One-Sample Rank Biserial (es_rank_biserial_os)

Author: P. Stikker
Website: https://peterStatistics.com
YouTube: https://www.youtube.com/stikpet
Version: 0.1 (2023-01-30)

Introduction

The *es_rank_biserial_os* function calculates a one-sample rank biserial coefficient.

This document contains the details on how to use the functions, and formulas used in them.

1 About the Function

1.1 Input parameters:

- **data**
  - Excel: a specific range with the numeric scores
  - Python: a pandas series with the numeric scores
  - R: a vector with the numeric scores

- *Optional parameters*
  - **hypMed**
    the hypothesized median. If not specified the midrange will be used.

1.2 Output

- **hypMed**
  The hypothesized median used
- **rb**
  The effect size measure

Note for *Excel:*

the array function *es_rank_biserial_os_arr* will require 2 rows and 2 columns.

1.3 Dependencies

- **Excel**
  - None, but you can run the **es_rank_biserial_os_addHelp** macro so that the function will be available with some help in the 'User Defined' category in the functions overview.
- **Python**
  The following libraries are needed:
  - pandas is needed for data entry and showing the results
- **R**
  None

2 Examples

2.1 Excel

	A	B	C	D	E	F	G
1	Teach_Motivate						
2	1						
3	2		0,3088235	=es_rank_biserial_os(A2:A21)			
4	5						
5	1		hypMed				
6	1		2	0,391813	=es_rank_biserial_os(A2:A21;C6)		
7	5						
8	3						
9	1		hyp. med.	rb			
10	5		3	0,308824			
11	1						
12	1		C9:D10 =>	=es_rank_biserial_os_arr(A2:A21)			
13	5						
14	1						
15	1						
16	3						
17	3						
18	3						
19	4						
20	2						
21	4						

2.2 Python

[1]:
```
from eff_size_rank_biserial_os import es_rank_biserial_os
import pandas as pd

dataList = [1, 2, 5, 1, 1, 5, 3, 1, 5, 1, 1, 5, 1, 1, 3, 3, 3, 4, 2, 4]
data = pd.Series(dataList)

es_rank_biserial_os(data)
```

[1]:

	mu	rb
0	3.0	0.308824

[2]:
```
es_rank_biserial_os(data, hypMed = 2.5)
```

[2]:

	mu	rb
0	2.5	0.057143

2.3 R

```
> source("eff_size_rank_biserial_os.R")
>
> data <- c(1, 2, 5, 1, 1, 5, 3, 1, 5, 1, 1, 5, 1, 1, 3, 3, 3, 4, 2, 4)
> es_rank_biserial_os(data)
  hypMed         rb
1      3 -0.3088235
> es_rank_biserial_os(data, hypMed=2)
  hypMed        rb
1      2 0.3918129
```

3 Details of Calculations

$$r_{rb} = \frac{4 \times \left| T - \frac{R^+ + R^-}{2} \right|}{n \times (n+1)} = \frac{|R^+ - R^-|}{R}$$

Symbols

- R^+ the sum of the ranks with a positive deviation from the hypothesized median
- R^- the sum of the ranks with a positive deviation from the hypothesized median
- T the minimum of R^+, R^-
- n the number of ranks with a non-zero difference with the hypothesized median
- R the sum of all ranks, i.e. $R^+ + R^-$

Note: A proof of the equivalence of the two formulas can be found in the appendix.

4 Sources

Unknown, the formula can be found in King and Minium. (2008, p. 403):

matched-pairs rank biserial correlation coefficient, (r_C) a measure of effect size for the Wilcoxon signed-ranks test

To calculate effect size, we use the **matched-pairs rank biserial correlation coefficient,** r_C:

FORMULA FOR r_C

$$r_C = \frac{4\left|T - \left(\frac{R_+ + R_-}{2}\right)\right|}{n(n+1)} \qquad (21.5)$$

Kerby's re-arrangement:

used as the test statistic (e.g., Glantz, 2005). Thus, the formula is now $r = W/S$. And of course, for a directional hypothesis, W can be stated as the difference between the favorable sums and unfavorable sums. The result

the total sum of ranks, which can be symbolized as S.

(Kerby, 2014, p. 5)

References

Kerby, D. S. (2014). The simple difference formula: An approach to teaching nonparametric correlation. *Comprehensive Psychology*, *3*, 1–9. https://doi.org/10.2466/11.IT.3.1

King, B. M., & Minium, E. W. (2008). *Statistical reasoning in the behavioral sciences* (5th ed.). John Wiley & Sons, Inc.

Appendix: Proof of Equivalence

Assuming $T = R^+$

$$\frac{4 \times \left|T - \frac{R^+ + R^-}{2}\right|}{n \times (n+1)} = \frac{4 \times \left|R^+ - \frac{R^+ + R^-}{2}\right|}{n \times (n+1)} = \frac{|4 \times R^+ - 2 \times (R^+ + R^-)|}{n \times (n+1)} = \frac{|2 \times R^+ - 2 \times R^-|}{n \times (n+1)}$$
$$= \frac{2 \times |R^+ - R^-|}{n \times (n+1)} = \frac{2}{n \times (n+1)} \times |R^+ - R^-| = \frac{1}{\frac{n \times (n+1)}{2}} \times |R^+ - R^-|$$

If we define R as the sum of all ranks we can write this as: $R = \sum ranks = \frac{n \times (n+1)}{2}$. Using this we get:

$$\frac{4 \times \left|T - \frac{R^+ + R^-}{2}\right|}{n \times (n+1)} = \frac{1}{\frac{n \times (n+1)}{2}} \times |R^+ - R^-| = \frac{1}{R} \times |R^+ - R^-| = \frac{|R^+ - R^-|}{R}$$

If we assume $T = R^-$ we get the same result:

$$\frac{4 \times \left|T - \frac{R^+ + R^-}{2}\right|}{n \times (n+1)} = \frac{4 \times \left|R^- - \frac{R^+ + R^-}{2}\right|}{n \times (n+1)} = \frac{|4 \times R^- - 2 \times (R^+ + R^-)|}{n \times (n+1)} = \frac{|2 \times R^- - 2 \times R^+|}{n \times (n+1)}$$
$$= \frac{2 \times |R^- - R^+|}{n \times (n+1)} = \frac{2}{n \times (n+1)} \times |R^- - R^+| = \frac{1}{\frac{n \times (n+1)}{2}} \times |R^- - R^+|$$
$$= \frac{1}{R} \times |R^- - R^+| = \frac{|R^- - R^+|}{R} = \frac{|R^+ - R^-|}{R}$$

Q.E.D.