

混合分布に従う 2 変数データの生成

青木繁伸

2020年3月17日

1 目的

指定した平均値行列 `mu` と標準偏差行列 `sigma`, 相関係数ベクトル `r` を持つ母集団から, 比率ベクトル `prob` で抽出される `n` 組の混合 2 変数正規乱数を発生させる。

2 使用法

```
import sys
sys.path.append("statlib")
from multi import mix2
mix2(n, mu, sigma, r, prob=None)
```

2.1 引数

<code>n</code>	データ組数
<code>mu</code>	平均値行列 (グループ数 × 2)
<code>sigma</code>	標準偏差行列 (グループ数 × 2)
<code>r</code>	相関係数ベクトル (グループ数)
<code>prob</code>	抽出割合 (グループ数) デフォルト (<code>None</code>) の場合は均等抽出

2.2 戻り値の名前

<code>"z"</code>	データ行列 (データ組数 × 2)
<code>"g"</code>	データの各列がどのグループからのものであるかを示す

3 使用例

```
n = 3000
mu = [[50, 60], [45, 50], [65, 48]]
sigma = [[8, 3], [7, 4], [2, 2]]
r = [0.8, 0.6, 0.3]
```

```
import sys
sys.path.append("statlib")
from multi import mix2

a = mix2(n, mu, sigma, r)
import matplotlib.pyplot as plt
dat = a["z"]
plt.scatter(dat[:,0], dat[:,1], c=a["g"], alpha=0.1)
```

