Nonparametric testing

1. Sign test
2. Wilcoxon signed-rank test
3. Mann-Whitney test
4. Kruskal-Wallis test

Basics of Health Intelligent Data Analysis

PhD Programme in Health Data Science

Cláudia Camila Dias

Pedro Pereira Rodrigues

Hypothesis test

With confidence intervals we can infer about a parameter in the population based on an estimate of that parameter.

Hypothesis tests are based on another different (but related) approach.

The idea now is to measure how much the results observed in the sample are compatible with a **hypothesis** about the population.

Paired Sample T-test

With two paired samples of individuals we want to know if the means of the two groups in the population are equal.

Assumption

The difference variable is usually distributed in the population.

But, what if it is not?

Nonparametric test - Paired Sample

Suppose the following example:

- 20 obese patients participated in a dietary assessment study

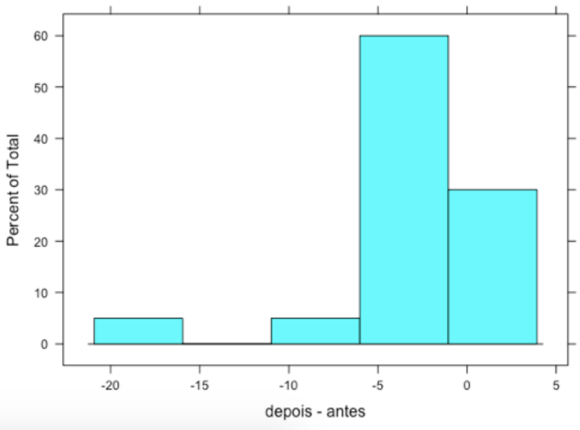
- Each individual was weighed before starting the diet and 3 months after starting the diet.

- The results were as follows:

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Before	103	104	141	146	147	106	107	110	109	114	116	117	122	114	118	125	125	118	125	125
After	100	95	121	142	143	101	102	105	109	109	113	112	117	117	117	125	125	120	122	122

Nonparametric test - Paired Sample

Is the difference variable (before and after) normally distributed across the population?



No, then :

- We need to use a nonparametric test

- Signal test and Wilcoxon test ate nonparametric tests applied to paired samples (such as the t test for paired samples)

Nonparametric test - Paired Sample

- Returning to the example, let's first check on each individual if there was weight loss or gain.

Patient ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Before	103	104	141	146	147	106	107	110	109	114	116	117	122	114	118	125	125	118	125	125
After	100	95	121	142	143	101	102	105	109	109	113	112	117	117	117	125	125	120	122	122
Sign	-	-	-	-	-	-	-	-	=	-	-	-	-	+	-	=	=	+	-	-

- 15 patients lost weight (sign -) and 2 gained weight (sign +)

- If the diet had no effect we should observe a similar number of weight increases and losses

Nonparametric test - Paired Sample

- So if the diet has no effect (H0) what is the probability of seeing in 20 individuals a result as (or more extreme than) 15 losses (sign -) and 2 increases (sign +)?

```
> binom.test(2, 17, p=0.5)

        Exact binomial test

data:  2 and 17
number of successes = 2, number of trials = 17, p-value = 0.00235
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.01457932 0.36440916
sample estimates:
probability of success
             0.1176471
```

This test is called the **sign test**

At a significance level of 0.05 we reject the null hypothesis and conclude that diet has effect.

Nonparametric test - Paired Sample

- **Sign test** ignores much of the information as it only takes into account whether weight loss or weight gain has occurred.

- Wouldn't it be interesting for a test to consider whether those who lost weight lost as many pounds as those who gained weight?

Signed-rank Wilcoxon test

Wilcoxon rank test

- For each individual we will now record the weight difference (absolute) and the sign of the difference (+ if it was weight gain, - if it was weight loss)

Patient ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Before	103	104	141	146	147	106	107	110	109	114	116	117	122	114	118	125	125	118	125	125
After	100	95	121	142	143	101	102	105	109	109	113	112	117	117	117	125	125	120	122	122
|B-A|	**3**	**9**	**20**	**4**	**4**	**5**	**5**	**5**	**0**	**5**	**3**	**5**	**5**	**3**	**1**	**0**	**0**	**2**	**3**	**3**
Sign	-	-	-	-	-	-	-	-	=	-	-	-	-	+	-	=	=	+	-	-

Wilcoxon rank test

- We sort all the differences (excluding the zeros) and record the rank of each

|A-D|	1	2	3	3	3	3	3	4	4	5	5	5	5	5	5	9	20
Rank	**1**	**2**	**5**	**5**	**5**	**5**	**5**	**8.5**	**8.5**	**12.5**	**12.5**	**12.5**	**12.5**	**12.5**	**12.5**	**14**	**15**
Sign	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-

- In the case of ties we give the average position. For example there are 5 individuals with difference (positive or negative) of 3kg. These would correspond to positions 3, 4, 5, 6 and 7; as they are ties, we assign them the average position (5th).

Wilcoxon rank test

|A-D|	1	2	3	3	3	3	3	4	4	5	5	5	5	5	5	9	20
Rank	**1**	**2**	**5**	**5**	**5**	**5**	**5**	**8.5**	**8.5**	**12.5**	**12.5**	**12.5**	**12.5**	**12.5**	**12.5**	**14**	**15**
Sign	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-

- The sum of the ranks of the 15 individuals who lost weight (sign -) is:

  1+5+5+5+5+8.5+8.5+12.5+12.5+12.5+12.5+12.5+12.5+14+15 = **146**

  i.e., the weight losses (sign -) have a mean rank =146/15 = **9.73**

- The sum of the ranks of the 2 patients who gain weight is: 2+5= **7**

  i.e., the weight increases (sign +) have a mean rank of 7/2 = **3.5**

If the diet had no effect one would expect to have an identical average rank for weight losses and increases!

Wilcoxon rank test

- The p value is then calculated as the probability of observing a difference as (or more extreme than) 9.73 vs 3.5 if the diet had no effect.

```
> wilcox.test(depois - antes, exact=F, correct=F)

        Wilcoxon signed rank test

data:  depois - antes
V = 7, p-value = 0.0009128
alternative hypothesis: true location is not equal to 0
```

Independent Sample T-test

With two independent samples of individuals we want to know if the means of the two groups in the population are equal.

Assumption:

The difference variable is usually distributed in the population.

But, what if it is not?

Nonparametric test: ***Mann-Whitney U test***

Mann-Whitney test

If the groups are not paired we can use Mann-Whitney U test

The values are sorted regardless of the group to which they belong. The sum of ranks, R1, is calculated for one of the groups

- The following statistics are calculated

$$U = n_A n_B + \frac{n_A (n_A + 1)}{2} - R_1.$$

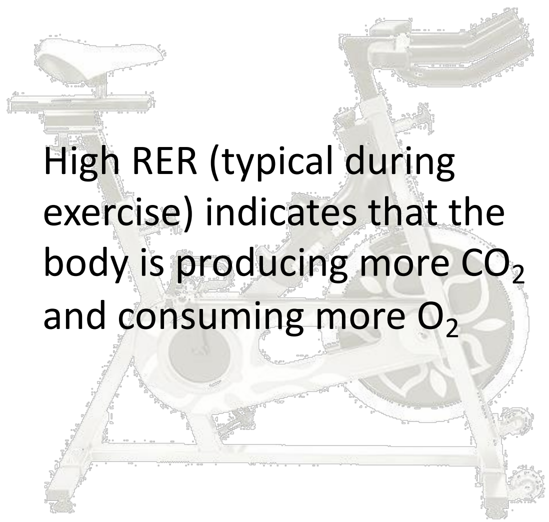
If $U > \frac{n_A n_B}{2}$, we use the statistic $U' = n_A n_B - U$

Using a normal approximation, we obtain the p value associated with the statistic.

Mann-Whitney test - example

The aim is to study the effect of caffeine on muscle metabolism as measured by the ratio of Co2 and O2 (RER) exchanges.

Placebo %RER
105
119
100
97
96
101
94
95
98



High RER (typical during exercise) indicates that the body is producing more CO_2 and consuming more O_2

Caffeine %RER
96
99
94
89
96
93
88
105
88

Mann-Whitney test - example

- As in the Wilcoxon test we will sort all values by assigning their rank and recording which group they belong to (P-placebo or C-caffeine)

RER	88	88	89	93	94	94	95	96	96	96	97	98	99	100	101	105	105	119
Rank	1.5	1.5	3	4	5.5	5.5	7	9	9	9	11	12	13	14	15	16.5	16.5	18
Group	C	C	C	C	P	C	P	P	C	C	P	P	C	P	P	P	C	P

- If there was no effect from caffeine the C's and P's in the table should be mixed.

- If there is an effect, C's should concentrate on lower ranks and P's on higher ranks.

Mann-Whitney test - example

What is the probability of observing a distribution of P's and C's as (or more extreme than) this if caffeine has no effect?

H_0: There are no differences in the (RER) of the two groups.

```
> wilcox.test(rer ~ grupo, exact=F, correct=F)

        Wilcoxon rank sum test

data:  rer by grupo
W = 18, p-value = 0.04615
alternative hypothesis: true location shift is not equal to 0
```

One-Way ANOVA

With more than two independent samples of individuals we want to know if the means of the groups in the population are equal.

Assumption:

The variable is usually distributed in the population.

The variances are equal in all groups in the population.

But, what if it is not?

Nonparametric Test: **Kruskal-Wallis Test**

Kruskal-Wallis test

Example: Weights in Kg of 3 groups of individuals from different ethnic groups.

Group 1: 72; 75; 73; 67; 76; 71; 71; 70; 78; 64

Group 2: 64; 74; 63; 69; 70; 62; 69; 65; 68; 73

Group 3: 58; 59; 61; 63; 66; 53; 68; 69; 61; 57

All values are arranged in ascending order so that each value has an assigned position.

group	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**2**	**3**	**1**	**2**	**2**	**3**	**1**	**2**	**...**
weight	53	57	58	59	61	61	62	63	63	64	64	65	66	67	68	...

...	**3**	**2**	**2**	**3**	**1**	**2**	**1**	**1**	**1**	**1**	**2**	**2**	**1**	**1**	**1**
...	68	69	69	69	70	70	71	71	72	73	73	74	75	76	78

Kruskal-Wallis test

The statisitc is calculated

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i}{n_i} - 3(N+1)$$

N = total number of individuals
$\mathbf{n_i}$ = total number of individuals in group i
$\mathbf{R_i}$ = sum of positions in group I

which follows a Chi-square distribution with k-1 degrees of freedom.

Kruskal-Wallis test

grupo	peso	ordem
3	53	1
3	57	2
3	58	3
3	59	4
3	61	5.5
3	61	5.5
2	62	7
2	63	8.5
3	63	8.5
1	64	10.5
2	64	10.5
2	65	12
3	66	13
1	67	14
2	68	15.5
...	...	...

...	...	...
3	68	15.5
2	69	18
2	69	18
3	69	18
1	70	20.5
2	70	20.5
1	71	22.5
1	71	22.5
1	72	24
1	73	25.5
2	73	26
2	74	27
1	75	28
1	76	29
1	78	30

Group 3 :
(1+2+3+4+5.5+5.5+8.5+13+15.5+18)/10 = 7.6

```
> kruskal.test(peso ~ grupo)

        Kruskal-Wallis rank sum test

data:  peso by grupo
Kruskal-Wallis chi-squared = 14.758, df = 2, p-value = 0.0006242
```