

STAT-S632

Assignment 7

John Koo

Problem 1

```
library(ggplot2)
import::from(magrittr, `>`)
library(lme4)
import::from(pbkrtest, KRmodcomp)

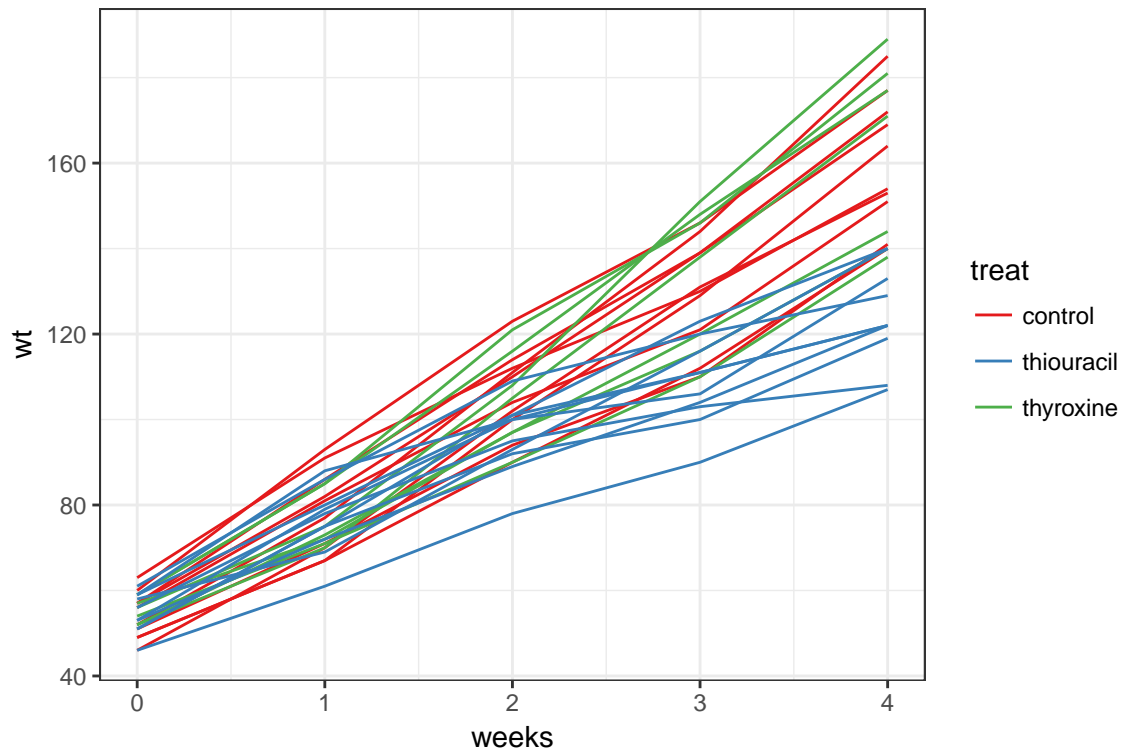
theme_set(theme_bw())

rat.test.df <- faraway::ratdrink
summary(rat.test.df)
```

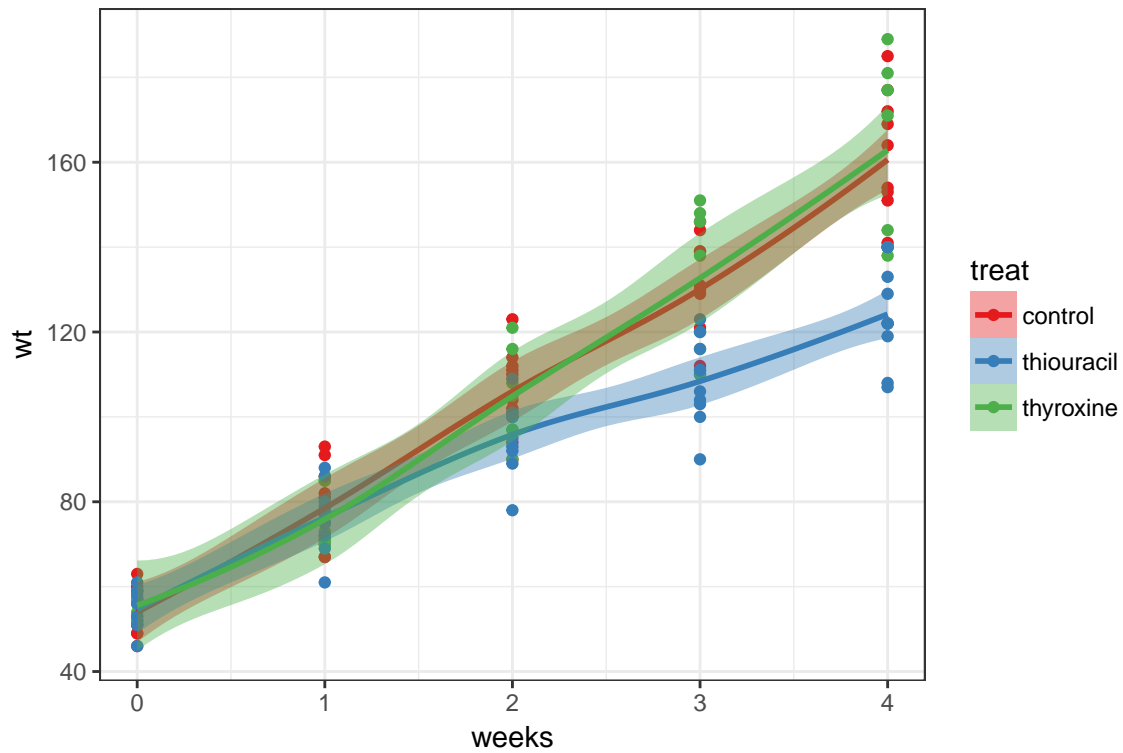
wt		weeks		subject		treat	
Min.	: 46.0	Min.	:0	1	: 5	control	:50
1st Qu.:	71.0	1st Qu.:	1	2	: 5	thiouracil:	50
Median	:100.0	Median	:2	3	: 5	thyroxine	:35
Mean	:100.8	Mean	:2	4	: 5		
3rd Qu.:	122.5	3rd Qu.:	3	5	: 5		
Max.	:189.0	Max.	:4	6	: 5		
						(Other):	105

Part a

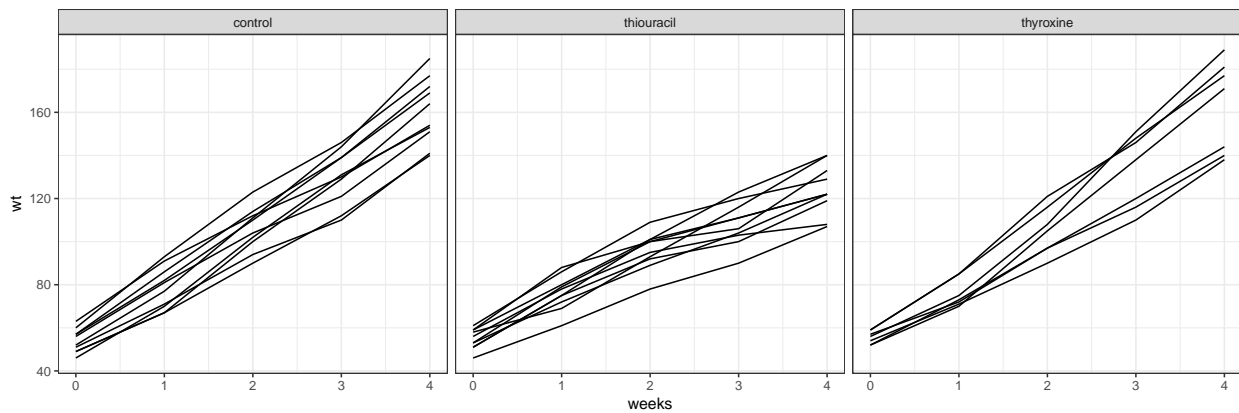
```
ggplot(rat.test.df) +
  geom_line(aes(x = weeks, y = wt, group = subject, colour = treat)) +
  scale_colour_brewer(palette = 'Set1')
```



```
ggplot(rat.test.df) +  
  geom_point(aes(x = weeks, y = wt, group = subject, colour = treat)) +  
  scale_colour_brewer(palette = 'Set1') +  
  scale_fill_brewer(palette = 'Set1') +  
  stat_smooth(aes(x = weeks, y = wt, colour = treat, fill = treat))
```



```
ggplot(rat.test.df) +
  geom_line(aes(x = weeks, y = wt, group = subject)) +
  facet_wrap(~treat)
```



Weight increases over time, but it appears that the rats treated with thlouracil gained less weight.

Part b

```
mixed.mod <- lmer(wt ~ weeks * treat + (weeks | subject), data = rat.test.df)
summary(mixed.mod)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: wt ~ weeks * treat + (weeks | subject)
Data: rat.test.df
```

REML criterion at convergence: 878.7

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.83136	-0.54991	0.04003	0.58231	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
weeks	26.4800	1.2661	20.915
treatthiouracil	4.7800	2.9610	1.614
treatthyroxine	-0.7943	3.2628	-0.243
weeks:treatthiouracil	-9.3700	1.7905	-5.233
weeks:treatthyroxine	0.6629	1.9730	0.336

Correlation of Fixed Effects:

	(Intr) weeks	trtthr	trtthy	wks:trtthr
weeks	-0.250			
treatthircl	-0.707	0.177		
treatthyrrn	-0.642	0.160	0.454	
wks:trtthrc	0.177	-0.707	-0.250	-0.113
wks:trtthyr	0.160	-0.642	-0.113	-0.250
				0.454

- Intercept - The average weight at week 0 for the baseline (control) group
- `weeks:thiouracil` - On average, each week the rats in the thiouracil group gained 9.4 units less in weight compared to the baseline (control) group.
- $\hat{\sigma}_{intercept}$ - Since this is greater than $\hat{\sigma}_{\epsilon}$, the subject effect is not negligible. We can also compute the ICC, which is ≈ 0.632 .

Part c

The D matrix has the form $diag(d, d, \dots, d)$ where each of the 27 (one for each subject) ds is a 2×2 variance-covariance matrix of the form $d = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$. So $D \in \mathbb{R}^{54 \times 54}$.

Part d

```
week.mod <- lmer(wt ~ weeks + (weeks | subject), data = rat.test.df)
KRmodcomp(mixed.mod, week.mod)
```

F-test with Kenward-Roger approximation; computing time: 0.76 sec.

large : wt ~ weeks * treat + (weeks | subject)

small : wt ~ weeks + (weeks | subject)

	stat	ndf	ddf	F.scaling	p.value
Ftest	8.7124	4.0000	26.8141	0.94552	0.0001215 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The treatment effect is significant, but we can also see whether the interaction term is.

```
no.inter.mod <- lmer(wt ~ weeks + treat + (weeks | subject),
  data = rat.test.df)
KRmodcomp(mixed.mod, no.inter.mod)
```

F-test with Kenward-Roger approximation; computing time: 0.56 sec.

```
large : wt ~ weeks * treat + (weeks | subject)
small : wt ~ weeks + treat + (weeks | subject)
      stat      ndf      ddf F.scaling  p.value
Ftest 18.319   2.000 24.000          1 1.478e-05 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

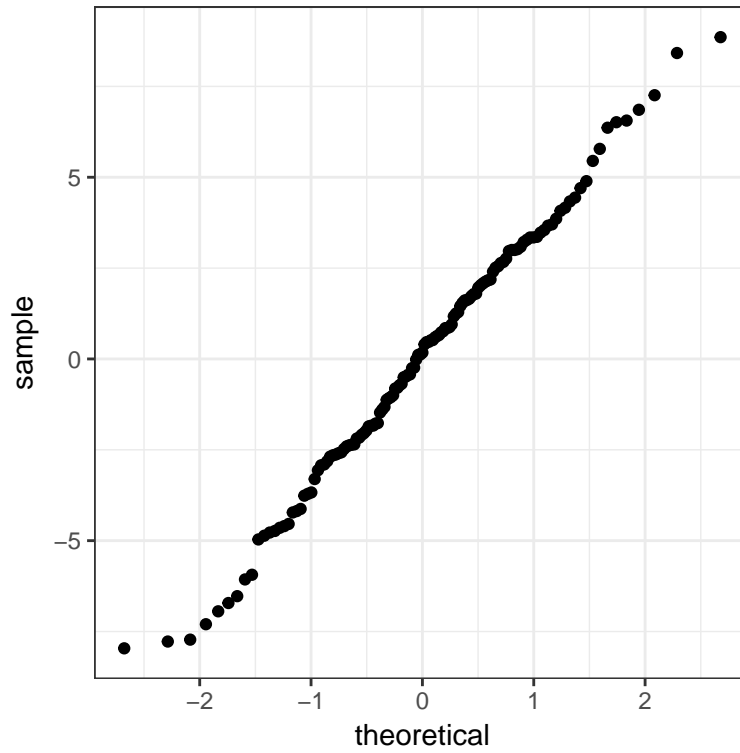
And we can see that the interaction term is also significant.

Part e

```
diag.df <- ggplot2::fortify(mixed.mod)
summary(diag.df)
```

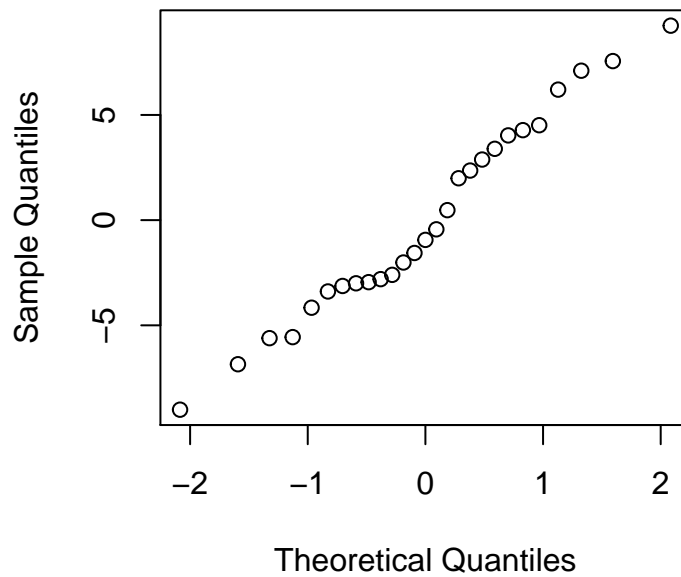
wt	weeks	subject	treat
Min. : 46.0	Min. :0	1 : 5	control :50
1st Qu.: 71.0	1st Qu.:1	2 : 5	thiouracil:50
Median :100.0	Median :2	3 : 5	thyroxine :35
Mean :100.8	Mean :2	4 : 5	
3rd Qu.:122.5	3rd Qu.:3	5 : 5	
Max. :189.0	Max. :4	6 : 5	
		(Other):105	
.fitted	.resid	.sresid	
Min. : 46.03	Min. : -7.963	Min. : -1.83136	
1st Qu.: 73.81	1st Qu.: -2.391	1st Qu.: -0.54991	
Median : 98.05	Median : 0.174	Median : 0.04003	
Mean :100.81	Mean : 0.000	Mean : 0.00000	
3rd Qu.:126.28	3rd Qu.: 2.532	3rd Qu.: 0.58230	
Max. :182.44	Max. : 8.855	Max. : 2.03660	

```
ggplot(diag.df) +
  stat_qq(aes(sample = .resid))
```



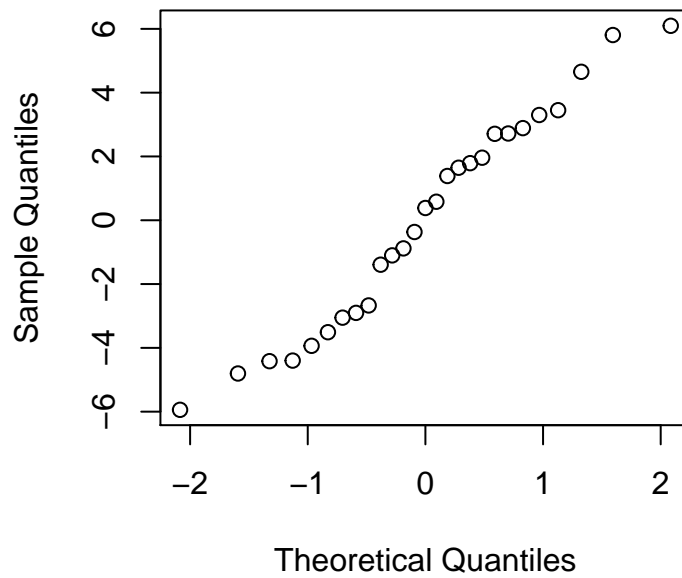
```
qqnorm(ranef(mixed.mod)$subject$`(Intercept)`)
```

Normal Q-Q Plot

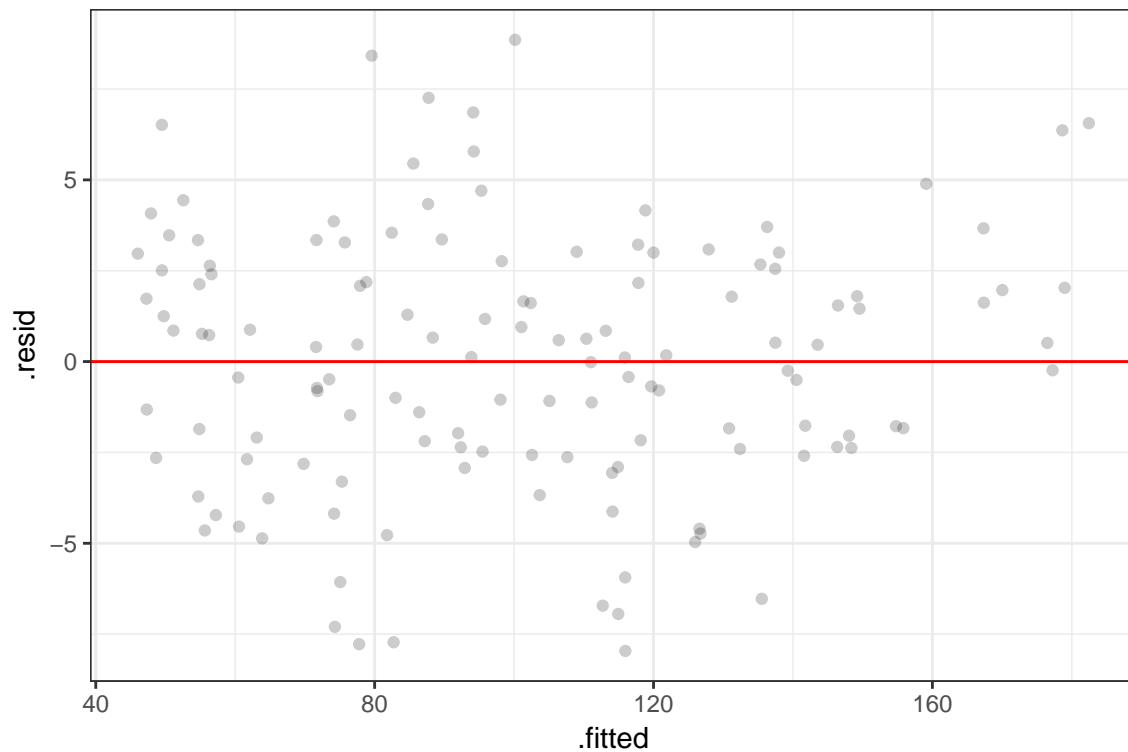


```
qqnorm(ranef(mixed.mod)$subject$weeks)
```

Normal Q-Q Plot



```
ggplot(diag.df) +  
  geom_point(aes(x = .fitted, y = .resid), alpha = .2) +  
  geom_abline(slope = 0, colour = 'red')
```



We don't have too much evidence of non-normality or heteroscedacity of the residuals. There also isn't much

evidence that the random effects are not normally distributed. We also don't see any outliers.

Part f

```
confint(mixed.mod, method = 'boot')
```

	2.5 %	97.5 %
.sig01	3.2777812	8.0905862
.sig02	-0.5891325	0.4450289
.sig03	2.6306787	5.0739311
.sigma	3.6695575	4.9643565
(Intercept)	49.0724758	56.6565110
weeks	24.1463073	28.9486434
treatthiouracil	-0.8000334	10.6085799
treatthyroxine	-7.5459914	5.1905991
weeks:treatthiouracil	-12.8247658	-5.8477773
weeks:treatthyroxine	-3.3880935	4.3061266

Based on the confidence intervals, the random effects are significant. However, looking at the fixed effects, we can see that since both the marginal and interaction terms for **thyroxine** have confidence intervals that contain 0, they may not differ significantly from the control group.