

Excel にある一変量統計関数

青木繁伸

2020年3月17日

1 目的

Excel にある一変量統計関数を R で定義する。

2 使用法

```
import sys
sys.path.append("statlib")
from univ import avedev    etc.
```

<code>avedev(x)</code>	平均偏差
<code>average(x)</code>	平均値
<code>count(x)</code>	サンプルサイズ
<code>devsq(x)</code>	変動
<code>geomean(x)</code>	幾何平均
<code>harmean(x)</code>	調和平均
<code>stdev(x)</code>	不偏標準偏差
<code>stdevp(x)</code>	標準偏差
<code>var(x)</code>	不偏分散
<code>varp(x)</code>	分散
<code>skew(x, ddof=1)</code>	歪度
<code>kurt(x, ddof=1)</code>	尖度
<code>trimmean(x, p)</code>	トリム平均
<code>large(x, k)</code>	大きい方から k 番目のデータ
<code>small(x, k)</code>	小さい方から k 番目のデータ

2.1 引数

<code>x</code>	データベクトル
<code>method</code>	不偏推定量を求めるときには <code>ddof=1</code> (デフォルト)
<code>p</code>	両側から取り除くデータの割合
<code>k</code>	何番目か

2.2 戻り値の名前

戻り値は、それぞれの関数で計算される統計量

3 使用例

```
import sys
sys.path.append("statlib")
from univ import avedev, average, count, devsq, geomean, harmean,
    stdev, stdevp, var, varp, skew, kurt, trimmean, large, small

from scipy.stats import describe

x = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20]

nobs0, minmax0, mean0, variance0, skewness0, kurtosis0 = describe(x,
    ddof=0)
nobs1, minmax1, mean1, variance1, skewness1, kurtosis1 = describe(x,
    ddof=1)
```

```
avedev(x)
```

```
3.619834710743801
```

```
print(average(x), mean0, mean1)
```

```
6.818181818181818 6.818181818181818 6.818181818181818
```

```
print(count(x), nobs0, nobs1)
```

```
11 11 11
```

```
devsq(x)
```

```
273.6363636363636
```

```
from scipy.stats import gmean
print(geomean(x), gmean(x))
```

```
5.183433701120851 5.183433701120851
```

```
from scipy.stats import hmean
print(harmean(x), hmean(x))
```

3.692553616624484 3.692553616624484

```
print(stdevp(x), stdev(x))
```

4.987587899761892 5.231026320296655

```
import numpy as np
print(np.std(x), np.std(x, ddof=1))
```

4.987587899761892 5.231026320296655

```
from scipy.stats import tstd
print(tstd(x, ddof=0), tstd(x, ddof=1)) # ddof=1 がデフォルト(?), トリム指定がない場合は全数対象
```

4.987587899761892 5.231026320296655

```
print(varp(x), var(x))
```

24.876033057851238 27.363636363636363

```
print(variance0, variance1) # ddof=1 がデフォルト
```

24.876033057851238 27.363636363636363

```
print(skew(x, ddof=0), skew(x))
```

1.422437762739346 1.6576281239248674

```
print(skewness0, skewness1) # ddof の補正はない
```

1.4224377627393465 1.4224377627393465

```
import scipy.stats as st
print(st.skew(x), st.skew(x, bias=False)) # bias=False は ddof=1 と同じ意味
```

1.4224377627393465 1.657628123924868

```
print(kurt(x, ddof=0), kurt(x))
```

1.7438878158077706 3.739813026346284

```
print(kurtosis0, kurtosis1) # ddof の補正はない
```

1.7438878158077733 1.7438878158077733

```
from scipy.stats import kurtosis
print(kurtosis(x), kurtosis(x, bias=False)) # bias=False は ddof=1 と同じ意味
```

1.7438878158077733 3.7398130263462876

```
from scipy.stats import tmean
print(trimmean(x, 0.2), tmean(x, limits=(2, 10))) # 指定方法が異なる
```

6.0 6.0

```
large(x, 3)
```

9

```
small(x, 4)
```

4