ブートストラップ検定の例（07 年度版資料から）

標本の大きさ(sample size)が 10 である下記のデータがある。

10.9	7.1	8.8	8.4	7.6	9.7	9.0	4.9	5.0	7.6

これは独立かつ同一の正規分布からのランダム標本であると仮定する。

この標本における、標本平均は 7.9 であり、標本標準偏差は 1.907 である。ゆえに、標本平均を母集団平均の推定量としたとき、標準誤差(standard error)は

$$1.907\big/\sqrt{10} = 0.603$$

となる。

① ブートストラップ標準誤差の導出

(A) パラメトリック・ブートストラップ【図１，図３(A)】

平均が 7.9、標準偏差が 1.907 である正規乱数を 10 個発生させ、その標本平均を求める。これを B 回繰り返す。そして計算された標本平均の B 個のデータから、（標本）標準偏差を求める。これが（パラメトリック）ブートストラップ標準誤差となる。

B=59 として計算した結果、$se(\hat{\mu})_{pb} = 0.582$ がえられた。

(B) ノンパラメトリック・ブートストラップ【図１，図３(B)】

上記の 10 個のデータからなる標本から、無作為復元抽出で 10 個データを取り出し、その標本平均を求める。これを B 回繰り返す。そして計算された標本平均の B 個のデータから、（標本）標準偏差を求める。これが（ノンパラメトリック）ブートストラップ標準誤差となる。

B=59 として計算した結果、$se(\hat{\mu})_{nb} = 0.611$ がえられた。

② ブートストラップ検定

$H_0 : \mu = 6.0$ vs $H_1 : \mu > 6.0$ の検定を考える。有意水準を 5%とする。

(A) パラメトリック・ブートストラップ【図４】

平均が 6.0、標準偏差が 1.907 である正規乱数を 10 個発生させ、その標本平均を求める。これを B 回繰り返す。そして計算された標本平均の B 個のデータからなる経験分布の上側 5%点を求める。B=59 として計算した結果、上側 5%点は、B 個のデータを大きい方から並べて３番目の 6.717 である。元の標本の大きさ 10 のデータから計算した標本平均（=母平均の推定量）は、7.9 であるので、経験分布の上側 5%点よりも大きい。ゆえに、5%水準で帰無仮説を棄却する。

(B) ノンパラメトリック・ブートストラップ【図４】

上記の 10 個のデータから、それぞれ標本平均 7.9 を引き、その 10 個の残差データからなる標本を作成する。その標本から無作為復元抽出で 10 個データを取り出し、それに帰無仮説の平均 6.0 を加え、その標本平均を求める。これを B 回繰り返す。そそして計算された標本平均の B 個のデータからなる経験分布の上側 5%点を求める。B=59 として計算した結果、上側 5%点は、B 個のデータを大きい方から並べて３番目の 6.9 である。元の標本の大きさ 10 のデータから計算した標本平均（=母平均の推定量）は、7.9 であるので、経験分布の上側 5%点よりも大きい。ゆえに、5%水準で帰無仮説を棄却する。

図１． Scilab による計算結果

```
         ___________________________________________

                         scilab-4.1.2
                   Copyright (c) 1989-2007
               Consortium Scilab (INRIA, ENPC)
         ___________________________________________


Startup execution:
  loading initial environment

--> x  =
    10.9
    7.1
    8.8
    8.4
    7.6
    9.7
    9.
    4.9
    5.
    7.6
 mu  =
    7.9
 sig  =
    1.9072959
 n  =
    10.
 b  =
    59.
pb  =
    0.
    0.
    0.
    0.
    0.
    0.
    0.
[More (y or n ) ?]

-->mupb
 mupb  =
    7.9302346
-->sigpb
 sigpb  =
    0.5817199

-->munb
 munb  =
    7.9220339
-->signb
 signb  =
    0.6110415
```

図２．図１の計算のための Scilab のコード

```
clear
x={10.9;7.1;8.8;8.4;7.6;9.7;9.0;4.9;5.0;7.6}
mu=mean(x)
sig=stdev(x)
n=10
b=59
//parametric bootstrap
pb=zeros(b,1)
for i=1:b, y=zeros(n,1), y=mu+sig*rand(n,1,'normal'), pb(i,1)=mean(y), end
mupb=mean(pb)
sigpb=stdev(pb)
//nonparamatric bootstrap
nb=zeros(b,1)
for i=1:b, y=zeros(n,1), y=sample(n,x), nb(i,1)=mean(y), end
munb=mean(nb)
signb=stdev(nb)
histplot(10,pb);
clf
histplot(10,nb);
```

図３A.　パラメトリック・ブートストラップによる標本平均の分布

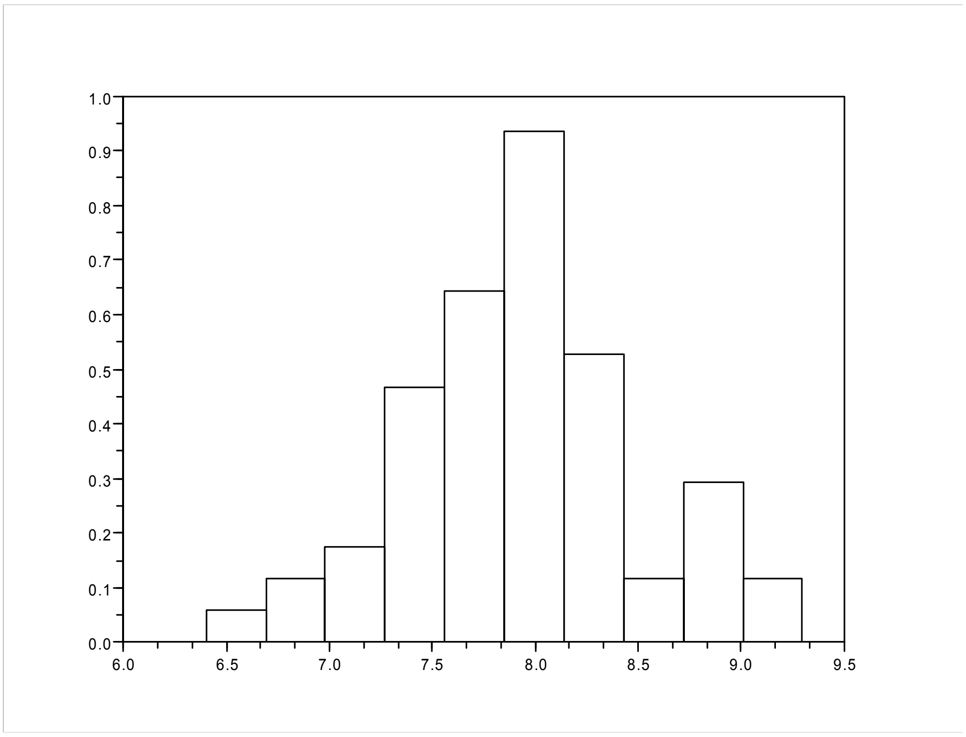

図３B.　ノンパラメトリック・ブートストラップによる標本平均の分布

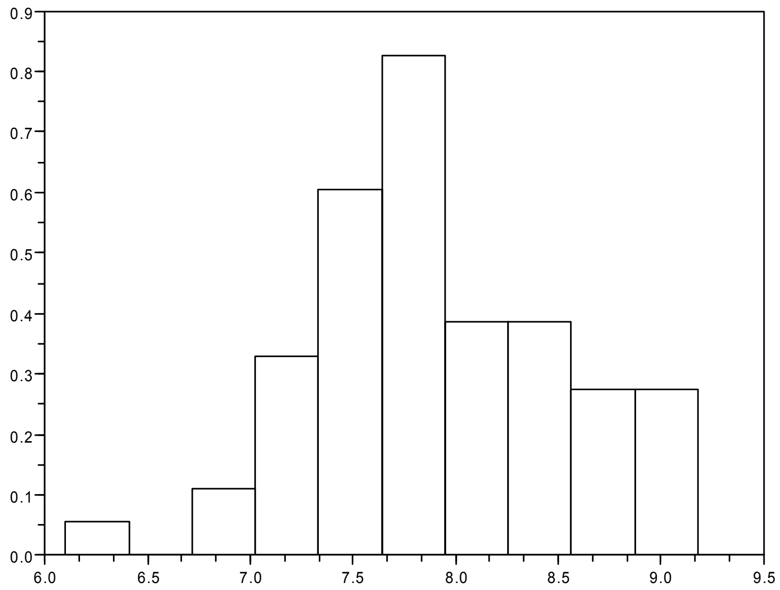

図４. ブートストラップ検定のための経験分布の計算結果（左）とその Scilab コード（右）

```
-->decpb(1:10)
 ans  =

    7.6947878
    6.7342402
    6.7172983
    6.7027829
    6.6880651
    6.6273286
    6.6162382
    6.4716517
    6.4329194
    6.4076237

-->decnb(1:10)
 ans  =

    7.09
    6.92
    6.9
    6.87
    6.86
    6.85
    6.83
    6.74
    6.74
    6.68
```

```
clear
x={10.9;7.1;8.8;8.4;7.6;9.7;9.0;4.9;5.0;7.6}
mu=mean(x)
sig=stdev(x)
format(10);mu
format(10);sig
n=10
b=59
mu0=6.0
//parametric bootstrap
pb=zeros(b,1)
for i=1:b, y=zeros(n,1), y=mu0+sig*rand(n,1,'normal'),
  pb(i,1)=mean(y), end
mupb=mean(pb)
sigpb=stdev(pb)
decpb=sort(pb)
//nonparamatric bootstrap
nb=zeros(b,1)
for i=1:b, y=zeros(n,1), u=zeros(n,1), u=sample(n,x)-mu,
   y=mu0+u, nb(i,1)=mean(y), end
munb=mean(nb)
signb=stdev(nb)
decnb=sort(nb)
```

Scilab について

Scilab はフランスの INRIA（国立計算機科学・自動制御研究所）および ENPC（国立土木学校）で開発された Matlab とほぼ同一のコード記法をもつ数値計算ソフトです。オープンソースのソフトですが、高機能であり、Linux, Windows, Mac-OS/X の各プラットフォーム上で利用可能です。最新版は 2009 年 4 月にリリースされた Scilab-5.1.1 になります。

インストールするには、Scilab コンソーシアムのホームページ (URL: http://www.scilab.org/) からファイルをダウンロードしてください。また、日本語のマニュアルなどは、広島大学大学院工学研究科複雑システム工学専攻複雑システム制御研究室（大野修一先生）のホームページ（URL: http://www.ecl.hiroshima-u.ac.jp/~ohno/scilab/introscilab/introscilab.html）から入手することができます。

日本語でかかれた解説書には、

櫻井鉄也 (2003),『MATLAB/Scilab で理解する数値計算』, 東京大学出版会

上坂吉則 (2007), 『MATLAB+Scilab プログラミング事典』, ソフトバンククリエイティブ

（共に附属図書館総合図書館学習用図書コーナーに配架）

赤間世紀 (2009), 『Scilab 入門講座』, 電波新聞社

大野修一 (2009), 『Scilab 入門』, CQ 出版社

があります。