Homework 6

John Kaspers

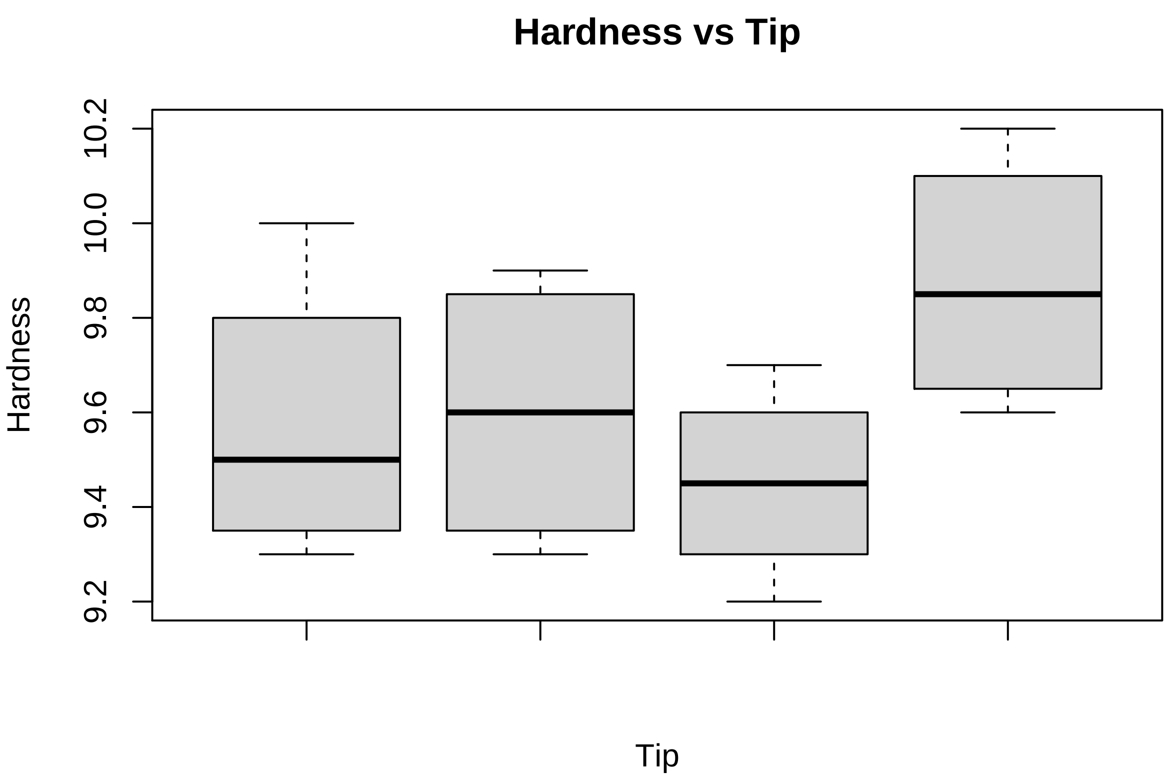
10/17/2020

Questions graded for completion

Problem 4

```
Tip1 <- c(9.3, 9.4, 9.6, 10.0)
Tip2 <- c(9.4, 9.3, 9.8, 9.9)
Tip3 <- c(9.2, 9.4, 9.5, 9.7)
Tip4 <- c(9.7, 9.6, 10.0, 10.2)

boxplot(Tip1, Tip2, Tip3, Tip4, xlab = "Tip", ylab = "Hardness", main = "Hardness vs Tip")
```




```
coupon <- data.frame(Tip1, Tip2, Tip3, Tip4)
coupon
```

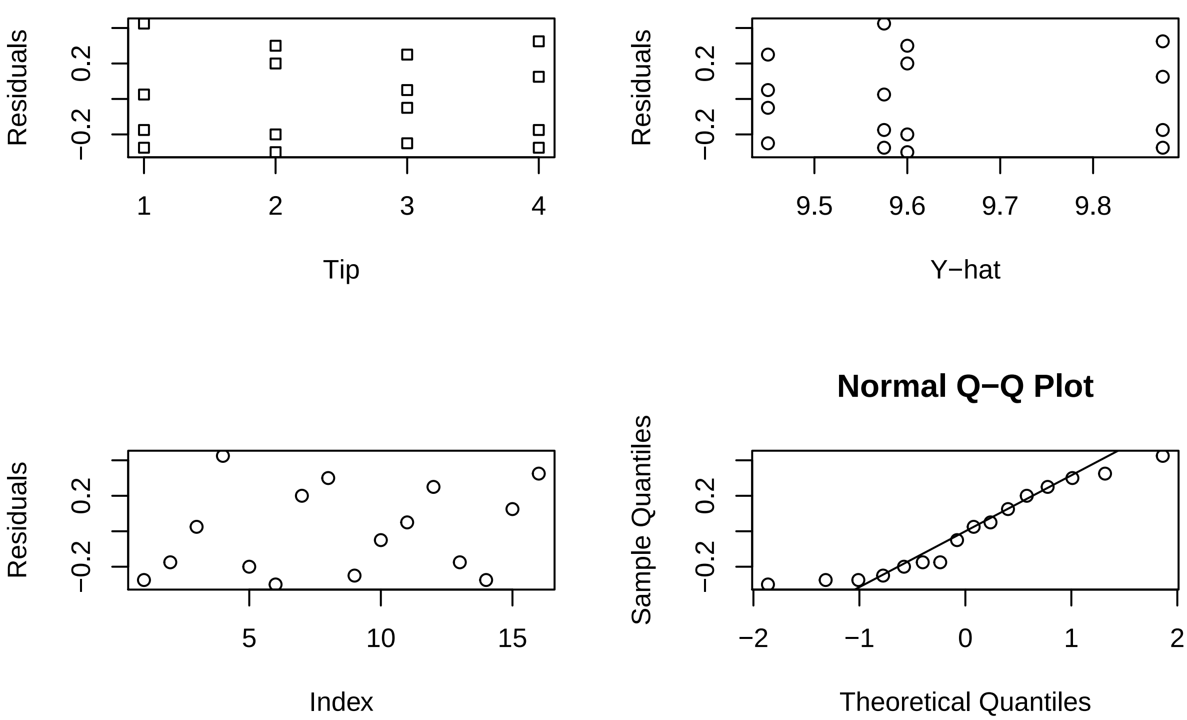
```
##   Tip1 Tip2 Tip3 Tip4
## 1  9.3  9.4  9.2  9.7
## 2  9.4  9.3  9.4  9.6
## 3  9.6  9.8  9.5 10.0
## 4 10.0  9.9  9.7 10.2
```

```
tip = sort(rep(1:4,4))
hardness = c(9.3, 9.4, 9.6, 10.0, 9.4, 9.3, 9.8, 9.9, 9.2, 9.4, 9.5, 9.7, 9.7, 9.6, 10.0, 10.2)
Tip = as.factor(tip)
model <- lm(hardness~Tip)
summary.aov(model)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Tip          3  0.385 0.12833   1.702   0.22
## Residuals   12  0.905 0.07542
```

Running Diagnostics plots

```
model_fit = model$fitted
residuals = model$residuals
par(mfrow=c(2,2))
stripchart(residuals~Tip,vertical=T,xlab="Tip",ylab="Residuals") # Testing for constant variance
plot(model_fit,residuals,xlab="Y-hat",ylab="Residuals")#Testing for outliers
plot(residuals,ylab="Residuals") #against index; this could be run order
qqnorm(residuals)#is sample drawn from a normal distribution?
qqline(residuals)
```



Answer Assuming the relationship between tip and hardness is linear, based on the normal QQ-line plot which is approximately linear, we can say there appears to be a linear relationship between tip and hardness.

Looking at the upper right plot, there appears to be no outliers. Looking at the lower left plot, there is no linear relationship between run order, and there appears to be a constant variance.

Determining if Tukey HSD is appropriate

```
m3=aov(hardness~Tip)#usual ANOVA
TukeyHSD(m3, "Tip")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = hardness ~ Tip)
##
## $Tip
##       diff      lwr     upr     p adj
## 2-1  0.025 -0.55152 0.60152 0.9991931
## 3-1 -0.125 -0.70152 0.45152 0.9157208
## 4-1  0.300 -0.27652 0.87652 0.4431278
## 3-2 -0.150 -0.72652 0.42652 0.8653399
## 4-2  0.275 -0.30152 0.85152 0.5135079
## 4-3  0.425 -0.15152 1.00152 0.1815685
```

Because all six confidence intervals contain zero, we can say with 95% confidence that there is no difference in mean hardship readings across the four tip types.

Problem 5.

4.7 CRD vs RCBD

```
#Finding p-value from f-statistic
pf(46.3588, 4, 20, lower.tail = FALSE)
```

```
## [1] 7.668296e-10
```

```
#Thus, because we have a p-value of 1.184659e-09, which is less than our alpha level of 0.05.
```

Problem 6

```
#(a)
#We must divide 0.05 by 30 because it's a bonferroni

b <- 1:20
dfe <- (6-1)*(b-1)
qtvalue <- qt(1-0.05/30, dfe)
```

```
## Warning in qt(1 - 0.05/30, dfe): NaNs produced
```

```
msd <- qtvalue*sqrt(200/b)

mydata<- data.frame(b, dfe, qtvalue, msd)
mydata
```

```
##      b dfe  qtvalue       msd
## 1    1   0      NaN       NaN
## 2    2   5 5.247417 52.474169
## 3    3  10 3.827337 31.250080
## 4    4  15 3.483677 24.633315
## 5    5  20 3.330641 21.064826
## 6    6  25 3.244255 18.730717
## 7    7  30 3.188806 17.044884
## 8    8  35 3.150211 15.751053
## 9    9  40 3.121804 14.716323
## 10  10  45 3.100023 13.863724
## 11  11  50 3.082793 13.145075
## 12  12  55 3.068824 12.528420
## 13  13  60 3.057269 11.991595
## 14  14  65 3.047553 11.518668
## 15  15  70 3.039269 11.097843
## 16  16  75 3.032123 10.720174
## 17  17  80 3.025895 10.378734
## 18  18  85 3.020419 10.068063
## 19  19  90 3.015566  9.783789
## 20  20  95 3.011237  9.522366
```

```
#Therefore, we would need six blocks for a pairwise comparison using Bonferroni's method
```

#part b *Answer* Fewer because Tukey's HSD focuses on differences of means, while Bonferroni compares each pair with another pair. Bonferroni is more conservative.

#Part c ## How many blocks would be needed if Scheffe's method is used to get a set of simultaneous C.I.s for various contrasts and the width of the C.I. is at most 20 units?

```
alpha=0.05 #significance level

b <- 1:20
a = 4
dfe <- (6-1)*(b-1)
scheffe <- sqrt((a-1)*qf(1-alpha, a-1, dfe))
```

```
## Warning in qf(1 - alpha, a - 1, dfe): NaNs produced
```

```
msd <- scheffe*sqrt(200/b)

mydata1<- data.frame(b, dfe, scheffe, msd)
mydata1
```

```
##      b dfe  scheffe       msd
## 1    1   0      NaN       NaN
## 2    2   5 4.028443 40.284431
```

```
## 3   3  10 3.335385 27.233306
## 4   4  15 3.140405 22.206020
## 5   5  20 3.048799 19.282296
## 6   6  25 2.995617 17.295204
## 7   7  30 2.960884 15.826588
## 8   8  35 2.936420 14.682100
## 9   9  40 2.918259 13.756805
## 10 10  45 2.904244 12.988172
## 11 11  50 2.893100 12.336218
## 12 12  55 2.884027 11.773990
## 13 13  60 2.876497 11.282549
## 14 14  65 2.870147 10.848137
## 15 15  70 2.864721 10.460481
## 16 16  75 2.860029 10.111730
## 17 17  80 2.855933  9.795770
## 18 18  85 2.852326  9.507754
## 19 19  90 2.849125  9.243783
## 20 20  95 2.846265  9.000682
```

```
##Thus 5 blocks would be needed if Scheffe's method is used and a width of confidence interval is at mo
```

#Part d ### Power analysis

```
#Carry out power analysis
alpha=0.05#significance level
a=6#number of treatments
D=10 #desired maximum difference in means
sigsq=100 #our estimate of sigma^2
b=2:40 #the number of blocks
Fcrit=qf(1-alpha,a-1,(a-1)*(b-1))#value at which we reject H0
lam=b*(D^2)/(2*sigsq)#non-centrality parameter (ncp)
beta=pf(Fcrit,a-1,(a-1)*(b-1),ncp=lam)#Type 2 error
power=1-beta
cbind(b,Fcrit,beta,power)
```

```
##        b    Fcrit       beta      power
##  [1,]  2 5.050329 0.92991849 0.07008151
##  [2,]  3 3.325835 0.90475204 0.09524796
##  [3,]  4 2.901295 0.87694229 0.12305771
##  [4,]  5 2.710890 0.84694932 0.15305068
##  [5,]  6 2.602987 0.81516126 0.18483874
##  [6,]  7 2.533555 0.78194979 0.21805021
##  [7,]  8 2.485143 0.74767276 0.25232724
##  [8,]  9 2.449466 0.71267063 0.28732937
##  [9,] 10 2.422085 0.67726260 0.32273740
## [10,] 11 2.400409 0.64174343 0.35825657
## [11,] 12 2.382823 0.60638121 0.39361879
## [12,] 13 2.368270 0.57141603 0.42858397
## [13,] 14 2.356028 0.53705947 0.46294053
## [14,] 15 2.345586 0.50349469 0.49650531
## [15,] 16 2.336576 0.47087724 0.52912276
## [16,] 17 2.328721 0.43933610 0.56066390
## [17,] 18 2.321812 0.40897528 0.59102472
```

```
## [18,] 19 2.315689 0.37987545 0.62012455
## [19,] 20 2.310225 0.35209590 0.64790410
## [20,] 21 2.305318 0.32567651 0.67432349
## [21,] 22 2.300888 0.30063975 0.69936025
## [22,] 23 2.296868 0.27699274 0.72300726
## [23,] 24 2.293205 0.25472913 0.74527087
## [24,] 25 2.289851 0.23383102 0.76616898
## [25,] 26 2.286771 0.21427074 0.78572926
## [26,] 27 2.283931 0.19601248 0.80398752
## [27,] 28 2.281305 0.17901381 0.82098619
## [28,] 29 2.278869 0.16322706 0.83677294
## [29,] 30 2.276603 0.14860062 0.85139938
## [30,] 31 2.274491 0.13507998 0.86492002
## [31,] 32 2.272517 0.12260878 0.87739122
## [32,] 33 2.270667 0.11112964 0.88887036
## [33,] 34 2.268932 0.10058489 0.89941511
## [34,] 35 2.267299 0.09091723 0.90908277
## [35,] 36 2.265761 0.08207023 0.91792977
## [36,] 37 2.264310 0.07398880 0.92601120
## [37,] 38 2.262937 0.06661949 0.93338051
## [38,] 39 2.261638 0.05991083 0.94008917
## [39,] 40 2.260406 0.05381348 0.94618652
```

So we would need 35 blocks would be needed if a difference of at least 10 units between treatments means should be detected with probability 0.9.