Analysis of Quantitative data
Non parametric statistics

Anne Segonds-Pichon

v2020-09

Comparison between 2 groups
Non-Parametric data

Non-parametric test:

Mann-Whitney = Wilcoxon rank test

- Non-parametric equivalent of the *t*-test (and not).
- Not meeting the assumptions for parametric tests is not enough to switch to a non-parametric approach.
  - Like always, data exploration is key.

- **How does the Mann-Whitney test work?**

Group 1	Group 2
5	8
7	9
3	6

→

Real values	Ranks
3	1
5	2
6	3
7	4
8	5
9	6

→

	Group 1	Group 2
	2	5
	4	6
	1	3
Sum	7	14

- Statistic of the Mann-Whitney test: **U (W)**
  U_1 = 7-6 = 1 and **U_2** = 14-6 = 8
  - Smallest of the 2 Us: U_1 + sample size → **p-value**

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2}$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2}$$

Where:
- R = sum of ranks
- n = sample size.

<u>Exercise</u>: smelly T-shirt.xlsx

- Hypothesis: Group body odour is less disgusting when associated with an in-group member versus an out-group member.
- Study: Two groups of Cambridge University students are presented with one of two smelly, worn T-shirts with university logos.

- **Question**: Can Cambridge students tell the difference between worn smelly T-shirts from Oxford or Cambridge?

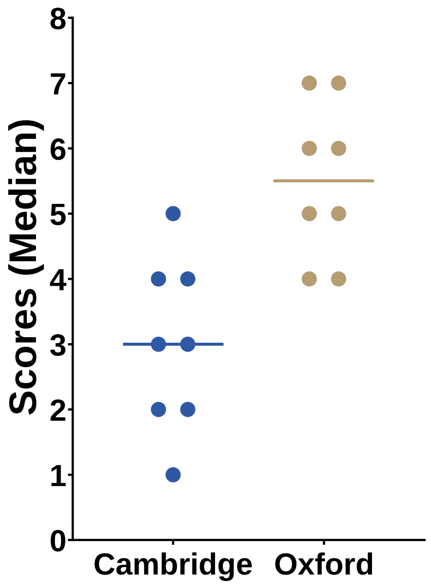
Disgust score: 1 to 7, with 7 the most disgusting

  - Explore the data with an appropriate combination of 2 graphs
  - Answer the question with a non-parametric approach
  - What do you think about the design?

Exercise: smelly T-shirt.xlsx

- **Question**: Can Cambridge students tell the difference between worn smelly T-shirts from Oxford or Cambridge?

Disgust score: 1 to 7, with 7 the most disgusting



	Mann-Whitney test	
1	Table Analyzed	smelly teeshirt
2		
3	Column B	Oxford
4	vs.	vs.
5	Column A	Cambridge
6		
7	Mann Whitney test	
8	P value	0.0037
9	Exact or approximate P value?	Exact
10	P value summary	**
11	Significantly different (P < 0.05)?	Yes
12	One- or two-tailed P value?	Two-tailed
13	Sum of ranks in column A,B	41 , 95
14	Mann-Whitney U	5
15		

Answer:

- Cambridge students can tell the difference between Oxford and Cambridge ($U = 5$, $p = 0.0037$).
- A paired design would have been better.

Non-parametric test:

Wilcoxon's signed-rank

- Non-parametric equivalent of the paired *t*-test (ish).
- **How does the test work?**

2+3=5/2=2.5: average rank

Before	After	Differences
9	3	-6
7	4	-3
10	4	-6
8	5	-3
5	6	1
8	2	-6
7	7	0
9	4	-5
10	5	-5

Abs. Diff.	Ranking	Ranks
0		
1	1	1
3	2	2.5
3	3	2.5
5	4	4.5
5	5	4.5
6	6	7
6	7	7
6	8	7

	Negative ranks	Positives ranks
		1
	-2.5	
	-2.5	
	-4.5	
	-4.5	
	-7	
	-7	
	-7	
Sum	-35	1

- Statistic of the Wilcoxon's signed-rank test: Sum of signed ranks = **W**
  - Here: W = -35 + 1 = -34
  - Statistic W + sample size → **p-value**

Exercise: botulinum.xlsx

	Before	After
1	9	3
2	7	4
3	10	4
4	8	5
5	9	6
6	8	2
7	7	4
8	9	4
9	10	5

A group of 9 disabled children with muscle spasticity (or extreme muscle tightness limiting movement) in their right upper limb underwent a course of injections with botulinum toxin to reduce spasticity levels.
A neurologist (blinded) assessed levels of spasticity pre- and post-treatment for all 9 children using a 10-point ordinal scale.

Higher ratings indicated higher levels of spasticity.

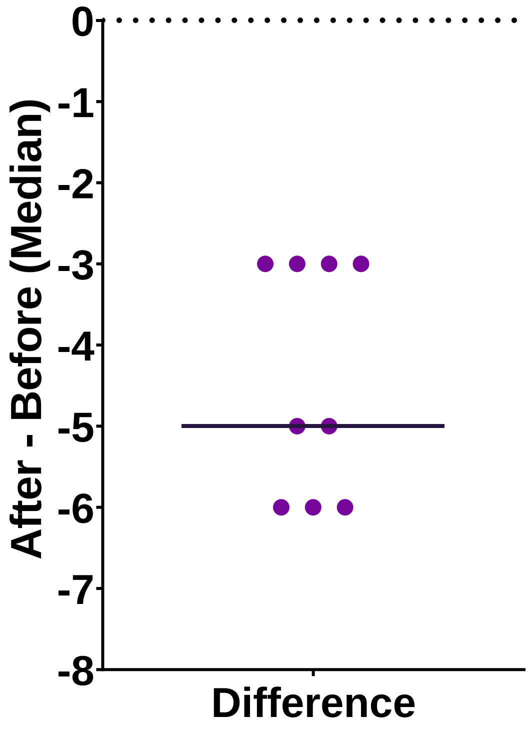
- **Question**: do botulinum toxin injections reduce muscle spasticity levels?
    - Score: 1 to 10, with 10 the highest spasticity

Exercise: botulinum.xlsx

	Before	After
1	9	3
2	7	4
3	10	4
4	8	5
5	9	6
6	8	2
7	7	4
8	9	4
9	10	5

- **Question**: do botulinum toxin injections reduce muscle spasticity levels?

	Wilcoxon test	
1	Table Analyzed	botulinum
2		
3	Column B	after
4	vs.	vs.
5	Column A	before
6		
7	Wilcoxon matched-pairs signed rank test	
8	P value	0.0039
9	Exact or approximate P value?	Exact
10	P value summary	**
11	Significantly different (P < 0.05)?	Yes
12	One- or two-tailed P value?	Two-tailed
13	Sum of positive, negative ranks	0 , -45
14	Sum of signed ranks (W)	-45
15	Number of pairs	9



Answer: There was a significant difference pre- and post- treatment in ratings of muscle spasticity (W = -45, p = 0.0039).

Comparison between more than 2 groups
One factor
Non-Parametric data

Kruskal-Wallis and Friedman tests

- Non-parametric equivalents of the One-Way ANOVA
  - Also based on ranks
  - **Kruskal-Wallis**: independent measures
  - **Friedman**: repeated measures

- Statistic associated with **Kruskal-Wallis** is **H**
- Statistic associated with **Friedman** is **Q** or **T1** or **FM**
- The statistics have a Chi^2 distribution
  - Kruskal-Wallis H = Friedman statistic = One-Way ANOVA F

- Post-hoc test associated with Kruskal-Wallis and Friedman: **Dunn's test**
  - The Dunn's test works pretty much like the Mann-Whitney test.

Kruskal-Wallis test: Example

- Creatine, a supplement popular among body builders
- Three groups: No creatine; Once a day; and Twice a day.
- **Question**: does the average weight gain depend on the creatine group to which people were assigned?

Kruskal-Wallis

Example: creatine.xlsx

Actual values

No	Once	Twice
63	0	2239
-261	-652	171
-153	4724	40
-13	-2	1395
965	0	
	-86	

Ranks

No	Once	Twice
10	7.5	14
2	1	11
3	15	9
5	6	13
12	7.5	
	4	
32	**41**	**47**

$$H = \left[\frac{12}{n(n+1)} \sum_{j=1}^{c} \frac{T_j^2}{n_j}\right] - 3(n+1)$$

$$\mathbf{H} = \left[\frac{12}{15(15+1)}\left(\frac{32^2}{5} + \frac{41^2}{6} + \frac{47^2}{4}\right)\right] - 3(15+1) = \mathbf{3.868}$$

Where:

•n = sum of sample sizes for all samples,

•c = number of samples,

•T_j = sum of ranks in the j^{th} sample,

•n_j = size of the j^{th} sample.

Friedman test: Example

- An auction house is putting three violins, A, B, and C, up for bidding. Ten violinists are blindfolded are asked to rate the instruments and each player plays the violins in a randomly determined sequence (BCA, ACB, etc.).
- After each violin is played, the violinist rates the instrument on a 10-point scale of overall excellence (1=lowest, 10=highest).
- **<u>Question</u>**: which violin is the best according to the 10 violinists?

Friedman test

Example: violin.xlsx

Actual values

Violinists	Violin A	Violin B	Violin C
1	9	7	6
2	9.5	6.5	8
3	5	7	4
4	7.5	7.5	6
5	9.5	5	7
6	7.5	8	6.5
7	8	6	6
8	7	6.5	4
9	8.5	7	6.5
10	6	7	3

Ranks

Violinists	Violin A	Violin B	Violin C
1	3	2	1
2	3	1	2
3	2	3	1
4	2.5	2.5	1
5	3	1	2
6	2	3	1
7	3	1.5	1.5
8	3	2	1
9	3	2	1
10	2	3	1
Sum	77.5	67.5	57

$$\text{Q or T1 or FM} = \frac{n(k-1)\left[\sum_{i=1}^{k}\frac{R_i^2}{n} - C_F\right]}{\sum r_{ij}^2 - C_F}$$

$$C_F = \left(\frac{1}{4}\right)nk(k+1)^2$$

Where:

- n = sum of sample sizes for all samples,
- k = number of samples,
- R_j = sum of ranks in the j^{th} sample,
- r_{ij} = rank i of the j^{th} sample.

Kruskal-Wallis and Friedman tests

- **Have a go!**

Exercise: creatine.xlsx

- **Question**: does the average weight gain depend on the creatine group to which people were assigned?

Exercise: violin.xlsx

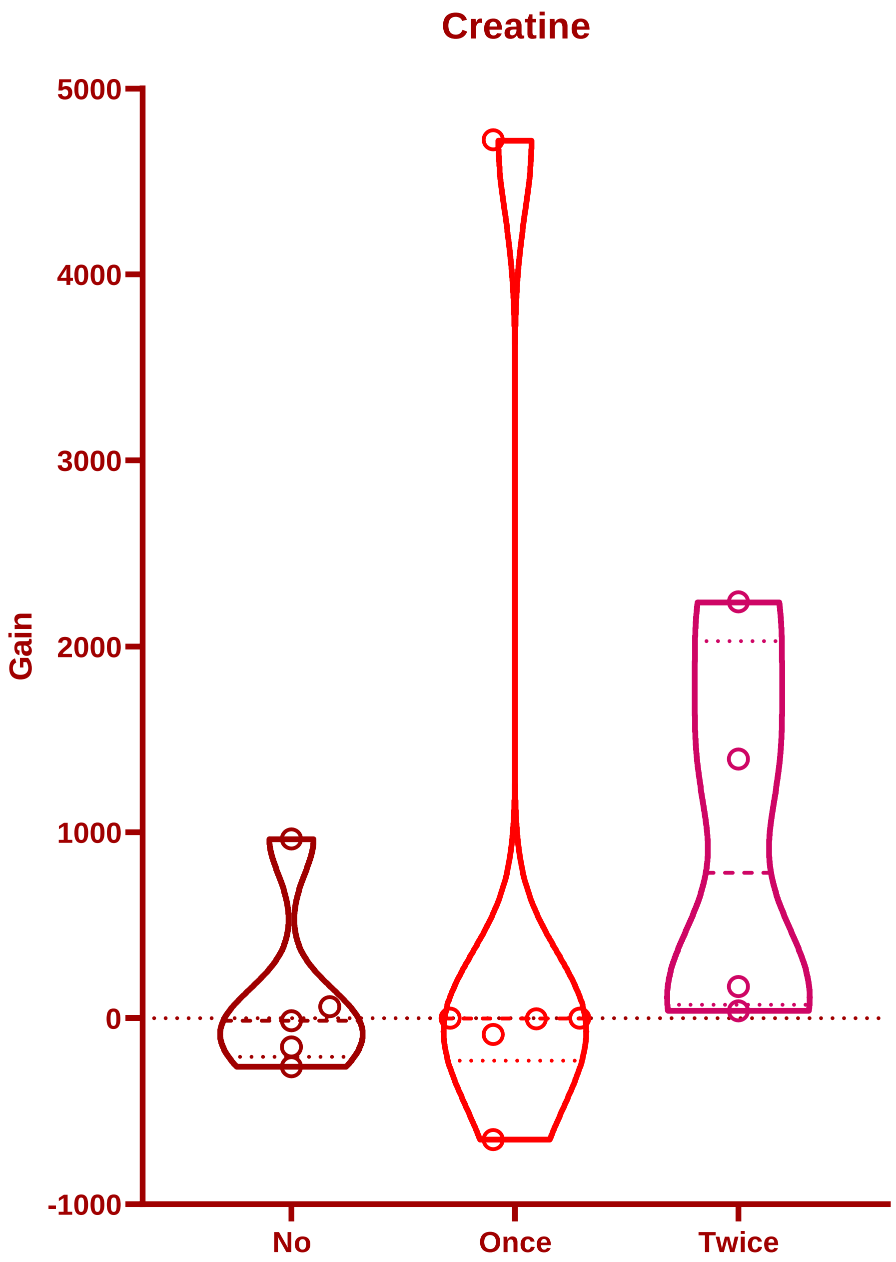
- **Question**: which violin is the best according to the 10 violinists?

Kruskal-Wallis

Example: creatine.xlsx

Results

	Kruskal-Wallis test ANOVA results	
1	Table Analyzed	Creatine
2		
3	**Kruskal-Wallis test**	
4	P value	0.1458
5	Exact or approximate P value?	Exact
6	P value summary	ns
7	Do the medians vary signif. (P < 0.05)?	No
8	Number of groups	3
9	Kruskal-Wallis statistic	3.868
10		
11	**Data summary**	
12	Number of treatments (columns)	3
13	Number of values (total)	15
14		

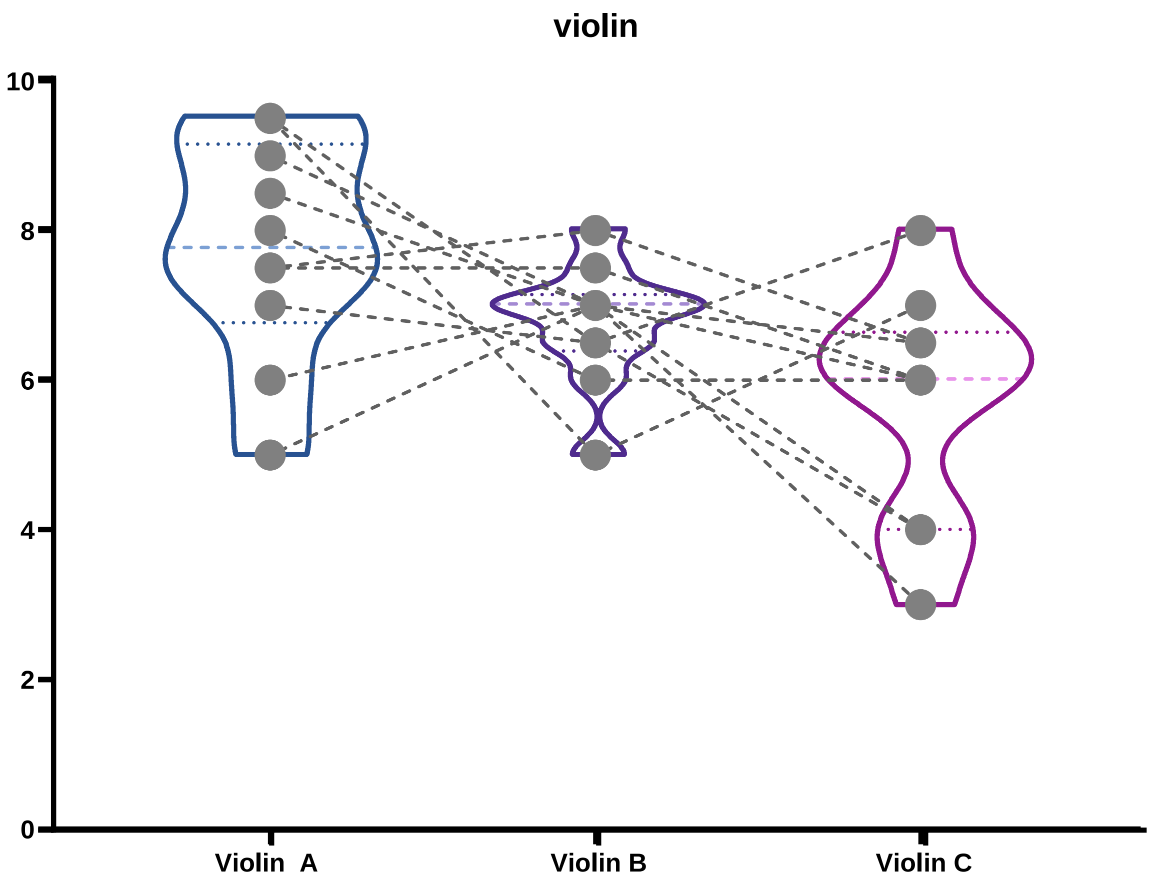


Friedman

Example: violin.xlsx

Results

Table Analyzed	violin
Friedman test	
P value	0.0033
Exact or approximate P value?	Exact
P value summary	**
Are means signif. different? (P < 0.05)	Yes
Number of groups	3
Friedman statistic	10.47
Data summary	
Number of treatments (columns)	3
Number of subjects (rows)	10



Dunn's multiple comparisons test	Rank sum diff.	Significant?	Summary	Adjusted P Value
Violin A vs. Violin B	5.500	No	ns	0.6563
Violin A vs. Violin C	14.00	Yes	**	0.0052
Violin B vs. Violin C	8.500	No	ns	0.1720

Association between 2 continuous variables

Linear relationship

Non-Parametric data

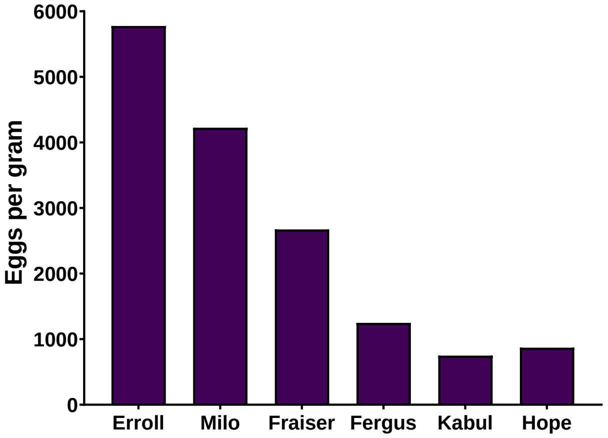
Non-Parametric:

Spearman Correlation Coefficient

- Only really useful for ranks (either one or both variables)
- **ρ (rho)** is the equivalent of **r** and calculated in a similar way

- **Example**: **Dominance.xslx**
  - Six male colobus monkeys ranked for dominance
  - Question: is social dominance associated with parasitism?
    - Eggs of *Trichirus* nematode per gram of monkey faeces

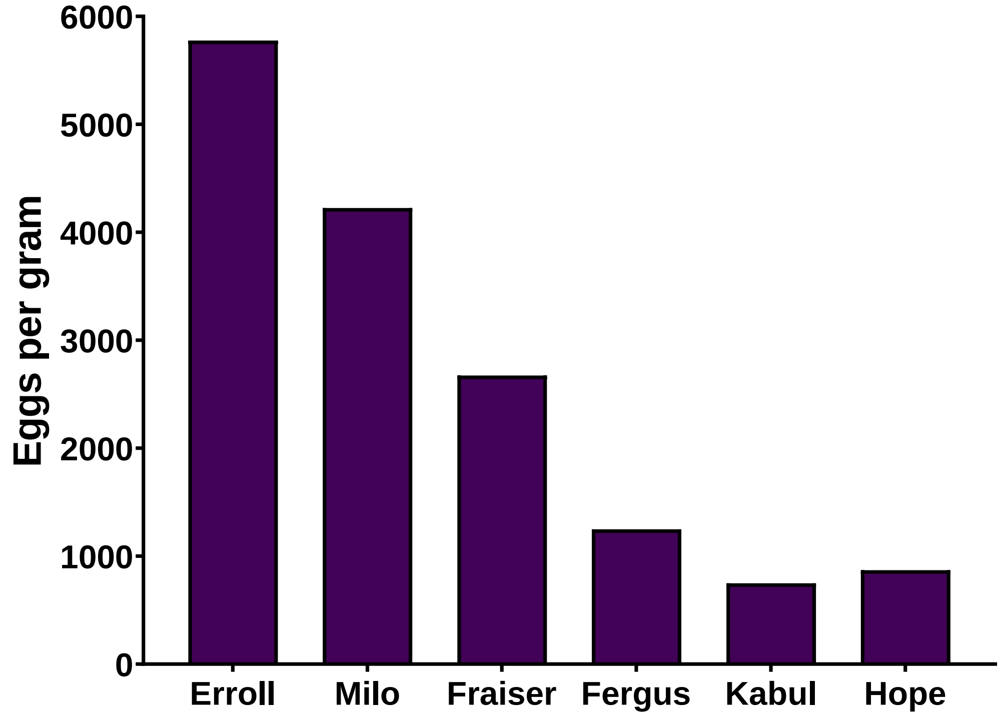
Monkey	Dominance	Eggs.per.gram
Erroll	1	5777
Milo	2	4225
Fraiser	3	2674
Fergus	4	1249
Kabul	5	749
Hope	6	870



Non-Parametric:

Spearman Correlation Coefficient

	Correlation	Dominance vs. Eggs per gram
1	**Spearman r**	
2	r	-0.9429
3	95% confidence interval	
4		
5	**P value**	
6	P (two-tailed)	0.0167
7	P value summary	*
8	Exact or approximate P value?	Exact
9	Significant? (alpha = 0.05)	Yes
10		
11	**Number of XY Pairs**	6
12		



- **Answer**: the relationship between dominance and parasitism is significant (ρ = -0.94, p = 0.017) with high ranking males harbouring a heavier burden.