

Brown-Forsythe 検定

青木繁伸

2019 年 3 月 7 日

1 目的

Brown-Forsythe 検定を行う。

普通は、R にも用意されている `Bartlett.test` 関数を使う方がよいと思われる。

2 使用法

原データを用いる場合

```
import sys
sys.path.append("statlib")
from xtest import Brown_Forsythe_test
Brown_Forsythe_test(x, g=None, verbose=True)
```

二次データを用いる場合

```
import sys
sys.path.append("statlib")
from xtest import Brown_Forsythe_test2
Brown_Forsythe_test2(n, m, u, verbose=True)
```

2.1 引数

<code>x</code>	データベクトル (リスト) または群ごとのデータの二重リスト
<code>g</code>	<code>x</code> がデータベクトルの場合はそれぞれのデータの所属群を表すベクトル (リスト) <code>x</code> が二重リストの場合は省略される
<code>n</code>	各群のサンプルサイズを指定するベクトルまたはリスト
<code>m</code>	各群の平均値を指定するベクトルまたはリスト
<code>u</code>	各群の不偏分散を指定するベクトルまたはリスト
<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>"F"</code>	検定統計量 (F 分布にしたがう)
<code>"df1"</code>	第 1 自由度
<code>"df2"</code>	第 2 自由度
<code>"pvalue"</code>	p 値

3 使用例

3.1 使用例 1

データと所属群の2つのベクトル（リスト）で与える。

```
import sys
sys.path.append("statlib")
from xtest import Brown_Forsythe_test

a = Brown_Forsythe_test([3, 2, 4, 2, 3, 5, 4, 3, 3, 1, 3, 4, 5, 4, 5],
                        [1, 1, 2, 1, 3, 3, 1, 3, 1, 2, 2, 2, 1, 2, 2])
```

```
Brown-Forsythe test of homogeneity of variances
F = 0.19824, df1 = 2, df2 = 9.5098, p value = 0.82348
```

群の指定は数値でも文字列（1文字に限らない）でもよい。

```
b = Brown_Forsythe_test([3, 2, 4, 2, 3, 5, 4, 3, 3, 1, 3, 4, 5, 4, 5],
                        ["a", "a", "b", "a", "c", "c", "a", "c", "a", "b",
                         "b", "b", "a", "b", "b"])
```

```
Brown-Forsythe test of homogeneity of variances
F = 0.19824, df1 = 2, df2 = 9.5098, p value = 0.82348
```

3.2 使用例 2

複数の群のデータを二重リストで与える。

使用例 1 と同じデータを群ごとに整理して与える。

```
c = Brown_Forsythe_test([[3, 2, 2, 4, 3, 5],
                          [4, 1, 3, 4, 4, 5],
                          [3, 5, 3]])
```

```
Brown-Forsythe test of homogeneity of variances
F = 0.19824, df1 = 2, df2 = 9.5098, p value = 0.82348
```

3.3 使用例 3

`Brown_Forsythe_test2()` により、サンプルサイズ、平均値、不偏分散しかわからない二次データに基づいて検定を行うことができる。

```
import scipy as sp

w = [[3, 2, 2, 4, 3, 5],
      [4, 1, 3, 4, 4, 5],
      [3, 5, 3]]
n = list(map(len, w))
```

```
m = list(map(sp.mean, w))
u = list(map(lambda x: sp.var(x, ddof=1), w))
```

各群のサンプルサイズ (n), 平均値 (m), 不偏分散 (u) がわかっているならば, 元データは分からなくても検定できる。

```
print("n =", n)
```

```
n = [6, 6, 3]
```

```
print("m =", m)
```

```
m = [3.1666666666666665, 3.5, 3.6666666666666665]
```

```
print("u =", u)
```

```
u = [1.3666666666666667, 1.9, 1.3333333333333333]
```

```
from xtest import Brown_Forsythe_test2
```

```
d = Brown_Forsythe_test2(n, m, u) # 元データは不要
```

```
Brown-Forsythe test of homogeneity of variances
```

```
F = 0.19824, df1 = 2, df2 = 9.5098, p value = 0.82348
```

4 参考文献

1) Mehmet Mendes and Akin Pala: Evaluation of Four Tests When Normality and Homogeneity of Variance Assumptions are Violated, Journal of Applied Sciences 4(1): 38-42, 2004.

2) Brown, M. B., and A. B. Forsythe: The small sample behavior of some statistics which test the equality of several means., Technometrics, 16: 129-132, 1974.