

```
# IRDR04 Research Tools Module
# Week 5 Tutorial
# 30/10/18 9-11am @ Birkbeck, Malet Street 414/415 Public Cluster
# Tutor: Dr Alison Fairbrass, UCL Centre for Biodiversity & Environment Research, alison.fairbrass@gmail.com
```

```
# # # Walk-through # # #
```

```
### Unpaired t-test
```

```
# Group 1: Evacuees
set.seed(10)
age <- rnorm(n=235, mean=30, sd=11)
age <- round(age)
group <- 1:length(age)
group <- rep("Group 1",length(group))
df1 <- data.frame(age, group)
```

```
# Group 2: Non-evacuees
set.seed(10)
age <- rnorm(n=230, mean=60, sd=12)
age <- round(age)
group <- 1:length(age)
group <- rep("Group 2",length(group))
df2 <- data.frame(age, group)
```

```
# Combine data into dataframe
df <- rbind(df1, df2)
```

```
# Checking distribution of data
hist(df$age[df$group=="Group 1"], main = "Group 1", xlab = "Age")
hist(df$age[df$group=="Group 2"], main = "Group 2", xlab = "Age")
```

```
# Checking if variance of groups is equal
# Using the Levene test
# For this test, the null hypothesis is that all populations variances are equal;
# and the alternative hypothesis is that they differ.
library(car)
leveneTest(age ~ group, data = df)
```

```
# From the output, it can be seen that the p-value of 0.1147 is not less than
# the significance level of 0.05. This means that there is no evidence to suggest
# that the variance in age is statistically significantly different for
# the evacuees and non-evacuees.
```

```
# now use t.test
t.test(df$age[df$group=="Group 2"], df$age[df$group=="Group 1"])
```

```
# Mann-Whitney test
# Investigating differences in men and womens weights
# We want to know, if the median womens weight differs from the median mens weight?
women_weight <- c(38.9, 61.2, 73.3, 21.8, 63.4, 64.6, 48.4, 48.8, 48.5)
men_weight <- c(67.8, 60, 63.4, 76, 89.4, 73.3, 67.3, 61.3, 62.4)
# Create a data frame
df <- data.frame(
  group = rep(c("Woman", "Man"), each = length(women_weight)),
  weight = c(women_weight, men_weight)
)
```

```
# Checking distribution of data
hist(df$weight[df$group=="Woman"], main = "Woman", xlab = "Weight")
hist(df$weight[df$group=="Man"], main = "Men", xlab = "Weight")
```

```
# Use a boxplot to visualise groups
boxplot(weight~group, data = df)
```

```
# Compute Wilcoxon test
wilcox.test(df$weight[df$group=="Woman"], df$weight[df$group=="Man"])
wilcox.test(women_weight, men_weight, exact=F)
```

```
# It will give a warning message, saying that cannot compute exact p-value with ties
# It comes from the assumption of a Wilcoxon test that the responses are continuous.
# You can suppress this message by adding another argument exact = FALSE, but the result
# will be the same.
```

```
# For reporting calculate the median and confidence intervals for the data
sort(women_weight)
women_median <- median(women_weight)
```

```
# Lower CI
n <- length(women_weight)
rank <- (n/2) - ((1.96*(sqrt(n))))/2)
rank <- round(rank)
LowerCI <- sort(women_weight)[rank]
# Upper CI
rank <- 1 + (n/2) + ((1.96*(sqrt(n))))/2)
rank <- round(rank)
UpperCI <- sort(women_weight)[rank]
```

```
paste("Median womens weight: ", women_median)
paste("Lower Confidence Interval: ", LowerCI)
paste("Upper Confidence Interval: ", UpperCI)
```