

一元配置分散分析 (exact test)

青木繁伸

2020年3月17日

1 目的

正確な p 値を計算する一元配置分散分析 (正確確率検定; 並べ替え検定 permutation test) である。データによっては計算量が多くなり実用的な時間内で計算が終了できないこともあるので、そのような場合にはモンテカルロ法による近似計算もできる。

なお、ここで取り上げる一元配置分散分析は、各群の分散が等しいことを仮定しないものである (ウェルチの方法)。

2 使用法

```
import sys
sys.path.append("statlib")
from exact import exact_oneway_test
exact_oneway_test(x, g=None, permutation=True, hybrid=False, loop=10000, verbose=True)
```

2.1 引数

| | |
|--------------------------|--|
| <code>x</code> | リストまたは二重リスト |
| <code>g</code> | <code>x</code> が二重リストでない場合は、グループを表すリスト |
| <code>permutation</code> | 並べ替え検定を行う (デフォルトで <code>True</code>) |
| <code>hybrid</code> | シミュレーションを行う (デフォルトは <code>False</code>) |
| <code>loop</code> | シミュレーションの回数 |
| <code>verbose</code> | 必要最小限のプリント出力をする |

2.2 戻り値の名前

| | |
|-----------------------|--------------------------|
| <code>"obsP"</code> | 観察されたデータの一元配置分散分析の p 値 |
| <code>"exactP"</code> | 正確な p 値 |
| <code>"simP"</code> | シミュレーションによる p 値 |

3 使用例

3.1 正確な検定

```
x = [[36.7, 52.4, 65.8], [45.7, 61.9, 65.3], [52.6, 76.6, 81.3]]

import sys
sys.path.append("statlib")
from exact import exact_oneway_test

a = exact_oneway_test(x)
```

Oneway ANOVA

F = 1.0225, df1 = 2, df2 = 3.8705, p value = 0.44004

p value = 0.41071

1680 tables are inspected.

```
dat = [36.7, 52.4, 45.7, 65.8, 61.9, 52.6, 65.3, 81.3, 76.6]
grp = [1, 1, 2, 1, 2, 3, 2, 3, 3 ]

a = exact_oneway_test(dat, grp)
```

Oneway ANOVA

F = 1.0225, df1 = 2, df2 = 3.8705, p value = 0.44004

p value = 0.41071

1680 tables are inspected.

3.2 シミュレーションによる検定

```
x = [[36.7, 52.4, 65.8], [45.7, 61.9, 65.3], [52.6, 76.6, 81.3]]

import numpy as np
np.random.seed(123) # 乱数の種 (通常は不要)
a = exact_oneway_test(x, hybrid=True, loop=1000)
```

Oneway ANOVA

F = 1.0225, df1 = 2, df2 = 3.8705, p value = 0.44004

simulated p value = 0.38300

based on 1000 times simulation.

```
dat = [36.7, 52.4, 45.7, 65.8, 61.9, 52.6, 65.3, 81.3, 76.6]
grp = [1, 1, 2, 1, 2, 3, 2, 3, 3 ]
```

```
import numpy as np
np.random.seed(123) # 乱数の種 (通常は不要)
a = exact_oneway_test(dat, grp, hybrid=True, loop=1000)
```

Oneway ANOVA

F = 1.0225, df1 = 2, df2 = 3.8705, p value = 0.44004
simulated p value = 0.38300
based on 1000 times simulation.

3.3 二群の平均値の差の検定 (t 検定)

二群の平均値の差の検定 (t 検定) も行える。

```
a = exact_oneway_test([[3,2,4,5], [4,2,5,7,9]])
```

Difference of means (two independent samples)

t = 1.3869, df = 5.9615, p value = 0.21509
p value = 0.22222
48 tables are inspected.

```
import numpy as np
np.random.seed(123) # 乱数の種 (通常は不要)
b = exact_oneway_test([[3,2,4,5], [4,2,5,7,9]], hybrid=True)
```

Difference of means (two independent samples)

t = 1.3869, df = 5.9615, p value = 0.21509
simulated p value = 0.22520
based on 10000 times simulation.