

適合度の検定 (exact test)

青木繁伸

2020年3月17日

1 目的

カイ二乗分布を用いる適合度の検定の exact test である。

データの大きさが固定された、可能性のある全ての度数分布表を発生させ、その度数分布表についてカイ二乗分布を用いる適合度検定の検定統計量を求め、観察された度数分布表における検定統計量より大きい場合に、その度数分布表の生起確率（多項分布に基づく）を累積したものを、正確な p 値とするものである。

なお、度数分布表によっては計算量が多くなり実用的な時間内に計算が終わらない可能性があるため、注意されたい。**Python** では、**R** による計算時間の約 5 倍の時間が必要である。

2 使用法

```
import sys
sys.path.append("statlib")
from exact import goodness_of_fitness_test
goodness_of_fitness_test(o, p = None, verbose = True)
```

2.1 引数

o	度数
p	理論比（省略した場合は全てのカテゴリーの理論比が等しいと仮定する） 合計が 1 でなくても良い $p=c(9, 3, 3, 1)$ のように指定できる。
verbose	必要最小限のプリント出力をする

2.2 戻り値の名前

"chisq"	χ^2 近似検定統計量
"df"	自由度
"pvalue"	p 値
"exactpvalue"	正確な p 値

3 使用例

```
import sys
sys.path.append("statlib")
from exact import goodness_of_fitness_test

a = goodness_of_fitness_test([3, 5, 7])
```

```
Exact goodness of fitness test
chisq = 1.6, df = 2, p value = 0.44933
exact p value = 0.50148
```

```
a = goodness_of_fitness_test([15, 8, 10, 1], [9/16, 3/16, 3/16, 1/16])
```

```
Exact goodness of fitness test
chisq = 3.9608, df = 3, p value = 0.26573
exact p value = 0.25827
```

```
a = goodness_of_fitness_test([15, 8, 10, 1], p = [9, 3, 3, 1])
```

```
Exact goodness of fitness test
chisq = 3.9608, df = 3, p value = 0.26573
exact p value = 0.25827
```