HW3 STAT512 Fall2014

Yet Nguyen

September 24, 2014

1.

Which pairs of diets are significantly different? Establish this by performing 6 pair-wise t-tests (although, as we know, this doesnt control overall error rate), reporting a p-value for each.

Sol:

```
#1####
mse <- 400
t <- 4
diet <- c("C1", "R5", "R10", "T")
# treatment mean
mean_trt <- c(158, 240, 282, 279)

mean_overall <- mean(mean_trt)
# least square estimate of treatment with zero sum constrain
ls_trt <- mean_trt - mean_overall

# 6 contrast for pair comparisons
 C <- matrix(c(1, -1, 0, 0,
               1, 0, -1, 0,
               1, 0, 0, -1,
               0, 1, -1, 0,
               0, 1, 0, -1,
               0, 0, 1, -1
               ), byrow = T,
             nrow = 6)



#  least square estimate of the constrasts

Ct <- as.vector(C%*%ls_trt)
```

```r
# their corresponding variance
se_Ct <- unlist(lapply(1:nrow(C),
                       function(i) sqrt(mse * sum(C[i,]^2)/t)))

# test statistics
t_test <- abs(Ct/se_Ct)

# df
df<- (t-1)*(t-2)

# pvalues for each test - two sided test

pvalues <- unlist(lapply(1:length(t_test), function(i){
  2*(1 - pt(t_test[i], df))
}))


pair <- c("C1-R5", "C1-R10", "C1- T",
          "R5-R10", "R5-T", "R10-T")
as.data.frame(cbind(pair = pair,pvalue =  round(pvalues, digits = 5)))

##     pair  pvalue
## 1  C1-R5 0.00115
## 2 C1-R10 0.00012
## 3  C1- T 0.00014
## 4 R5-R10 0.02496
## 5   R5-T 0.03295
## 6  R10-T 0.83903
```

2.

Suppose the experiment could actually have been carried out over 8 weeks (rather than 4), but that only 4 animals were available. The experiment could then be planned according to a more general row-column design, testing 4 treatments within a block structure of 4 rows (animals) and 8 columns (weeks). ONE way to do this is by direct application of one of the replicated Latin Square strategies we talked about in class. Suppose this had, in fact, been done, and the 5 statistics reported to you above had been calculated for the data collected from this larger experiment. Repeat the part (1); does this change the answers?

Sol:

```r
#2 ####

# This bdesign is Version 2 of the replicate LSD
```

```r
r <- 2 # number of replicates

df2 <-  r*t^2-1 - (r-1) - (t-1) - (t-1) - r*(t-1) # df of error

# test statistics are the same, but df is not
pvalues2 <- unlist(lapply(1:length(t_test), function(i){
  2*(1 - pt(t_test[i], df2))
}))

pair <- c("C1-R5", "C1-R10", "C1- T",
          "R5-R10", "R5-T", "R10-T")
as.data.frame(cbind(pair = pair,pvalue =  round(pvalues2, digits = 5)))
```

```
##     pair  pvalue
## 1  C1-R5   2e-05
## 2 C1-R10       0
## 3  C1- T       0
## 4 R5-R10 0.00821
## 5   R5-T 0.01296
## 6  R10-T 0.83439
```

3.

Return now to the original Latin Square design. Suppose that after the conclusion of the experiment, it was discovered that diet T was actually contaminated, and that none of the resulting data could be used (not even as reflecting an altered and uninteresting diet). That is, the remaining treatments might have been assigned to units as follows: Use R, or another computer package that supports numerical linear algebra, to determine whether this design is Condition E equivalent to a completely randomized design for t = 3 and each $n_i = 4$.

Sol:

```r
# 3####
library(MASS)
# X1 matrix for LSD
X1_LSD <- matrix(c(1,0,0,0,1,0,0,0,
                   1,0,0,0,0,0,1,0,
                   1,0,0,0,0,0,0,1,
                   0,1,0,0,0,1,0,0,
                   0,1,0,0,0,0,1,0,
                   0,1,0,0,0,0,0,1,
                   0,0,1,0,1,0,0,0,
                   0,0,1,0,0,1,0,0,
                   0,0,1,0,0,0,1,0,
```

```
                  0,0,0,1,1,0,0,0,
                  0,0,0,1,0,1,0,0,
                  0,0,0,1,0,0,0,1),
                byrow = T, nrow = 12)



# X2 for LSD
X2_LSD <- matrix(c(0,1,0,
                   1,0,0,
                   0,0,1,
                   0,0,1,
                   0,1,0,
                   1,0,0,
                   1,0,0,
                   0,1,0,
                   0,0,1,
                   0,0,1,
                   1,0,0,
                   0,1,0),
                 byrow = T,
                 nrow = 12)

H1_LSD <- X1_LSD %*%ginv(t(X1_LSD)%*%X1_LSD)%*%t(X1_LSD)
H1_LSD %*% X2_LSD
```

```
##          [,1]   [,2]   [,3]
##  [1,] 0.3333 0.3333 0.3333
##  [2,] 0.3333 0.3333 0.3333
##  [3,] 0.3333 0.3333 0.3333
##  [4,] 0.3333 0.3333 0.3333
##  [5,] 0.3333 0.3333 0.3333
##  [6,] 0.3333 0.3333 0.3333
##  [7,] 0.3333 0.3333 0.3333
##  [8,] 0.3333 0.3333 0.3333
##  [9,] 0.3333 0.3333 0.3333
## [10,] 0.3333 0.3333 0.3333
## [11,] 0.3333 0.3333 0.3333
## [12,] 0.3333 0.3333 0.3333
```

```
# X1 for CRD
X1_CRD <- matrix(rep(1, 12), nrow = 12)
# X2 for CRD is the same as X2 for LSD
X2_CRD <- X2_LSD
H1_CRD <- X1_CRD %*%ginv(t(X1_CRD)%*%X1_CRD)%*%t(X1_CRD)
H1_CRD %*% X2_CRD
```

```
##          [,1]   [,2]   [,3]
##  [1,] 0.3333 0.3333 0.3333
##  [2,] 0.3333 0.3333 0.3333
##  [3,] 0.3333 0.3333 0.3333
##  [4,] 0.3333 0.3333 0.3333
##  [5,] 0.3333 0.3333 0.3333
##  [6,] 0.3333 0.3333 0.3333
##  [7,] 0.3333 0.3333 0.3333
##  [8,] 0.3333 0.3333 0.3333
##  [9,] 0.3333 0.3333 0.3333
## [10,] 0.3333 0.3333 0.3333
## [11,] 0.3333 0.3333 0.3333
## [12,] 0.3333 0.3333 0.3333
```

The result above implies that two designs are equivalent.