STATS 201 Lab Class 3

LiYuzhuo-2019210146

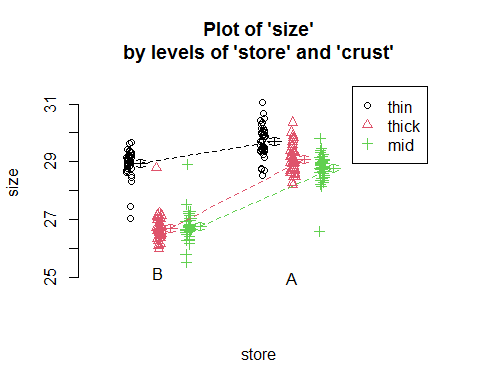
# Pizza size

# Code and output

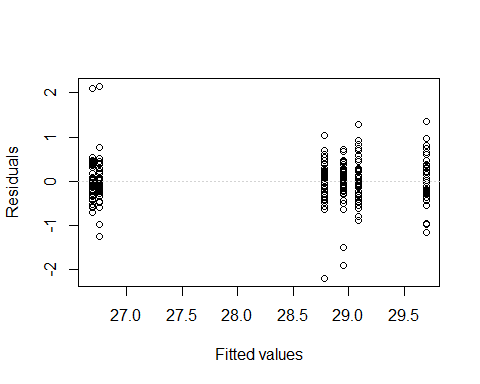
require(s20x)

## Loading required package: s20x

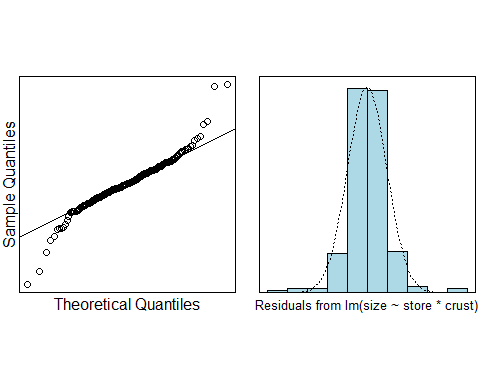
## Loading in the data.  
pizza.df = read.table(file = "pizza.txt", header = TRUE)  
  
pizza.df$store = as.factor(pizza.df$store)  
pizza.df$crust = as.factor(pizza.df$crust)  
  
## Plot the data.  
interactionPlots(size ~ store + crust, pizza.df)



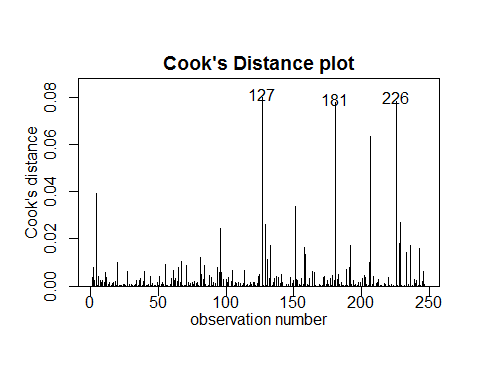
## Analyse the data.  
size.fit = lm(size ~ store \* crust, pizza.df)  
eovcheck(size.fit)



normcheck(size.fit)



cooks20x(size.fit)



summary(size.fit)

##   
## Call:  
## lm(formula = size ~ store \* crust, data = pizza.df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.20209 -0.24224 -0.00099 0.27235 2.14667   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 28.78209 0.07574 380.035 < 2e-16 \*\*\*  
## storeB -2.02876 0.10774 -18.830 < 2e-16 \*\*\*  
## crustthick 0.30721 0.10711 2.868 0.00449 \*\*   
## crustthin 0.91842 0.10982 8.363 4.86e-15 \*\*\*  
## storeB:crustthick -0.37054 0.15333 -2.417 0.01641 \*   
## storeB:crustthin 1.28093 0.15475 8.277 8.54e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4966 on 242 degrees of freedom  
## Multiple R-squared: 0.8487, Adjusted R-squared: 0.8455   
## F-statistic: 271.4 on 5 and 242 DF, p-value: < 2.2e-16

anova(size.fit)

## Analysis of Variance Table  
##   
## Response: size  
## Df Sum Sq Mean Sq F value Pr(>F)   
## store 1 180.957 180.957 733.69 < 2.2e-16 \*\*\*  
## crust 2 123.307 61.653 249.97 < 2.2e-16 \*\*\*  
## store:crust 2 30.455 15.228 61.74 < 2.2e-16 \*\*\*  
## Residuals 242 59.687 0.247   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary2way(size.fit, page = "interaction")

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = fit)  
##   
## $`Comparisons within c("store", "crust")`  
## diff lwr upr p adj  
## A:thick-A:mid 0.30720930 -0.0004735879 0.6148922 0.0506094  
## A:thin-A:mid 0.91841980 0.6029462339 1.2338934 0.0000000  
## B:thick-B:mid -0.06333333 -0.3785250980 0.2518584 0.9924384  
## B:thin-B:mid 2.19934959 1.8861327930 2.5125664 0.0000000  
## A:thin-A:thick 0.61121049 0.2957369316 0.9266841 0.0000010  
## B:thin-B:thick 2.26268293 1.9456216810 2.5797442 0.0000000  
##   
## $`Comparisons between c("store", "crust")`  
## diff lwr upr p adj  
## B:mid-A:mid -2.0287597 -2.338269 -1.7192508 0  
## B:thick-A:thick -2.3993023 -2.712701 -2.0859035 0  
## B:thin-A:thin -0.7478299 -1.066942 -0.4287177 0

# Methods and Assumption Checks

We have two explanatory factors, store and crust, and one numeric response size. The interaction plot indicated different slopes between the levels. So we fitted a two-way ANOVA model with interaction between store and crust. The interaction term was significant .

The residuals were fine, there were no problems with normality and no unduly influential points. We have independence from taking a random sample.

Our model is where store=0 when the pizza was purchased from store A, and 1 for pizza purchased from store B, and .

Our model explained 84.87% of variability in pizza’s size.

# Executive Summary

We are interested in find out whether a store actually made the larger pizzas than their rival’s.

From the function summary2way, we can concluded that when the pizza has the same crust, store A actually make bigger pizzas than store B. When the crust level is mid, thick, thin, the diameter of the pizza purchased from store A is 2.03cm, 2.40cm, 0.75cm larger than pizza purchased from store B.

Obviously, the type of crust is related to the size of the pizza .

And from the function anova, the difference in pizza size between stores is depend on the crust type(the p-value of store:crust is nearly zero).