

信頼区間の意味

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats
```

正規母集団から標本抽出

```
In [2]: m=12                # 母平均の設定
s=3                # 母標準偏差の設定
n=10               # 母集団から取り出すサンプル数
nsample=np.random.normal(m, s, size=n) # 正規母集団  $N(m, s^2)$  から取り出した n 個のサンプル
```

```
In [3]: nsample
```

```
Out[3]: array([ 8.33933859, 10.21211615, 16.14986579, 13.96338894, 10.45877938,
 6.17681229,  7.11673718, 17.99008649, 12.20453506, 12.97482804])
```

```
In [4]: sample_mean = np.mean(nsample)
sample_mean
```

```
Out[4]: 11.558648791639692
```

信頼係数 1- alpha の信頼区間

```
In [5]: # 標準正規分布の上側  $\alpha/2$  点
normal = stats.norm() # 標準正規分布を準備
alpha = 0.05
normal.isf(alpha/2) # 上側  $\alpha/2$  点 = 両側  $\alpha$  点
```

```
Out[5]: 1.9599639845400545
```

```
In [6]: LCI = sample_mean - normal.isf(alpha/2)*np.sqrt(s**2/n) # 信頼区間の下端
UCI = sample_mean + normal.isf(alpha/2)*np.sqrt(s**2/n) # 信頼区間の上端
LCI, UCI
```

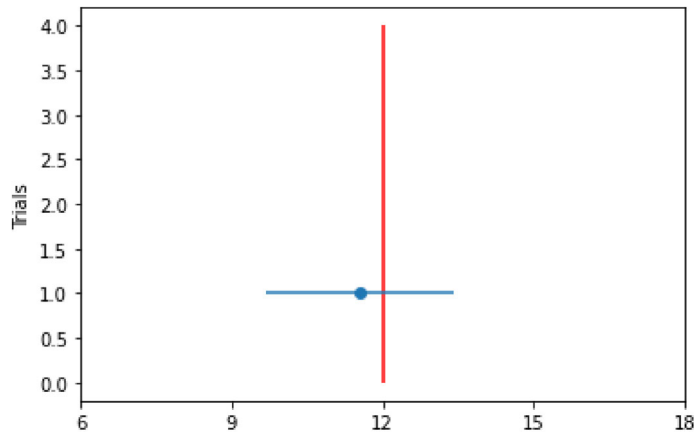
```
Out[6]: (9.699263694726007, 13.418033888553378)
```

```
In [7]: # 信頼区間は母平均を捕まえているか?
LCI <= m <= UCI
```

```
Out[7]: True
```

```
In [8]: # 信頼区間の図示
fig=plt.figure(figsize=(6,4))
ax=fig.add_subplot(1,1,1)
ax.vlines(m,0,4, color='red') # 母平均
ax.hlines(1,LCI,UCI)         # 信頼区間
ax.scatter(sample_mean, 1)   # 標本平均
ax.set_xticks([m-2*s,m-s, m, m+s, m+2*s])
# ax.set_xticklabels(['Population mean'])
ax.set_ylabel('Trials')
```

Out[8]: Text(0, 0.5, 'Trials')



信頼区間の現れ方

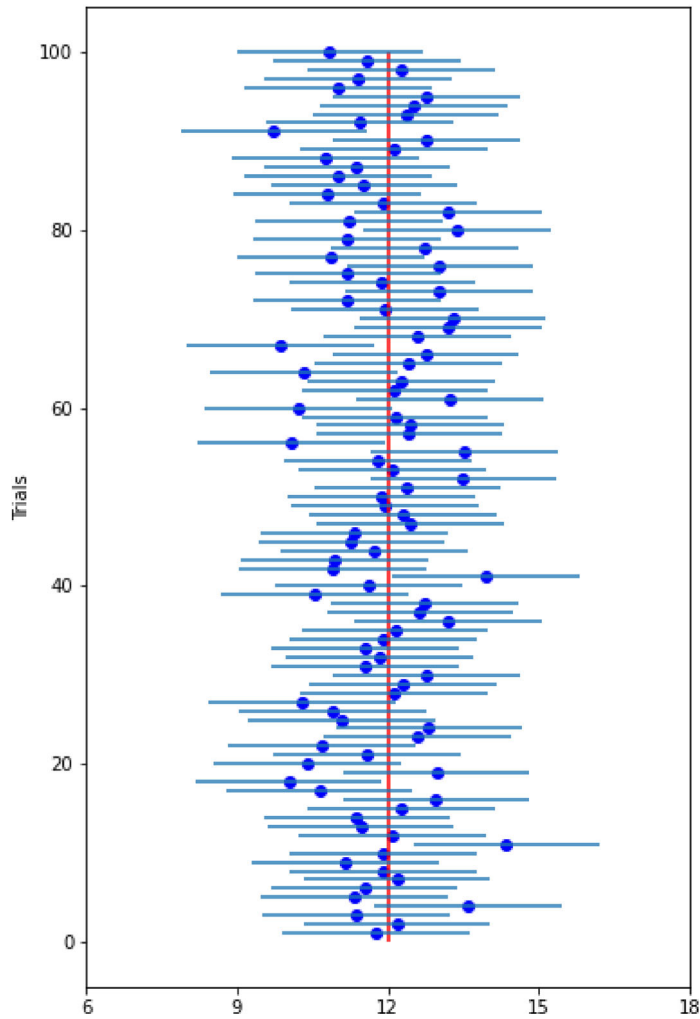
```
In [9]: # 同様の標本抽出を T 回繰り返して信頼区間がどのように現れるかを図示する
T=100 # トライアルの回数
SM_list=[]
for i in range(T):
    nsample=np.random.normal(m, s, size=n) # 正規母集団 N(m, s^2) から取り出した n 個のサンプル
    sample_mean = np.mean(nsample)
    SM_list = SM_list+[sample_mean]
# SM_list
```

```

In [10]: fig=plt.figure(figsize=(6,10))
ax=fig.add_subplot(1,1,1)
ax.vlines(m,0,T, color='red') # 母平均
for i in range(T):
    LCI = SM_list[i] - normal.isf(alpha/2)*np.sqrt(s**2/n) # 信頼区間の下端
    UCI = SM_list[i] + normal.isf(alpha/2)*np.sqrt(s**2/n) # 信頼区間の上端
    ax.hlines(i+1,LCI,UCI)
    ax.scatter(SM_list[i], i+1, color='blue')
ax.set_xticks([m-2*s,m-s, m, m+s, m+2*s])
ax.set_ylabel('Trials')

```

Out[10]: Text(0, 0.5, 'Trials')



```
In [11]: # 信頼区間は母平均を捕まえているか？ 母平均が信頼区間に含まれているトライアル数の割合
ST=0
for i in range(T):
    LCI = SM_list[i] - normal.isf(alpha/2)*np.sqrt(s**2/n) # 信頼区間の下端
    UCI = SM_list[i] + normal.isf(alpha/2)*np.sqrt(s**2/n) # 信頼区間の上端
    if LCI <= m <= UCI:
        ST=ST+1
    else:
        ST=ST+0
ST/T
```

Out[11]: 0.94

In []: