

情報 I

4-1-12 数学 I 連携 統計基礎

統計用 statistics モジュール

試料PDFをテキストコピーしてコーディングしてください



(12-0) 統計値算出(数学 I)を表計算で実行



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W											
1	<table border="1"><caption>ヒストグラムデータ</caption><thead><tr><th>値</th><th>頻度</th></tr></thead><tbody><tr><td>10.00</td><td>2</td></tr><tr><td>26.00</td><td>4</td></tr><tr><td>42.00</td><td>6</td></tr><tr><td>58.00</td><td>4</td></tr><tr><td>74.00</td><td>4</td></tr></tbody></table>																				値	頻度	10.00	2	26.00	4	42.00	6	58.00	4	74.00	4		
値																					頻度													
10.00																					2													
26.00																					4													
42.00																					6													
58.00																					4													
74.00																					4													
2																					データ個数	"=COUNT(A1:T1)"												
3																					合計	"=SUM(A1:T1)"												
4																					最小	"=MIN(A1:T1)"												
5																					最大	"=MAX(A1:T1)"												
6																					平均値	"=AVERAGE(A1:T1)"												
7																					分散	"=VARP(A1:T1)"												
8																					標準偏差	"=STDEV.P(A1:T1)"												
9																					中央値	"=MEDIAN(A1:T1)"												
10	最頻値	"=MODE(A1:T1)"																																
11																																		
12																																		
13																																		
14																																		
15																																		



(12-1) 統計値算出

```
import statistics # 統計モジュールstatistics導入
a=[2,4,5,5,8] # 配列
print(len(a)) # データ個数
print(sum(a)) # 合計
print(min(a)) # 最小
print(max(a)) # 最大
print(statistics.median(a)) # 中央値
print(statistics.mean(a)) # 平均
print(statistics.p variance(a)) # 分散
print(statistics.pstdev(a)) # 標準偏差
print(statistics.mode(a)) # 最頻値
```





```
import statistics
a=[2, 4, 5, 5, 8]
print(len(a))
print(sum(a))
print(min(a))
print(max(a))
print(statistics.median(a))
print(statistics.mean(a))
print(statistics.pvariance(a))
print(statistics.pstdev(a))
print(statistics.mode(a))
```

```
5
24
2
8
5
4.8
3.76
1.9390
5
```

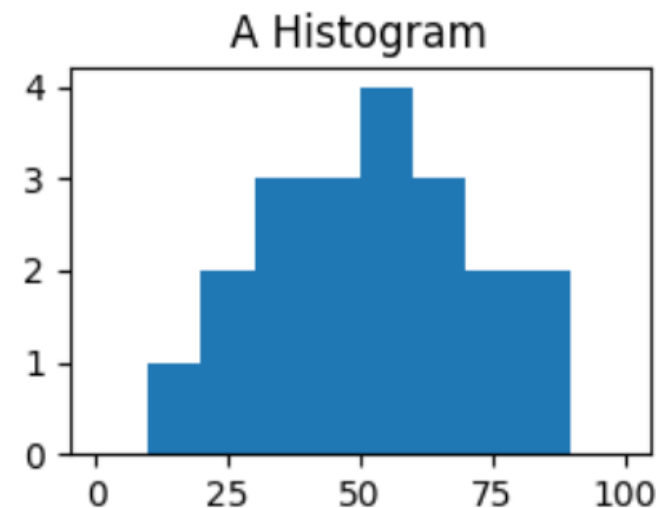


(12-2) 統計値算出 表計算で実施した基礎統計をPythonでプログラミング

http://strnun.fool.jp/pov-ray_strnun/A.csv よりA.csvをDLして配列a=[]にインポート

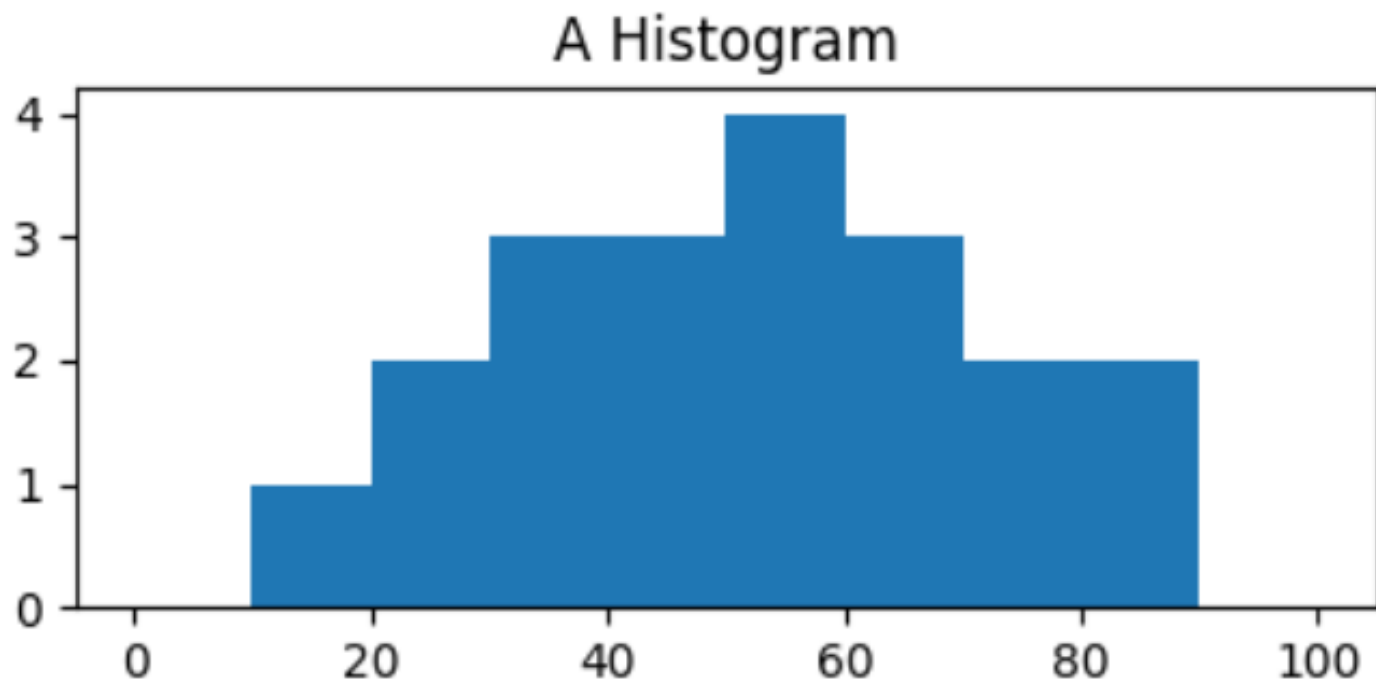
```
import statistics # 統計ライブラリ導入
a=[42,25,36,38,55,14,65,67,78,63,39,59,57,86,53,75,48,45,86,29] # 配列a
print(len(a)) # データ個数
print(sum(a)) # 合計
print(min(a)) # 最小
print(max(a)) # 最大
print(statistics.mean(a)) # 平均
print(statistics.p variance(a)) # 分散
print(statistics.pstdev(a)) # 標準偏差
print(statistics.median(a)) # 中央値
print(statistics.mode(a)) # 最頻値
fig = plt.figure(figsize =(3,2)) # グラフサイズ
plt.hist(a, bins = [0, 10, 20, 30,40, 50, 60, 70,80, 90, 100]) # 階級
plt.title("A Histogram") # タイトル
plt.show()
```

☐→	20
	1060
	14
	86
	53
	379.2
	19.473
	54.0
	86



(12-3) ライブラリnumpy、matplotlibを使うヒストグラム

```
import numpy as np # numpy導入
import matplotlib.pyplot as plt # matplotlib導入
a = np.array([42,25,36,38,55,14,65,67,78,63,39,59,57,86,53,75,48,45,86,29]) # numpy配列
fig = plt.figure(figsize=(5,2)) # matplotlib作図サイズ
plt.hist(a, bins=[0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]) # matplotlib作図区間
plt.title("A Histogram") # matplotlib作図タイトル
plt.show()
```



(13-4) ライブラリmatplotlibを使う箱ひげ図

```
import matplotlib.pyplot as plt
```

```
a = [42,25,36,38,55,14,65,67,78,63,39,59,57,86,53,75,48,45,86,29]
```

```
b = [45,54,72,33,64,36,42,56,51,65,66,55,49,28,61,42,52,54,56,79]
```

```
fig = plt.figure(figsize =(5,2))
```

```
plt.boxplot([a,b], labels=["a","b"])
```

```
plt.title("boxplot")
```

```
plt.show()
```

