression estimates using standard non-informative or weakly informative prior (as in Week 5 Lecture note slide 16-17), which is not the case here. However, we notice that the mean of the fitted values from the `lm()` model is analogous to the mean scores in Bayesian regression (≈ 86.79), and the fitted values from the two approaches do not differ greatly, i.e. the estimates of the intercept and slope are comparable to results from `lm()`.

```
## [1] 86.79724
```

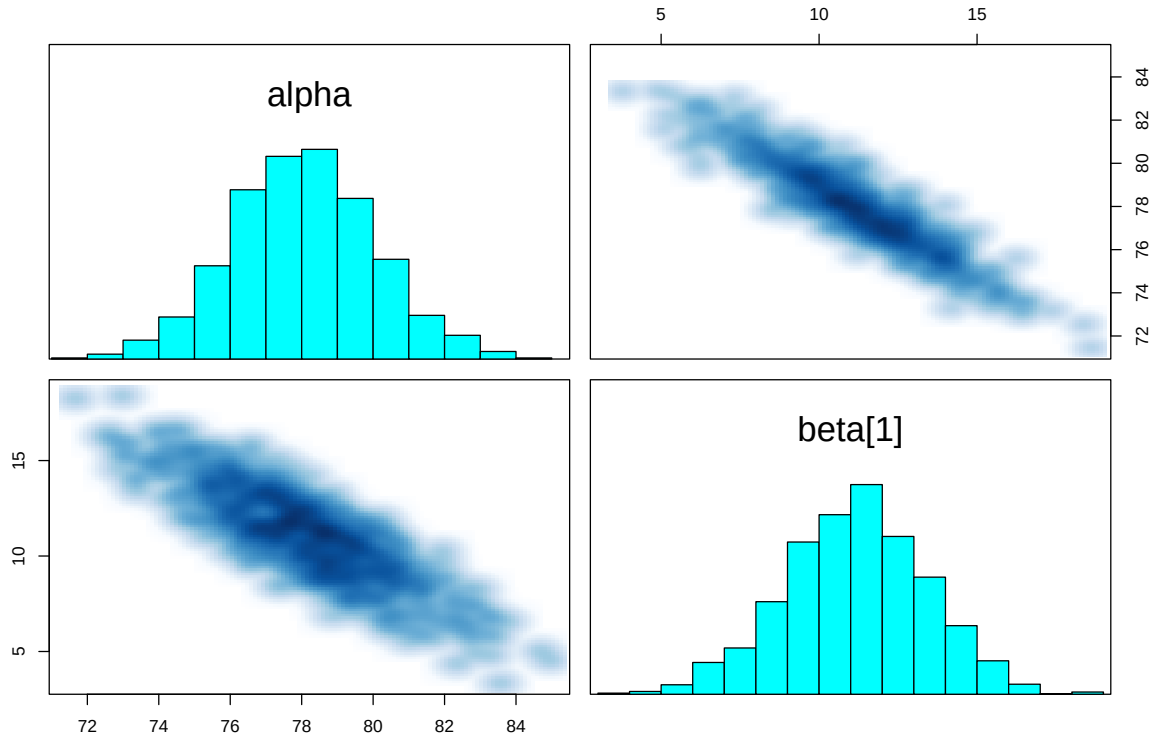
```
## [1] 86.79393
```

```
##   mom_hs true_score fitted_lm fitted_bayes
## 1      1         65  99.67839     89.20647
## 2      1         98  80.30825     89.20647
## 3      1         85  96.21717     89.20647
## 4      1         83  86.46153     89.20647
## 5      1        115  82.37230     89.20647
## 6      0         98  91.61716     77.94793
```

```
##     mom_hs true_score fitted_lm fitted_bayes
## 429      0         93  71.46292     77.94793
## 430      0         94  77.57284     77.94793
## 431      1         76  82.52155     89.20647
## 432      0         50  83.66179     77.94793
## 433      1         88  84.87986     89.20647
## 434      1         70  81.46199     89.20647
```

b) Do a `pairs` plot to investigate the joint sample distributions of the slope and intercept. Comment briefly on what you see. Is this potentially a problem?

We see that α and β don't look reasonably centred, which may induce the opposite change in the intercept and make it hard to interpret the intercepts and hard to sample as the chain converges fast to stationarity. It is noteworthy that the effect of β is cancelled out when $X = 0$, resulting in the underestimation of scores of kids' whose moms didn't finish high school and the overestimation of their counterparts as indicated in tables above.



Question 4: Add in mother's IQ as a covariate and rerun the model. Please mean center the covariate before putting it into the model. Interpret the coefficient on the (centered) mum's IQ.

Multiple Bayesian regression: Score $= 82.349488 + 5.664857$`mom_hs` $+ 0.565145$ `centered_mom_iq`

The coefficient $\hat{\beta}_2 = 0.565145$ is the posterior mean $E(\beta_2|score)$, which is is the same as the OLS estimates of β_2 where

$$\begin{bmatrix} 1 & x_1 & x_2 \end{bmatrix} \begin{bmatrix} 82.349488 \\ 5.664857 \\ 0.565145 \end{bmatrix}$$

gives the expected scores for each kid ($x_1 =$ `mom_hs` and $x_2 =$ `centered_mom_iq`)

Interpretation: If we observe two kids whose mom both finished or didn't finish high school, 1 unit difference in `centered_mom_iq` would make the kid's score to be expected to differ by 0.565145 points.

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 2e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.2 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
```

```
## Chain 1: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 1: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 1: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 1: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 1: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 1: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 1: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 1: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 1: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 1: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 1: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 1: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.143 seconds (Warm-up)
## Chain 1:                0.102 seconds (Sampling)
## Chain 1:                0.245 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 1.8e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.18 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 2: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 2: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 2: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 2: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 2: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 2: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 2: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 2: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 2: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 2: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 2: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 0.177 seconds (Warm-up)
## Chain 2:                0.093 seconds (Sampling)
## Chain 2:                0.27 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 1.8e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.18 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 3: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 3: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 3: Iteration: 300 / 1000 [ 30%]  (Warmup)
```

```
## Chain 3: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 3: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 3: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 3: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 3: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 3: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 3: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 3: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 0.166 seconds (Warm-up)
## Chain 3:                0.106 seconds (Sampling)
## Chain 3:                0.272 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 4).
## Chain 4:
## Chain 4: Gradient evaluation took 2.3e-05 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.23 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:   1 / 1000 [  0%]  (Warmup)
## Chain 4: Iteration: 100 / 1000 [ 10%]  (Warmup)
## Chain 4: Iteration: 200 / 1000 [ 20%]  (Warmup)
## Chain 4: Iteration: 300 / 1000 [ 30%]  (Warmup)
## Chain 4: Iteration: 400 / 1000 [ 40%]  (Warmup)
## Chain 4: Iteration: 500 / 1000 [ 50%]  (Warmup)
## Chain 4: Iteration: 501 / 1000 [ 50%]  (Sampling)
## Chain 4: Iteration: 600 / 1000 [ 60%]  (Sampling)
## Chain 4: Iteration: 700 / 1000 [ 70%]  (Sampling)
## Chain 4: Iteration: 800 / 1000 [ 80%]  (Sampling)
## Chain 4: Iteration: 900 / 1000 [ 90%]  (Sampling)
## Chain 4: Iteration: 1000 / 1000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 0.155 seconds (Warm-up)
## Chain 4:                0.098 seconds (Sampling)
## Chain 4:                0.253 seconds (Total)
## Chain 4:
```

```
##               mean     se_mean         sd       2.5%        25%        50%
## alpha   82.2309093 0.068625815 1.93468978 78.4335014 80.9231525 82.2015530
## beta[1]  5.8019423 0.077292572 2.19351281  1.3746266  4.3338504  5.8139298
## beta[2]  0.5671571 0.001743236 0.05974439  0.4543341  0.5264749  0.5677017
##                75%      97.5%     n_eff     Rhat
## alpha   83.5137805 86.0535878  794.7813 1.002524
## beta[1]  7.2124467 10.1915073  805.3874 1.004055
## beta[2]  0.6051073  0.6866374 1174.5777 1.003463
```

Question 5: Confirm the results from Stan agree with `lm()`

Multiple linear regression: Score $= 82.12214 + 5.95012$ `mom_hs` $+ 0.56391$`centered_mom_iq`

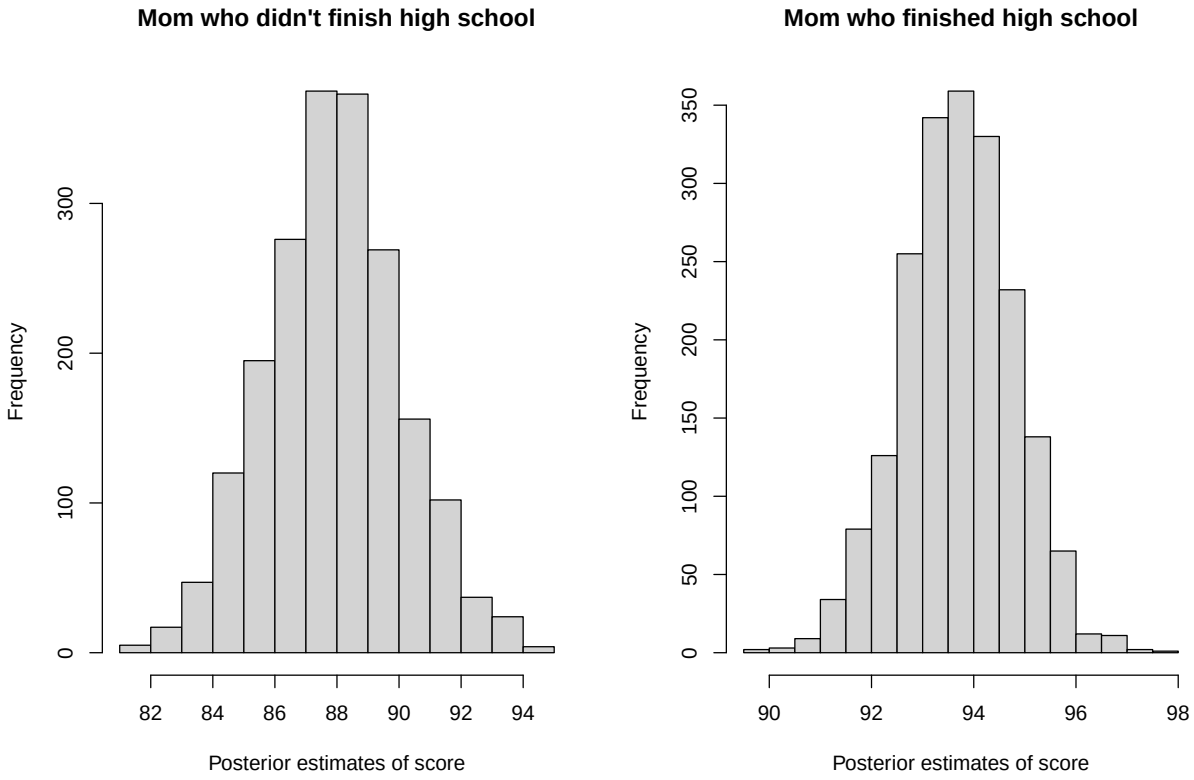
```
##
## Call:
```

```
## lm(formula = kid_score ~ mom_hs + centered_mom_iq, data = kidiq)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -52.873 -12.663   2.404  11.356  49.545
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     82.12214    1.94370  42.250  < 2e-16 ***
## mom_hs           5.95012    2.21181   2.690  0.00742 **
## centered_mom_iq  0.56391    0.06057   9.309  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 18.14 on 431 degrees of freedom
## Multiple R-squared:  0.2141, Adjusted R-squared:  0.2105
## F-statistic: 58.72 on 2 and 431 DF,  p-value: < 2.2e-16
```

Question 6: Plot the posterior estimates of scores by education of mother for mothers who have an IQ of 110.

```
x_new <- 110
post_samples <- extract(fit3)
alpha_hat <- post_samples[["alpha"]]
beta1_hat <- post_samples[["beta"]][,1]
beta2_hat <- post_samples[["beta"]][,2]
lin_pred0 <- alpha_hat + beta1_hat*0 + beta2_hat*(x_new - mean(kidiq$mom_iq))
lin_pred1 <- alpha_hat + beta1_hat*1 + beta2_hat*(x_new - mean(kidiq$mom_iq))

par(mfrow = c(1, 2))
hist(lin_pred0, xlab = "Posterior estimates of score", main = "Mom who didn't finish high school")
hist(lin_pred1, xlab = "Posterior estimates of score", main = "Mom who finished high school")
```

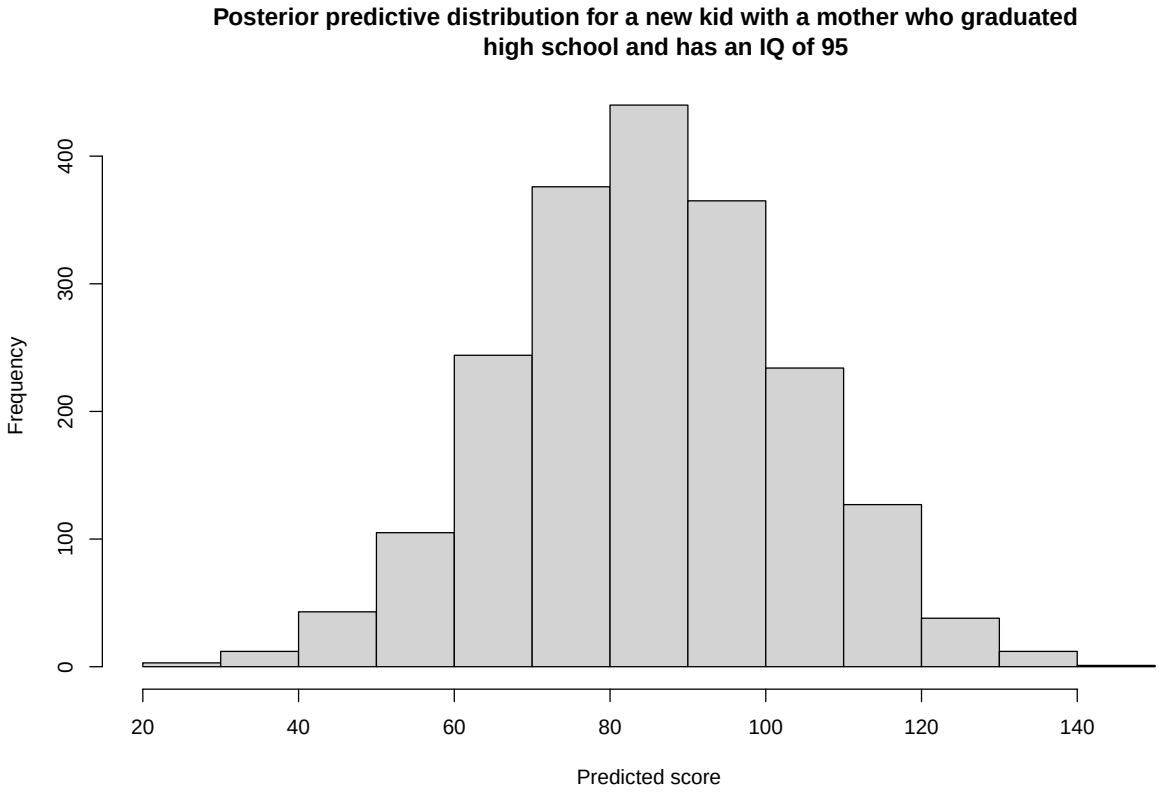


Question 7: Generate and plot (as a histogram) samples from the posterior predictive distribution for a new kid with a mother who graduated high school and has an IQ of 95.

```
sigma <- post_samples[["sigma"]]
lin_pred <- alpha_hat + beta1_hat*1 + beta2_hat*(95 - mean(kidiq$mom_iq))
y_new <- rnorm(n = length(sigma),mean = lin_pred, sd = sigma)
y_new[1:20]
```

```
##  [1]  88.83280  73.54334  82.43068  95.81967  88.97460  84.18608  98.57978
##  [8]  78.50909 111.86562 112.63430  89.03989  77.36989 112.45669  83.53003
## [15]  68.26775  81.46076  57.46071  98.01110  96.09713  98.14854
```

```
hist(y_new, main = "Posterior predictive distribution for a new kid with a mother who graduated
     high school and has an IQ of 95", xlab = "Predicted score")
```


Posterior predictive distribution for a new kid with a mother who graduated high school and has an IQ of 95
Frequency
0
100
200
300
400
20
40
60
80
100
120
140
Predicted score