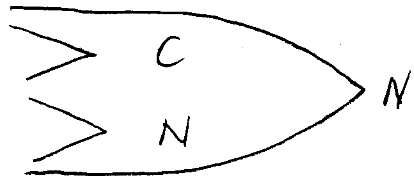
Stat 461 HW9 Jiaqi Li

1.

	Factor	R/F
i =	State	R
j =	Enviroment	F
t =	Lake	R

N/C



Thus, we use the <u>Mixed Model</u>: $(\alpha\beta)_{ij} \overset{iid}{\sim} N(0, \sigma^2_{Int})$

$$Y_{ijt} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijt}, \quad \alpha_i \overset{iid}{\sim} N(0, \sigma^2_{State}), \quad \varepsilon_{ijt} \overset{iid}{\sim} N(0, \sigma^2).$$

Y_{ijt} = the phosphorous level of t^{th} lake in j^{th} enviroment of state i.

i = New York, Pennsylvania, Vermount.

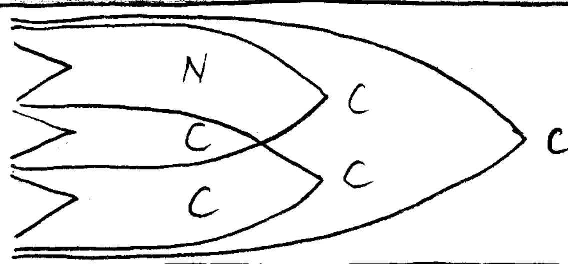
j = agriculture land, forest

t = 1, 2, 3, 4

2.

	Factor	R/F
i =	Type	F
j =	Road	R
k =	Paint	F
t =	Year (replicate)	R

NC



Thus, we use the <u>Mixed Model</u>:

$$Y_{ijkt} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \varepsilon_{ijkt}, \quad \beta_{j(i)} \overset{iid}{\sim} N(0, \sigma^2_{Road}), \quad \varepsilon_{ijkt} \overset{iid}{\sim} N(0, \sigma^2)$$

Y_{ijkt} = number of crashes on road j of type i with or without paint.

i = mountainous, city

j = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

k = with, without

t = 1, 2, 3.

3.

Factor	R/F	N/C
i = Store	R	C
j = Display	F	C
t = Week (replicate)	R	C

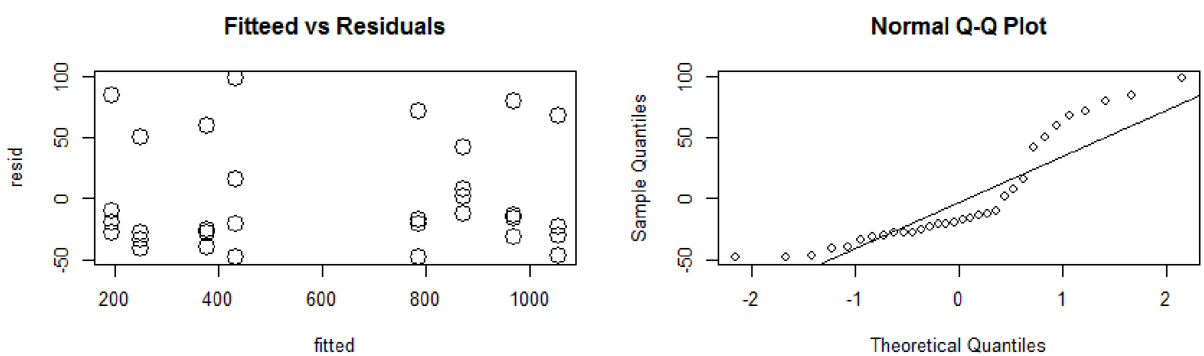
$Y_{ijt} = \mu + \alpha_i + \beta_j + \varepsilon_{ijt}$, $\alpha_i \overset{iid}{\sim} N(0, \sigma_{store}^2)$, $\varepsilon_{ijt} \overset{iid}{\sim} N(0, \sigma^2)$

Y_{ijt} = sales in t^{th} week of toys with/without display in store i.

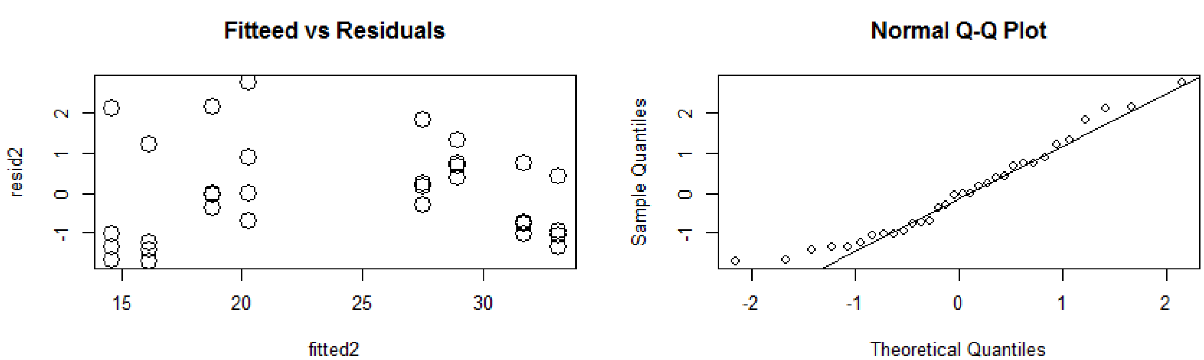
$i = 1, 2, 3, 4$

j = with, without.

$t = 1, 2, 3, 4$



Due to the graphs, we can see that constant variance is satisfied, but the normality is not satisfied. Thus, we will use transformation by taking square root of the response and obtain:



Now, we still have approximately constant variance and normality is satisfied. Thus, we will use this model for further study.

```
> Anova(model2,type="III")
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: sqrt(sales)
             Chisq Df Pr(>Chisq)
(Intercept) 40.736  1  1.743e-10 ***
display     83.374  1  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since, there is only one fixed factor and the interaction term is not considered in the model because it is a random variable, we only look at the main effect line.

The main effect for display tests:

$$H_0: \beta_{yes}=\beta_{no}$$

We reject this null hypothesis since p-value is smaller than 0.05. We conclude that there are difference between “display” and “not display”. Then, we will do pairwise comparison.

```
> difflsmeans(model2,"display")
Differences of LSMEANS:
                 Estimate Standard Error   DF t-value Lower CI Upper CI p-val
ue
display No - Yes     -4.2          0.455 27.0   -9.13    -5.08    -3.22  <2e-
16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

From the information above, we can observe that:

Toys with window display has larger effect on the mean sales of toys than toys without window display.

Then, we will check ANOVA table test for the random effects:

```
> rand(model2)
Analysis of Random effects Table:
      Chi.sq Chi.DF p.value
store   81.8      1  <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The null hypothesis test is:

$$H_0: \sigma_{store}^2= 0$$

We reject this null hypothesis test since the p-value is smaller than 0.05. We conclude that there are difference between mean sales of each store.

Thus, we conclude that toys with window display have more sales than toys without window display. In addition, different stores also have effects on sales of toys.

R code:

```
install.packages("lsmeans")
install.packages("car")
install.packages("multcompView")
install.packages("lme4")
install.packages("lmerTest")

library(lsmeans)
library(car)
library(multcompView)
library(lme4)
library(lmerTest)
options(contrasts = c("contr.sum", "contr.poly"))

displays=read.table("displays.dat",header=TRUE)
displays$store=as.factor(displays$store)
displays$week=as.factor(displays$week)
displays

df=data.frame(store=as.factor(displays$store), sales=displays$sales,display=displays$display)

model1=lmer(sales~(1|store)+display, data=df)

fitted=fitted(model1)
resid=residuals(model1)
par(mfrow=c(2,2))
plot(fitted,resid,main="Fitteed vs Residuals",pch=1,cex=2)
qqnorm(resid)
qqline(resid)

model2=lmer(sqrt(sales)~(1|store)+display, data=df)

fitted2=fitted(model2)
resid2=residuals(model2)
par(mfrow=c(2,2))
plot(fitted2,resid2,main="Fitteed vs Residuals",pch=1,cex=2)
qqnorm(resid2)
qqline(resid2)

Anova(model2,type="III")

difflsmeans(model2,"display")

rand(model2)
```