Data2x02 Study Sheet

# Wilcoxon rank sum test

### Assumptions

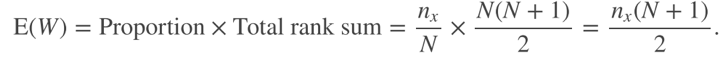
* Two samples from distinct populations
* Same distribution but differ by some shift (diff in means)

### Ranks

**Wilcoxon signed-rank test**, the ranks are summed over positive side of the differences

For two sample **Wilcoxon rank-sum test**, the ranks are summed over one of the samples.

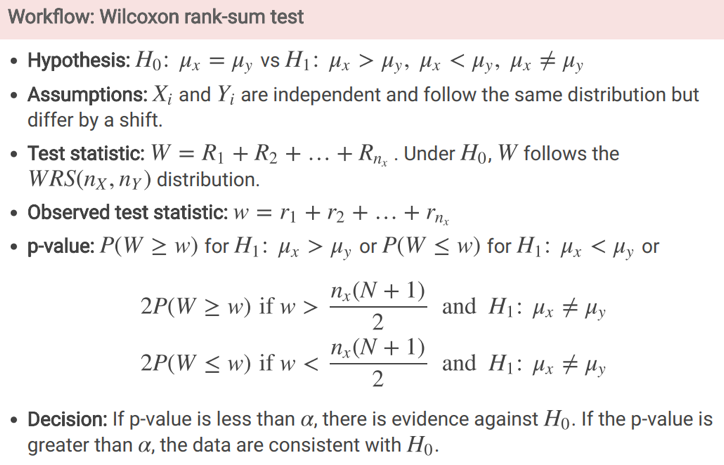
* Rank from smallest to largest
* For tied ranks, rank them the average of what the ranks would have been
* Sum ranks of y or x pop to find test statistic
* You compare this to the expected value to find the p value



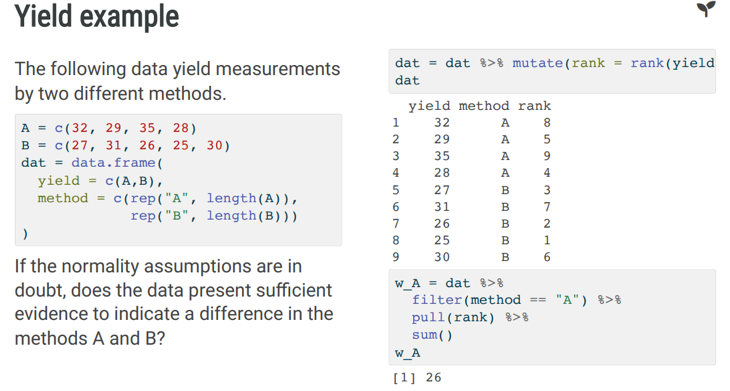
nx = the num of values in sample x

N = number of values in both populations

### Workflow



### Example



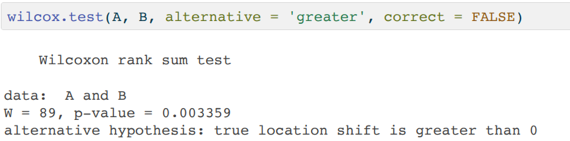
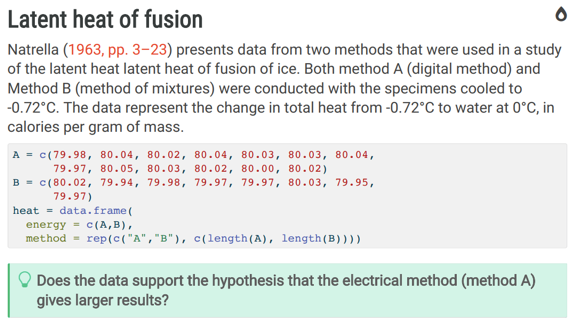
# 

In practice you would just use the wilcox.test function:

Wilcox.test(A, B)

* By default, its looking in the lower tail

### Example 2



# Permutation Test

### Assumption

* Normality
* Weight against group from dataset dat
* Var.equal=TRUE gives us traditional t test instead of welch’s

# Graphical user interface, text, application, email Description automatically generated

Result of t.test stored in variable, using that we can access teststatistic, pvalue etc, conf.int etc.

### Process:

Text

Description automatically generated

* The sample function will sample from the data group
* Then t\_null[i] – a list will store the t test results

Distribution of t\_null test statistics

Chart

Description automatically generated

* Because for this specific example we are checking for any change, we are looking in both tails, therefore we take the abs value
* What proportion of test statistics are >=(more extreme than) the test statistic we found earlier and stored in a variable

Graphical user interface, text, application

Description automatically generated

The permutation test is shuffling all your data, sampling from it and then generating a probability that your test statistic is as or more extreme than it

# Bootstrapping

## Confidence Intervals

### Estimation

* A population parameter is unknown
* Use the sample statistics to generate confidence interval

### Hypothesis Testing

* Hypothesis regarding population parameter
* Test statistics are generated which will support or reject the hypothesis

### General info on Confidence intervals

* We should avoid reporting just a point estimate for a sample
* We should always include a measure of variability

Text

Description automatically generated

* The point estimate of a parameter does not show its variability across samples
* To show such estimation precision, we should find an *interval estimate*

Text

Description automatically generated

Suppose a 95% confidence interval for the mean -mu is (a,b)

### DOES NOT MEAN:

* 95% of means are in (a,b), that is P(a<mu<b) = 0.95 since mu is a fixed but unknown parameter.
* P(a<Xbar<b) =0.95 where Xbar is the sample mean since the CI is for the true mean mu not the sample mean Xbar

### DOES MEAN:

* If we draw a large number of random samples and compute for each sample a 95% CI, about 95% of these CI’s will contain mu
* It can be described as a **range of plausible values** for the population parameter

### Distributional Assumptions

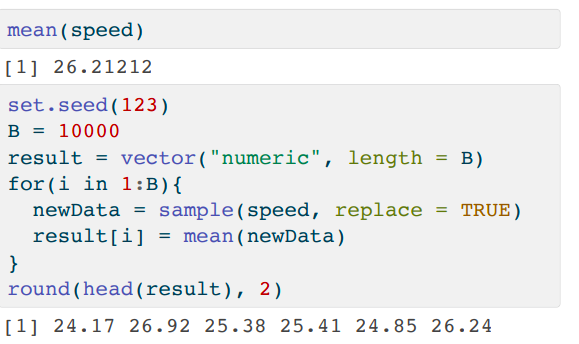
Q: What happens if your data does not follow a normal distribution?

A: Use **Bootstrap Resampling** to empirically model the distribution of the data

### Bootstrapping

Bootstrapping allows us to make inferences about the population where no information is available about the population

The classic approach to bootstrapping is to **repeatedly resample** from the population with replacement

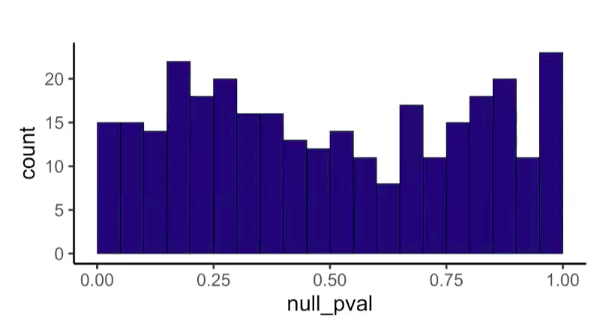


Chart, histogram

Description automatically generated

# Multiple Testing

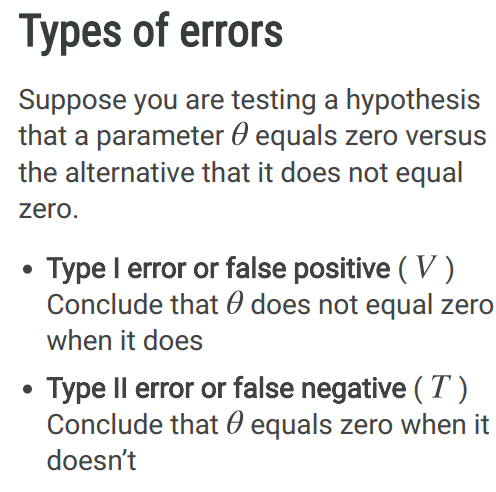
* If there is no association between x and y, you would expect to see uniformly distributed p values



* If there is a relationship, expect a distribution that reflects it

How can we tell if something is actually significant or if it is significant by chance?

* For every experiment at alpha= 0.05, we accept there is a 5% chance we reject the null hypothesis when it is true



Table

Description automatically generated

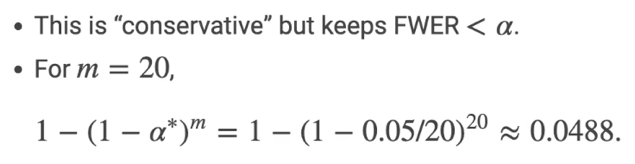
Graphical user interface, text, application, email

Description automatically generated

The **Bonferroni correction** is the oldest multiple testing correction.

Given that the number of false positives for **m** tests is **m x alpha** then consider defining a new threshold for significance:Graphical user interface, text, application

Description automatically generated A picture containing diagram

Description automatically generated

# ANOVA

Anova stands for analysis of variance and are a type of t test where we compare the means of three groups (usually a control and two test groups)

## Example

Chart, box and whisker chart

Description automatically generated

Workflow:

Graphical user interface, text, application

Description automatically generated

How is it different

* So far, our two sample t tests have been a one sample t test on the differences between two samples

Text

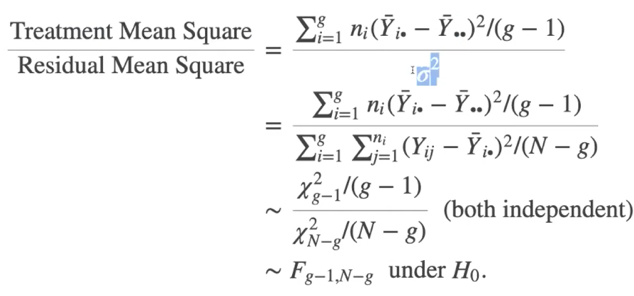
Description automatically generated

Change in Subscript notation:

* Subscript i indexes groups
* Subscript j indexes individuals
* SUBij is the subscript for group I individual j

General

* Follow normal distribution
* Same variance
* The test statistic is called the F statistic:



Graphical user interface, text

Description automatically generated

Numeric variable against categorical variable (weight ~ group)

Graphical user interface, text, application

Description automatically generatedGraphical user interface, text

Description automatically generated with medium confidence

# ANOVA CONTRASTS