

## R HELP SHEET: Mann-Whitney U test

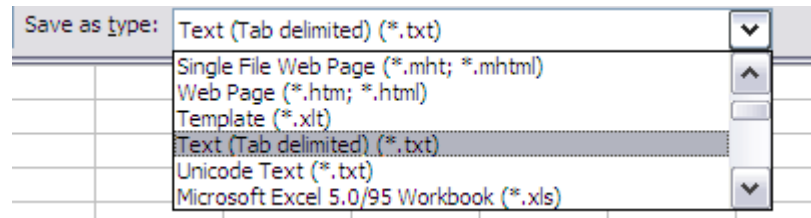
### CONTENTS

1. Creating a tab delimited data file using Excel
2. Conducting a Mann-Whitney U test
3. Identifying the key elements of the output
4. Additional notes

### 1. Creating a tab delimited data file using Excel

Open Excel and type data from your dependent variable into one column and data from your independent variable into another column with appropriate headings at the top of each column (e.g., *bmd*, *sex*), then save the file as a **Text (Table delimited)(\*.txt)** with an appropriate name (e.g., *RData\_Bones*) file using **Save as type**. (Note, the example below has *sex* coded as numbers 1 and 2 but it could also be written in words e.g., *female* and *male*.)

	A	B
1	bmd	sex
2	0.972	1
3	0.732	1
4	0.874	1
5	0.943	1
6	1.024	1
7	0.755	1
8	0.779	1
9	1.007	1
10	0.816	1
11	0.755	1
12	0.871	1
13	0.721	1
14	0.727	1
15	0.796	1
16	0.612	1
17	0.775	1
18	0.849	1
19	0.773	1
20	0.649	1
21	0.865	1
22	0.905	2
23	1.016	2
24	0.873	2



### 2. Conducting a Mann-Whitney U test

The text in green after the hash (#) sign is just **notes** to help you remember what's in the output: it does not get R to actually "do" anything. The text in blue is **R code** with stars representing words that are specific to the example: you need to replace this with text specific to your data as shown in the output in section 3.

To get R to conduct a Mann-Whitney U test:

Open an **R-Editor** window by selecting **File** then **New script**.

Type in (or copy and paste) the notes and code below.

Replace the stars with appropriate text as indicated in notes.

Highlight everything and press **Ctrl R**.

#Importing data from tab delimited file

\*(replace stars with an appropriate object name e.g., bones )

```
****<-read.table(file.choose(),header=TRUE)
```

```
attach(****)
```

```
names(****)
```

#Conducting a Mann-Whitney U test (variable layout) (replace stars with appropriate text e.g., bmd, sex)

```
wilcox.test(***~***, conf.int=TRUE)
```

### 3. Identifying the key elements of the output

Following the instructions above will produce the following output in the **R Console** window: the **key elements** are annotated in orange.

```
> #Importing data from tab delimited file (replace stars with an appropriate object name e.g.,bones )
> bones<-read.table(file.choose(),header=TRUE)
> attach(bones)
> names(bones)
[1] "bmd" "sex"
>
> #Conducting a Mann-Whitney U test (replace stars with sample headings e.g.,bmd,sex)
> wilcox.test(bmd~sex,conf.int=TRUE)

Wilcoxon rank sum test with continuity correction

data: bmd by sex
W = 120.5, p-value = 0.03259
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
-0.130028399 -0.004941922
sample estimates:
difference in location
-0.07041078

Warning messages:
1: In wilcox.test.default(x = c(0.972, 0.732, 0.874, 0.943, 1.024, :
  cannot compute exact p-value with ties
2: In wilcox.test.default(x = c(0.972, 0.732, 0.874, 0.943, 1.024, :
  cannot compute exact confidence intervals with ties
>
Note: Using this route, the U value that R reports is the one associated with the sum of ranks for the second variable on the list so it depends on the order as to whether it's the higher or lower value.
```

In summary the key information from the test is

**U=120.5, n<sub>1</sub>=20, n<sub>2</sub>=20, P=0.032**

And the unstandardized effect size (estimated difference between the populations) is

**-0.070 (female–male; 95%CI [-0.130, -0.005])**

### 4. Additional notes

a.To find the sample sizes n<sub>1</sub> and n<sub>2</sub> use the following code:

**#To find sample sizes**

**tapply(\*\*\*\*,\*\*\*\*,length)**

For example: **tapply(sex,sex,length)**

b.If your data are in “sample layout” (data from dependent variable in different columns with categories of independent variable as column headings) rather than “variable layout” (data from dependent variable in one column and from independent in another) use the following code:

**#Conducting a Mann-Whitney U test (sample layout)**

**wilcox.test(\*\*\*\*,\*\*\*\*,conf.int=TRUE)**

For example: **wilcox.test(female,male,conf.int=TRUE)**

And to find sample sizes use length command.

For example: **length (female)**