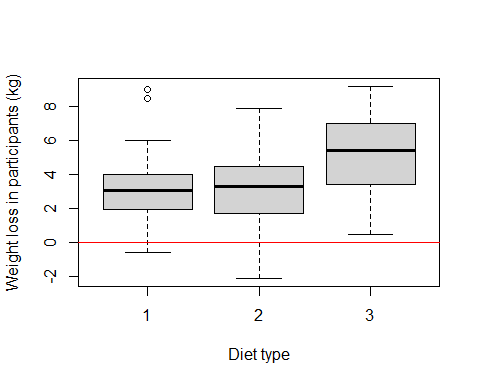
Stats\_Project\_613

Elizabeth Marge

4/25/2021

## Reading into the csv   
  
diet = read.csv("diet\_1.csv")  
diet$weight.loss = diet$initial.weight - diet$final.weight   
diet$gender = factor(diet$gender,levels=c("Female","Male"))  
boxplot(weight.loss~diet.type,data=diet,col="light gray",  
 ylab = "Weight loss in participants (kg)", xlab = "Diet type")  
abline(h=0,col="red")



## Fisher  
## Welch  
## Diet Kruskal   
  
diet.fisher = aov(weight.loss~diet.type,data=diet)  
diet.welch = oneway.test(weight.loss~diet.type,data=diet)  
diet.kruskal = kruskal.test(weight.loss~diet.type,data=diet)  
  
summary(diet.fisher)

## Df Sum Sq Mean Sq F value Pr(>F)   
## diet.type 1 45.8 45.78 7.639 0.00716 \*\*  
## Residuals 76 455.5 5.99   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

print(diet.welch)

##   
## One-way analysis of means (not assuming equal variances)  
##   
## data: weight.loss and diet.type  
## F = 6.0467, num df = 2.000, denom df = 49.893, p-value = 0.004453

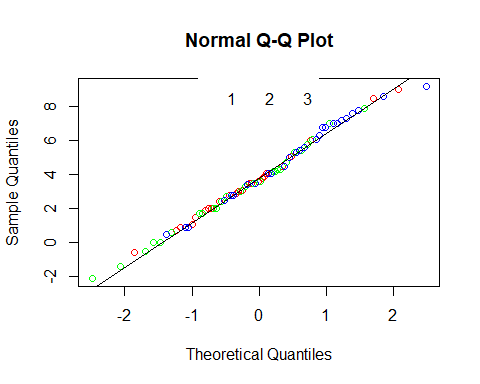
print(diet.kruskal)

##   
## Kruskal-Wallis rank sum test  
##   
## data: weight.loss by diet.type  
## Kruskal-Wallis chi-squared = 10.437, df = 2, p-value = 0.005416

## One-way Anova   
  
summary(aov(weight.loss~diet.type,data=diet))

## Df Sum Sq Mean Sq F value Pr(>F)   
## diet.type 1 45.8 45.78 7.639 0.00716 \*\*  
## Residuals 76 455.5 5.99   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

diet$diet.type = factor(diet$diet.type,levels=c("1","2","3"))  
col\_group = rainbow(nlevels(diet$diet.type))  
qqnorm(diet$weight.loss,col=col\_group[as.numeric(diet$diet.type)])  
qqline(diet$weight.loss)  
legend("top",legend=levels(diet$diet.type),col=col\_group,ncol=3,box.lwd=NA)



## Two-way t-tests   
  
t.test(weight.loss~diet.type,data=diet[diet$diet.type!="1",],var.equal = TRUE)

##   
## Two Sample t-test  
##   
## data: weight.loss by diet.type  
## t = -3.1693, df = 52, p-value = 0.00256  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -3.4658892 -0.7785553  
## sample estimates:  
## mean in group 2 mean in group 3   
## 3.025926 5.148148

t.test(weight.loss~diet.type,data=diet[diet$diet.type!="2",],var.equal = TRUE)

##   
## Two Sample t-test  
##   
## data: weight.loss by diet.type  
## t = -2.8348, df = 49, p-value = 0.006644  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -3.1582988 -0.5379975  
## sample estimates:  
## mean in group 1 mean in group 3   
## 3.300000 5.148148

t.test(weight.loss~diet.type,data=diet[diet$diet.type!="3",],var.equal = TRUE)

##   
## Two Sample t-test  
##   
## data: weight.loss by diet.type  
## t = 0.40798, df = 49, p-value = 0.6851  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -1.075930 1.624079  
## sample estimates:  
## mean in group 1 mean in group 2   
## 3.300000 3.025926

## assigning mean and median   
  
mean = tapply(diet$weight.loss,diet$diet.type,mean)  
median = tapply(diet$weight.loss,diet$diet.type,median)  
mean

## 1 2 3   
## 3.300000 3.025926 5.148148

median

## 1 2 3   
## 3.05 3.30 5.40

## residual mean   
## residuals median  
  
diet$residual.mean = (diet$weight.loss - mean[as.numeric(diet$diet.type)])  
diet$residual.median = (diet$weight.loss - median[as.numeric(diet$diet.type)])  
diet[1:25,]

## Person gender Age Height initial.weight diet.type final.weight weight.loss  
## 1 1 <NA> 22 159 58 1 54.2 3.8  
## 2 2 <NA> 46 192 60 1 54.0 6.0  
## 3 3 <NA> 55 170 64 1 63.3 0.7  
## 4 4 <NA> 33 171 64 1 61.1 2.9  
## 5 5 <NA> 50 170 65 1 62.2 2.8  
## 6 6 <NA> 50 201 66 1 64.0 2.0  
## 7 7 <NA> 37 174 67 1 65.0 2.0  
## 8 8 <NA> 28 176 69 1 60.5 8.5  
## 9 9 <NA> 28 165 70 1 68.1 1.9  
## 10 10 <NA> 45 165 70 1 66.9 3.1  
## 11 11 <NA> 60 173 72 1 70.5 1.5  
## 12 12 <NA> 48 156 72 1 69.0 3.0  
## 13 13 <NA> 41 163 72 1 68.4 3.6  
## 14 14 <NA> 37 167 82 1 81.1 0.9  
## 15 15 <NA> 39 168 71 1 71.6 -0.6  
## 16 16 <NA> 31 158 72 1 70.9 1.1  
## 17 17 <NA> 40 173 74 1 69.5 4.5  
## 18 18 <NA> 50 160 78 1 73.9 4.1  
## 19 19 <NA> 43 162 80 1 71.0 9.0  
## 20 20 <NA> 25 165 80 1 77.6 2.4  
## 21 21 <NA> 52 177 83 1 79.1 3.9  
## 22 22 <NA> 42 166 85 1 81.5 3.5  
## 23 23 <NA> 39 166 87 1 81.9 5.1  
## 24 24 <NA> 40 190 88 1 84.5 3.5  
## 25 25 <NA> 41 171 60 2 60.0 0.0  
## residual.mean residual.median  
## 1 0.500000 0.75  
## 2 2.700000 2.95  
## 3 -2.600000 -2.35  
## 4 -0.400000 -0.15  
## 5 -0.500000 -0.25  
## 6 -1.300000 -1.05  
## 7 -1.300000 -1.05  
## 8 5.200000 5.45  
## 9 -1.400000 -1.15  
## 10 -0.200000 0.05  
## 11 -1.800000 -1.55  
## 12 -0.300000 -0.05  
## 13 0.300000 0.55  
## 14 -2.400000 -2.15  
## 15 -3.900000 -3.65  
## 16 -2.200000 -1.95  
## 17 1.200000 1.45  
## 18 0.800000 1.05  
## 19 5.700000 5.95  
## 20 -0.900000 -0.65  
## 21 0.600000 0.85  
## 22 0.200000 0.45  
## 23 1.800000 2.05  
## 24 0.200000 0.45  
## 25 -3.025926 -3.30

## shapiro test   
  
shapiro.test(diet$residual.mean)

##   
## Shapiro-Wilk normality test  
##   
## data: diet$residual.mean  
## W = 0.99208, p-value = 0.9161