A4Q4

Yuqi Gu

7/23/2024

a.i)

```r
frogs.dat <- read.csv("~/Desktop/STAT 332/Assignments/A4/frogs.csv")

frogs.dat$thickness <- as.factor(frogs.dat$thickness)
frogs.dat$dimensions <- as.factor(frogs.dat$dimensions)
frogs.dat$colour <- as.factor(frogs.dat$colour)
frogs.dat$person <- as.factor(frogs.dat$person)

frogs.aov <- aov(jump~thickness+colour + thickness:colour, data = frogs.dat)
summary(frogs.aov)
```

```
##                   Df Sum Sq Mean Sq F value   Pr(>F)
## thickness          2  25121   12561  10.995 3.19e-05 ***
## colour             1   1235    1235   1.081    0.300
## thickness:colour   2   1585     793   0.694    0.501
## Residuals        174 198779    1142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
qf(0.95, 2, 174) #Test for interaction
```

```
## [1] 3.047906
```

```r
qf(0.95, 2, 174) #Test for main effects of paper thickness
```

```
## [1] 3.047906
```

```r
qf(0.95, 1, 174) #Test for main effects of colour
```

```
## [1] 3.895458
```

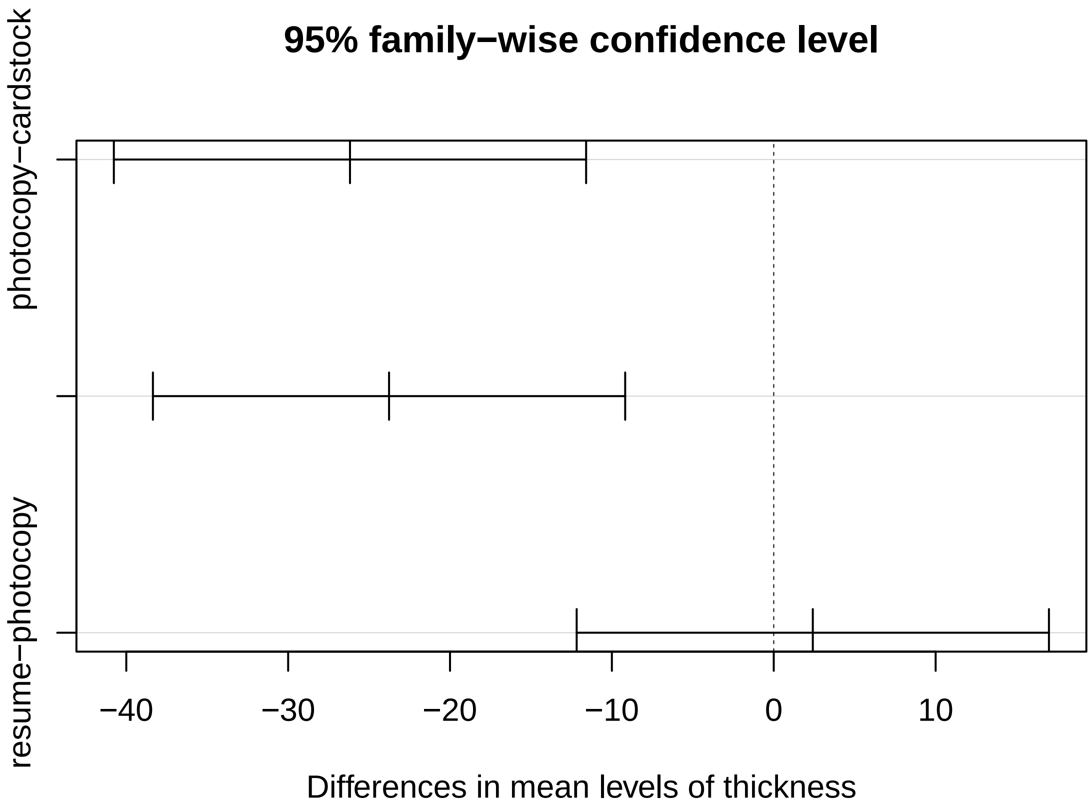
Since $0.694 < 3.047906$ and the corresponding p-value is 0.501, larger than 0.05, we do not reject the null hypothesis. Consequently, the interactions between colour and paper thickness are not significant for jump distance. Thus, the jump distance stay the same if the interactions between colour and paper thickness are different. Since $10.995 > 3.047906$ and the corresponding p-value is 3.19e-05, which is much smaller than 0.05, we reject H0. Thus, we conclude that the effect of paper thickness on jump distance is significantly important. Thus, at the 5% level, the jump distance differs if the paper thickness are different. Since 1.081 < 3.895458 and the corresponding p-value is 0.300, larger than 0.05, we do not reject H0. Thus, we conclude that the effect of colour on jump distance is not significantly important. Thus, at the 5% level, the jump distance stay the same if the paper colours are different.

a.ii)

```r
TukeyHSD(frogs.aov, "thickness", conf.level = 0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = jump ~ thickness + colour + thickness:colour, data = frogs.dat)
##
## $thickness
##                           diff       lwr        upr     p adj
## photocopy-cardstock -26.180000 -40.76771 -11.592286 0.0001059
## resume-cardstock    -23.766667 -38.35438  -9.178953 0.0004821
## resume-photocopy      2.413333 -12.17438  17.001047 0.9191973
```

```r
plot(TukeyHSD(frogs.aov, "thickness"))
```



Cardstock jumps the furthest. Since the difference between photocopy-cardstock is -26.180000, the mean jump distance of photocopy is 26.18 cm shorters than that of cardstock. Since the difference between resume-cardstock is -23.766667, the mean jump distance of resume is 23.766667 cm shorters than that of cardstock. Since the difference between resume-photocopy is 2.413333, the mean jump distance of resume is 2.413333 cm longers than that of photocopy. Moreover, the confidence intervals for photocopy-cardstock and resume-cardstock does not contain 0 and both of them are negative. This means that cardstock jumps further than photocopy and resume. The confidence interval for resume-photocopy contains 0, which means that the jump distance of resume and photocopy do not have differences. These are also shown in the plot of 95% Confidence interval of differences in mean levels of thickness.

a.iii)

```r
tmp <- model.tables(frogs.aov,type = "means")
tmpthickness <- tmp[[1]][[2]]

thickness.tab <- data.frame(names (tmp$tables[[2]]), c(tmp$tables[[2]]) )
colnames(thickness.tab) <- c(names (tmp$tables[2]), "means")
rownames(thickness.tab)<-NULL

MSE <- 1142
J <- 2
n_0 <- 180/(3*2)
lh <- c(75, 97, 181)
ci <- lh-mean(lh)
L_hat <- thickness.tab$means[1]*ci[3]+thickness.tab$means[2]*ci[1]+thickness.tab$means[3]*ci[2]
var <- MSE*((ci[1])^2/(J*n_0)+(ci[2])^2/(J*n_0)+(ci[3])^2/(J*n_0))

t <- qt(0.975, 174)
lwr <- L_hat - t*sqrt(var)
upr <- L_hat + t*sqrt(var)
lwr
```

```
## [1] 926.9865
```

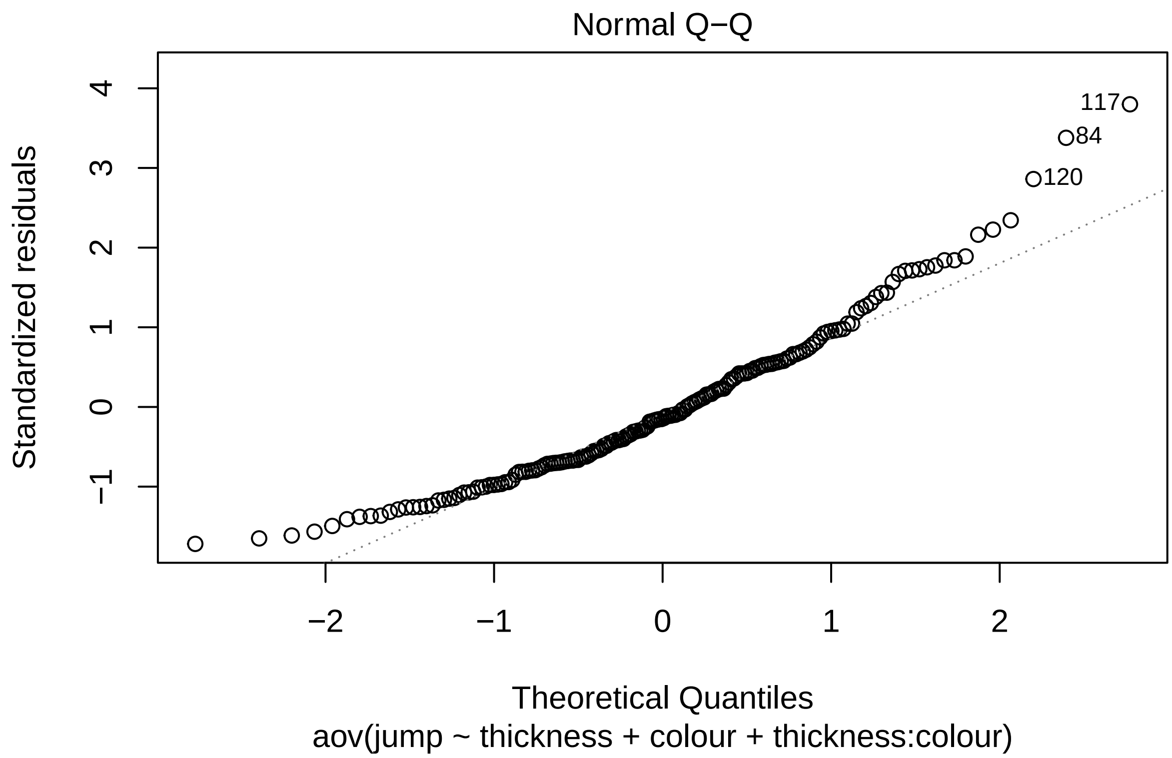
```r
upr
```

```
## [1] 2289.396
```

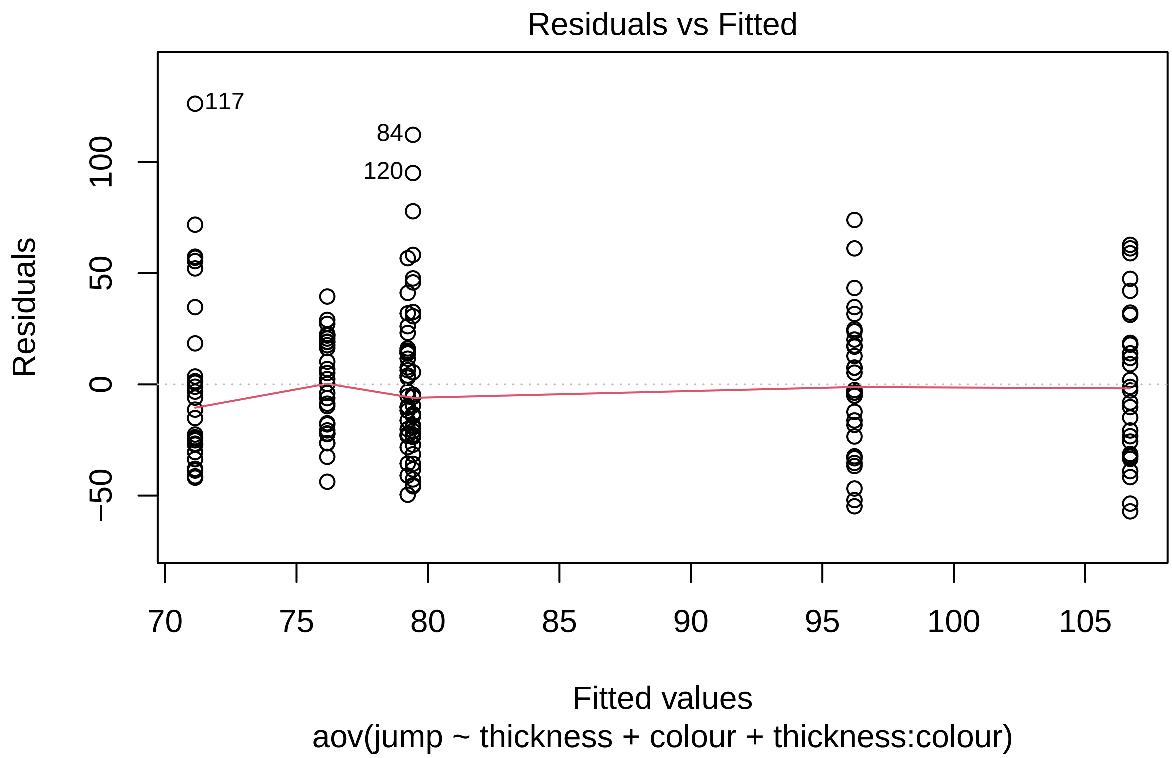
The corresponding 95% CI is [926.9865, 2289.396]. It does not contain 0 and so we reject the null hypothesis that there is no linear relationship in paper thickness. Thus, there is a linear relationship in paper thickness.

a.iv)

```r
plot(frogs.aov, which = 2)
```




```
plot(frogs.aov, which = 1)
```



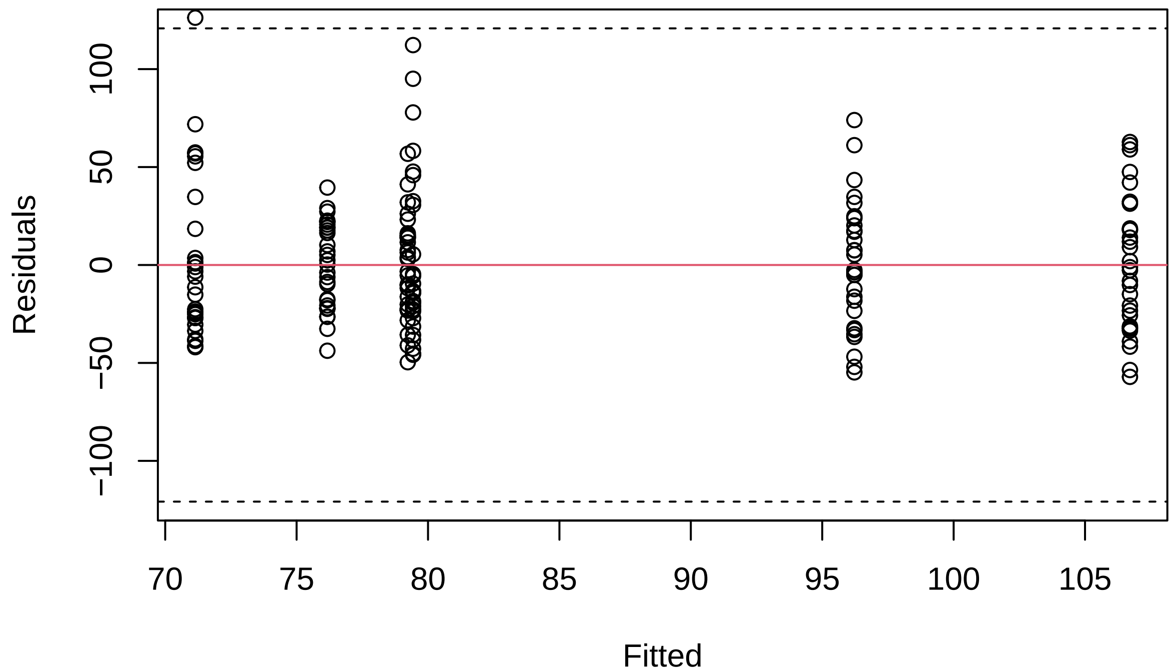
Looking at the QQ plot, the middle of the observations is close to a straight line but the points on the two sides are not perfectly close to the straight line. Thus, the points do not fall (more a less) on a straight line generally. Thus, the assumption of normality is not met perfectly.

Looking at the plot of residuals versus fitted values, since about half the points are above the red line at 0 and the other half are below it, the assumption $E(\epsilon_{ijk}) = 0$ appears respected. The equal variance assumption is also met generally.

```
frogs.res <- data.frame(res=residuals(frogs.aov), fitted = fitted(frogs.aov))

mse <- sum(residuals(frogs.aov)^2/frogs.aov$df)
tmp <- qnorm(1-0.05/(2*180), sd = sqrt(mse*(1-1/(180/3/2))))

plot(res~fitted, data = frogs.res, xlab="Fitted" , ylab = "Residuals",
     ylim = c(-tmp, tmp))
abline(h=0, col=2)
abline(h=tmp, lty=2)
abline(h=-tmp, lty=2)
```



There is one outlier by observing the plot above. Specifically, this outlier is above 120.8025, which is the upper dashed line in the plot.

b.i)

```
frogs.aovii <- aov(jump~thickness+dimensions + thickness:dimensions, data = frogs.dat)
summary(frogs.aovii)
```

```
##                       Df Sum Sq Mean Sq F value   Pr(>F)    
## thickness              2  25121   12561   27.38 4.84e-11 ***
## dimensions             2  95722   47861  104.32  < 2e-16 ***
## thickness:dimensions   4  27428    6857   14.95 1.70e-10 ***
## Residuals            171  78450     459                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qf(0.95, 4, 171) #Test for interaction
```

```
## [1] 2.424502
```

```
qf(0.95, 2, 171) #Test for main effects of paper thickness
```

```
## [1] 3.048833
```

```r
qf(0.95, 2, 171) #Test for main effects of dimensions
```

```
## [1] 3.048833
```

Since 14.95 > 2.424502 and the corresponding p-value is 1.70e-10, which is much smaller than 0.05, we reject the null hypothesis. Consequently, the interactions between dimensions and paper thickness are significant for jump distance. Thus, the jump distance differs if the interactions between dimensions and paper thickness are different. Since 27.38 > 3.048833 and the corresponding p-value is 4.84e-11, which is much smaller than 0.05, we reject H0. Thus, we conclude that the effect of paper thickness on jump distance is significantly important. Thus, at the 5% level, the jump distance differs if the paper thickness are different. Since 104.32 > 3.048833 and the corresponding p-value is < 2e-16, which is much smaller than 0.05, we reject H0. Thus, we conclude that the effect of paper dimensions on jump distance is significantly important. Thus, at the 5% level, the jump distance differs if the paper dimensions are different.

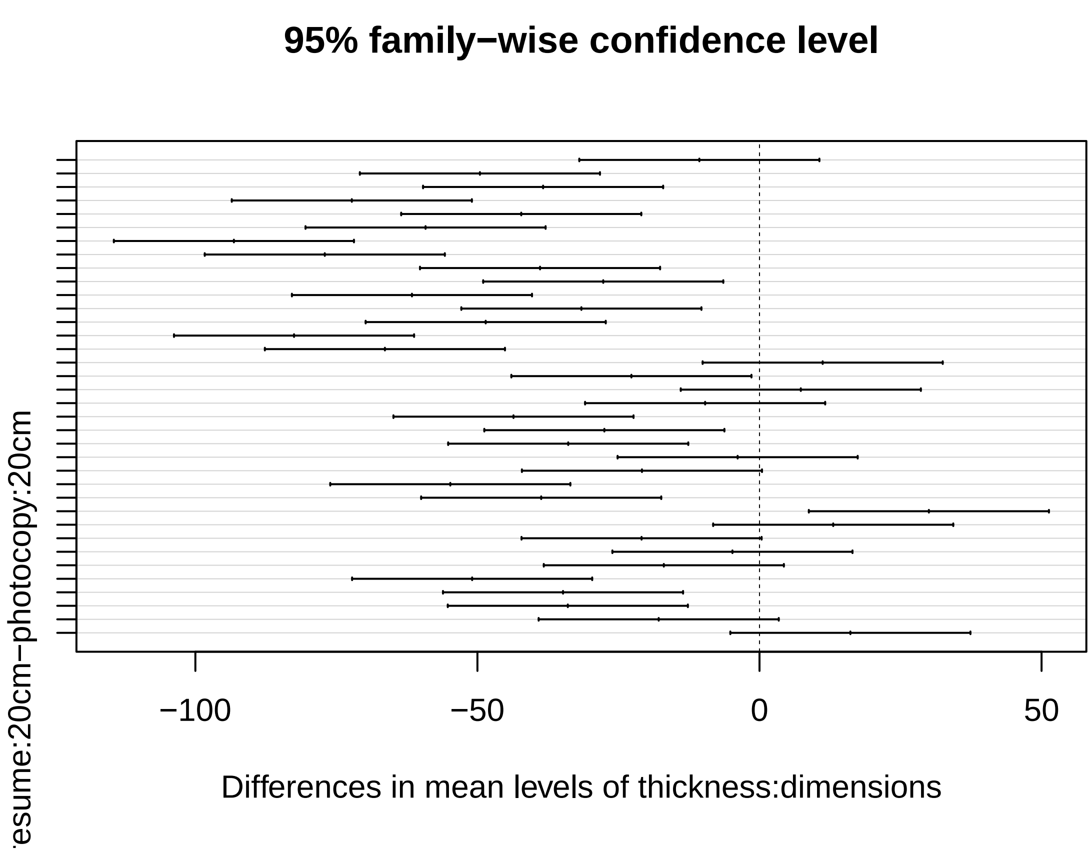
b.ii)

```r
TukeyHSD(frogs.aovii, conf.level = 0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = jump ~ thickness + dimensions + thickness:dimensions, data = frogs.dat)
##
## $thickness
##                             diff        lwr       upr    p adj
## photocopy-cardstock -26.180000 -35.425701 -16.93430 0.000000
## resume-cardstock    -23.766667 -33.012368 -14.52097 0.000000
## resume-photocopy      2.413333  -6.832368  11.65903 0.810949
##
## $dimensions
##                 diff       lwr       upr p adj
## 15cm-10cm -30.88000 -40.12570 -21.63430     0
## 20cm-10cm -56.40167 -65.64737 -47.15597     0
## 20cm-15cm -25.52167 -34.76737 -16.27597     0
##
## $`thickness:dimensions`
##                                   diff         lwr         upr     p adj
## photocopy:10cm-cardstock:10cm -10.660  -31.941273  10.6212728 0.8176589
## resume:10cm-cardstock:10cm    -49.560  -70.841273 -28.2787272 0.0000000
## cardstock:15cm-cardstock:10cm -38.360  -59.641273 -17.0787272 0.0000022
## photocopy:15cm-cardstock:10cm -72.265  -93.546273 -50.9837272 0.0000000
## resume:15cm-cardstock:10cm    -42.235  -63.516273 -20.9537272 0.0000001
## cardstock:20cm-cardstock:10cm -59.195  -80.476273 -37.9137272 0.0000000
## photocopy:20cm-cardstock:10cm -93.170 -114.451273 -71.8887272 0.0000000
## resume:20cm-cardstock:10cm    -77.060  -98.341273 -55.7787272 0.0000000
## resume:10cm-photocopy:10cm    -38.900  -60.181273 -17.6187272 0.0000015
## cardstock:15cm-photocopy:10cm -27.700  -48.981273  -6.4187272 0.0021243
## photocopy:15cm-photocopy:10cm -61.605  -82.886273 -40.3237272 0.0000000
## resume:15cm-photocopy:10cm    -31.575  -52.856273 -10.2937272 0.0002136
## cardstock:20cm-photocopy:10cm -48.535  -69.816273 -27.2537272 0.0000000
```

```
## photocopy:20cm-photocopy:10cm -82.510 -103.791273 -61.2287272 0.0000000
## resume:20cm-photocopy:10cm    -66.400  -87.681273 -45.1187272 0.0000000
## cardstock:15cm-resume:10cm     11.200  -10.081273  32.4812728 0.7734955
## photocopy:15cm-resume:10cm    -22.705  -43.986273  -1.4237272 0.0268596
## resume:15cm-resume:10cm         7.325  -13.956273  28.6062728 0.9761905
## cardstock:20cm-resume:10cm     -9.635  -30.916273  11.6462728 0.8877521
## photocopy:20cm-resume:10cm    -43.610  -64.891273 -22.3287272 0.0000000
## resume:20cm-resume:10cm       -27.500  -48.781273  -6.2187272 0.0023744
## photocopy:15cm-cardstock:15cm -33.905  -55.186273 -12.6237272 0.0000477
## resume:15cm-cardstock:15cm     -3.875  -25.156273  17.4062728 0.9997125
## cardstock:20cm-cardstock:15cm -20.835  -42.116273   0.4462728 0.0601273
## photocopy:20cm-cardstock:15cm -54.810  -76.091273 -33.5287272 0.0000000
## resume:20cm-cardstock:15cm    -38.700  -59.981273 -17.4187272 0.0000017
## resume:15cm-photocopy:15cm     30.030    8.748727  51.3112728 0.0005506
## cardstock:20cm-photocopy:15cm  13.070   -8.211273  34.3512728 0.5942723
## photocopy:20cm-photocopy:15cm -20.905  -42.186273   0.3762728 0.0584323
## resume:20cm-photocopy:15cm     -4.795  -26.076273  16.4862728 0.9986308
## cardstock:20cm-resume:15cm    -16.960  -38.241273   4.3212728 0.2375112
## photocopy:20cm-resume:15cm    -50.935  -72.216273 -29.6537272 0.0000000
## resume:20cm-resume:15cm       -34.825  -56.106273 -13.5437272 0.0000258
## photocopy:20cm-cardstock:20cm -33.975  -55.256273 -12.6937272 0.0000455
## resume:20cm-cardstock:20cm    -17.865  -39.146273   3.4162728 0.1790989
## resume:20cm-photocopy:20cm     16.110   -5.171273  37.3912728 0.3026021
```

```r
plot(TukeyHSD(frogs.aovii, which = "thickness:dimensions"))
```



b.iii)

Based on the result of part b.ii), cardstock:10cm and photocopy:10cm are the groups jump the furthest. By observing the differences of the mean jump distances of photocopy:10cm-cardstock:10cm, resume:10cm-cardstock:10cm, cardstock:15cm-cardstock:10cm, photocopy:15cm-cardstock:10cm, resume:15cm-cardstock:10cm, cardstock:20cm-cardstock:10cm, photocopy:20cm-cardstock:10cm, resume:20cm-cardstock:10cm, we find that all of them are negative. These indicate that the mean jump distance of the group of cardstock:10cm is larger than the other 8 groups. However, the confidence intervals of the second to the eighth are negative and do not contain 0 but the confidence interval in the first line contains 0 and p-value is 0.8176589, which is much larger than 0.05. Thus, the jump distance of cardstock:10cm and photocopy:10cm should not have any differences. Also, by observing the differences, confidence intervals of resume:10cm-photocopy:10cm, cardstock:15cm-photocopy, photocopy:15cm-photocopy:10cm, resume:15cm-photocopy:10cm, cardstock:20cm-photocopy:10cm, photocopy:20cm-photocopy:10cm, resume:20cm-photocopy:10cm, we find that all of them are negative. Thus, photocopy:10cm jumps further than these seven groups. Thus, the jump distance of cardstock:10cm and photocopy:10cm should not have any differences and they both jump further than the other seven groups.

c.i)

```r
frogs.aoviii <- aov(jump~thickness+dimensions + thickness:dimensions + person, data = frogs.dat)
summary(frogs.aoviii)
```

```
##                       Df Sum Sq Mean Sq F value   Pr(>F)    
## thickness              2  25121   12561   29.96 7.14e-12 ***
## dimensions             2  95722   47861  114.17  < 2e-16 ***
## person                 1   7182    7182   17.13 5.48e-05 ***
## thickness:dimensions   4  27428    6857   16.36 2.35e-11 ***
## Residuals            170  71268     419                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
qf(0.95, 4, 170) #Test for interaction
```

```
## [1] 2.424815
```

```r
qf(0.95, 2, 170) #Test for main effects of paper thickness
```

```
## [1] 3.049149
```

```r
qf(0.95, 2, 170) #Test for main effects of dimensions
```

```
## [1] 3.049149
```

After consider person as a blocking factor, we find that since 16.36 > 2.424815 and the corresponding p-value is 2.35e-11, which is much smaller than 0.05, we reject the null hypothesis. Consequently, the interactions between dimensions and paper thickness are significant for jump distance. Thus, the jump distance differs if the interactions between dimensions and paper thickness are different. Since 29.96 > 3.049149 and the corresponding p-value is 7.14e-12, which is much smaller than 0.05, we reject H0. Thus, we conclude that the effect of paper thickness on jump distance is significantly important. Thus, at the 5% level, the jump distance differs if the paper thickness are different. Since 114.17 > 3.049149 and the corresponding p-value is < 2e-16, which is much smaller than 0.05, we reject H0. Thus, we conclude that the effect of paper dimensions on jump distance is significantly important. Thus, at the 5% level, the jump distance differs if the paper dimensions are different.

Thus, they are the same as in b.i).

c.ii)

```
tmpiii <- model.tables(frogs.aoviii, type="mean")
tmpperson <- tmpiii[[1]][[4]]

person.tab <- data.frame(names (tmpiii$tables[[4]]), c(tmpiii$tables[[4]]) )
colnames(person.tab) <- c(names (tmpiii$tables[4]), "means")
rownames(person.tab)<-NULL

MSE <- 419
n <- 90
D_hat <- person.tab$means[1]-person.tab$means[2]
var <- 2*MSE/n

tiii <- qt(0.975, 170)
lwriii <- D_hat - tiii*sqrt(var)
upriii <- D_hat + tiii*sqrt(var)
lwriii
```

```
## [1] -18.65687
```

```
upriii
```

```
## [1] -6.609796
```

The corresponding 95% CI is [-18.65687,-6.609796]. Because it does not contain 0 and the upper bound is negative, we conclude that Christian is not better than Michelle at making the paper frogs jump further.