One-Sample Student t-Test (ts_student_t_os)

Author: P. Stikker
Website: https://peterStatistics.com
YouTube: https://www.youtube.com/stikpet
Version: 0.1 (2023-01-31)

Introduction

The *ts_student_t_os* function calculates a one-sample Student t-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**
  - 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*
  - **mu**
    the hypothesized mean. If not specified the midrange will be used.

1.2 Output

- **mu**
  The hypothesized mean
- **sample.mean**
  The mean from the sample data
- **statistic**
  The t-value
- **df**
  The degrees of freedom
- **p-value**
  The two-tailed significance of the test (p-value)
- **testUsed**
  The test used

Note for *Excel:*

the array function *ts_student_t_os_arr* will require 2 rows and 6 columns.

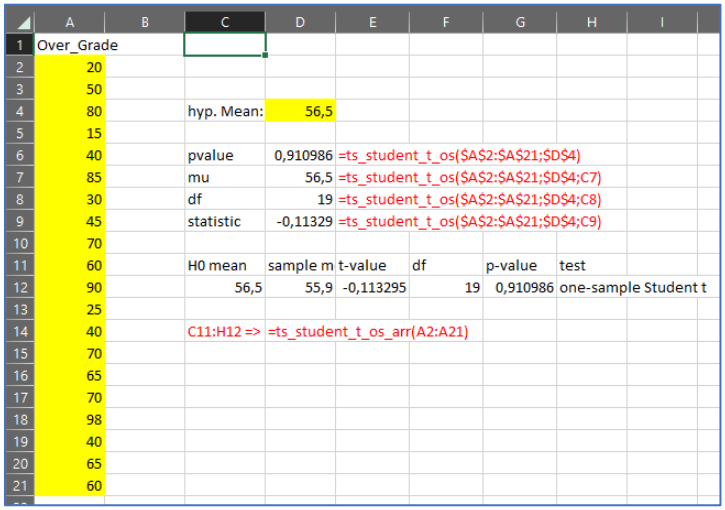
1.3 Dependencies

- **Excel**
  - None, but you can run the **ts_student_t_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	H	I
1	Over_Grade								
2	20								
3	50								
4	80		hyp. Mean:	56,5					
5	15								
6	40		pvalue	0,910986	=ts_student_t_os($A$2:$A$21;$D$4)				
7	85		mu	56,5	=ts_student_t_os($A$2:$A$21;$D$4;C7)				
8	30		df	19	=ts_student_t_os($A$2:$A$21;$D$4;C8)				
9	45		statistic	-0,11329	=ts_student_t_os($A$2:$A$21;$D$4;C9)				
10	70								
11	60		H0 mean	sample m	t-value	df	p-value	test	
12	90		56,5	55,9	-0,113295	19	0,910986	one-sample Student t	
13	25								
14	40		C11:H12 =>	=ts_student_t_os_arr(A2:A21)					
15	70								
16	65								
17	70								
18	98								
19	40								
20	65								
21	60								

2.2 Python

```
[1]: from test_student_t_os import ts_student_t_os

     #import pandas as pd
     #from statistics import mean, stdev
     #from scipy.stats import t

     data = [20, 50, 80, 15, 40, 85, 30, 45, 70, 60, 90, 25,
             40, 70, 65, 70, 98, 40, 65, 60]

     ts_student_t_os(data)
```

[1]:

	mu	sample mean	statistic	df	p-value	test used
0	56.5	55.9	-0.113295	19	0.910986	one-sample Student t

```
[2]: ts_student_t_os(data, mu=70)
```

[2]:

	mu	sample mean	statistic	df	p-value	test used
0	70	55.9	-2.662423	19	0.015385	one-sample Student t

2.3 R

```
> source("test_student_t_os.R")
> data <- c(20, 50, 80, 15, 40, 85, 30, 45, 70, 60, 90, 25, 40, 70, 65, 70, 98, 40, 65, 60)
> ts_student_t_os(data)
   mu sample.mean    t.value df      pVal             testUsed
t 56.5        55.9 -0.1132946 19 0.9109855 one-sample Student t
> ts_student_t_os(data, mu=70)
  mu sample.mean   t.value df       pVal             testUsed
t 70        55.9 -2.662423 19 0.01538494 one-sample Student t
```

3 Details of Calculations

The two-sided p-value can be obtained by:

$$sig = 2 \times \big(1 - T(|t|, df)\big)$$

With:

$$t = \frac{\bar{x} - \mu_{H0}}{SE}$$

$$df = n - 1$$

$$SE = \frac{s}{\sqrt{n}}$$

Symbols:

- T the cumulative distribution function of the t-distribution
- $\bar{x}$ the sample mean
- μ_{H0} the hypothesized mean in the population
- SE the standard error (i.e. the standard deviation of the sampling distribution)
- df the degrees of freedom
- n the sample size (i.e. the number of scores)
- s the unbiased sample standard deviation

4 Sources

The Student t test (Student, 1908) was described by Gosset under the pseudo name Student.

References

Student. (1908). The probable error of a mean. *Biometrika*, *6*(1), 1–25.

https://doi.org/10.1093/biomet/6.1.1