

カテゴリーデータの再カテゴリー化

青木繁伸

1 目的

カテゴリーデータの再カテゴリー化を行う。

R に用意されている `factor` 関数の `labels` 引数を指定することで同じことができるが、ちょっと注意が必要なので、間違いが少ない方法を紹介する。

2 使用法

```
from recode2 import recode2
recode2(x, y)
```

2.1 引数

`x` 変数データベクトル (主に離散変数)

`y` 再カテゴリー化を定義する $n \times 3$ 行列

i 行の 1 列目 2 列目を $y[i,1], y[i,2]$ としたとき、 $y[i,1] \leq x \leq y[i,2]$ である x を $y[i,3]$ に置き換える。

2.2 戻り値

再カテゴリー化した x

3 使用例

```
y = [[1, 3, 1],
      [4, 7, 5],
      [8, 10, 9]]
x = [7, 4, 3, 1, 8, 6, 5, 5, 8, 1, 1, 10, 7, 6, 3, 5, 10, 8, 3, 3]

import sys
sys.path.append("statlib")
from recode2 import recode2

z = recode2(x, y)
print(z)
```

```
[ 7 4 3 1 8 6 5 5 8 1 1 10 7 6 3 5 10 8 3 3]
```