

マハラノビスの距離による基準群への帰属確率

青木繁伸

2020年3月17日

1 目的

基準となる群の重心からのマハラノビスの距離を計算することにより、あるデータポイントがその基準群に帰属する確率を計算する。

Rには、`mahalanobis`関数がある。使い方と出力の違いに注意のこと。

2 使用法

```
import sys
sys.path.append("statlib")
from multi import Mahalanobis_distance
Mahalanobis_distance(dat, x, verbose=True)
```

2.1 引数

<code>dat</code>	基準群のデータ行列
<code>x</code>	所属確率を計算したいデータ行列
<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>"d2"</code>	マハラノビスの距離
<code>"pvalue"</code>	所属確率

3 使用例

```
dat = [[1, 2, 5, 3], # データ行列
        [7, 4, 8, 5],
        [5, 4, 7, 1],
        [2, 3, 5, 4],
        [9, 5, 4, 7],
        [2, 1, 4, 2],
```

```

    [5, 4, 7, 4],
    [2, 3, 5, 7],
    [4, 1, 8, 2],
    [3, 2, 5, 8],
    [3, 5, 4, 8]]

# (1, 2, 5, 3) と (3, 1, 4, 1) から dat の重心へのマハラノビス距離を求
める

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

import sys
sys.path.append("statlib")
from multi import Mahalanobis_distance

a = Mahalanobis_distance(dat, x)

    d2    p value
0  2.000707  0.735629
1  8.637195  0.070836

```