t 検定

青木繁伸

1　目的

母平均の検定，2 群の平均値の差の検定を行う。

2　使用法

原データを用いる場合

```
from t_test import t_test
t_test(x, y=None, alternative="two.sided", mu=0.0, paired=False, var_equal=False,
   conf_level=0.95, verbose=True)
```

二次データを用いる場合

```
from t_test2 import t_test2
t_test2(nx, mx, ux, ny, my, uy, var_equal = False, verbose=True)
```

2.1　引数

x	データベクトル
y	データベクトル
alternative	対立仮説: "two.sided"（デフォルト），"less"，"greater"
mu	対応のある検定の場合の母平均（デフォルトは 0）
paired	対応のあるデータの場合に True を指定する（デフォルトは False）
var_equal	等分散性を仮定するかどうか（デフォルトは False）
conf_level	信頼区間を求めるときの信頼率（デフォルトは 0.95）
verbose	必要最小限のプリント出力をする（デフォルトは
nx	第一群のデータ個数
mx	第一群の平均値
ux	第一群の不偏分散
ny	第二群のデータ個数（一標本の場合は None）
my	第二群の平均値（一標本の場合は None）
uy	第二群の不偏分散（一標本の場合は None）True）

2.2　戻り値の名前

"t"	検定統計量（t にしたがう）
"df"	自由度
"p value"	p 値
"conf int"	信頼区間

"conf level"　信頼率
"estimate"　標本平均
"mu"　母平均
"alternative"　両側検定・片側検定の種別
"method"　検定手法名

3 使用例

```
x = [1, 2, 3, 2, 5, 4]
y = [2, 2, 1, 5, 4, 2]
```

3.1 独立 2 標本，等分散を仮定しない

```
import sys
sys.path.append("statlib")
from t_test import t_test

a = t_test(x, y)
```

```
Welch Two Sample t-test
t = 0.19389,  df = 9.9949,  p-value = 0.8501
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.74874 2.08208
sample estimates:
mean of x mean of y 2.83333333 2.66666667
```

二次データを使っても同じ結果になることを確認する。二次データの有効数字が少ないと p 値に影響が出る。

```
from t_test import t_test2
import scipy as sp

a = t_test2(len(x), sp.mean(x), sp.var(x, ddof=1), len(y), sp.mean(y),
    sp.var(y, ddof=1))
```

```
Welch Two Sample t-test
t = 0.19389,  df = 9.9949,  p-value = 0.8501
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.74874 2.08208
sample estimates:
mean of x mean of y 2.83333333 2.66666667
```

3.2 独立 2 標本，等分散を仮定する

```
a = t_test(x, y, var_equal=True)
```

```
Two Sample t-test
t = 0.19389,  df = 10,  p-value = 0.8501
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.74861 2.08194
sample estimates:
mean of x mean of y 2.83333333 2.66666667
```

3.3 独立 2 標本，等分散を仮定しない，片側検定

```
a = t_test(x, y, alternative="less")
```

```
Welch Two Sample t-test
t = 0.19389,  df = 9.9949,  p-value = 0.5749
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval: -inf 1.72471
sample estimates:
mean of x mean of y 2.83333333 2.66666667
```

```
a = t_test(x, y, alternative="greater")
```

```
Welch Two Sample t-test
t = 0.19389,  df = 9.9949,  p-value = 0.4251
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval: -1.39138 inf
sample estimates:
mean of x mean of y 2.83333333 2.66666667
```

3.4 対応のある標本

```
a = t_test(x, y, paired=True)
```

```
Paired t-test
t = 0.21035,  df = 5,  p-value = 0.8417
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.87007 2.2034
sample estimate:
mean of the differences
 0.166667
```

```
import scipy as sp
diff = sp.ravel(x)-sp.ravel(y)
a = t_test(diff)
```

```
One Sample t-test
t = 0.21035,  df = 5,  p-value = 0.8417
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.87007 2.2034
```

```
sample estimate:
mean of the differences
 0.166667
```

```
a = t_test2(len(diff), sp.mean(diff), sp.var(diff, ddof=1))
```

```
One Sample t-test
t = 0.21035,  df = 5,  p-value = 0.8417
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.87007 2.2034
sample estimate:
mean of the differences
 0.166667
```

3.5 対応のある標本，片側検定

```
a = t_test(x, y, paired=True, alternative="less")
```

```
Paired t-test
t = 0.21035,  df = 5,  p-value = 0.5792
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval: -inf 1.76324
sample estimate:
mean of the differences
 0.166667
```

```
a = t_test(x, y, paired=True, alternative="greater")
```

```
Paired t-test
t = 0.21035,  df = 5,  p-value = 0.4208
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval: -1.42991 inf
sample estimate:
mean of the differences
 0.166667
```

```
print(a)
```

```
{'t': 0.21035158095583562, 'df': 5, 'p value': 0.4208483515552009,
'conf int': (-1.4299051010319612, inf), 'conf level': 0.95,
'estimate': 0.16666666666666666, 'mu': 0, 'alternative': 'greater',
'method': 'Paired t-test'}
```