Cohen h2 for one-sample (es_cohen_h2_os)

Author: P. Stikker
Website: https://peterStatistics.com
YouTube: https://www.youtube.com/stikpet
Version: 0.1 (2023-01-05)

Introduction

The *es_cohen_h2_os* function (and *es_cohen_h2_os_arr* in VBA) calculates an effect size known as Cohen's h_2 for a one-sample test. This effect size measure can be used with a one-sample binomial test, Wald, or Score test.

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**
  The data to be used. Note for Python this needs to be a pandas data series.

- *Optional parameters*
  - **codes** (default is none)
    Two codes for the two categories to be compared. For example, if the data has a list of scores with "national" and "international", the codes used can be exactly those: "national", "international".
    This makes it possible to also use a nominal data set (with more than two categories) and then select the two for this test to be used, and keep it in line with a one-sample binomial, Wald, or score test.

  - **p0** (default is 0.5)
    The hypothesized proportion for the first category (as in codes or found in data).

  - **out** (default is "value") – only applies to VBA non-array function
    Choice what to show as result. Either:
    - "value": show the effect size value
    - "qual": show the qualification

1.2 Output:

- The **value**, and the **classification**. Except for the non-array version in VBA (Excel) which will only show the requested output via the ‘out’ parameter.

- The array version in VBA (*es_cohen_h2_os_arr*) requires **two rows** and **two columns**.

1.3 Dependencies

- **Excel**
  None.
  You can run the **es_cohen_h2_os_addHelp** macro so that the function will be available with some help in the ‘User Defined’ category in the functions overview.

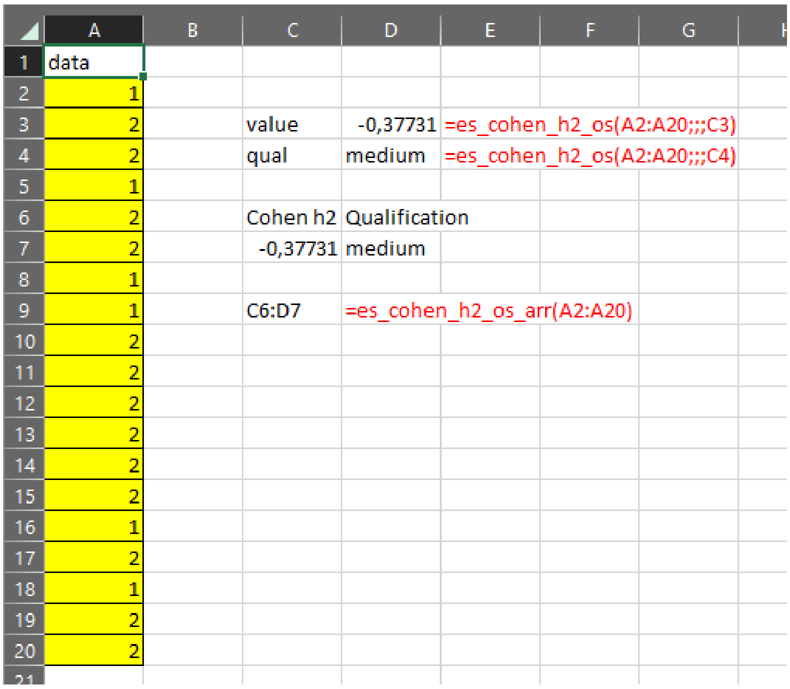
- **Python**
  The following additional libraries will have to be installed:
  - *pandas*
    the data input needs to be a pandas data series, and the output is also a pandas dataframe.

  - *math*
    the build in library math from Python is needed for the asin function.

- **R**
  No other libraries required.

2 Examples

2.1 Excel



	A	B	C	D	E	F	G
1	data						
2	1						
3	2		value	-0,37731	=es_cohen_h2_os(A2:A20;;;C3)		
4	2		qual	medium	=es_cohen_h2_os(A2:A20;;;C4)		
5	1						
6	2		Cohen h2	Qualification			
7	2		-0,37731	medium			
8	1						
9	1		C6:D7	=es_cohen_h2_os_arr(A2:A20)			
10	2						
11	2						
12	2						
13	2						
14	2						
15	2						
16	1						
17	2						
18	1						
19	2						
20	2						

2.2 Python

```
[2]: #example
dataList = ['Female', 'Male', 'Male', 'Female', 'Male', 'Male', 'Female', 'Female',
            'Male', 'Male', 'Male', 'Male', 'Male', 'Male', 'Female', 'Male',
            'Female', 'Male', 'Male']
data = pd.Series(dataList)

[3]: es_cohen_h2_os(data)
```

[3]:

	Cohen h2	Classification
0	0.37731	Medium

```
[4]: codes = ["Female", "Male"]
es_cohen_h2_os(data, codes)
```

[4]:

	Cohen h2	Classification
0	-0.37731	Medium

2.3 R

```
> data <- c("Female", "Male", "Male", "Female", "Male", "Male", "Female",
+           "Female", "Male", "Male", "Male", "Male", "Male", "Male",
+           "Female", "Male", "Female", "Male", "Male")
> es_cohen_h2_os(data)
        h2 classify
1 -0.37731   medium
> 
```

3 Details of Calculations

3.1 The Effect Size

For one-sample:

$$h_2 = \phi_1 - \phi_{h_0}$$

With

$$\phi_i = 2 \times \arcsin\left(\sqrt{p_i}\right)$$

$$p_i = RF_i = \frac{F_i}{n}$$

$$n = \sum_{i=1}^{k} F_i$$

Symbols:

F_i the (absolute) frequency (count) of category i

n the sample size, i.e. the sum of all frequencies

p_i the proportion of cases in category i

p_{h_0} the expected proportion

RF_i the relative frequency of category i

3.2 Interpretation

To convert an h_2 to h use:

$$h = h_2 \times \sqrt{2}$$

Table 1

Rule of thumb for Cohen h interpretation

Cohen's h	*Interpretation*
0.00 < 0.20	Negligible
0.20 < 0.50	Small
0.50 < 0.80	Medium
0.80 or more	Large

Note. Adapted from Cohen (1988, pp. 184–185)

4 Source

Cohen's g can be found in *Statistical power analysis for the behavioral sciences* (2nd ed) (Cohen, 1988), on page 147.

(5.2.1) $$\mathbf{g} = \mathbf{P} - .50 \text{ or } .50 - \mathbf{P} \quad \text{(directional)},$$

and

$$\mathbf{g} = |\mathbf{P} - .50| \quad \text{(nondirectional)}.$$

(Cohen, 1988, p. 147)

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.