

EX3

Yizhen

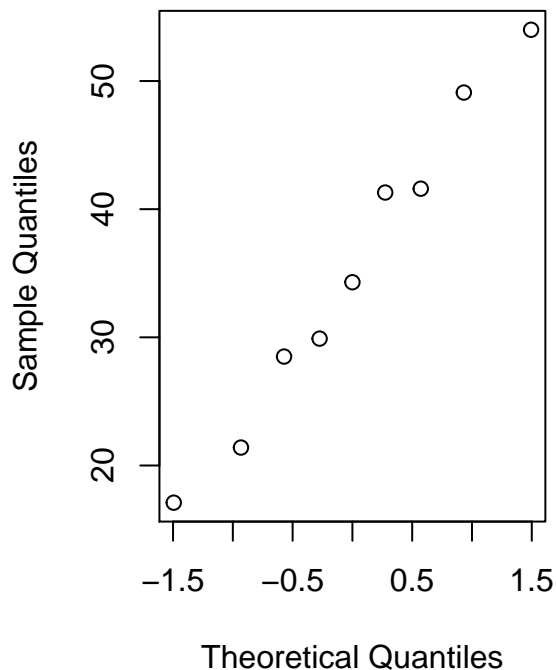
Exercise 3. (cow)

a) repeated measures - exchangeable

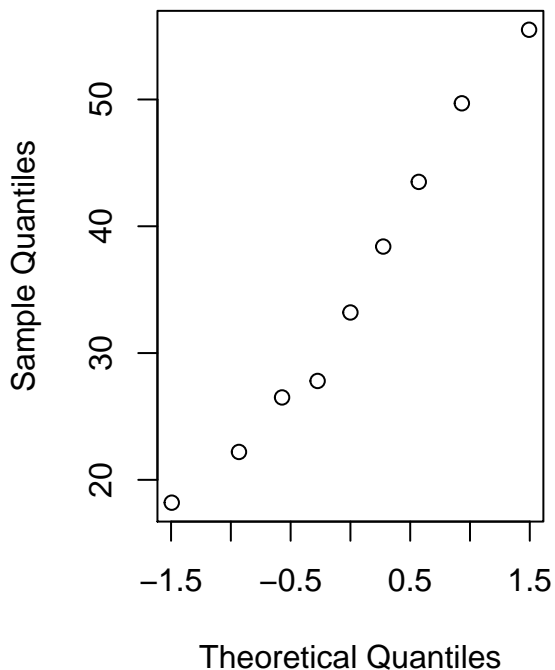
First we evaluate normality to determine which test to use. From the QQ-plots below we conclude that both treatment samples are normally distributed.

```
cow = read.table("cow.txt", header = TRUE);  
  
#There are two factors influence the milk production  
  
mm = subset(cow, treatment=="A")["milk"]  
sf = subset(cow, treatment=="B")["milk"]  
mm1 = as.numeric(unlist(mm))  
sf1 = as.numeric(unlist(sf))  
  
par(mfrow=c(1,2))  
qqnorm(mm1)  
qqnorm(sf1)
```

Normal Q-Q Plot



Normal Q-Q Plot



From the ANOVA results below, we can conclude that within-cow variation(see variable “id”) the milk production differs. Because the p-value for id is less than 0.05, therefore, H_0 is rejected. Furthermore, from the summary we can conclude that most of the cows(except id4) are different from the cow with id1. Then, we could see p-value for per is less than 0.05 so we reject H_0 which means whether a cow is going

through the first period or second seems to make a difference. Furthermore, because the p-value of treatment is 0.51654, so we do not reject H_0 which means the treatment A does not significantly differ from treatment B. Also this could be seen from second table (treatmentB). Therefore there is no significant difference in milk production. However, it is important to note that this is not the appropriate way of testing the cross-over design.

```
cow$id = factor(cow$id)
cow$per = factor(cow$per)
cowanova = lm(milk~id+per+treatment,data = cow)
anova(cowanova); summary(cowanova)
```

```
## Analysis of Variance Table
##
## Response: milk
##           Df Sum Sq Mean Sq F value    Pr(>F)
## id          8 2467.47  308.434 124.4832 7.494e-07 ***
## per         1   24.50   24.500   9.8881  0.01628 *
## treatment   1    1.16    1.156   0.4666  0.51654
## Residuals   7   17.34    2.478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Call:
## lm(formula = milk ~ id + per + treatment, data = cow)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2600 -0.4375  0.0000   0.4375  2.2600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  30.3000     1.2444   24.349 5.02e-08 ***
## id2          23.0000     1.5741   14.612 1.68e-06 ***
## id3          11.1500     1.5741    7.084 0.000196 ***
## id4          -1.3500     1.5741   -0.858 0.419480
## id5          -7.0500     1.5741   -4.479 0.002870 **
## id6          23.4500     1.5741   14.898 1.47e-06 ***
## id7          13.5500     1.5741    8.608 5.69e-05 ***
## id8           4.9000     1.5741    3.113 0.017011 *
## id9         -11.2000     1.5741   -7.115 0.000191 ***
## per2         -2.3900     0.7466   -3.201 0.015046 *
## treatmentB   -0.5100     0.7466   -0.683 0.516536
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.574 on 7 degrees of freedom
## Multiple R-squared:  0.9931, Adjusted R-squared:  0.9832
## F-statistic: 100.6 on 10 and 7 DF,  p-value: 1.349e-06
```

b) modelling the cow effect as a random effect (use the function lmer).

```
library(lme4)
```

```
## Loading required package: Matrix

cowlmer = lmer(milk~treatment+per+(1|id), data=cow, REML = FALSE)
summary(cowlmer)
```

```
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: milk ~ treatment + per + (1 | id)
## Data: cow
##
##      AIC      BIC    logLik deviance df.resid
##    117.5    122.0    -53.8    107.5      13
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.54103 -0.36774  0.02717  0.27827  1.71498
##
## Random effects:
## Groups Name Variance Std.Dev.
## id      (Intercept) 136.118  11.667
## Residual              1.927   1.388
## Number of obs: 18, groups: id, 9
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  36.5722     3.9335   9.298
## treatmentB   -0.5100     0.6585  -0.775
## per2         -2.3900     0.6585  -3.630
##
## Correlation of Fixed Effects:
##              (Intr) trtmnB
## treatmentB -0.093
## per2       -0.093  0.111
```

```
cowlmerTreatment = lmer(milk~per+(1|id), data=cow, REML = FALSE)
anova(cowlmerTreatment, cowlmer)
```

```
## Data: cow
## Models:
## cowlmerTreatment: milk ~ per + (1 | id)
## cowlmer: milk ~ treatment + per + (1 | id)
##              Df    AIC    BIC logLik deviance Chisq Chi Df
## cowlmerTreatment  4 116.09 119.65 -54.045  108.09
## cowlmer           5 117.51 121.96 -53.755  107.51 0.5807    1
##              Pr(>Chisq)
## cowlmerTreatment
## cowlmer           0.446
```

Based on the p-value, we could know that treatment is not important. The result is the same as that in a).

```
cowlmerPer = lmer(milk~treatment+(1|id), data=cow, REML = FALSE)
anova(cowlmerPer, cowlmer)
```

```
## Data: cow
## Models:
## cowlmerPer: milk ~ treatment + (1 | id)
## cowlmer: milk ~ treatment + per + (1 | id)
##              Df    AIC    BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## cowlmerPer      4 116.09 119.65 -54.045  108.09
## cowlmer          5 117.51 121.96 -53.755  107.51 0.5807    1
```

```
## cowlmerPer 4 123.62 127.19 -57.812 115.62
## cowlmer 5 117.51 121.96 -53.755 107.51 8.1151 1 0.00439 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the p-value, we could know that whether a cow is going through the first treatment or second is important.

c)

```
attach(cow)
t.test(milk[treatment=="A"],milk[treatment=="B"],paired=TRUE)

##
## Paired t-test
##
## data: milk[treatment == "A"] and milk[treatment == "B"]
## t = 0.22437, df = 8, p-value = 0.8281
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -2.267910 2.756799
## sample estimates:
## mean of the differences
## 0.2444444
```

From this result that p-value is 0.8281 which means the treatment is important. However, from previous analysis treatment should not be important. Given the design that it is inappropriate to use paired t-test. Because during our analysis, the variable “per” has a significant effect on milk production, it may not be wise to ignore this factor when comparing two groups.