

Homework 7

Name: Zixin Ouyang

1.

```
> Views<-read.table("/Users/Constance/Downloads/viewcounts.dat",header=TRUE)
> viewmod<-lm(log(log(Views))~Channel, data=Views)
```

(a)

```
> anova(viewmod)
```

Analysis of Variance Table

Response: log(log(Views))

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Channel	6	1.2644	0.210740	2.938	0.03054 *
Residuals	21	1.5063	0.071728		

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

(b)

```
> library(lme4)
> views.reml<-lmer(log(log(Views))~(1|Channel),data=Views)
> summary(views.reml)
```

Linear mixed model fit by REML ['lmerMod']

Formula: log(log(Views)) ~ (1 | Channel)

Data: Views

REML criterion at convergence: 15.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.9910	-0.5421	-0.2887	0.7762	1.6743

Random effects:

Groups	Name	Variance	Std.Dev.
Channel	(Intercept)	0.03475	0.1864
Residual		0.07173	0.2678

Number of obs: 28, groups: Channel, 7

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	1.51004	0.08676	17.41

(c)

```
> 0.03475/(0.03475+0.0717)
```

```
[1] 0.3264443
```

(d)

```
> views.ml<-lmer(log(log(Views))~(1|Channel),data=Views,REML=FALSE)
> summary(views.ml)
```

Linear mixed model fit by maximum likelihood ['lmerMod']

Formula: log(log(Views)) ~ (1 | Channel)

Data: Views

AIC	BIC	logLik	deviance	df.resid
18.1	22.1	-6.1	12.1	25

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.0124	-0.5562	-0.2586	0.7570	1.7529

Random effects:

Groups	Name	Variance	Std.Dev.
Channel	(Intercept)	0.02723	0.1650
	Residual	0.07173	0.2678

Number of obs: 28, groups: Channel, 7

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	1.51004	0.08032	18.8

The maximum likelihood estimate of the variance component associated with channels: 0.02723.

(e)

```
> nullmod<-lm(log(log(Views))~1,data=Views)
> llrts<-as.numeric(2*(logLik(views.ml)-logLik(nullmod)))
> llrts
```

```
[1] 2.544656
```

```
> pchisq(llrts,1,lower=FALSE)
```

```
[1] 0.110668
```

(f)

```
> set.seed(1)
> lrstats <- numeric(10000)
> for(i in 1:10000){
+ y <- unlist(simulate(nullmod))
+ nullsim <- lm(y ~ 1)
+ altsim <- lmer(y ~ (1|Channel), data=Views, REML=FALSE)
+ lrstats[i] <- as.numeric(2 * (logLik(altsim) - logLik(nullsim)))
+ }
```

```
> pval <- mean(lrstats >= llrts)
```

```
> pval
```

```
[1] 0.0324
```

```
> se.pval <- sqrt(pval*(1-pval)/10000)
```

```
> se.pval
```

```
[1] 0.0017706
```

(g)

```
> ranef(views.reml)
$Channel
(Intercept)
A -0.05758216
B -0.08152907
C -0.15467586
D -0.06656573
E -0.06993830
F 0.18547217
G 0.24481894
```

2.

```
> library(faraway)
```

(a) The whole-plot factor is recipe. The whole plots are batches.

(b) The split-plot factor is temp. The split plots are parts of each batch.

(c)

```
> choccake$batches<-choccake$recipe:choccake$batch
> choccake$batches
 [1] 1:1 1:1 1:1 1:1 1:1 1:1 1:2 1:2 1:2 1:2 1:2 1:2 1:3 1:3 1:3
[16] 1:3 1:3 1:3 1:4 1:4 1:4 1:4 1:4 1:4 1:5 1:5 1:5 1:5 1:5 1:5
[31] 1:6 1:6 1:6 1:6 1:6 1:6 1:7 1:7 1:7 1:7 1:7 1:7 1:8 1:8 1:8
[46] 1:8 1:8 1:8 1:9 1:9 1:9 1:9 1:9 1:9 1:10 1:10 1:10 1:10 1:10 1:10
[61] 1:11 1:11 1:11 1:11 1:11 1:11 1:12 1:12 1:12 1:12 1:12 1:12 1:13 1:13 1:13
[76] 1:13 1:13 1:13 1:14 1:14 1:14 1:14 1:14 1:14 1:14 1:15 1:15 1:15 1:15 1:15 1:15
[91] 2:1 2:1 2:1 2:1 2:1 2:1 2:2 2:2 2:2 2:2 2:2 2:2 2:3 2:3 2:3
[106] 2:3 2:3 2:3 2:4 2:4 2:4 2:4 2:4 2:4 2:5 2:5 2:5 2:5 2:5 2:5
[121] 2:6 2:6 2:6 2:6 2:6 2:6 2:7 2:7 2:7 2:7 2:7 2:7 2:8 2:8 2:8
[136] 2:8 2:8 2:8 2:9 2:9 2:9 2:9 2:9 2:9 2:10 2:10 2:10 2:10 2:10 2:10
[151] 2:11 2:11 2:11 2:11 2:11 2:11 2:12 2:12 2:12 2:12 2:12 2:12 2:13 2:13 2:13
[166] 2:13 2:13 2:13 2:14 2:14 2:14 2:14 2:14 2:14 2:15 2:15 2:15 2:15 2:15 2:15
[181] 3:1 3:1 3:1 3:1 3:1 3:1 3:2 3:2 3:2 3:2 3:2 3:2 3:3 3:3 3:3
[196] 3:3 3:3 3:3 3:4 3:4 3:4 3:4 3:4 3:4 3:5 3:5 3:5 3:5 3:5 3:5
[211] 3:6 3:6 3:6 3:6 3:6 3:6 3:7 3:7 3:7 3:7 3:7 3:7 3:8 3:8 3:8
[226] 3:8 3:8 3:8 3:9 3:9 3:9 3:9 3:9 3:9 3:10 3:10 3:10 3:10 3:10 3:10
[241] 3:11 3:11 3:11 3:11 3:11 3:11 3:12 3:12 3:12 3:12 3:12 3:12 3:13 3:13 3:13
[256] 3:13 3:13 3:13 3:14 3:14 3:14 3:14 3:14 3:14 3:15 3:15 3:15 3:15 3:15 3:15
45 Levels: 1:1 1:2 1:3 1:4 1:5 1:6 1:7 1:8 1:9 1:10 1:11 1:12 1:13 1:14 ... 3:15
```

(d)

```
> choccake.reml<-lmer(breakang~recipe*factor(temp)+(1|batches),data=choccake)
> anova(choccake.reml)
Analysis of Variance Table

              Df Sum Sq Mean Sq F value
recipe          2   10.19      5.09  0.2488
factor(temp)     5 2100.30    420.06 20.5199
recipe:factor(temp) 10  205.98     20.60  1.0062
```

(e)

```
> drop1(choccake.reml,test="Chisq")
Single term deletions

Model:
breakang ~ recipe * factor(temp) + (1 | batches)
              Df    AIC    LRT Pr(>Chi)
<none>                1719.0
recipe:factor(temp) 10 1709.6 10.53  0.3953
```

The test statistic value is 10.53 and p-value is 0.3953. We cannot reject the null hypothesis. Therefore, we will drop the interaction term.

(f)

```
> drop1(update(choccake.reml, .~.-recipe:factor(temp)), test="Chisq")
```

Single term deletions

Model:

```
breakang ~ recipe + factor(temp) + (1 | batches)
```

	Df	AIC	LRT	Pr(Chi)
<none>		1709.6		
recipe	2	1706.1	0.530	0.7672
factor(temp)	5	1785.7	86.106	<2e-16 ***

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

The test statistic of recipe is 0.530 and p-value is 0.7672. The test statistic of temp is 86.106 and p-value is almost zero. The effect of recipe is not significant while the effect of temp is significant.

(g)

```
> choccake.aov<-aov(breakang~recipe*factor(temp)+Error(batches),data=choccake)
```

```
> summary(choccake.aov)
```

Error: batches

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
recipe	2	135	67.54	0.249	0.781
Residuals	42	11403	271.49		

Error: Within

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(temp)	5	2100	420.1	20.520	<2e-16 ***
recipe:factor(temp)	10	206	20.6	1.006	0.439
Residuals	210	4299	20.5		

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

(h) The p-value of temp is less than 0.05, so only temp effect is significant.

3.

```
> library(alr4)
```

(a)

```
> shlogit<-glm(cbind(Y,m-Y)~Intensity, family=binomial,data=shocks)
> summary(shlogit)
```

Call:

```
glm(formula = cbind(Y, m - Y) ~ Intensity, family = binomial,
     data = shocks)
```

Deviance Residuals:

1	2	3	4	5	6
-2.2507	0.3892	-0.1466	1.1080	0.3234	-1.6679

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-3.3010	0.3238	-10.20	<2e-16 ***
Intensity	1.2459	0.1119	11.13	<2e-16 ***

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

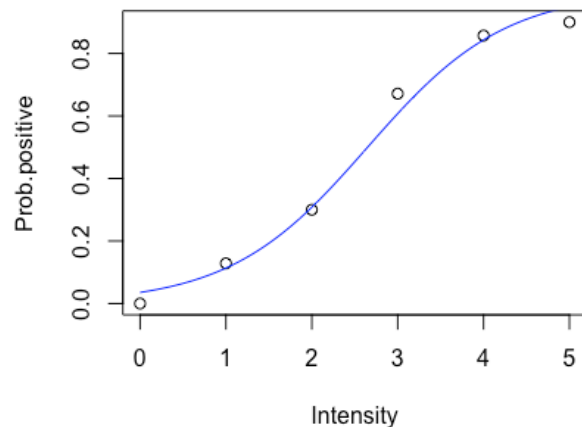
(Dispersion parameter for binomial family taken to be 1)

Null deviance: 250.4866 on 5 degrees of freedom
Residual deviance: 9.3526 on 4 degrees of freedom
AIC: 34.093

Number of Fisher Scoring iterations: 4

(b)

```
> with(shocks, plot(Intensity, Y/m,xlab="Intensity",ylab="Prob.positive"))
> curve(predict(shlogit, data.frame(Intensity=x), type="response"), add=TRUE,lty=1, col="blue")
```



(c)

```
> fitted(shlogit)
1      2      3      4      5      6
0.0355357 0.1135383 0.3080695 0.6074909 0.8432637 0.9492448
> residuals(shlogit)
1      2      3      4      5      6
-2.2506765 0.3892406 -0.1465914 1.1079581 0.3234081 -1.6679077
```

(d)

```
> prob<-predict(shlogit, newdata=data.frame(Intensity=1.5), type="response")
> odds<-prob/(1-prob)
> odds
      1
0.2388
```

(e)

```
> pchisq(deviance(shlogit), df.residual(shlogit), lower=FALSE)
[1] 0.05286523
```

We fail to reject the null hypothesis. So there is no lack of fit.

(f)

```
> X.2 <- sum(residuals(shlogit,type="pearson")^2)
> X.2
[1] 7.583229
> pchisq(X.2, df.residual(shlogit), lower=FALSE)
[1] 0.1080947
```

We fail to reject the null hypothesis. So there is no lack of fit.

(g)

```
> shprobit <- glm(cbind(Y,m-Y)~Intensity, family=binomial(link=probit),data=shocks)
> summary(shprobit)
```

Call:

```
glm(formula = cbind(Y, m - Y) ~ Intensity, family = binomial(link = probit),
    data = shocks)
```

Deviance Residuals:

1	2	3	4	5	6
-1.9941	0.3296	-0.2709	1.3687	0.6558	-1.7768

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.91099	0.16794	-11.38	<2e-16 ***
Intensity	0.71459	0.05673	12.60	<2e-16 ***

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 250.4866 on 5 degrees of freedom
Residual deviance: 9.6189 on 4 degrees of freedom
AIC: 34.359

Number of Fisher Scoring iterations: 5

```
> pchisq(deviance(shprobit), df.residual(shprobit), lower=FALSE)
[1] 0.04736097
```

We reject the null hypothesis and conclude that there is lack of fit based on based deviance.