Nonparametric Statistics

Leah Wright, Tyler Ross, Taylor Brown

Before we get to nonparametric statistics, what are **parametric statistics**?

- These statistics estimate and test population means, while holding certain ***assumptions*** about the distribution of the population where the sample data come from
- These assumptions include:
  - Normally distributed population
  - No outliers
  - Large sample sizes
  - Random independent samples
  - Interval or ratio measurements
  - Homogeneity of variance

What happens when our data violate these assumptions?

1. Ignore the violations
2. Transform the data
3. Permutation test
4. Choose a nonparametric test as an alternative

What are Nonparametric Statistics?

Compared with parametric statistics, they:

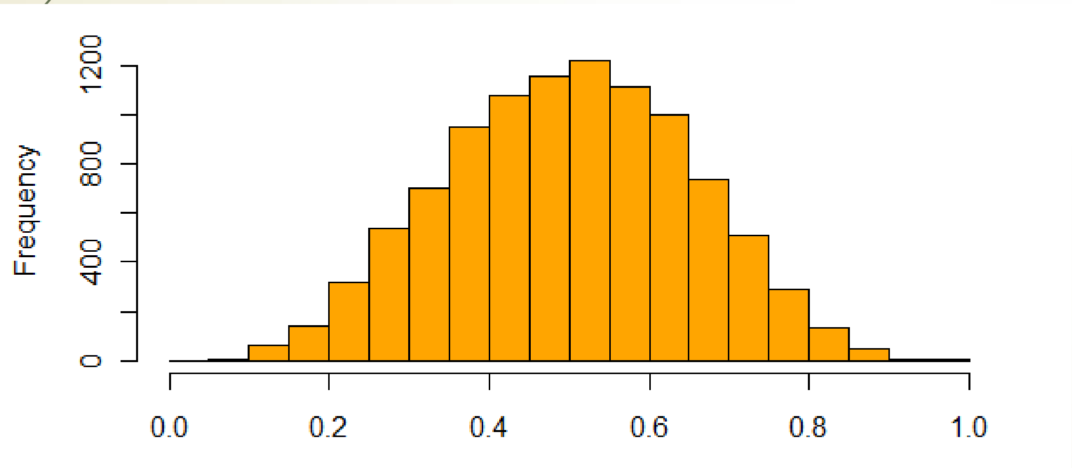
- Make fewer assumptions
- Rank data to replace actual numerical values
- *Do not* rely upon parameter estimations, as parametric statistics do.

Why not use *only* nonparametric statistics?

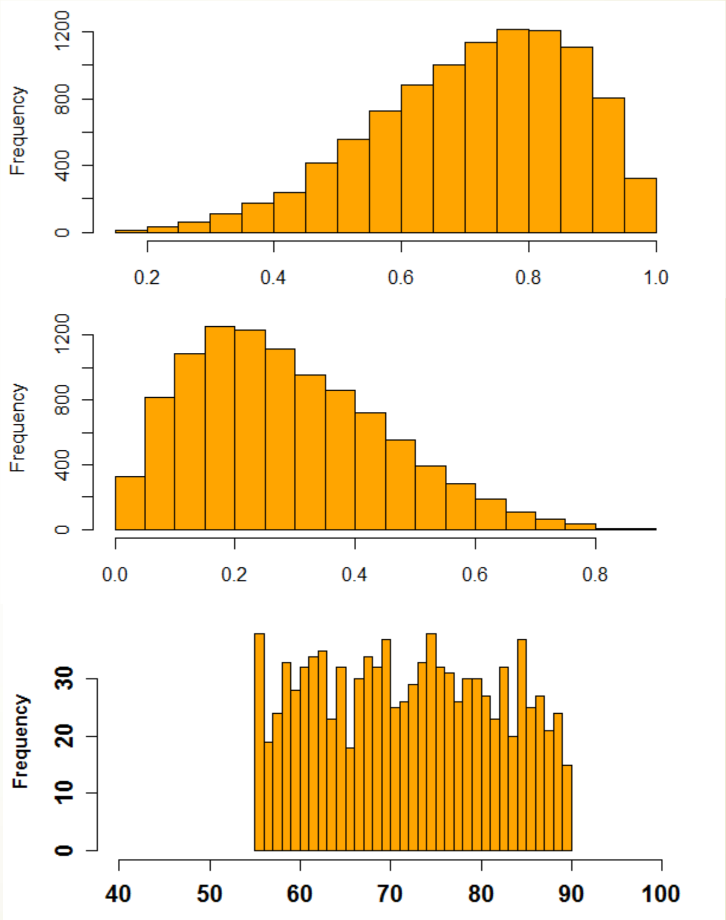
- They are conservative
- Less statistical power than parametric tests
- More likely to produce type II errors

How to determine if data violate normality assumption

1. Plot histogram



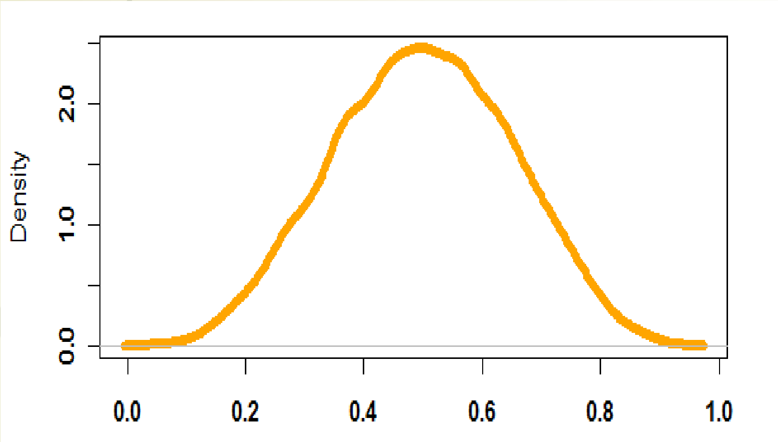
Normal

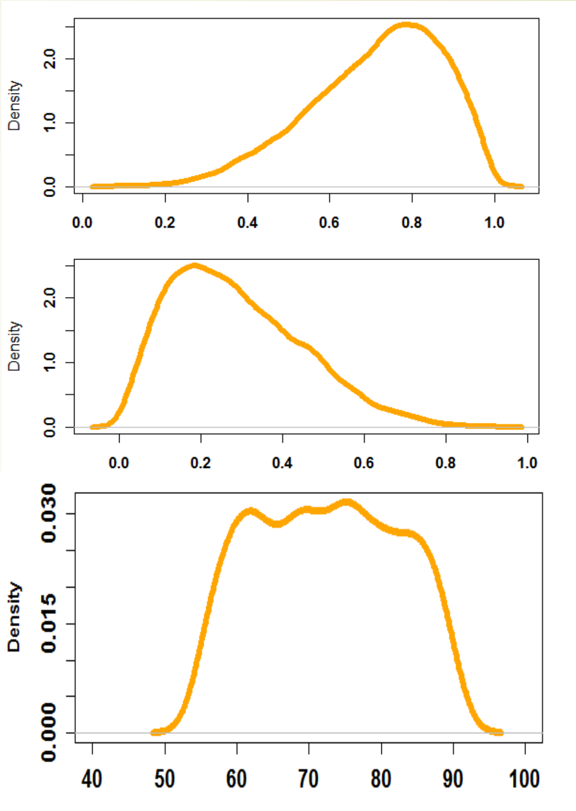


Non-normal

How to determine if data violate normality assumption

2. Plot density plot





Normal

Paranormal

Non-normal

How to determine if data violate normality assumption

3. Formal statistical tests
   - I.e. Shapiro Wilk Test
     - Estimates mean and standard deviation of sample population, then tests goodness of fit to data of normal distribution with the same mean and standard deviation
     - *In R*- shapiro.test(x)

Parametric Tests and their Nonparametric alternatives

- When parametric assumptions not met, use corresponding nonparametric test:

Parametric	Nonparametric
One-way (independent) ANOVA	Kruskal-Wallis test
One-way repeated measures ANOVA	Friedman test
Pearson Product-Moment Correlation	Spearman's Rank-Order Correlation
Paired samples t-test	The Sign test
Two sample t-test	Mann-Whitney U test

Analysis of Variance: A (very) Brief Review

- Comparison of three or more *means;*
- Partitions and compares variability – within group variation and between group variation;
- Assumptions:
  - k independent samples (not necessarily equal) taken from…
  - k **normally** distributed populations
  - with equal variance

Kruskal-Wallis: A Nonparametric Alternative

- No assumption of normality; however, the following assumptions do apply:
  1. Completely randomized design (i.e. subjects/organisms are randomly assigned to treatments)
  2. Distributions of the treatments have approximately the same shape and the same spread

Kruskal-Wallis: A Nonparametric Alternative

- Unlike analysis of variance, observations are ranked relative to one another across the different trials:

59	62	81	65	87	42	31

Kruskal-Wallis: A Nonparametric Alternative

- Unlike analysis of variance, observations are ranked relative to one another across the different trials:

59 (3)   62 (4)   81 (6)   65 (5)   87 (7)   42 (2)   31 (1)

59 (1)   **62** (2.5)   **62** (2.5)   71 (3)

↓

[2 + 3 / 2 = 2.5]

Kruskal-Wallis: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk)

H_a: not all medians are equal (i.e. at least one md ≠)

Test statistic:

$$H = \frac{12}{n(n+1)}\left[\sum_{i=1}^{k}\frac{T_{R_i}^2}{n_i}\right] - 3(n+1)$$

where T_{R_i} is the sum of the ranks assigned to observations in treatment i. H_O is rejected if $H > \chi_{\alpha}^{2}$(k -1)

Kruskal-Wallis: A Nonparametric Alternative

Example:

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

I	II	III	IV
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88
81	72	83	
69	79	76	
	90		
n = 6	**n = 7**	**n = 6**	**n = 4**

Kruskal-Wallis: A Nonparametric Alternative

Example:

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

I	II	III	IV
65 (3)	75 (9)	59 (1)	94 (23)
87 (19)	69 (5.5)	78 (11)	89 (21)
73 (8)	83 (17.5)	67 (4)	80 (14)
79 (12.5)	81 (15.5)	62 (2)	88 (20)
81 (15.5)	72 (7)	83 (17.5)	
69 (5.5)	79 (12.5)	76 (10)	
	90 (22)		
n = 6	**n = 7**	**n = 6**	**n = 4**
$T_{R_1} = 63.5$	$T_{R_2} = 89$	$T_{R_3} = 45.5$	$T_{R_4} = 78$

Kruskal-Wallis: A Nonparametric Alternative

Example:

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

$$H = \frac{12}{n(n+1)}\left[\sum_{i=1}^{k} \frac{T_{R_i}^2}{n_i}\right] - 3(n+1)$$

$$H = \frac{12}{23(24)}\left[\frac{(63.5)^2}{6} + \frac{(89)^2}{7} + \frac{(45.5)^2}{6} + \frac{(78)^2}{4}\right] - 3(24)$$

$$H = 7.78$$

Kruskal-Wallis: A Nonparametric Alternative

Example:

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

H_0 **is rejected if** $H > \chi^2_\alpha$**(k -1)**

H	$\chi^2_{0.10}(3)$
7.78	**6.25**

Kruskal-Wallis: A Nonparametric Alternative

Example:

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

Thus, since H = 7.78 > 6.25, one should reject H_o and conclude that at 10% level of significance, there is evidence to say that the four teaching techniques differ.

Kruskal-Wallis: A Nonparametric Alternative

Example (in R):

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

```
> Teaching_Techniques <- read.delim("C:/Users/Tyler/Desktop/Teaching_Techniques.txt")
>   View(Teaching_Techniques)
> show(Teaching_Techniques)
   Treatment Student_Score
1          1            65
2          1            87
3          1            73
4          1            79
5          1            81
6          1            69
7          2            75
8          2            69
9          2            83
10         2            81
11         2            72
12         2            79
13         2            90
14         3            59
15         3            78
```

Kruskal-Wallis: A Nonparametric Alternative

Example (in R):

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

```
> show(Teaching_Techniques)

> kruskal.test(Student_Score ~ Treatment, data = Teaching_Techniques)

```

```
> kruskal.test(Student_Score ~ Treatment, data = Teaching_Techniques)

 Kruskal-Wallis rank sum test

data:  Student_Score by Treatment
Kruskal-Wallis chi-squared = 7.7905, df = 3, p-value = 0.05055
```

Kruskal-Wallis: A Nonparametric Alternative

Example:

1. Students are randomly assigned to four different treatment groups (different teaching techniques), and test scores are measured
2. Distributions of the four treatments have approximately the same shape and the same spread.

Again, since p ≈ 0.05 one should reject H_o and conclude that at 10% level of significance, there is evidence to say that the four teaching techniques differ.

Friedman Test: A Nonparametric Alternative

- No assumption of normality; however, the following assumptions do apply:
  1. One group is measured on three or more different occasions
  2. Group is a random sample of the population
  3. Dependent variable is ordinal (e.g. 10-point scale) or continuous

Friedman Test: A Nonparametric Alternative

- Like the Kruskal-Wallis test, observations are ranked relative to one another across the different trials:

Patient	Treatment 1	Treatment 2	Treatment 3
1	209	88	109
2	412	388	142
3	315	451	155
4	389	325	121
5	210	126	75
6	136	118	49
7	178	227	101
n = 7			

Friedman Test: A Nonparametric Alternative

- Like the Kruskal-Wallis test, observations are ranked relative to one another across the different trials:

Patient	Treatment 1	Treatment 2	Treatment 3
1	209 (3)	88 (1)	109 (2)
2	412 (3)	388 (2)	142 (1)
3	315 (2)	451 (3)	155 (1)
4	389 (3)	325 (2)	121 (1)
5	210 (3)	126 (2)	75 (1)
6	136 (3)	118 (2)	49 (1)
7	178 (2)	227 (3)	101 (1)
n = 7	**19**	**15**	**8**

Friedman Test: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk) – identical effects

H_a: not all medians are equal (i.e. at least one md ≠) – different effects

Test statistic:

$$FM = \left[\frac{12}{\left(N \cdot k \cdot (k+1)\right)}\right] \cdot \sum R^2 - [3 \cdot N \cdot (k+1)]$$

where N = # subjects, k = # of trials, and R = the total ranks for each column. H_O is rejected if $FM > FM_{critical\ value}$*

* If your k is over 5, or your n is over 13, use the chi square critical value table

Friedman Test: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk) – identical effects

H_a: not all medians are equal (i.e. at least one md ≠) – different effects

$$FM = \left[\frac{12}{\left(N \cdot k \cdot (k+1)\right)}\right] \cdot \sum R^2 - [3 \cdot N \cdot (k+1)]$$

$$FM = \left[\frac{12}{\left(7 \cdot 3 \cdot (3+1)\right)}\right] \cdot (19^2 + 15^2 + 8^2) - [3 \cdot 7 \cdot (3+1)]$$

$$FM = 8.86$$

Friedman Test: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk) – identical effects

H_a: not all medians are equal (i.e. at least one md ≠) – different effects

H_O is rejected if $FM > FM_{critical\ value}$

FM	$FM_{critical\ value}$
8.86	**7.14**

Friedman Test: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk) – identical effects

H_a: not all medians are equal (i.e. at least one md ≠) – different effects

Thus, since FM = 8.86 > 7.14 , one should reject H_o and conclude that at 5% level of significance, there is evidence to say that the three treatment effects differ.

Friedman Test: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk) – identical effects

H_a: not all medians are equal (i.e. at least one md ≠) – different effects

```
Treatment <- matrix(c(209, 88, 109,
                      412, 388, 142,
                      315, 451, 155,
                      389, 325, 121,
                      210, 126, 75,
                      136, 118, 49,
                      178, 227, 101),
                    nrow = 7,
                    byrow = "true",
                    dimnames = list(1:7, c("Treatment_1", "Treatment_2", "Treatment_3")))

Treatment

friedman.test(Treatment)
```

Friedman Test: A Nonparametric Alternative

H_0: treatment medians are equal (i.e. md1 = md2… = mdk) – identical effects

H_a: not all medians are equal (i.e. at least one md ≠) – different effects

```
Treatment <- matrix(c(209, 88, 109,
                      412, 388, 142,
                      315, 451, 155,
                      389, 325, 121,
                      210, 126, 75,
                      136, 118, 49,
                      178, 227, 101),
                    nrow = 7,
                    byrow = "true",
                    dimnames = list(1:7, c("Treatment_1", "Treatment_2", "Treatment_3")))

Treatment

friedman.test(Treatment)
```

```
> friedman.test(Treatment)

 Friedman rank sum test

data:  Treatment
Friedman chi-squared = 8.8571, df = 2, p-value = 0.01193
```

Spearman's Rank-Order Correlation

- When to use??
  - Nonparametric version of the Pearson product-moment correlation
  - Spearman's correlation coefficient (ρ or r_s) measures strength and direction of association between two ranked variables

Spearman's Rank-Order Correlation

Assumptions??

1. Need two variables that are either ordinal, interval, or ratio
2. Although Pearson product-moment correlation would likely be used on interval or ratio data as well, use Spearman when Pearson's assumptions are violated
3. Observations are independent

Spearman's Rank-Order Correlation

What are *Pearson*'s assumptions?

1. Variables are interval or ratio measurements
2. Variables are approximately normally distributed
3. Possibility of a linear relationship
4. Outliers are few or removed
5. Homoscedasticity of the data – variance along the line of best fit remains the same

Spearman's Rank-Order Correlation

So what's the difference?

- Spearman's correlation determines the *strength* and *direction* of the monotonic relationship between the two variables, rather than…
- The strength and direction of the linear relationship between them, as in Pearson's correlation

Spearman's Rank-Order Correlation

Monotonicity

statistics.laerd.com

But monotonicity is not strictly an assumption – you can run Spearman's correlation on a non-monotonic relationship to determine if there is a <u>monotonic component</u>

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- Start with your table of data values, pairing the corresponding values with their observation (e.g. Student 1 got 75% in English, and 70% in Math)

Student	English grade	Math grade
1	75	70
2	87	95
3	64	70
4	42	53
5	93	86
6	75	60

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- To rank the data, you first have to rearrange your table…

English grade	Math grade	English rank	Math rank
75	70		
87	95		
64	70		
42	53		
93	86		
75	60		

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- Rank *both* variables, but separately
- Take the average rank of two values when they are tied
- Rank highest to lowest

English grade	Math grade	English rank	Math rank
75	70	3.5	3.5
87	95	2	1
64	70	5	3.5
42	53	6	6
93	86	1	2
75	60	3.5	5

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- For now let's use different data that have *no ties*

English grade	Math grade	English rank	Math rank
78	72	3	3
87	95	2	1
64	69	5	4
42	53	6	6
93	86	1	2
74	60	4	5

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- There are <u>two methods</u> to calculate Spearman's correlation depending on whether:

1. **Data do not have tied ranks**

   $\rho = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$, where d_i = difference in paired ranks

   and

   n = number of cases

2. **Data <u>do</u> have tied ranks (modified Pearson's equation)**

   $\rho = \frac{\Sigma_i(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma_i(x_i - \bar{x})^2 \Sigma_i(y_i - \bar{y})^2}}$, where *i* = paired score

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- Now add two more columns: one for *d* and one for d^2

English grade	Math grade	English rank	Math rank	d	d^2
78	72	3	3	0	0
87	95	2	1	1	1
64	69	5	4	1	1
42	53	6	6	0	0
93	86	1	2	1	1
74	60	4	5	1	1

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

- Then calculate Σd_i^2

= 4

English grade	Math grade	English rank	Math rank	d	d²
78	72	3	3	0	0
87	95	2	1	1	1
64	69	5	4	1	1
42	53	6	6	0	0
93	86	1	2	1	1
74	60	4	5	1	1

d = difference between the ranks

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

$$\rho = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)}$$

$\Sigma d_i^2 = 4$

$n = 6$

$\rho = 0.8857$

English grade	Math grade	English rank	Math rank	d	d^2
78	72	3	3	0	0
87	95	2	1	1	1
64	69	5	4	1	1
42	53	6	6	0	0
93	86	1	2	1	1
74	60	4	5	1	1

d = difference between the ranks

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

What values can Spearman correlation coefficient, r_s (or ρ) take?

- Values from +1 to -1
  - r_s of +1 indicates perfect association of ranks
  - r_s of 0 indicates no association between ranks
  - r_s of -1 indicates perfect *negative* association of ranks

Spearman's Rank-Order Correlation

Steps in calculating Spearman's Correlation Coefficient

What does this mean for our example?

$\rho = 0.8857$

Indicates *very strong,* positive association of the ranks

Spearman's Rank-Order Correlation

How you report Spearman's correlation coefficient depends on whether or not you've determined the statistical significance of the coefficient

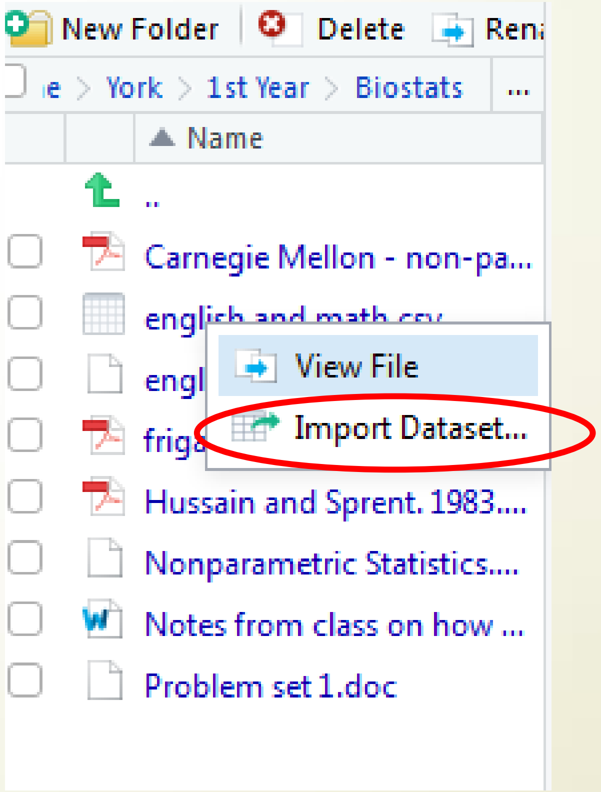
- *Without* the significance test:
  - ρ = 0.89 or r_s = 0.89
- *With* significance test:
  - ρ(4) = 0.89, P = P-value

…where df = N-2, where N = number of pairwise cases

Spearman's Rank-Order Correlation

… in R

- Make Excel spreadsheet of data
- Save as ".csv" file
- Then…



Spearman’s Rank-Order Correlation

… in R

- Name dataframe
- Note the column title names

Untitled1 × english_and_math ×

Filter

	English grade	Math grade
1	78	72
2	87	95
3	64	69
4	42	53
5	93	86
6	74	60

Spearman's Rank-Order Correlation

… in R

Untitled1 × | english_and_math ×

Filter

	English grade	Math grade
1	78	72
2	87	95
3	64	69
4	42	53
5	93	86
6	74	60

Showing 1 to 6 of 6 entries

Console | Terminal ×

~/

```
> cor(english_and_math$`English grade`, english_and_math$`Math grade`, method = "spearman")
[1] 0.8857143
>
```

Yep, ρ = 0.8857

Spearman's Rank-Order Correlation

… in R

cor(dataframe$column1, dataframe$column2, method = "spearman")

At the end, before the bracket, you can also add "…spearman", **use="pairwise.complete.obs"**) if you're not sure that all x's have y's and vice versa

Untitled1 | english_and_math

Filter

	English grade	Math grade
1	78	72
2	87	95
3	64	69
4	42	53
5	93	86
6	74	60

Showing 1 to 6 of 6 entries

Console | Terminal

~/

```
> cor.test(english_and_math$`English grade`, english_and_math$`Math grade`, method = "spearman", use="pairwise.complete.obs", exact=FALSE)

        Spearman's rank correlation rho

data:  english_and_math$`English grade` and english_and_math$`Math grade`
S = 4, p-value = 0.01885
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.8857143

>
```

Spearman’s Rank-Order Correlation

Null hypothesis:

H_0 : There is no [monotonic] association between English grades and Math grades (ρ = 0)

*Statistical significance does not indicate the *strength* of Spearman’s correlation. Using α=0.05 and getting a significant P-value means there is a <5% chance that the strength of the relationship happened by chance; i.e. that the null hypothesis is true.

Two sample T-test: A brief review

- Used to compare two means
- tests H_0 that means from *two independent* groups are equal, $H_{0:}U_1=U_2$

- *Assumes that :*
  1. Both populations have ***normal distribution***
  2. **Random** samples from population
  3. Both populations have similar standard deviation and variance

The Mann-Whitney U: A nonparametric alternative

- Compares two *independent* samples
- Ranks data
  - Changes data from interval ->ordinal
- Tests whether there is a difference between medians of both populations

Mann-Whitney U

- Assigning ranks allows us to discard normality assumption. However, test *does assume*:

1. Random samples
2. Independence of observations
3. Distribution of both samples have the same shape

Calculating Mann-Whitney U

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 n_2 - U_1$$

Where:

n_1 and n_2 = Sample size of each group

R_1 and R_2 = Rank sums for each treatment

U_1 *and* U_2 = Mann Whitney U test statistics

Mann-Whitney U

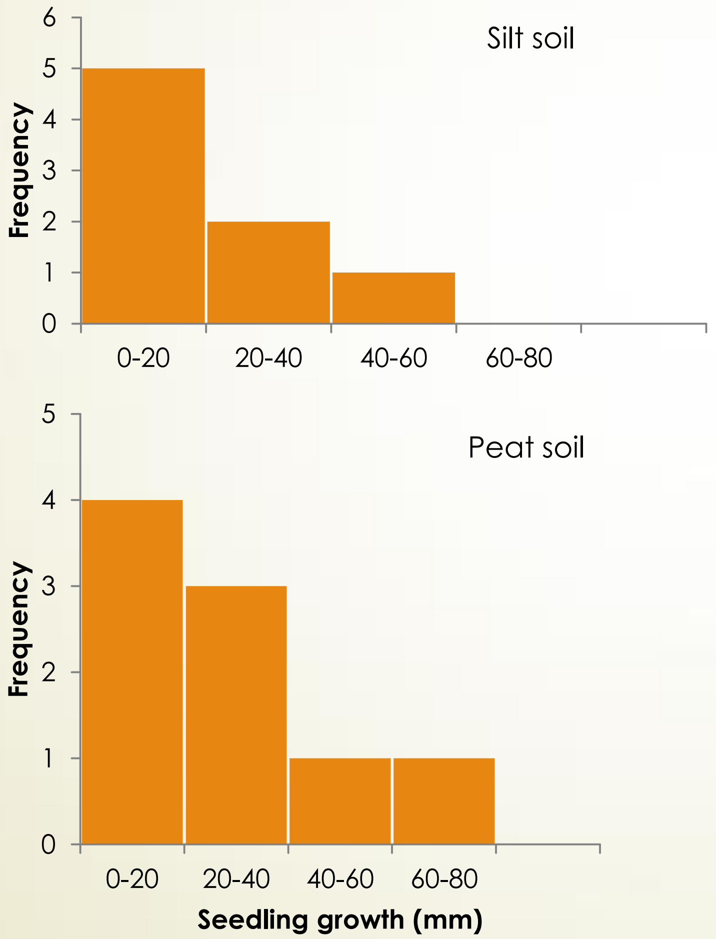
Consider this example:

- Red maple tree (*Acer rubrum*) seeds are allowed to germinate and grow for three weeks in two different types of soil. Growth progress for each seedling was measured as follows:

Height of seedlings (mm) grown in silt soil	Height of seedlings (mm) grown in peat soil
10	12.2
10.5	11.7
13.2	12.4
19.7	15.7
16.6	22.8
23	24.5
26.4	27
45.3	41
	62
n_1 =8	n_2 =9

Mann-Whitney U

First, see if distributions of both groups have the same shape



- Both frequency distributions are not normal, but they have similar shapes; both show a positive skew.

Mann-Whitney U

Hypotheses

H_0 = The two samples are equal

- The seedlings grown in either soil type do not differ in growth height

H_1 = The two samples are not equal

- The seedlings grow taller in one type of soil than the other

Our null hypothesis is rejected if calculated U test statistic ≥ critical value for U ($\alpha = 0.05$)

Mann-Whitney U

Steps in calculating Mann-Whitney U test statistic

Silt soil	Peat soil	Rank for silt soil	Rank for peat soil
10	12.2	1	4
10.5	11.7	2	3
13.2	12.4	6	5
19.7	15.7	9	7
16.6	22.8	8	10
23	24.5	11	12
26.4	27	13	14
45.3	41	16	15
	62		17
n_1 =8	n_2 =9	$\boldsymbol{R_1 = 66}$	$\boldsymbol{R_2 = 87}$

Rank Sums for each group

Mann-Whitney U

Steps in calculating Mann-Whitney U test statistic

- Plug in our sample sizes and rank sums to calculate U_1, and then U_2

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= 8(9) + \frac{8(8+1)}{2} - 66$$

$$=42$$

$$U_2 = n_1 n_2 - U_1$$

$$= 8(9) - 46$$

$$=30$$

Now, select the larger of U_1 and U_2 as your test statistic

Mann-Whitney U

Steps in calculating Mann-Whitney U test statistic

- Consult table of critical values and compare calculated U to critical value

$$\alpha = 0.05$$

$$U_1 = 42 < U_{critical} = 57$$

Mann-Whitney U

Conclusions

$U_1 = 42 < U_{critical} = 57$

- Thus, our calculated U is **less** than the critical value, so we cannot reject the null hypothesis at 5% significance level. There is **not** enough evidence to suggest growth progress is significantly different in the two types of soil.

Wilcoxon Mann-Whitney U

Example in SAS

1. Load data into SAS

Log - (Untitled)

```
32   datalines;

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.
NOTE: The data set WORK.SEEDS has 17 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


36      ;
```

Editor - Untitled1 *

```
data Seeds;
input Soil Growth @@;
datalines;
   1 10     1 10.5   1 13.2   1 19.7   1 16.6   1 23     1 26.4
   1 45.3   2 12.2   2 11.7   2 12.4   2 15.7   2 22.8
   2 24.5   2 27     2 41     2 62
   ;
```

Wilcoxon Mann-Whitney U

Example in SAS

2. Run NPAR1WAY, with the option Wilcoxon

```
Log - (Untitled)

      cpu time            0.00 seconds


  proc npar1way wilcoxon correct=no data=Seeds;
     class Soil;
     var Growth;
     exact wilcoxon;
  run;

NOTE: Writing HTML Body file: sashtml.htm
```

Wilcoxon Mann-Whitney U

Example in SAS

3. Interpret output of test

Wilcoxon Scores (Rank Sums) for Variable Growth Classified by Variable Soil					
Soil	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	8	66.0	72.0	10.392305	8.250000
2	9	87.0	81.0	10.392305	9.666667

Wilcoxon Two-Sample Test	
Statistic (S)	66.0000
Normal Approximation	
Z	-0.5774
One-Sided Pr < Z	0.2819
Two-Sided Pr > |Z|	0.5637
t Approximation	
One-Sided Pr < Z	0.2859
Two-Sided Pr > |Z|	0.5717
Exact Test	
One-Sided Pr <= S	0.3029
Two-Sided Pr >= |S - Mean|	0.6058

The p-value =0.6058, since p-value is larger than 0.05, we conclude again that the growth rate between both treatments are not significantly different.

One sample T-test: A brief review

- Tests the null hypothesis that the mean of observations is the same as known/hypothesized value

- This test assumes:
  1. Normal distribution
  2. Samples are random

The Sign Test: A Nonparametric Alternative

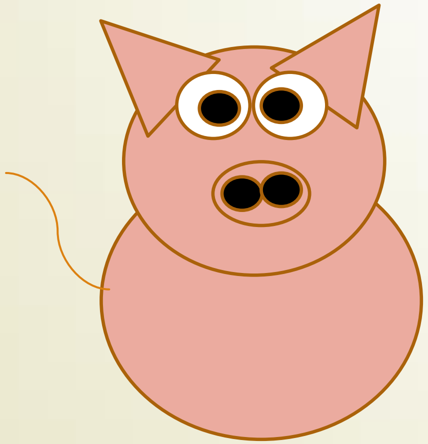
- Can use this test when distribution is neither normal *nor* symmetrical
- Used to test whether median of population is different from some hypothesized value
- Measurements scored "+" if falling above hypothesized median, or "-" if below hypothesized value
  - If H0 true, expect half values to be assigned "+" and half values to be assigned "-"
- *Assumes* samples are random
- Low statistical power
  - increases with larger sample size (n >5)

The Sign-test

Consider this example:

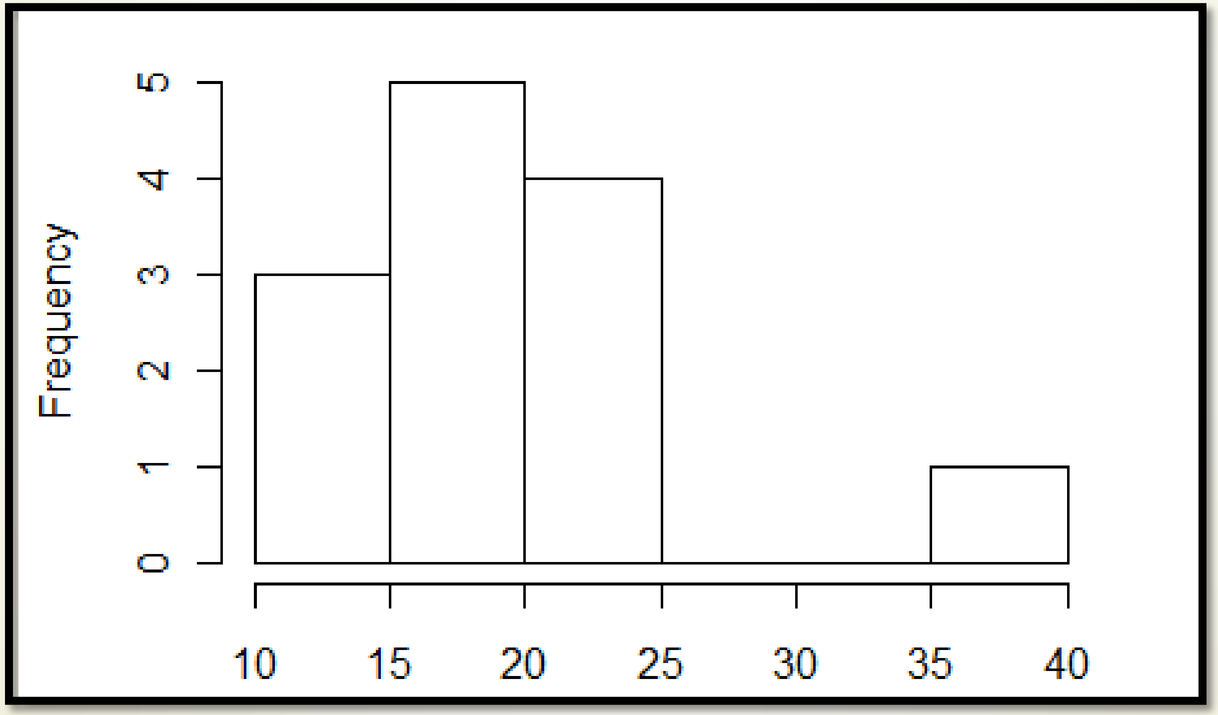
- Pigs raised their first year on an organic farm were randomly selected and weighed. Farmers wish to know if their rate of growth is different than the known weight gained during this time in factory farmed pigs, η= 21.

Organic pig weight (kg)	10.1	13.1	17	12.2	18.6	19.5	20.5	22.1	39	21.7	18.4	18.6	20.2	15.8



The Sign-test

Quick check to visualize distribution of data



Doesn't appear to be normal, nor symmetrical

The Sign-test

Hypotheses

- H_0: $\eta = \eta_0$
  - The population median (η) equals the hypothesized median ($\eta_0=21$)
- H1: $\eta \neq \eta_0$
  - The population median (η) differs from the hypothesized median ($\eta_0=21$).

The Sign-test

Steps in calculating the sign-test statistic

Organic pig weight (kg)	10.1	13.1	17	12.2	18.6	19.5	20.5	22.1	39	21.7	18.4	18.6	20.2	15.8
Above(+) or below(-) 21	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**+**	**+**	**+**	**-**	**-**	**-**	**-**

Total "-"	**Total"+"**
11 out of 14	3 out of 14

The Sign-test

Steps in calculating the sign-test statistic

- If H_0 is correct, we would expect half of our values to fall above 21, and half of our values to fall below 21

- We use binomial distribution to get our p-value
  - Want to know the probability of seeing our observation of 3 “+" out of 14 total observations, when the probability of observing “+" is 0.5

The Sign-test

Steps in calculating the sign-test statistic

$$Pr(x \leq 3) = \sum_{i=0}^{3} \binom{14}{i} (0.5)^i (0.5)^{14-i}$$

=0.0286

Two sided test,
P=2(0.0286)= 0.0572

The Sign-test

Conclusions

- $p=0.0572 > 0.05$, we can conclude that there is no significant difference between the median weights of factory farmed versus organically farmed pigs at the 5% significance level.