Biostat 200C Homework 4

Due 11:59PM May 23rd

Jiahao Tian

2. Calculate the three estimates for the `pulp` example in class, check if your results match with the R output.

Solution:

```
data(pulp)
help(pulp)
pulp <- as_tibble(pulp) %>%
  print(n = Inf)
```

```
## # A tibble: 20 x 2
##    bright operator
##     <dbl> <fct>
##  1   59.8 a
##  2   60   a
##  3   60.8 a
##  4   60.8 a
##  5   59.8 a
##  6   59.8 b
##  7   60.2 b
##  8   60.4 b
##  9   59.9 b
## 10   60   b
## 11   60.7 c
## 12   60.7 c
## 13   60.5 c
## 14   60.9 c
## 15   60.3 c
## 16   61   d
## 17   60.8 d
## 18   60.6 d
## 19   60.5 d
## 20   60.5 d
```

```
#mu hat
mean(pulp$bright)
```

```
## [1] 60.4
```

```
aovmod <- aov(bright ~ operator, data = pulp) %>%
  summary()

#sigma (alpha) hat
(aovmod[1][[1]][[3]][1] - aovmod[1][[1]][[3]][2]) / 5
```

```
## [1] 0.06808333
```

```
#sigma (epsilon) hat
aovmod[1][[1]][[3]][2]
```

```
## [1] 0.10625
```

Q2. ELMR Exercise 11.1 (p251)

The `ratdrink` data consist of 5 weekly measurements of body weight for 27 rats. The first 10 rats are on a control treatment while 7 rats have thyroxine added to their drinking water and 10 rats have thiouracil added to their water.

```
help("ratdrink")
```

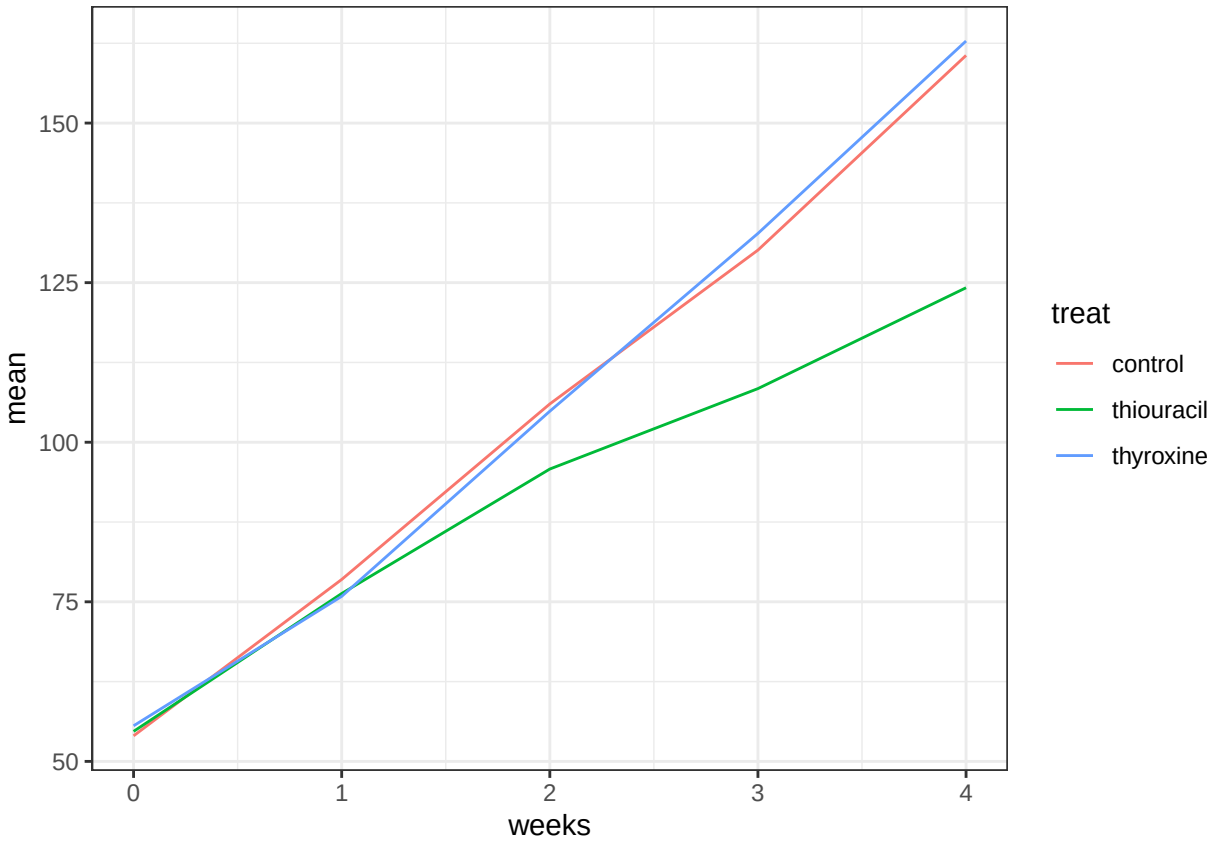
1. Plot the data showing how weight increases with age on a single panel, taking care to distinguish the three treatment groups. Now create a three-panel plot, one for each group. Discuss what can be seen.

Solution:

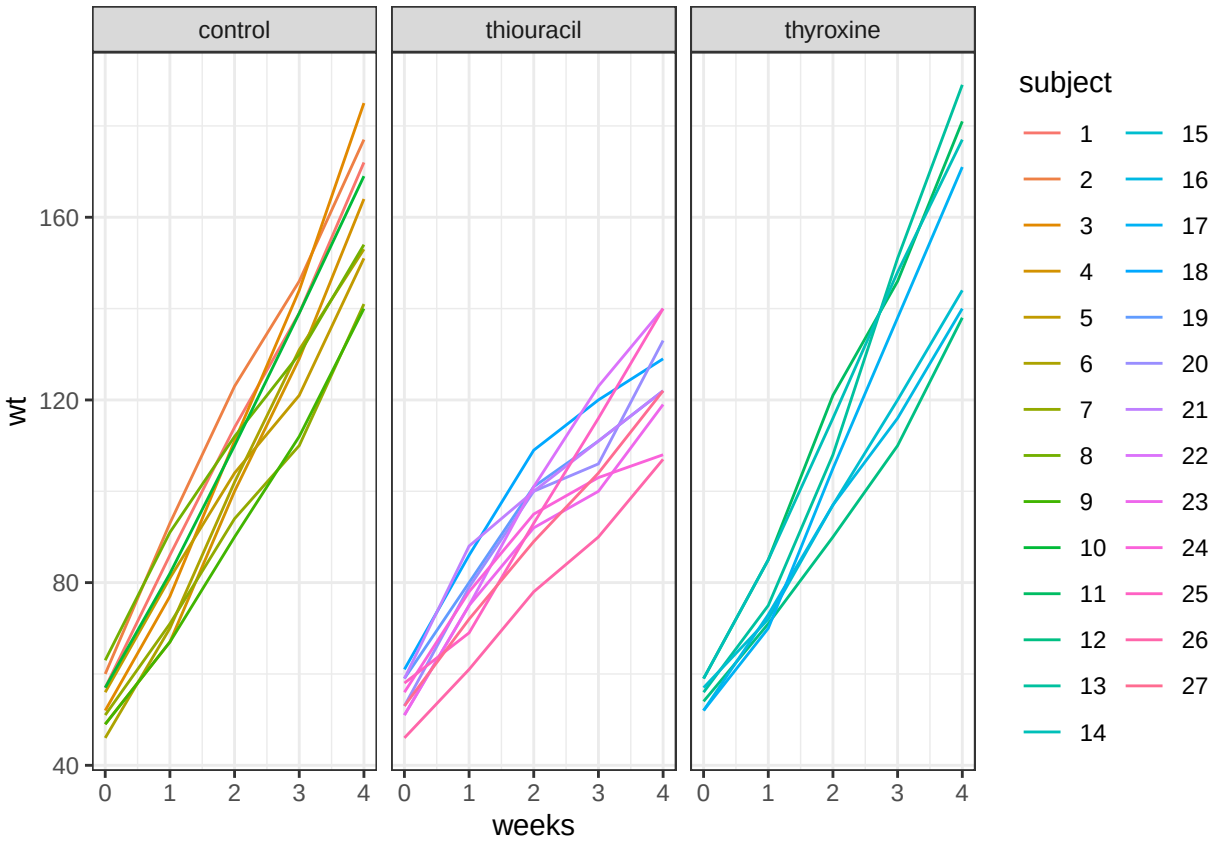
```
data("ratdrink")
ratdrink <- as_tibble(ratdrink)
```

```
ratdrink %>%
  group_by(weeks, treat) %>%
  summarise(mean = mean(wt)) %>%
  ggplot() +
  geom_line(mapping = aes(x = weeks, y = mean, color = treat)) +
  theme_bw()
```

```
## `summarise()` has grouped output by 'weeks'. You can override using the
## `.groups` argument.
```




```
ratdrink %>%
  ggplot() +
  geom_line(mapping = aes(weeks, wt, group = subject, color=subject)) +
  facet_wrap(~treat) +
  theme_bw()
```



From the plot, in the control group, the rats' weight increases follow when the weeks increase. In the thiouracil group, rats' weights do not increase that much compared to the other two groups. In the thyroxine group, rats' weights were increased as in the control group, but there is a gap.

2. Fit a linear longitudinal model with a random slope and intercept for each rat. Each treatment group should have a different mean line. Give interpretation for the following estimates:

- The fixed effect intercept term.
- The interaction between thiouracil and week.
- The intercept random effect SD (standard deviation).

```
summary(ratdrink)
```

```
##        wt             weeks      subject           treat
##  Min.   : 46.0   Min.   :0   1      :  5   control   :50
##  1st Qu.: 71.0   1st Qu.:1   2      :  5   thiouracil:50
##  Median :100.0   Median :2   3      :  5   thyroxine :35
##  Mean   :100.8   Mean   :2   4      :  5
##  3rd Qu.:122.5   3rd Qu.:3   5      :  5
##  Max.   :189.0   Max.   :4   6      :  5
##                              (Other):105
```

Solution:

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
library(Matrix)
mmod <- lmer(wt ~ treat * weeks + (1 + weeks | subject), data = ratdrink)
summary(mmod)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: wt ~ treat * weeks + (1 + weeks | subject)
##    Data: ratdrink
##
## REML criterion at convergence: 878.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.83136 -0.54991  0.04003  0.58230  2.03660
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  subject  (Intercept) 32.49    5.700
##           weeks       14.14    3.760    -0.13
##  Residual             18.90    4.348
## Number of obs: 135, groups:  subject, 27
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)            52.8800     2.0937  25.256
## treatthiouracil         4.7800     2.9610   1.614
## treatthyroxine         -0.7943     3.2628  -0.243
## weeks                  26.4800     1.2661  20.915
## treatthiouracil:weeks  -9.3700     1.7905  -5.233
## treatthyroxine:weeks    0.6629     1.9730   0.336
##
## Correlation of Fixed Effects:
##             (Intr) trtthr trtthy weeks  trtthr:
## treatthircl -0.707
## treatthyrxn -0.642  0.454
## weeks       -0.250  0.177  0.160
## trtthrcl:wk  0.177 -0.250 -0.113 -0.707
## trtthyrxn:w  0.160 -0.113 -0.250 -0.642  0.454
```

Interpretation: 1. The average weight of a rat at week 0 is 52.88 among all treatment groups. 2. As the average weight increases each week, the average weight in thiouracil group is 9.39 lower than the control group. 3. The average weight for each individul at week 0 has a standard deviation of 5.7 within group.

3. Check whether there is a significant treatment effect.

Solution:

```
library(pbkrtest)
mmod2 <- lmer(wt ~ weeks + (1 + weeks | subject), data = ratdrink, REML = TRUE)
KRmodcomp(mmod, mmod2)
```

```
## large : wt ~ treat * weeks + (1 + weeks | subject)
## small : wt ~ weeks + (1 + weeks | subject)
##          stat     ndf     ddf F.scaling   p.value
## Ftest  8.7125  4.0000 26.8141   0.94552 0.0001215 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the Kenward-Roger approach F-test results, shows that the treatment has a significant effect p- value $< .001$.

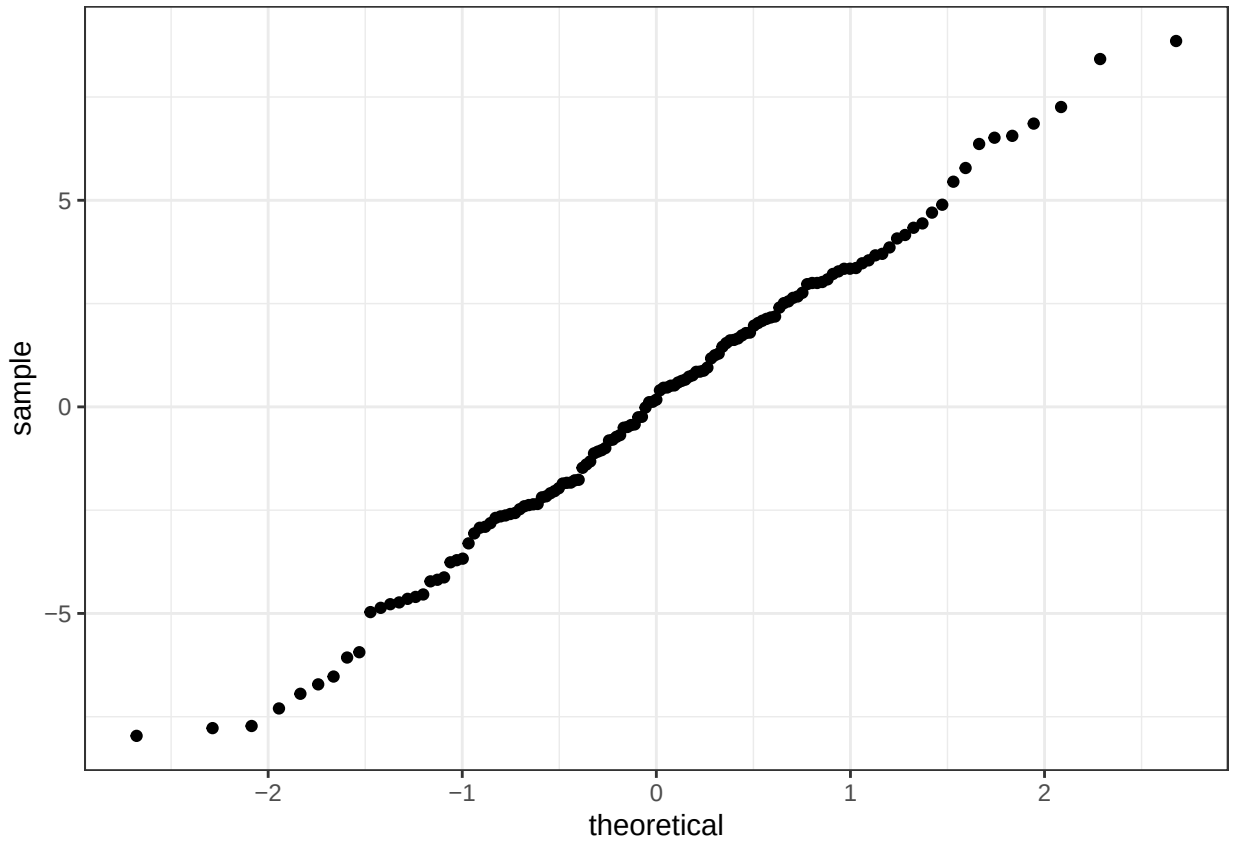
4. Construct diagnostic plots showing the residuals against the fitted values and a QQ plot of the residuals. Comment on the plots.

Solution:

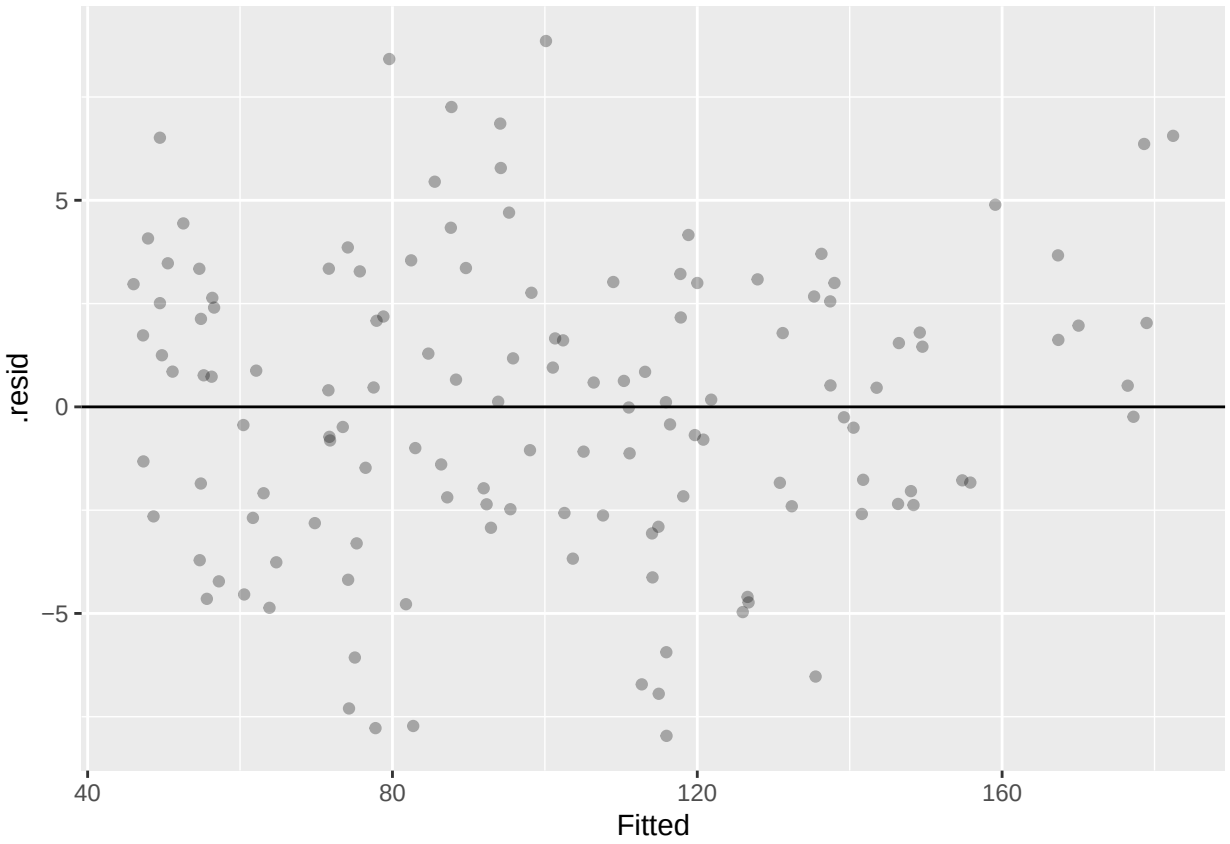
```
library(broom.mixed)
(diagd <- augment(mmod))
```

```
## # A tibble: 135 x 15
##       wt treat   weeks subject .fitted .resid  .hat .cooksd .fixed   .mu .offset
##    <dbl> <fct>   <int> <fct>     <dbl>  <dbl> <dbl>   <dbl>  <dbl> <dbl>   <dbl>
##  1    57 control     0 1          56.3  0.730 0.447 0.00686   52.9  56.3       0
##  2    86 control     1 1          84.7  1.29  0.236 0.00593   79.4  84.7       0
##  3   114 control     2 1         113.   0.848 0.183 0.00175  106.  113.        0
##  4   139 control     3 1         142.  -2.59  0.289 0.0338   132.  142.        0
##  5   172 control     4 1         170.   1.97  0.552 0.0937   159.  170.        0
##  6    60 control     0 2          60.4 -0.440 0.447 0.00249   52.9  60.4       0
##  7    93 control     1 2          89.6  3.36  0.236 0.0403    79.4  89.6       0
##  8   123 control     2 2         119.   4.16  0.183 0.0420   106.  119.        0
##  9   146 control     3 2         148.  -2.04  0.289 0.0209   132.  148.        0
## 10   177 control     4 2         177.  -0.240 0.552 0.00139  159.  177.        0
## # ... with 125 more rows, and 4 more variables: .sqrtXwt <dbl>, .sqrtrwt <dbl>,
## #   .weights <dbl>, .wtres <dbl>
```

```
diagd %>%
  ggplot(mapping = aes(sample = .resid)) +
  stat_qq() +
  theme_bw()
```




```
diagd %>%
  ggplot() +
  geom_point(mapping = aes(x = .fitted, y = .resid), alpha = 0.3) +
  geom_hline(yintercept = 0) +
  labs(x = "Fitted", ylab = "Residuals")
```



From the plots above, the residuals follow a normal distribution, and there is no evidence to show that there are outliers. And from the QQ plot, it shows the linearity and no outliers, also close to the 45 degree. Overall, it is a good model.

5. Construct confidence intervals for the parameters of the model. Which random effect terms may not be significant? Is the thyroxine group significantly different from the control group?

Solution:

sig02 is a not significant term, becuase it contains 0. From CI of `weeks` and interaction term `treatthyroxine:weeks` which both contains 0, so thyroxine group is not significantly different from the control group.

```
confint(mmod, method = "boot")
```

```
## Computing bootstrap confidence intervals ...

##
## 1 message(s): boundary (singular) fit: see help('isSingular')
## 6 warning(s): Model failed to converge with max|grad| = 0.00269961 (tol = 0.002, component 1) (and o

##                            2.5 %     97.5 %
## .sig01                 3.4308367  7.9653919
## .sig02                -0.5768396  0.5068243
## .sig03                 2.6794266  4.8756118
## .sigma                 3.6884578  4.9656792
```

```
## (Intercept)            48.6931778 57.0247266
## treatthiouracil        -1.3163973 10.2186251
## treatthyroxine         -6.9991472  5.8847526
## weeks                  24.0220964 29.1229641
## treatthiouracil:weeks -13.2832733 -5.8043360
## treatthyroxine:weeks   -3.2981947  4.2118382
```