# ✪ Psychological Statistics ✪

# Week 05: *Comparing means of 2 groups (Nonparametric)*


- Edited by Prof. **Changwei Wu**
- Graduate Institute of Mind, Brain and Consciousness (***GIMBC***), Taipei Medical University


In [7]:

```r
### [ Loading the required libraries ]

library("pastecs")
library("car")
library("tidyverse")
library("rstatix")
library("effsize")
library("ggpubr")
```

---

## (1) SIGN Test

---

> ### Usage: sign_test {rstatix}

In [28]:

```r
### [ Example 6.8 ]

E608<-data.frame(length=c(24.1, 22.9, 23.0, 26.1, 25.0, 30.8, 27.1, 23.2,
                          22.8, 23.7, 24.6, 30.3, 23.9, 21.8, 28.1, 25.4, 31.2, 30.9

# Question: Are the 18 samples clicking froglets (median length of 30 mm)

E608 %>% sign_test(length ~ 1, mu = 30)

# PASWR::SIGN.test(E608$length, md=30)
```

A rstatix_test: 1 × 7

| | .y. | group1 | group2 | n | statistic | df | p |
|---|---|---|---|---|---|---|---|
| | <chr> | <chr> | <chr> | <int> | <dbl> | <dbl> | <dbl> |
| **1** | length | 1 | null model | 18 | 4 | 18 | 0.0309 |

## Example: the weight of 10 mice before and after the treatment

## *[ Hypothesis ]* The weight of mice is changed after treatment. *(2-tailed)*

- Null hypothesis ***$H_0$ :*** Weight(before) = Weight(after)
- Alternative hyp. ***$H_1$ :*** Weight(before) ≠ Weight(after)

In [38]:

```
### [ Step.1 ] Load data

data("mice2", package = "datarium")
#head(mice2, 3) # wide form

# Transform into long data:
mice2.long <- mice2 %>%
  gather(key = "group", value = "weight", before, after)

head(mice2.long, 3)

mice2.long %>%
  group_by(group) %>%
  get_summary_stats(weight, type = "mean_sd")
```

A data.frame: 3 × 3

| | id | group | weight |
|---|---|---|---|
| | <int> | <chr> | <dbl> |
| **1** | 1 | before | 187.2 |
| **2** | 2 | before | 194.2 |
| **3** | 3 | before | 231.7 |

A tibble: 2 × 5

| group | variable | n | mean | sd |
|---|---|---|---|---|
| <chr> | <chr> | <dbl> | <dbl> | <dbl> |
| after | weight | 10 | 400.04 | 30.087 |
| before | weight | 10 | 200.56 | 20.028 |

In [ ]:

```
### [ Step.2 ] Assumption check

# → No specific assumption of normal distribution
```

In [38]:

```
### [ Step.3 ] Nonparametric 2-sample comparison - Wilcoxon Signed-rank Test

(stat.test <- mice2.long  %>% sign_test(weight ~ group) %>% add_significance())
```

A rstatix_test: 1 × 9

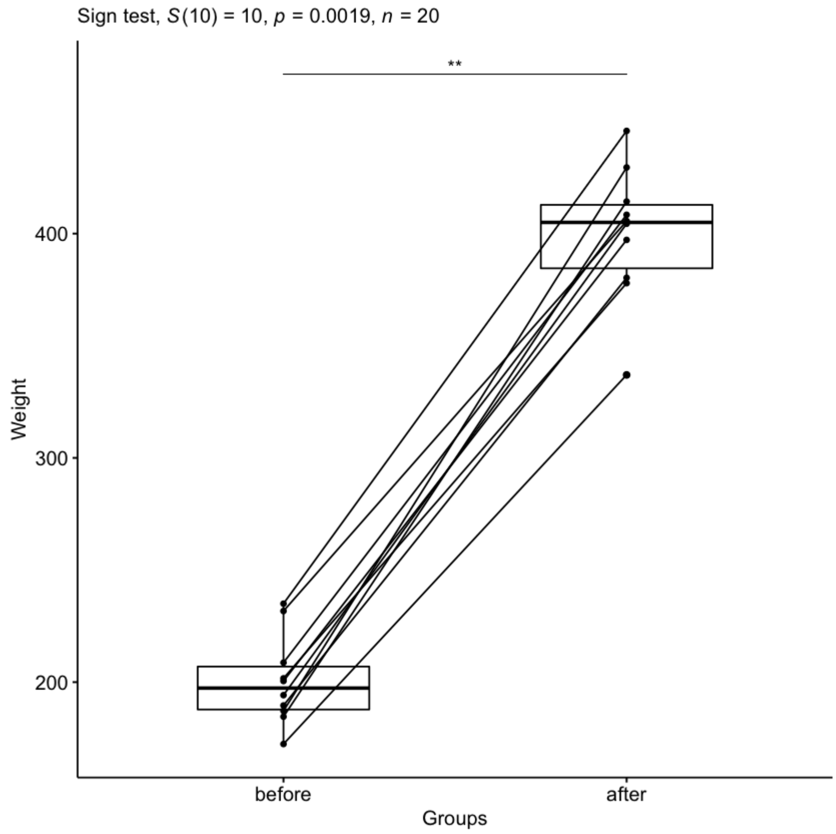
| .y. | group1 | group2 | n1 | n2 | statistic | df | p | p.signif |
|---|---|---|---|---|---|---|---|---|
| <chr> | <chr> | <chr> | <int> | <int> | <dbl> | <dbl> | <dbl> | <chr> |
| weight | after | before | 10 | 10 | 10 | 10 | 0.00195 | ** |

In [ ]:

```
### [ Step.4 ] Effect size

# → no statistics in sign test, so no effect for this test.
```

In [39]:

```
### [ Step.5 ] Visualization & Report

bxp <- ggpaired(mice2.long, x = "group", y = "weight",
         order = c("before", "after"),
         ylab = "Weight", xlab = "Groups")

stat.test <- stat.test %>% add_xy_position(x = "group")
bxp + stat_pvalue_manual(stat.test, tip.length = 0) +
  labs(subtitle = get_test_label(stat.test, detailed= TRUE))
```



### ~ *Report* ~

- The median weight of the mice before treatment (200±20g) is significantly different from the median weight after treatment (400±30g) using ***sign test, p = .002***.

---

# (2) Wilcoxon Signed-Rank Test

---

> ### Usage: wilcox_test {rstatix}

In [32]:

```
### [ Example 6.13 ]
E613<-data.frame(height=c(171, 175, 177, 178, 180, 182, 190, 192, 195, 202))

E613 %>% wilcox_test(height ~ 1, mu=178, alternative = "two.sided")
#E613 %>% wilcox_test(height ~ 1, mu=178, alternative = "greater")
#E613 %>% wilcox_test(height ~ 1, mu=178, alternative = "less", exact=F)
```

A rstatix_test: 1 × 6

| | .y. | group1 | group2 | n | statistic | p |
|---|---|---|---|---|---|---|
| | <chr> | <chr> | <chr> | <int> | <dbl> | <dbl> |
| **1** | height | 1 | null model | 10 | 36 | 0.124 |

In [36]:

```
### [ Example 7.13 ]
E713<-data.frame(height=c(171, 175, 177, 178, 180, 182, 190, 192, 195, 202),
                 armspan=c(173, 178, 182, 182, 188, 185, 186, 198, 193, 202))

E713L <- E713 %>% gather(key = "type", value = "length", height, armspan)

E713L %>% wilcox_test(length ~ type, paired = TRUE)
```

A rstatix_test: 1 × 7

| | .y. | group1 | group2 | n1 | n2 | statistic | p |
|---|---|---|---|---|---|---|---|
| | <chr> | <chr> | <chr> | <int> | <int> | <dbl> | <dbl> |
| **1** | length | armspan | height | 10 | 10 | 38 | 0.0748 |

## Example: Wednesday vs. Sunday (repeated measure)

## *[ Hypothesis ]* Ecstasy users' depression level changes between Wednesday & Sunday. *(2-tailed)*

- Null hypothesis $H_0$ : BDI(Wednesday) = BDI(Sunday)
- Alternative hyp. $H_1$ : BDI(Wednesday) ≠ BDI(Sunday)

In [28]:

```
### [ Step.1 ] Load data

drugData <- read.delim("Drug.dat", header = TRUE)
drugData$BDIchange <- drugData$wedsBDI-drugData$sundayBDI
drugData %>% head(5)
```

A data.frame: 5 × 4

| | drug | sundayBDI | wedsBDI | BDIchange |
|---|---|---|---|---|
| | <fct> | <int> | <int> | <int> |
| 1 | Ecstasy | 15 | 28 | 13 |
| 2 | Ecstasy | 35 | 35 | 0 |
| 3 | Ecstasy | 16 | 35 | 19 |
| 4 | Ecstasy | 18 | 24 | 6 |
| 5 | Ecstasy | 19 | 39 | 20 |

In [30]:

```
### [ Step.2 ] Assumption check

by(drugData$BDIchange, drugData$drug, stat.desc, basic = FALSE, norm = TRUE)
leveneTest(drugData$BDIchange, drugData$drug)

alcoholData<-subset(drugData, drug == "Alcohol")
ecstasyData<-subset(drugData, drug == "Ecstasy")
```

```
drugData$drug: Alcohol
      median         mean      SE.mean CI.mean.0.95          var
std.dev
 -7.50000000  -6.30000000   2.09788253   4.74573999  44.01111111   6.6
3408706
    coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   nor
mtest.W
 -1.05302969   1.23907117   0.90174219   0.98664006   0.36973617   0.8
2795980
  normtest.p
  0.03161929
------------------------------------------------------------
drugData$drug: Ecstasy
      median         mean      SE.mean CI.mean.0.95          var
std.dev
  14.0000000   12.4000000    2.5307004    5.7248420   64.0444444    8.
0027773
    coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   nor
mtest.W
   0.6453853   -0.4140842   -0.3013525   -1.3686700   -0.5128991    0.
9087803
  normtest.p
   0.2727175
```

A anova: 2 × 3

| | Df | F value | Pr(>F) |
|---|---|---|---|
| | <int> | <dbl> | <dbl> |
| **group** | 1 | 0.7931655 | 0.3848943 |
| | 18 | NA | NA |

In [31]:

```
### [ Step.3 ] Repeated-measure Wilcoxon test

(ecstasyModel<-wilcox.test(ecstasyData$wedsBDI, ecstasyData$sundayBDI, paired = TRUE
t.test(ecstasyData$wedsBDI, ecstasyData$sundayBDI, paired = TRUE)

#(alcoholModel<-wilcox.test(alcoholData$wedsBDI, alcoholData$sundayBDI, paired = TRU
```

```
Warning message in wilcox.test.default(ecstasyData$wedsBDI, ecstasyDat
a$sundayBDI, :
“cannot compute exact p-value with ties”
Warning message in wilcox.test.default(ecstasyData$wedsBDI, ecstasyDat
a$sundayBDI, :
“cannot compute exact p-value with zeroes”


        Wilcoxon signed rank test with continuity correction

data:  ecstasyData$wedsBDI and ecstasyData$sundayBDI
V = 36, p-value = 0.01403
alternative hypothesis: true location shift is not equal to 0


        Paired t-test

data:  ecstasyData$wedsBDI and ecstasyData$sundayBDI
t = 4.8998, df = 9, p-value = 0.0008478
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  6.675158 18.124842
sample estimates:
mean of the differences
                   12.4
```

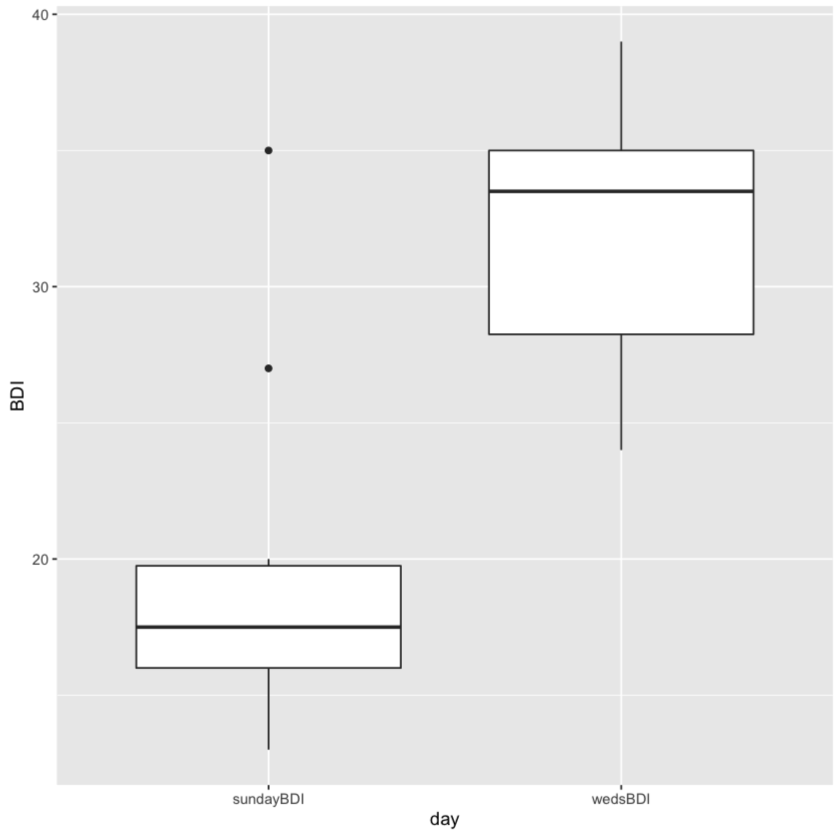
In [32]:

```
### [ Step.4 ] Effect size

rFromWilcox<-function(wilcoxModel, N){
  z<- qnorm(wilcoxModel$p.value/2)
  r<- z/ sqrt(N)
  cat(wilcoxModel$data.name, "Effect Size, r = ", r)
}

rFromWilcox(ecstasyModel, 20)
#rFromWilcox(alcoholModel, 20)
```

```
ecstasyData$wedsBDI and ecstasyData$sundayBDI Effect Size, r =  -0.549
2942
```

In [36]:

```
### [ Step.5 ] Visualization & Report

ecstasy.long <- ecstasyData %>%
  gather(key = "day", value = "BDI", sundayBDI, wedsBDI)
#ecstasy.long %>% head(5)

ggplot(ecstasy.long, aes(day, BDI)) +
geom_boxplot()
```




### *~ Report ~*

- For ecstasy users, depression levels were significantly higher on Wednesday (median=33.5) than on Sunday (median=17.5), ***p = .012, effect size(r) = 0.56***.

---

# (3) Wilcoxon Rank-Sum Test (Mann-Whitney U Test)

---

## Example 1: Ecstasy vs. Alcohol (independent measure)

## *[ Hypothesis ]* Depression level of ecstasy users is different from depression level of alcohol bingers. *(2-tailed)*

- Null hypothesis $H_0$ : BDI(Ecstasy) = BDI(Alcohol)
- Alternative hyp. $H_1$ : BDI(Ecstasy) ≠ BDI(Alcohol)

In [37]:

```
### [ Step.1 ] Load data

drugData<-read.delim("Drug.dat", header = TRUE)
#drugData %>% head(5)

# Transform into long data:
drug.long <- drugData %>%
  gather(key = "day", value = "BDI", sundayBDI, wedsBDI)
drug.long %>% head(5)

drug.long %>% group_by(drug) %>%
    get_summary_stats(BDI, type = "mean_sd")
```

A data.frame: 5 × 3

| | drug | day | BDI |
|---|---|---|---|
| | <fct> | <chr> | <int> |
| 1 | Ecstasy | sundayBDI | 15 |
| 2 | Ecstasy | sundayBDI | 35 |
| 3 | Ecstasy | sundayBDI | 16 |
| 4 | Ecstasy | sundayBDI | 18 |
| 5 | Ecstasy | sundayBDI | 19 |

A tibble: 2 × 5

| drug | variable | n | mean | sd |
|---|---|---|---|---|
| <fct> | <chr> | <dbl> | <dbl> | <dbl> |
| Alcohol | BDI | 20 | 13.25 | 6.544 |
| Ecstasy | BDI | 20 | 25.80 | 8.483 |

In [19]:

```
### [ Step.2 ] Assumption check

#----- (a) Outliers -----#
drug.long %>% group_by(drug) %>%
    identify_outliers(BDI)

#----- (b) Normality -----#
drug.long %>% group_by(drug) %>%
    shapiro_test(BDI)

#----- (c) Homogeneity -----#
drug.long %>% levene_test(BDI ~ drug)
```

A data.frame: 0 × 5

| drug | day | BDI | is.outlier | is.extreme |
|---|---|---|---|---|
| <fct> | <chr> | <int> | <lgl> | <lgl> |

A tibble: 2 × 4

| drug | variable | statistic | p |
|---|---|---|---|
| <fct> | <chr> | <dbl> | <dbl> |
| Alcohol | BDI | 0.9445645 | 0.29200160 |
| Ecstasy | BDI | 0.9122415 | 0.07033191 |

A tibble: 1 × 4

| df1 | df2 | statistic | p |
|---|---|---|---|
| <int> | <int> | <dbl> | <dbl> |
| 1 | 38 | 2.724389 | 0.1070684 |

In [21]:

```
### [ Step.3 ] Mann-Whitney U test

(TwoSamp.test <- drug.long %>%
     t_test(BDI ~ drug, paired = FALSE) %>% add_significance())

(TwoSamp.test <- drug.long %>%
     wilcox_test(BDI ~ drug, paired = FALSE) %>% add_significance())
```

A rstatix_test: 1 × 9

| .y. | group1 | group2 | n1 | n2 | statistic | df | p | p.signif |
|---|---|---|---|---|---|---|---|---|
| <chr> | <chr> | <chr> | <int> | <int> | <dbl> | <dbl> | <dbl> | <chr> |
| BDI | Alcohol | Ecstasy | 20 | 20 | -5.238569 | 35.70097 | 7.39e-06 | **** |

A rstatix_test: 1 × 8

| .y. | group1 | group2 | n1 | n2 | statistic | p | p.signif |
|---|---|---|---|---|---|---|---|
| <chr> | <chr> | <chr> | <int> | <int> | <dbl> | <dbl> | <chr> |
| BDI | Alcohol | Ecstasy | 20 | 20 | 53 | 7.21e-05 | **** |

In [23]:

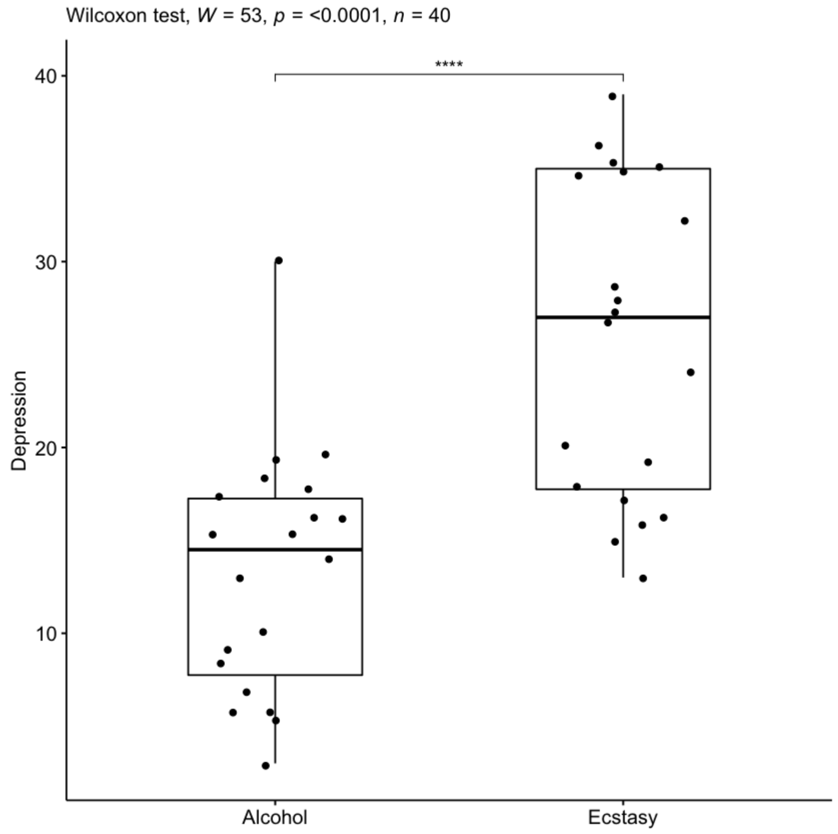
```
### [ Step.4 ] Effect size

drug.long %>% wilcox_effsize(BDI ~ drug)
```

A rstatix_test: 1 × 7

| | .y. | group1 | group2 | effsize | n1 | n2 | magnitude |
|---|---|---|---|---|---|---|---|
| | <chr> | <chr> | <chr> | <dbl> | <int> | <int> | <ord> |
| 1 | BDI | Alcohol | Ecstasy | 0.6297233 | 20 | 20 | large |

In [24]:

```
### [ Step.5 ] Visualization & Report

TwoSamp.test <- TwoSamp.test %>% add_xy_position(x="drug")

ggboxplot(drug.long, x="drug", y="BDI", width = .5,
                  add = "jitter", ylab = "Depression", xlab = FALSE)+
    stat_pvalue_manual(TwoSamp.test, tip.length=0.01)+
    labs(subtitle = get_test_label(TwoSamp.test, detailed = TRUE))
```




## ~ *Report* ~

- Depression levels in ecstasy users (Mean±SD=25.8±8.5) were significantly more depressed than alcohol users (Mean±SD=13.3±6.5), U = 53, p < .001, effect size = 0.63.

## Example 2: Ozone in Gardens (in pphm)

## *[ Hypothesis ]* The two gardens have different ozone concentrations.

***(2-tailed)***

- Null hypothesis $\boldsymbol{H_0}$ **:** Ozone(Garden 1) = Ozone(Garden 2)
- Alternative hyp. $\boldsymbol{H_1}$ **:** Ozone(Garden 1) ≠ Ozone(Garden 2)

In [39]:

```
### [ Step.1 ] Load data

ozone.data <- read.csv("ozone.csv")
Garden<-reshape2::melt(ozone.data, variable.names(names(ozone.data)), value.name="oz
names(Garden)[1]="label"
Garden %>% head(5)
```

A data.frame: 5 × 2

| | label | ozone |
|---|---|---|
| | <fct> | <int> |
| **1** | gardenA | 3 |
| **2** | gardenA | 4 |
| **3** | gardenA | 4 |
| **4** | gardenA | 3 |
| **5** | gardenA | 2 |

In [40]:

```
### [ Step.2 ] Assumption check

#----- Normality -----#
Garden %>% group_by(label) %>%
    shapiro_test(ozone)

#----- Homogeneity -----#
Garden %>% levene_test(ozone ~ label)
```

A tibble: 2 × 4

| label | variable | statistic | p |
|---|---|---|---|
| <fct> | <chr> | <dbl> | <dbl> |
| gardenA | ozone | 0.9528768 | 0.7025892 |
| gardenB | ozone | 0.9528768 | 0.7025892 |

A tibble: 1 × 4

| df1 | df2 | statistic | p |
|---|---|---|---|
| <int> | <int> | <dbl> | <dbl> |
| 1 | 18 | 1.980957e-32 | 1 |

In [44]:

```
### [ Step.3 ] Mann-Whitney U test

(TwoSamp.test <- Garden %>%
     wilcox_test(ozone ~ label, paired = FALSE) %>% add_significance())

#(TwoSamp.test <- Garden %>%
#      t_test(ozone ~ label, paired = FALSE) %>% add_significance())

### [ Step.4 ] Effect size

Garden %>% wilcox_effsize(ozone ~ label)
```

A rstatix_test: 1 × 8

| .y. | group1 | group2 | n1 | n2 | statistic | p | p.signif |
|---|---|---|---|---|---|---|---|
| <chr> | <chr> | <chr> | <int> | <int> | <dbl> | <dbl> | <chr> |
| ozone | gardenA | gardenB | 10 | 10 | 11 | 0.00299 | ** |

A rstatix_test: 1 × 7

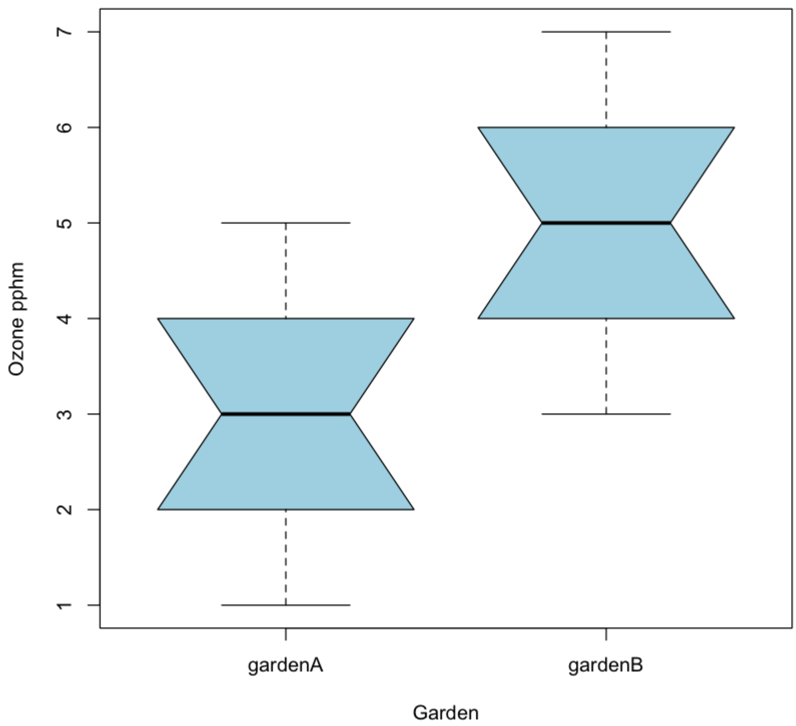
| | .y. | group1 | group2 | effsize | n1 | n2 | magnitude |
|---|---|---|---|---|---|---|---|
| | <chr> | <chr> | <chr> | <dbl> | <int> | <int> | <ord> |
| **1** | ozone | gardenA | gardenB | 0.672498 | 10 | 10 | large |

In [43]:

```
### [ Step.5 ] Visualization

boxplot(ozone~label, notch=T, xlab="Garden", ylab="Ozone pphm",col="lightblue")

Garden %>% group_by(label) %>%
    get_summary_stats(ozone, type = "mean_sd")
```

A tibble: 2 × 5

| label | variable | n | mean | sd |
|---|---|---|---|---|
| <fct> | <chr> | <dbl> | <dbl> | <dbl> |
| gardenA | ozone | 10 | 3 | 1.155 |
| gardenB | ozone | 10 | 5 | 1.155 |




## ~ *Report* ~

- Ozone levels in gardenB (Mean±SD=5±1.2) were significantly higher than gardenA (Mean±SD=3±1.2), * **U = 11, p = .003, effect size = 0.67***.