

独立性の検定

青木繁伸

1 目的

R の `chisq.test` 関数を移植したもので、いわゆる χ^2 検定（分布の検定、独立性の検定）を行う。
独立性の検定の場合には、残差分析の結果（残差および標準化残差）を返す。

2 使用法

```
from chisq_test import chisq_test
chisq_test(x, correct=True, p=None, verbose=True)
```

2.1 引数

<code>x</code>	独立性の検定の場合は 2 次元配列、分布の検定の場合はベクトル
<code>correc</code>	独立性の検定で 2×2 分割表の場合にイエーツの補正を行う（デフォルトは <code>True</code> ）。
<code>p</code>	分布の検定の場合に各カテゴリーの理論比。デフォルトは <code>None</code> で、その場合には各カテゴリーの理論比は等しいものと計算され、一様性の検定を行う。
<code>verbose</code>	必要最小限のプリント出力をする（デフォルトは <code>True</code> ）。

2.2 戻り値の名前

<code>"statistic"</code>	検定統計量 (χ^2 分布にしたがう)
<code>"df"</code>	自由度
<code>"p value"</code>	p 値
<code>"E"</code>	各セルの期待値
<code>"residuals"</code>	残差
<code>"stdres"</code>	標準化残差
<code>"method"</code>	検定手法名

3 使用例

```
import sys
sys.path.append("statlib")
from chisq_test import chisq_test

x = [[22, 43, 54], [35, 21, 32]]
a = chisq_test(x)
```

Pearson's Chi-squared test

```
data:
[[22 43 54]
[35 21 32]]
Pearson's Chi-squared test
chisq = 11.777, df = 2, p-value = 0.0028
```

2×2 分割表で、イエーツの補正を行う。

```
x = [[2, 4], [5, 2]]
a = chisq_test(x)
```

```
Pearson's Chi-squared test with Yates' continuity correction
data:
[[2 4]
[5 2]]
Pearson's Chi-squared test with Yates' continuity correction
chisq = 0.66511, df = 1, p-value = 0.4148
Chi-squared approximation may be incorrect
```

2×2 分割表で、イエーツの補正を行わない。

```
x = [[2, 4], [5, 2]]
a = chisq_test(x, correct=False)
```

```
Pearson's Chi-squared test
data:
[[2 4]
[5 2]]
Pearson's Chi-squared test
chisq = 1.8866, df = 1, p-value = 0.1696
Chi-squared approximation may be incorrect
```

一様性の検定 (理論比が等しい) を行う。

```
x = [1, 3, 4]
a = chisq_test(x)
```

```
Chi-squared test for given probabilities
data: [1 3 4]
Chi-squared test for given probabilities
chisq = 1.75, df = 2, p-value = 0.4169
Chi-squared approximation may be incorrect
```

理論比を与えて分布の検定を行う。

```
x = [1, 3, 4]
a = chisq_test(x, p=[0.25, 0.25, 0.50])
```

```
Chi-squared test for given probabilities
data: [1 3 4]
Chi-squared test for given probabilities
```

```
chisq = 1, df = 2, p-value = 0.6065
Chi-squared approximation may be incorrect
```

残差分析を行う。

```
x = [[4,5,2,0], [0,7,6,1], [1,0,3,1]]
a = chisq_test(x)
```

Pearson's Chi-squared test

data:

```
[[4 5 2 0]
```

```
[0 7 6 1]
```

```
[1 0 3 1]]
```

Pearson's Chi-squared test

```
chisq = 11.344, df = 6, p-value = 0.0783
```

Chi-squared approximation may be incorrect

```
print(a["residuals"])
```

```
[[ 1.60018938  0.28603878 -1.01245685 -0.85634884]
```

```
[-1.52752523  0.59160798  0.38251843  0.06900656]
```

```
[ 0.18257419 -1.41421356  0.86164044  1.15470054]]
```

```
print(a["stdres"])
```

```
[[ 2.20265173  0.46401617 -1.59861608 -1.11382292]
```

```
[-2.29128785  1.04582503  0.65816809  0.0978076 ]
```

```
[ 0.21908902 -2.          1.18604324  1.30930734]]
```