

# 指定された相関係数行列を持つ変数の生成

青木繁伸

2020年3月17日

特定の相関係数行列を持つ多変量データを生成する。  
RのMASSパッケージのmvrnorm関数でもできる。

## 1 使用法

```
import sys
sys.path.append("statlib")
from multi import gendat
gendat(n, r)
```

### 1.1 引数

n サンプルサイズ  
r 1個の相関係数または相関係数行列またはその下三角行列のベクトル

### 1.2 戻り値

データ行列

## 2 使用例

### 2.1 2変数データの生成

第2引数に、1個の相関係数を指定する。

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

a = gendat(10, 0.2)

import numpy as np
print(np.corrcoef(a.T))
```

```
[[1. 0.2]
```

```
[0.2 1. ]]
```

## 2.2 3変数以上のデータ生成

下三角行列をベクトルで与える。

```
x = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6]
a = gendat(10, x)
print(np.corrcoef(a.T))
```

```
[[1.  0.1 0.2 0.4]
 [0.1 1.  0.3 0.5]
 [0.2 0.3 1.  0.6]
 [0.4 0.5 0.6 1. ]]
```

相関係数行列を与える。

```
r = [[1.0, 0.5, 0.4],
      [0.5, 1.0, 0.3],
      [0.4, 0.3, 1.0]]
a = gendat(2000, r)
print(np.corrcoef(a.T))
```

```
[[1.  0.5 0.4]
 [0.5 1.  0.3]
 [0.4 0.3 1. ]]
```

```
import matplotlib.pyplot as plt

plt.axes().set_aspect('equal', 'datalim')
plt.scatter(a[:, 0], a[:, 1], s=9, alpha=0.3)
plt.text(1.0, -3.0, "r = "+str(sp.corrcoef(a.T)[0, 1]))
plt.xlabel("x1")
plt.ylabel("x2")
```

