

# 中央値検定

青木繁伸

2020年3月17日

## 1 目的

中央値検定を行う。

中央値検定とは、独立  $k$  標本 ( $k \geq 2$ ) において、まず全体の中央値  $Med$  を求め、各群のデータにおいて  $Med$  以下と  $Med$  より大きいものに二分類し、 $2 \times k$  集計表にまとめる。次に、この集計表に対してピアソンの  $\chi^2$  検定を行う。これが各群の中央値に差があるかどうかの検定を行うことになる。

## 2 使用法

```
import sys
sys.path.append("statlib")
from xtest import median_test
median_test(x, g=None, 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>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>"chisq"</code>	検定統計量
<code>"df"</code>	自由度
<code>"pvalue"</code>	$p$ 値
<code>"median"</code>	中央値
<code>"tbl"</code>	中央値により分割した分割表
<code>"method"</code>	検定手法名

### 3 使用例

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

x = [3, 2, 4, 2, 3, 5, 4, 3, 3, 1, 3, 4, 5, 4, 5]
g = [1, 1, 2, 1, 3, 3, 1, 3, 1, 2, 2, 2, 1, 2, 2]
a = median_test(x, g)
```

```
median test
overall median: 3
[[4 2 2]
 [2 4 1]]
chisq = 1.6071, df = 2, p value = 0.44773
```

```
b = median_test([
    [83, 91, 94, 89, 89, 96, 91, 92, 90],
    [91, 90, 81, 83, 84, 83, 88, 91, 89, 84],
    [101, 100, 91, 93, 96, 95, 94],
    [78, 82, 81, 77, 79, 81, 80, 81]])
```

```
median test
overall median: 89
[[3 7 0 8]
 [6 3 7 0]]
chisq = 17.543, df = 3, p value = 0.00055
```