HW # 9 - SOLUTIONS

Ex 13.4.1

```
> x <-c(63,68,79,65,64,63,65,64,76,74,66,66,67,73,69,76)
> wilcox.test(x,mu=70,alternative="less")

        Wilcoxon signed rank test with continuity correction

data:  x
V = 48.5, p-value = 0.1622
alternative hypothesis: true location is less than 70
```

Since p-value is larger than 0.05, we accept the null hypothesis that the mean weight gain is not less than 70 grams

Ex 13.4.2

```
> x <-c(214,362,202,158,403,219,307,331)
> y <-c(232,276,224,412,562,203,340,313)
> wilcox.test(x,y,alternative="less",pair=TRUE)

        Wilcoxon signed rank test with continuity correction

data:  x and y
V = 9.5, p-value = 0.131
alternative hypothesis: true location shift is less than 0
```

Since p-value is larger than 0.05, we accept the null hypothesis that cortisol does not increase after a singing lesson.

Here is the paired t-test on the same data

```
> t.test(x,y,alternative="less",pair=TRUE)

        Paired t-test

data:  x and y
t = -1.1889, df = 7, p-value = 0.1366
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
    -Inf 27.1559
sample estimates:
mean difference
         -45.75
```

Ex 13.5.1

```
> x <-c(99,85,73,98,83,88,99,80,74,91,80,94,94,98,80)
> y <-c(78,74,69,79,57,78,79,68,59,91,89,55,60,55,79)
> library(nonpar)
> mediantest(x = x, y = y, exact=TRUE)

 Exact Median Test

 H0: The 2 population medians are equal.
```

HA: The 2 population medians are not equal.

Significance Level = 0.05
The p-value is 0.000142150287085559

Since the p-value is below 0.05, there is enough evidence to conclude that the population medians are different at a significance level of 0.05.

Ex 13.6.1

```
> hw1361 <- read.csv("C:/Users/dlabate/Desktop/Teaching/ma4310/EXR_C13_S06_01
.csv")
> hw1361$GROUP = factor(hw1361$GROUP)
> wilcox.test(WEIGHT ~ GROUP, alternative = "two.sided", data=hw1361)
        Wilcoxon rank sum test with continuity correction

data:  WEIGHT by GROUP
W = 712.5, p-value = 0.2398
alternative hypothesis: true location shift is not equal to 0
```

Since the p-value is larger than 0.05, there is not sufficient evidence to reject the null hypothesis. Hence, there is no significant difference in weight between the two groups.

REMARK. Here is the result of the t-test, for comparison

```
> t.test(WEIGHT ~ GROUP, alternative = "two.sided", data=hw1361)

        Welch Two Sample t-test

data:  WEIGHT by GROUP
t = 0.94956, df = 67.975, p-value = 0.3457
alternative hypothesis: true difference in means between group 1 and group 2
is not equal to 0
95 percent confidence interval:
 -12.02989  33.87302
sample estimates:
mean in group 1 mean in group 2
       223.8333        212.9118
```

Also in this case, since the p-value is larger than 0.05, there is not sufficient evidence to reject the null hypothesis.

Ex 11.8.1

```
> hw1381 <- read.csv("C:/Users/dlabate/Desktop/Teaching/ma4310/EXR_C13_S08_01
.csv")
> hw1381$GROUP = factor(hw1381$GROUP)
> kruskal.test(B12 ~ GROUP, data=hw1381)
```

```
        Kruskal-Wallis rank sum test

data:  B12 by GROUP
Kruskal-Wallis chi-squared = 11.381, df = 2, p-value = 0.003378
```

```
Since the p-value is less than 0.05, the populations are statistically differ
ent.

We now apply the post-hoc Dunn test
```

```
> library(FSA)
> dunnTest(B12 ~ GROUP, data=hw1381,method="bh")
Dunn (1964) Kruskal-Wallis multiple comparison
  p-values adjusted with the Benjamini-Hochberg method.

  Comparison          Z      P.unadj        P.adj
1      1 - 2 0.6343984 0.525820862 0.525820862
2      1 - 3 2.9404083 0.003277800 0.009833399
3      2 - 3 2.7463942 0.006025432 0.009038147
```

The post-hoc test shows that there is a statistically significant difference between the classes 1-3 and 2-3

Remark: This is the ANOVA test on the same data

```
> anova <- aov(B12 ~ GROUP,data = hw1381)
> summary(anova)
             Df   Sum Sq Mean Sq F value Pr(>F)
GROUP         2   119664   59832    0.64  0.528
Residuals   229 21394206   93424
```

Ex 13.11.1

```
> x <-c(163, 164, 156, 151, 152, 167, 165, 153, 155)
> y <-c(53.9, 57.4, 41.0, 40.0, 42.0, 64.4, 59.1, 49.9, 43.2)
> library(mblm)
> model.k = mblm(y ~ x)
> summary(model.k)

Call:
mblm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-5.7569 -2.2055  0.0000  0.6486  7.1973

Coefficients:
             Estimate      MAD V value Pr(>|V|)
(Intercept) -164.0574   67.7764       0  0.00391 **
x              1.3514    0.4263      45  0.00391 **

Residual standard error: 3.843 on 7 degrees of freedom

> model = lm(y ~ x)
```

```
> summary(model)

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.8611 -2.2362 -0.0612  0.4390  7.0140 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -159.8417    34.4437  -4.641 0.002368 ** 
x              1.3250     0.2172   6.099 0.000491 ***

Residual standard error: 3.839 on 7 degrees of freedom
Multiple R-squared:  0.8416,   Adjusted R-squared:  0.819 
F-statistic:  37.2 on 1 and 7 DF,  p-value: 0.0004914
```

Conclusion:

The Kendall-Theil regression line is: $y = -164.0574 + 1.3514\ x$

The least squares regression line is: $y = -159.8417 + 1.3250\ x$

Math 4310 – Fall 2023 Name: SOLUTION

Quiz #9

Please, write clearly and justify your work to receive credit. You need to report the R command you entered with the complete list of parameters. You also need to report the R output that you used to draw your conclusions.

1) A randomly selected group of singers were the subjects of a study about the possible beneficial effects of singing on well-being during a single singing lesson. The data below report the cortisol level (nmol/L) before and after the singing lesson. Use an appropriate non-parametric method to **test the hypothesis that cortisol increases after a singing lesson**. State the hypothesis testing problem and solve it using $\alpha = 0.05$.

Before:	214	301	221	197	198	205	188	321
After:	232	341	275	205	197	210	188	334

Solution. Let μ_b be the average cortisol level before and μ_a be the average cortisol level after.
We test the hypothesis $H_0 : \mu_b \geq \mu_a$ vs. $H_1 : \mu_b < \mu_a$

Data are paired. We apply the Wilcoxon paired signed-rank test.

```
> before <-c(214, 301, 221, 197, 198, 205, 188, 321)
> after <-c(232, 341, 275, 205, 197, 210, 188, 334)
> wilcox.test(before,after,alternative="less",pair=TRUE)
```

Wilcoxon signed rank test with continuity correction
data: before and after
V = 1, p-value = 0.01731
alternative hypothesis: true location shift is less than 0

Since p-value is less than 0.05, we accept the alternative hypothesis that cortisol increases after a singing lesson.

2) A randomly selected group amateur male singers (Group 1) and a randomly selected group amateur female singers (Group 2) are the subject of a study on cortisol level of male vs female singers. Data below report the cortisol levels measures in group 1 and group 2. Use an appropriate non-parametric method to **test the hypothesis that cortisol level is different in the two groups**. State the hypothesis testing problem and solve it using $\alpha = 0.05$.

Group 1:	214	301	221	197	198	205	188	321
Group 2:	314	205	275	197	232	332	341	339

Solution. Let μ_1 be the average cortisol level in Group 1 and μ_2 be the average cortisol level in Group 2
We test the hypothesis $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$

Data are independent. We apply the two-sample Mann–Whitney U test.

```
> group1 <-c(214, 301, 221, 197, 198, 205, 188, 321)
> group2 <-c(314, 205, 275, 197, 232, 332, 341, 339)
> wilcox.test(group1,group2,alternative="two.sided",pair=FALSE)
```

Wilcoxon rank sum test with continuity correction
data: group1 and group2
W = 16, p-value = 0.1031

alternative hypothesis: true location shift is not equal to 0

Since p-value is greater than 0.05, we accept the null hypothesis that cortisol level is the same in male and female amateur singers.

Remark: Comparison with parametric tests

1) Paired t.test

```
> t.test(before,after,alternative="less",pair=TRUE)
```

Paired t-test
data: before and after
t = -2.4425, df = 7, p-value = 0.0223

Also in this case, we reject the null hypothesis at significance level 0.05

2) t.test

```
> t.test(group1,group2,alternative="two.sided",pair=FALSE)
```

Welch Two Sample t-test
data: group1 and group2
t = -1.738, df = 13.58, p-value = 0.1048

Also in this case, we accept the null hypothesis at significance level 0.05