国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
一九八五年第四期　总第五十二期　No. 4 1985 Sum. 52

Bootstrap方法及其应用

茆　宁

摘　要　1977年埃弗伦教授提出了一种新的统计推断方法——Bootstrap方法。该方法适用于一大类估计问题，因而受到越来越多的统计学者的重视。本文介绍了该方法的基本思想及其应用实例，并讨论了存在的问题及今后研究的方向。

一、引　言

1977年，美国Stanford大学统计系教授Efron在总结、归纳前人研究成果的基础上提出了一种新的统计方法——Bootstrap方法。最近几年来，该方法引起了学术界的重视，我国也已初步开展这方面的研究。从最初的定义看，Bootstrap方法是一种非参数统计方法，但实际上不仅仅局限于非参数的范畴，对于参数统计推断的某些问题同样也是适用的。从另一方面看，Bootstrap方法实质上是一种再抽样(resampling)过程。再抽样的方法还有许多，如Jackknife方法、无穷小Jackknife方法、Delta方法、Half—Sampling方法等。本报告主要介绍Bootstrap方法的基本思想及其若干应用问题，其中包括作者个人的一些看法，在理论上不准备做详细的论述。其它的再抽样方法可阅读报告中列出的有关文献。

下面先简单介绍一下问题的基本提法。设随机子样$X=(X_1,\cdots,X_n)$来自未知的总体分布F；$R(X,F)$为某个预先选定的随机变量，它是X和F的函数。现要求根据子样观测值$X=x=(x_1,\cdots,x_n)$估计$R(X,F)$的分布特性，如均值、方差或分布密度函数等。例如，设$\theta=\theta(F)$是总体分布F的某分布参数（如均值、方差、相关系数），而$\hat{\theta}(X)$是关于θ的估计（如子样均值、子样方差、子样相关系数）。定义

$$R(X,F)=\hat{\theta}(X)-\theta(F) \tag{1.1}$$

现要由观测$X=x$估计$R(X,F)$的分布特性。注意此时$R(X,F)$的均值和方差分别为$\hat{\theta}$的偏和误差方差，正是我们所希望知道的。总体分布F一般是未知的，则上述问题为非参数问题。此外，也可能已知F的形式，但其某些分布参数未知，则此时是参数问题。

其次分析一下Bootstrap方法产生的历史背景。从原则上讲，人们熟悉的参数统计推断方法是在计算能力受到严重束缚（计算费用昂贵、计算速度缓慢）这一特定的历史条件下产生和发展的。在处理问题时总认为总体分布形式已知，特别通常假定是正态分布。在此基础上建立起一套“墨守陈规”的模式，即对任何一个具体问题，总是利用现

一九八五年元月二十六日收到

成的理论和统计数表（如F分布表，t分布表）进行推断。实际上，对总体分布做出假定一般出于计算上的考虑，真实分布通常是无法精确了解的。有人曾经这样说：对于“为什么要采用正态假设”这一问题的最诚实回答是“只有在正态假设下才能获得问题的解”。事实也确实如此。可见用参数方法解决问题在相当程度上受到主观因素的影响。因此，从五十年代开始，非参数统计方法逐渐得到发展并受到重视。特别在电子计算机高度发展的今天，有必要也完全有可能从正态假设的约束中解脱出来，借助于计算机这一强有力的工具，创造出新的具有更高的灵活性、现实性和可靠性的统计推断方法。在解决具体问题时，希望不依赖于统计表，而是一切从零出发，只依据观测信息作分析和判断。Bootstrap方法正是计算机时代的产物。Bootstrap可译作“自助（self—help）”，它说明该方法只依赖于给定的观测信息，不需要其它的假设和增加新的观测。

二、Bootstrap方法的主要内容

1. Bootstrap方法的基本步骤

首先重复一下问题的提法。考虑单子样场合，设随机子样$X=(X_1,\cdots,X_n)$来自未知的总体F，$R(X,F)$为一预先选定的随机变量，它是X和F的函数。现要求根据观测$x=(x_1,\cdots,x_n)$去估计$R(X,F)$的统计特性。

Bootstrap方法实质上为一再抽样过程，其基本步骤如下：

(1) 由子样观测值构造子样经验分布函数$\hat{F}_n$：$\hat{F}_n$在每点x_i处具有质量$\frac{1}{n}$，$i=1,\cdots,n$。 (2.1)

(2) 从$\hat{F}_n$中抽取简单子样$X_i^*=x_i^*$，$i=1,\cdots,n$。称$X^*=(X_1^*,\cdots,X_n^*)$为Bootstrap子样，$X_i^*\sim\hat{F}_n$，$i=1,\cdots,n$。

(3) 用$R^*=R(X^*,\hat{F}_n)$的分布来迫近$R(X,F)$的分布。R^*的分布称为Bootstrap分布。

例如由(1.1)定义的$R(X,F)$，知此时

$$R^*=R(X^*,\hat{F}_n)=\hat{\theta}(X^*)-\theta(\hat{F}_n) \tag{2.2}$$

可见有两处做了近似处理：一是用X^*代替了X，二是用$\hat{F}_n$的分布参数代替了F的分布参数。

2. 几点说明

(1) 从非参数统计的观点看，$\hat{F}_n$是F的非参数极大似然估计。它为一离散型分布，其可能值为$x_i(i=1,\cdots,n)$，且其均值和方差分别是

$$E_*[X^*]=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}, \tag{2.3}$$

$$Var_*[X^*]=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=s^2 \tag{2.4}$$

(2) 以上步骤是对一般非参数情况提出来的。若已知F的分布形式，则可对步骤(1)作适当修正，相应地称为参数Bootstrap方法。如若已知F为正态分布$N(\mu,\sigma^2)$，其中μ、σ^2未知，则可用$\bar{x}$和s^2代替未知参数μ和σ^2，用$N(\bar{x},s^2)$代替$\hat{F}_n$，从中抽取Boots-

trap子样。

(3) 若已知F为一连续分布，为了适当考虑F的连续性，可用$\hat{F}_n$与某一已知的连续分布G的卷积来代替$\hat{F}_n$，从中抽取Bootstrap子样，相应地称为平滑Bootstrap方法，即作

$$\hat{F}^c=\hat{F}_n*(cG) \tag{2.5}$$

其中$c\in[0,1]$。由$\hat{F}^c$中抽取x_i^c，$(i=1,\cdots,n)$，可这样得到：首先分别由$\hat{F}_n$和G中抽取子样x_i^*和y_i，$(i=1,\cdots,n)$，再作

$$x_i^c=\sqrt{1-c^2}x_i^*+cy_i,\ i=1,\cdots,n \tag{2.6}$$

即可。特别有$\hat{F}^0=\hat{F}_n$，$\hat{F}^1=G$.

(4) 设所考虑的随机变量$R(X,F)$关于X对称，即X的各分量X的排列不影响$R(X,F)$的取值。定义再抽样矢量

$$P^*=(P_1^*,\cdots,P_n^*) \tag{2.7}$$

它满足约束：$P_i^*\geqslant 0$，$\sum_{i=1}^{n}P_i^*=1$. Bootstrap方法即是考虑所有形如$\frac{N^*}{n}$的P^*，其中

$$N^*=(N_1^*,\cdots,N_n^*) \tag{2.8}$$

N_i^*为Bootstrap子样X^*中等于x_i的分量x_j^*的个数，$(i=1,\cdots,n)$，$\sum_{i=1}^{n}N_i^*=n$是显然的。

由定义，$P^*=\frac{N^*}{n}$服从多项分布，其均值和方差分别是

$$E[P^*]=\frac{e}{n} \tag{2.9}$$

$$V_{ar}[P^*]=\frac{I}{n^2}-\frac{e^Te}{n^3} \tag{2.10}$$

其中$e=(1,1,\cdots,1)$。由于在$R(X,F)$为对称时，P^*对于计算$R(X^*,\hat{F}_n)$已足够了，故可记

$$R^*=R(P^*)=R(X^*,\hat{F}_n) \tag{2.11}$$

(5) 在再抽样时，Bootstrap子样X^*的容量可以不为n，一般有

$$X^*=(X_1^*,\cdots,X_m^*),\ X_i^*\sim\hat{F}_n \tag{2.12}$$

对于$m\neq n$的場合有时具有重要的统计意义。例如可设$m=n^2$或$m=\sqrt{n}$，这样可以研究分析判断的结论在子样容量变化时的情况。在下面的讨论中我们均考虑$m=n$，但这不是唯一选择。

3. 获得Bootstrap分布的途径

显然，一旦观测子样x给定，$\hat{F}_n$便可唯一确定，Bootstrap子样也可得到，剩下的问题在于如何得到Bootstrap分布。该分布实质上是在给定x之下的条件分布。根据不同情况，可由以下途径获得。

途径1 直接的理论计算

例 1 设F为0-1分布，分布参数$\theta(F)=P\{X=1\}$，则

$$R(X,F)=\bar{X}-\theta(F),\ \bar{X}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

给定观测x后，Bootstrap子样$X^*=(X_1^*,\cdots,X_n^*)$的每个分量X_i^*显然也服从0-1分布，即

$$P_*\{X_i^*=1\}=\bar{x}=\theta(\hat{F}_n)$$
$$P_*\{X_i^*=0\}=1-\bar{x},\ i=1,\cdots,n$$

"$*$"表示概率运算关于$\hat{F}_n$进行。此时

$$R^*=R(X^*,\hat{F}_n)=\bar{X}^*-\bar{x},\ \bar{X}^*=\frac{1}{n}\sum_{i=1}^{n}X_i^*$$

于是有

$$E_*[R^*]=E_*[\bar{X}^*-\bar{x}]=\bar{x}-\bar{x}=0$$
$$Var_*[R^*]=E_*[(\bar{X}^*-\bar{x})^2]=\frac{1}{n}\bar{x}(1-\bar{x})$$

$\bar{x}$作为观测的函数已给定。该结果与Bootstrap子样无关。

途径2　用Monte—Carlo方法做迫近

在大多数场合直接计算是不可能的，因此必须借助于计算机这一有力工具进行统计模拟计算。即从$\hat{F}_n$中重复抽取N个Bootstrap子样$X^*(1)$，…，$X^*(N)$；计算相应的$R^*(1)=R(X^*(1),\hat{F}_n)$，…，$R^*(N)=R(X^*(N),\hat{F}_n)$；从而用$R^*(i)$，$i=1,\cdots,N$的频率曲线（即直方图）作为对Bootstrap分布的迫近，或者对$R(X^*,\hat{F}_n)$的统计特性（一、二阶矩等）进行估计。

例2　考虑参数估计的均方误差。设$f(x;\theta)$为总体的分布密度，其形式已知。则由观测x可对未知参数θ进行最大似然估计$\hat{\theta}_{ML}$。同时，也希望知道该估计的均方误差

$$\sigma^2=E[(\theta-\hat{\theta}_{ML})^2] \tag{2.13}$$

当然σ^2本身也可用ML方法估计，但在大多数场合很难得到问题的解。另一个可能的途径是由Fisher信息限给出σ^2的下界，但这通常也是十分困难的。用Bootstrap方法解决上述问题却十分简单：

(1) 用$f(x^*;\hat{\theta}_{ML})$代替$f(x;\theta)$，从中抽取N个相互独立的简单子样$x^*(i)$，$i=1,\cdots,N$，其中

$$x^*(i)=(x_1^*(i),\cdots,x_n^*(i))$$

同时计算相应的$\hat{\theta}_{ML}^*(i)$，$i=1,\cdots,N$。

(2) 用下式估计σ^2

$$\hat{\sigma^2}=\frac{1}{N-1}\sum_{i=1}^{n}(\hat{\theta}_{ML}^*(i)-\bar{\theta}_{ML}^*)^2 \tag{2.14}$$

$$\bar{\theta}_{ML}^*=\frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_{ML}(i) \tag{2.15}$$

途径3　用Taylor级数展开求R^*的一、二阶矩

如前所述，在$R(X,F)$为对称时，可记

$$R^*=R(P^*)=R(X^*,\hat{F}_n)$$

于是，对$R(P^*)$在P^*的均值$P^*=\frac{e}{n}$处作Taylor展开，便可得Bootstrap分布一、二阶矩

的近似公式。即

$$R(P^*)\cong R\left(\frac{e}{n}\right)+\left(P^*-\frac{e}{n}\right)U+\frac{1}{2}\left(P^*-\frac{e}{n}\right)V\left(P^*-\frac{e}{n}\right)^T \tag{2.16}$$

式中略去了二阶以上项。其中

$$U=\left.\begin{pmatrix}\dfrac{\partial R^*}{\partial P_1^*}\\ \vdots\\ \dfrac{\partial R^*}{\partial P_n^*}\end{pmatrix}\right|_{P^*=\frac{e}{n}} \tag{2.17}$$

$$V=\left.\begin{pmatrix}\dfrac{\partial^2 R^*}{\partial P_1^*\partial P_1^*}\cdots\dfrac{\partial^2 R^*}{\partial P_1^*\partial P_n^*}\\ \vdots\quad\ddots\quad\vdots\\ \dfrac{\partial^2 R^*}{\partial P_n^*\partial P_1^*}\cdots\dfrac{\partial^2 R^*}{\partial P_n^*\partial P_n^*}\end{pmatrix}\right|_{P^*=\frac{e}{n}} \tag{2.18}$$

注意到P^*满足约束$\sum_{i=1}^{n}P_i^*=1$，则$eU=0$，$eVe^T=0$. 从而由(2.16)有

$$E_*[R^*]\cong R\left(\frac{e}{n}\right)+\frac{1}{2}\mathrm{tr}\left[V\left(\frac{I}{n^2}-\frac{e^Te}{n^3}\right)\right]$$

$$=R\left(\frac{e}{n}\right)+\frac{1}{2n}\bar{V} \tag{2.19}$$

$$\bar{V}=\mathrm{tr}\left[V\left(\frac{I}{n}-\frac{e^Te}{n^2}\right)\right]=\mathrm{tr}\frac{V}{n}=\frac{1}{n}\sum_{i=1}^{n}V_{ii} \tag{2.20}$$

另外，略去(2.16)中二阶项，得近似式

$$Var_*[R^*]=E_*[(R^*-E_*[R^*])^2]=U^T\left(\frac{I}{n^2}-\frac{e^Te}{n^3}\right)U$$

$$=\frac{1}{n^2}U^TU=\frac{1}{n^2}\sum_{i=1}^{n}U_i^2 \tag{2.21}$$

(2.19)～(2.21)即为所求R^*的近似一、二阶矩。

例 3 定义

$$R(X,F)=\theta(\hat{F}_n)-\theta(F) \tag{2.22}$$

其意义在下节将专门讨论。此时

$$R^*=R(X^*,\hat{F}_n)=\theta(\hat{F}_n^*)-\theta(\hat{F}_n) \tag{2.23}$$

其中，$\hat{F}_n^*$为Bootstrap子样X^*的经验分布函数，即

$$\hat{F}_n^*:\text{ 在每点}x_i\text{处具有质量}P_i^*,\ i=1,\cdots,n. \tag{2.24}$$

因此，当$P^*=\frac{e}{n}$时，恒有$\hat{F}_n^*=\hat{F}_n$. 故$R\left(\frac{e}{n}\right)=0$.

且

$$E_*[\theta(\hat{F}_n^*)-\theta(\hat{F}_n)]\cong\frac{1}{n}\bar{V} \tag{2.25}$$

$$Var_*[\theta(\hat{F}_n^*)-\theta(\hat{F}_n)]\cong\frac{1}{n^2}\sum_{i=1}^{n}U_i^2 \tag{2.26}$$

可用上式来近似$E_F[\theta(\hat{F})-\theta(F)]$和$Var_F[\theta(\hat{F})-\theta(F)]$。

以上讨论了获得Bootstrap分布的三种途径。由于途径1的适用范围很小，而途径3只能得到R^*的近似一、二阶矩，故实际中最常用的是途径2．在计算机（特别是高速计算机）的支持下，它可以说是畅通无阻的。这也正是Bootstrap方法生命力之所在。

4. 多子样場合的推广

以上就单子样的場合讨论了Bootstrap方法的基本思想，对于多子样问题可以类推。以两个子样为例。设$X=(X_1,\cdots,X_m)$和$Y=(Y_1,\cdots,Y_n)$分别为来自未知总体F和G的两个相互独立的随机子样，$R((X,Y),(F,G))$为预先指定的某一随机变量。在给定观测$X=x$，$Y=y$之下估计$R((X,Y),(F,G))$的抽样分布特性可采取如下步骤：

(1) 分别构造X和Y的经验分布函数$\hat{F}_m$和$\hat{G}_n$；

(2) 分别从$\hat{F}_m$和$\hat{G}_n$中抽取Bootstrap子样：$X_i^*\sim\hat{F}_m$，$Y_j^*\sim\hat{G}_n$，$i=1,\cdots,m$，$j=1,\cdots,n$；

(3) 计算$R^*=R((X^*,Y^*),(\hat{F}_m,\hat{G}_n))$的分布，并以此作为$R((X,Y),(F,G))$真实分布的迫近。显然，上述步骤是单子样时Bootstrap方法的自然推广。

三、Bootstrap方法在统计推断问题中的应用

1.总体中位数(median)的估计

考虑单子样場合。设总体F是未知的一維分布，$\theta(F)$是F的中位数。对于子样$X=(X_1,\cdots,X_n)$，子样中位数为$t(X)=X_{(m)}$。这里为方便起见设$n=2m-1$为奇数，$X_{(1)}\leqslant\cdots\leqslant X_{(n)}$，为由子样$X$按大小排列的顺序统计量。在非参数場合通常用$t(X)$作为$\theta(F)$的估计。记

$$R(X,F)=t(X)-\theta(F)=X_{(m)}-\theta(F) \tag{3.1}$$

现要求对$R(X,F)$的分布特性进行统计推断。

仍记来自$\hat{F}_n$的Bootstrap子样$X^*=(X_1^*,\cdots,X_n^*)$，N_i^*为X_j^*等于$x_i(j=1,\cdots,n)$的数目，由前讨论知$N^*=(N_1^*,\cdots,N_n^*)$服从多项分布。又注意到对任意正整数l，$1\leqslant l\leqslant n$，事件$\{X_{(m)}^*>x_{(l)}\}$等于事件$\left\{\sum_{j=1}^{l}N_j^*\leqslant m-1\right\}$，而$N_{(l)}^*=\sum_{j=1}^{l}N_j^*$服从二项分布Binomial$\left(n,\frac{l}{n}\right)$．于是

$$P_*\{X_{(m)}^*>x_{(l)}\}=P_*\{N_{(l)}^*\leqslant m-1\}=P\left\{\text{Binomial}\left(n,\frac{l}{n}\right)\leqslant m-1\right\}$$

$$=\sum_{j=0}^{m-1}\binom{n}{j}\left(\frac{l}{n}\right)^j\left(\frac{n-l}{n}\right)^{n-j} \tag{3.2}$$

又因为

$$R^*=R(X^*,\hat{F}_n)=X^*_{(m)}-x_{(m)}$$

故有

$$\begin{aligned}P_*\{R^*=x_{(l)}-x_{(m)}\}&=P_*\{X^*_{(m)}=x_{(l)}\}\\&=P_*\{X^*_{(m)}>x_{(l-1)}\}-P_*\{X^*_{(m)}>x_{(l)}\}\\&=\sum_{j=0}^{m-1}\left[\binom{n}{j}\left(\frac{l-1}{n}\right)^j\left(\frac{n-l+1}{n}\right)^{n-j}-\binom{n}{j}\left(\frac{l}{n}\right)^j\left(\frac{n-l}{n}\right)^{n-j}\right]\end{aligned}\tag{3.3}$$

可见，只要给定子样容量n，即可得到$P_*\{R^*=x_{(l)}-x_{(m)}\}$的值，利用它可对$R^*$的Bootstrap分布进行估计。例如有

$$E_*[R^*]=\sum_{l=1}^{n}(x_{(l)}-x_{(m)})P_*\{R^*=x_{(l)}-x_{(m)}\}\tag{3.4}$$

$$E_*[(R^*)^2]=\sum_{l=1}^{n}(x_{(l)}-x_{(m)})^2P_*\{R^*=x_{(l)}-x_{(m)}\}\tag{3.5}$$

这可通过理论计算（途径1）得到。对于F具有连续有界的分布密度f时，可以证明估计是渐近一致的。

考虑对F的进一步假设，上述过程可作如下修正：

(1) 已知F为对称分布，可令分布$\hat{F}_{sym}$为

$$\hat{F}_{sym}:\ 在x_{(1)},\cdots,x_{(n)}和2x_{(m)}-x_{(1)},\cdots,2x_{(m)}-x_{(n)}处具有质量\frac{1}{2n-1}\tag{3.6}$$

则Bootstrap子样改由$\hat{F}_{sym}$中抽取。若记$Z_{(1)}\leqslant Z_{(2)}\leqslant\cdots\leqslant Z_{(2n-1)}$为来自$\hat{F}_{sym}$的顺序子样，则与(3.3)类似，对任意正整数$l$，$1\leqslant l\leqslant 2n-1$

$$\begin{aligned}P_*\{R^*=Z_{(l)}-x_{(m)}\}=&P\left\{\text{Binomial}\left(n,\frac{l-1}{2n-1}\right)\leqslant m-1\right\}\\&-P\left\{\text{Binomial}\left(n,\frac{l}{2n-1}\right)\leqslant m-1\right\}\end{aligned}\tag{3.7}$$

(2) 若已知F为连续分布，可考虑平滑Bootstrap．例如，可把Bootstrap子样改为

$$X^*_i=\bar{x}+c(x_{I_i}-\bar{x}+sZ_i),\ i=1,\cdots,n\tag{3.8}$$

其中$\bar{x}$，s^2如前定义；I_i为从整数集$\{1,\cdots,n\}$中独立随机抽取的整数；Z_i服从某已知的零均值、方差为σ_Z^2的连续分布（如$N(0,\sigma_Z^2)$）；而

$$c=(1+\sigma_Z^2)^{-\frac{1}{2}}\tag{3.9}$$

因$E[x_{I_i}]=\bar{x}$，$Var[x_{I_i}]=s^2$，易知

$$E_*[X^*_i]=\bar{x}+c(E[x_{I_i}]-\bar{x}+sE[Z_i])=\bar{x}$$

$$Var[X^*_i]=c^2[Var(x_{I_i})+s^2Var(Z_i)]=s^2$$

即X^*_i与来自$\hat{F}_n$的Bootstrap子样具有相同的一、二阶矩。但此时直接的理论计算很困难，需要通过途径2．即由(3.8)重复抽样$x^*(i)$，$(i=1,\cdots,N)$，其中$x^*(i)=(x^*_1(i),$

$\cdots$, $x_n^*(i)$)，计算得相应的$R^*(i)$，$(i=1,\cdots,n)$，则有

$$E_*[R^*]\cong\frac{1}{N}\sum_{i=1}^{N}R^*(i) \tag{3.10}$$

$$Var_*[R^*]\cong\frac{1}{N-1}\sum_{i=1}^{n}(R^*(i)-E_*[R^*])^2 \tag{3.11}$$

算例[1] 定义

$$R(X,F)=\frac{|t(X)-\theta(F)|}{\sigma(F)}$$

其中 $\sigma(F)$为总体F的方差。对于F为$N(0,1)$，$n=13$，此时$E_F[R]=0.95$. 若设F是未知的，由

$$R^*=R(X^*,\hat{F}_n)=\frac{|t(X^*)-\theta(\hat{F}_n)|}{s}$$

对于$N=100$，重复试验200次的平均结果见附表A. 可见，虽$n=13$不大，结果却比较好。

2. 回归问题

考虑一般的回归模型

$$X_i=g_i(\beta)+\varepsilon_i,\ i=1,\cdots,n \tag{3.12}$$

其中 β为$p\times1$未知参数矢量，$\varepsilon_i(i=1,\cdots,n)$来自未知分布$F$，且相互独立。设$F$为零均值，方差$\sigma_F^2$为未知。所谓回归问题，即由观测$x=(x_1,\cdots,x_n)$，求$\beta$的估计$\hat{\beta}$，使$\sum_{i=1}^{n}(x_i-g_i(\hat{\beta}))^2=\min_{\beta}\sum_{i=1}^{n}(x_i-g_i(\beta))^2$. 同样，希望得到有关$\hat{\beta}$抽样分布的一些性质。

运用Bootstrap方法解决此问题的步骤如下：

(1) 构造残差$\hat{\varepsilon_i}=x_i-g_i(\hat{\beta})$的经验分布函数

$$\hat{F}_n:\text{在每点}\hat{\varepsilon_i}-\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon_i}\text{处具有质量}\frac{1}{n},\ (i=1,\cdots,n) \tag{3.13}$$

(2) 在$\hat{F}_n$中抽取Bootstrap子样$\varepsilon^*=(\varepsilon_1^*,\cdots,\varepsilon_n^*)$，作

$$x_i^*=g_i(\hat{\beta})+\varepsilon_i^*,\ i=1,\cdots,n \tag{3.14}$$

并以此作为回归模型得到$\hat{\beta}^*$，它使

$$\sum_{i=1}^{n}(x_i^*-g_i(\hat{\beta}^*))^2=\min_{\hat{\beta}}\sum_{i=1}^{n}(x_i^*-g_i(\hat{\beta}))^2$$

(3) 将上述(2)重复独立进行N次，相应得到$\hat{\beta}^*(i)$，$(i=1,\cdots,n)$。用它们估计$\hat{\beta}^*$的Bootstrap分布，并以此作为$\hat{\beta}$分布的近似。

考虑线性回归这一特殊情况，此时有

$$X_i=C_i\beta+\varepsilon_i,\ i=1,\cdots,n$$

写成向量形式有

$$X=C\beta+\varepsilon$$

若记$C^TC=G$，显然，$\hat{\beta}=G^{-1}C^Tx$，且$E[\hat{\beta}]=\beta$，$Var[\hat{\beta}]=\sigma_F^2G^{-1}$. 由于$\sigma_F^2$是未知

的，其无偏估计为

$$\hat{\sigma}_F^2=\frac{1}{n-p}\sum_{i=1}^{n}(x_i-C_i\hat{\beta})^2$$

这些都是众所周知的。

另外同样有

$$E_*[\varepsilon_i^*]=0,\ Var[\varepsilon_i^*]=\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\varepsilon}_i-\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i\right)^2=\hat{\sigma}^2(\varepsilon_i^*)$$

则

$$\begin{cases}E_*[\hat{\beta}^*]=\hat{\beta}\\ Var_*[\hat{\beta}^*]=\hat{\sigma}^2(\varepsilon_i^*)G^{-1}\end{cases}\tag{3.15}$$

这里需要指出的是：在Efron原文[1]构造$\hat{F}_n$，$\hat{F}_n$：在每点$\hat{\varepsilon}_i$处具有质量$\frac{1}{n}$，$(i=1,\cdots,n)$，时，认为$\hat{F}_n$为零均值。事实上，按此定义，$E_*[\varepsilon_i^*]=\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i$，而不是零。这一点应该纠正。

3. 偏倚(Bias)的估计

先引进泛函统计量(Functional Statistic)的概念。设$\theta(F)$是与总体F有关的某参数，它是待估的。称

$$\hat{\theta}=\theta(\hat{F}_n)\tag{3.16}$$

为泛函统计量。它是$\theta(F)$的一个估计。由定义，$\theta(\hat{F}_n)$满足对称性条件。常见的泛函统计量有

(1) 对$\theta(F)=E_F[X]$，$\theta(\hat{F}_n)=\bar{x}$；

(2) 对$\theta(F)=Var_F[X]$，$\theta(\hat{F}_n)=s^2$；

(3) 对$\theta(F)=E_F[Z]/E_F[Y]$，其中F为二维分布，$X=(Y,Z)^T$服从F，$\theta(\hat{F}_n)=\bar{z}/\bar{y}$。

若设

$$R(X,F)=\theta(\hat{F}_n)-\theta(F)\tag{3.17}$$

则估计的偏倚为

$$\text{Bias}=E_F[R(X,F)]=E_F[\theta(\hat{F}_n)]-\theta(F)\tag{3.18}$$

下面考虑用Bootstrap方法估计偏倚。因为

$$R^*=R(X^*,\hat{F}_n)=\theta(\hat{F}_n^*)-\theta(\hat{F}_n)\tag{3.19}$$

于是，作Bootstrap子样X^*的经验分布函数

$$\hat{F}_n^*:\text{在每点}x_i\text{处具有质量}P_i^*,\ i=1,\cdots,n\tag{3.20}$$

从$\hat{F}_n^*$中抽取独立子样得到$\hat{\theta}^*=\theta(\hat{F}_n^*)$，重复进行$N$次相应有$\hat{\theta}^*(1),\cdots,\hat{\theta}^*(N)$.于是，偏倚的估计为

$$\hat{\text{Bias}}=\frac{1}{N}\sum_{j=1}^{N}\hat{\theta}^*(j)-\hat{\theta}\cong E_*[R^*]\tag{3.21}$$

在例3中曾经简单讨论过此问题。在那里由途径3得到了$E_*[R^*]$的近似公式，忽略二阶以上项后$E_*[R^*]=0$. 这说明$\hat{}$Bias是二阶以上量。

4. 标准偏差(Standard Deviation)估计

所谓标准偏差即误差方差的平方根。该问题已在例2中做了一般性的讨论。这里研究一个具体例子。

设总体F为二維分布，它是未知的。现随机子样$X=(X_1,\cdots,X_n)$来自F，$X_i=(U_i,V_i)^T$，U_i和V_i是相关的。要求根据观测$X=x$对U_i和V_i之间的相关系数ρ进行估计。在非参数情况下一般采用Pearrson相关系数

$$\hat{\rho}(x_1,\cdots,x_n)=\frac{\sum_{i=1}^{n}u_iv_i-\sum_{i=1}^{n}u_i\sum_{j=1}^{n}v_j/n}{\left\{\sum_{i=1}^{n}u_i^2-\left(\sum_{i=1}^{n}u_i\right)^2/n\right\}^{1/2}\left\{\sum_{i=1}^{n}v_i^2-\left(\sum_{i=1}^{n}v_i\right)^2/n\right\}^{1/2}} \tag{3.22}$$

这里我们感兴趣的是估计

$$\sigma(\hat{\rho})=\{E_F[(\rho-\hat{\rho})^2]\}^{\frac{1}{2}} \tag{3.23}$$

对此可运用二維分布时的Bootstrap方法：

(1) 根据观测$X=x$构造经验分布函数：

$$\hat{F}_n:\ \text{在每点}x_i=(u_i,v_i)^T\ \text{处具有质量}\frac{1}{n},\ i=1,\cdots,n \tag{3.24}$$

(2) 从$\hat{F}_n$中抽取Bootstrap子样$x^*=(x_1^*,\cdots,x_n^*)$，其中$x_i^*=(u_i^*,v_i^*)^T$，并计算$\hat{\rho}^*=\hat{\rho}(x_1^*,\cdots,x_n^*)$;

(3) 重复(2)N次得到$\hat{\rho}^*(i)$，$i=1,\cdots,N$，以此构造$\sigma(\hat{\rho})$的估计

$$\hat{\sigma}(\hat{\rho})=\left\{\frac{1}{N-1}\sum_{j=1}^{N}(\hat{\rho}^*(j)-\hat{\rho}^*(\cdot))^2\right\}^{\frac{1}{2}} \tag{3.25}$$

其中$\hat{\rho}^*(\cdot)=\frac{1}{N}\sum_{j=1}^{N}\hat{\rho}^*(j)$。当然也可以采用平滑Bootstrap方法，这里不再详述。

算例[2] 总体F为二維正态分布，$X_i=(U_i,V_i)^T$，$E[U_i]=E[V_i]=0$，$Var[U_i]=Var[V_i]=1$，$\rho(U_i,V_i)=\frac{1}{2}$。取$n=14$，则(3.23)真值$\sigma(\hat{\rho})=0.218$.下面对三种形式的Bootstrap过程进行模拟计算，N为128，重复试验200次取平均，结果见附表(B).

(1) F未知，由$\hat{F}_n$中抽取Bootstrap子样;

(2) 设F为連续型分布，考虑均匀平滑Bootstrap，即从$\hat{F}_1^{0.5}=\hat{F}_n*(0.5\hat{F}_U)$中抽取Bootstrap子样。其中$\hat{F}_U$为一均匀分布，其协方差阵等于

$$\hat{\Sigma}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^T$$

(3) 同(2)考虑正态平滑Bootstrap，即从$\hat{F}_2^{0.5}=\hat{F}_n*(0.5\hat{F}_N)$抽取Bootstrap子样，其中$\hat{F}_N=N_2(\bar{x},\hat{\Sigma})$。此时，因$F$的真实分布为正态分布，故采用正态平滑有一定的"利己(self—serving)"性。

由定义，$\hat{F}_1^{0.5}$、$\hat{F}_2^{0.5}$和$\hat{F}_n$具有相同相关系数。

5. 判别分析的错分率(Error Rate)估计

设两个独立的随机子样$X=(X_1,\cdots,X_m)$和$Y=(Y_1,\cdots,Y_n)$分别来自未知的K维连续分布F和G. 基于观测x和y，用线性判别分类的方法可将空间R^K分为两个互补的区域A和B，即

$$\begin{cases} B=\left\{z:\ (\bar{y}-\bar{x})^T s^{-1}\left(z-\dfrac{\bar{x}+\bar{y}}{2}\right)>\log\dfrac{m}{n}\right\} \\ A=R^K-B \end{cases} \tag{3.26}$$

对于新观测z，若$z\in B$，则认为z来自G；否则认为z服从F. 式中$\bar{x}=\dfrac{1}{m}\sum_{i=1}^{m}x_i$，

$\bar{y}=\dfrac{1}{n}\sum_{i=1}^{n}y_i$，$s=\dfrac{1}{m+n}\left[\sum_{i=1}^{m}(x_i-\bar{x})(x_i-\bar{x})^T+\sum_{i=1}^{n}(y_i-\bar{y})(y_i-\bar{y})^T\right]$。

显然，由于子样的随机性，上述分类会产生错误。对于分布F，定义错分率

$$\mathrm{error}_F=P_F\{X\in B\} \tag{3.27}$$

其直观的估计是

$$\widehat{\mathrm{error}}_F=\frac{\#\{x_i\in B\}}{m} \tag{3.28}$$

同样地，可定义error_G和$\widehat{\mathrm{error}}_G$. 由问题的对称性，只需考虑上两式即可。这里我们感兴趣的是

$$R((X,Y),(F,G))=\mathrm{error}_F-\widehat{\mathrm{error}}_F \tag{3.29}$$

的分布特性。为此运用两子样场合的Bootstrap方法：

(1) 在给定观测x和y条件下，构造子样经验分布函数：

$\hat{F}_m$：在每点x_i处具有质量$\dfrac{1}{m}$，$i=1,\cdots,m$ (3.30)

$\hat{G}_n$：在每点y_j处具有质量$\dfrac{1}{n}$，$j=1,\cdots,n$ (3.31)

(2) 从$\hat{F}_m$和$\hat{G}_n$中分别抽取独立的Bootstrap子样x_i^*和y_j^*，$i=1,\cdots,m$，$j=1,\cdots,n$；

(3) 用$\bar{x}^*$、$\bar{y}^*$和s^*代替$\bar{x}$、$\bar{y}$和s，由(3.26)计算B^*，同时计算

$$R^*=R((X^*,Y^*),(\hat{F}_m,\hat{G}_n))=\mathrm{error}_{\hat{F}_n}-\widehat{\mathrm{error}}_{\hat{F}_n}$$

$$=\frac{\#\{x_i\in B^*\}}{m}-\frac{\#\{x_i^*\in B^*\}}{m} \tag{3.32}$$

(4) 重复(2)、(3)N次获得$R^*(1)$，…，$R^*(N)$。用它们可做R^*的Bootstrap分布的迫近，并以此作为R真实分布的迫近。例如

$$E_*[R^*]\cong\frac{1}{N}\sum_{j=1}^{N}R^*(j) \tag{3.33}$$

$$V_{ar_*}[R^*]\cong\frac{1}{N-1}\sum_{j=1}^{N}(R^*(j)-E_*[R^*])^2 \tag{3.34}$$

算例[1] 设F：$X\sim N_2\left(\begin{pmatrix}-\frac{1}{2}\\0\end{pmatrix},\ I\right)$，$G$：$Y\sim N_2\left(\begin{pmatrix}\frac{1}{2}\\0\end{pmatrix},\ I\right)$分别$m=n=10$和$m=n=20$两种情况，而$N=100$．(3.29)式的真实均值和标准偏差由Monte—Carlo方法算出。重复试验100次的平均结果如附表(C)所示。

6. 比估计(Ratio Estimation)问题

设F为二维分布，即有$X_i=(Y_i,Z_i)^T\sim F$，$i=1,\cdots,n$，且X_i相互独立。比估计即为

$$\theta(F)=\frac{E_F[Y]}{E_F[Z]} \tag{3.35}$$

（设Y，Z均大于零）。通常用的统计量为

$$t(X)=\hat{\theta}(F)=\frac{\bar{Y}}{\bar{Z}}=\sum_{i=1}^{n}Y_i\Big/\sum_{i=1}^{n}Z_i \tag{3.36}$$

记

$$R(X,F)=\frac{\hat{\theta}(F)}{\theta(F)} \tag{3.37}$$

我们要估计$R(X,F)$的分布特性。

由前面定义可知$\hat{\theta}(F)=\theta(\hat{F}_n)$为泛函统计量。对此问题的Bootstrap解可采用途径2，具体作法无特别之处，故不再详述。由于$R(X,F)$关于X对称，下面试用途径3求出R^*的近似一、二阶矩。注意到

$$R^*=R(X^*,\hat{F}_n)=\frac{t(X^*)}{\theta(\hat{F}_n)}=\frac{\bar{Y}^*}{\bar{Z}^*}\ \frac{\bar{z}}{\bar{y}}\triangleq R(P^*) \tag{3.38}$$

$$\begin{cases}\bar{Y}^*=\sum_{i=1}^{n}P_i^*y_i,\ \bar{Y}^*\big|_{p^*=\frac{e}{n}}=\bar{y}\\ \bar{Z}^*=\sum_{i=1}^{n}P_i^*z_i,\ \bar{Z}^*\big|_{p^*=\frac{e}{n}}=\bar{z}\end{cases}$$

$$R\left(\frac{e}{n}\right)=\frac{\bar{y}}{\bar{z}}\cdot\frac{\bar{z}}{\bar{y}}=1$$

故 $$\frac{\partial\bar{Y}^*}{\partial P_i^*}=y_i,\ \frac{\partial\bar{Z}^*}{\partial P_i^*}=z_i$$

$$U_i=\frac{\partial R(P^*)}{\partial P_i^*}\Big|_{p^*=\frac{e}{n}}=\frac{y_i}{\bar{y}}-\frac{z_i}{\bar{z}}$$

$$V_{ij}=\frac{\partial^2R(P^*)}{\partial P_i^*\partial P_j^*}\Big|_{p^*=\frac{e}{n}}=2\frac{z_iz_j}{\bar{z}\bar{z}}-\frac{y_iz_j+y_jz_i}{\bar{y}\bar{z}}$$

$i,\ j=1,\cdots,n$．于是

$$\sum_{i=1}^{n}\frac{z_i}{\bar{z}}=\sum_{i=1}^{n}\frac{y_i}{\bar{y}}=n$$

$$E_*[R^*]=1+\frac{1}{2n}\sum_{i=1}^{n}V_{ii}=1-\frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{y_iz_i}{\bar{y}\bar{z}}-\frac{z_i^2}{\bar{z}\bar{z}}\right)$$

$$=1-\frac{1}{n^2}\left\{\sum_{i=1}^{n}\left(\frac{y_i}{\bar{y}}-1\right)\left(\frac{z_i}{\bar{z}}-1\right)-\sum_{i=1}^{n}\left(\frac{z_i}{\bar{z}}-1\right)^2\right\} \tag{3.39}$$

$$V_{ar_*}[R^*]=\sum_{i=1}^{n}\frac{U_i^2}{n^2}=\frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{y_i}{\bar{y}}-\frac{z_i}{\bar{z}}\right)^2 \tag{3.40}$$

根据以上结果还可以对$\hat{\theta}(F)$作适当的纠偏修正，得到$\theta(F)$的无偏估计

$$t'(X)=\frac{t(X)}{E_*[R^*]} \tag{3.41}$$

7. 非参数置信区间估计

问题是这样提出的：设F是未知的总体分布，$\theta=\theta(F)$为待估的未知分布参数。现以$\hat{\theta}=\theta(\hat{F}_n)$的Bootstrap分布为基础，给出$\theta$的置信区间估计的迫近。

下面所用的方法称作“百分值(percentile)”法。记$\hat{\theta}^*=\theta(\hat{F}_n^*)$，其Bootstrap分布的累积(cumulative)分布函数为

$$\widehat{\mathrm{CDF}}(t)=P_{rob_*}\{\hat{\theta}^*\leqslant t\} \tag{3.42}$$

对于Monte—Carlo模拟近似有

$$\widehat{\mathrm{CDF}}(t)\cong\frac{\#\{\hat{\theta}^*(i)\leqslant t\}}{N} \tag{3.43}$$

给定某一α，$0<\alpha<0.5$，定义

$$\hat{\theta}_{LOW}(\alpha)=\widehat{\mathrm{CDF}}^{-1}(\alpha),\quad \hat{\theta}_{UP}(\alpha)=\widehat{\mathrm{CDF}}^{-1}(1-\alpha) \tag{3.44}$$

简记$\hat{\theta}_{LOW}(\alpha)=\hat{\theta}_{LOW}$，$\hat{\theta}_{UP}(\alpha)=\hat{\theta}_{UP}$，有

$$\alpha=\widehat{\mathrm{CDF}}(\hat{\theta}_{LOW}),\quad 1-\alpha=\widehat{\mathrm{CDF}}(\hat{\theta}_{UP}) \tag{3.44'}$$

则$[\hat{\theta}_{LOW},\hat{\theta}_{UP}]$即为在置信水平$1-2\alpha$之下$\theta$的置信区间估计的迫近。

若$P_{rob_*}\{\hat{\theta}^*\leqslant\hat{\theta}\}=0.5$，即Bootstrap分布关于$\hat{\theta}$是对称的，则$\theta$的真值近似落在置信区间$[\hat{\theta}_{LOW},\hat{\theta}_{UP}]$的中间。若对称条件不满足，则可考虑偏倚校正(bias—corrected)百分值法，具体作法是：定义

$$z_0=\Phi^{-1}(\widehat{\mathrm{CDF}}(\hat{\theta})) \tag{3.45}$$

其中$\Phi(x)$为标准正态变量的累积分布函数

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$$

即有

$$\widehat{\mathrm{CDF}}(\hat{\theta})=\Phi(z_0)$$

则偏倚校正百分值法取z_α，满足$\Phi(z_\alpha)=1-\alpha$.

$$\hat{\theta}_{LOW}=\widehat{\mathrm{CDF}}^{-1}(\Phi(2z_0-z_\alpha)),\hat{\theta}_{UP}=\widehat{\mathrm{CDF}}^{-1}(\Phi(2z_0+z_\alpha)) \tag{3.46}$$

即

$$\widehat{\mathrm{CDF}}(\hat{\theta}_{LOW})=\Phi(2z_0-z_\alpha),\ \widehat{\mathrm{CDF}}(\hat{\theta}_{UP})=\Phi(2z_0+z_\alpha) \tag{3.46}'$$

则$[\hat{\theta}_{LOW},\hat{\theta}_{UP}]$为$\theta$在置信水平$1-2\alpha$之下置信区间的近似表示。特殊地，当对称条件成立，显然$z_0=0$则(3.44)即为(3.46)。

算例[6] 总体F为二维正态分布，即$X_i=(U_i,V_i)^T\sim F$，$i=1,\cdots,n$，$\rho(U_i,V_i)=0.5$，$n=15$. 若设F未在68%的置信水平下，$N=1000$，重复试验10次平均，关于P的置信区间非参数估计的结果见附表(D)。

四、结 论 与 展 望

与经典的参数统计理论相比，Bootstrap方法有两个鲜明的特点：

1. 对于解决某一具体问题，Bootstrap方法完全从问题的提法本身出发，仅依赖于给定的观测数据，而不附加任何其它人为的假设。与此相反，经典的参数方法则过分依赖于使问题能得到理论解的标准分布模型的小集合及现成的数理统计图表，这在许多場合难免有些牵强附会。因此Bootstrap方法具有更强的现实性。

2. Bootstrap方法借助于计算机这一有力工具，使其具有较高的灵活性和实用性。我们已经看到，无论$R(X,F)$的形式如何复杂，相应的Bootstrap算法都是比较简单的，也易于在计算机上实现。当然，从传统的观点看，Bootstrap方法在计算上存在着巨大的浪费，它需要将通常的统计计算重复上千次甚至更多。这在计算机沒有高度发展的时代是不可想象的，但在当今做到这点毫不困难。

作为一种新的统计方法，Bootstrap方法尚在发展中，许多问题有待于研究解决。这里不妨列举几条：

1. 算法的收敛性问题：Efron对于一般情况考虑了Bootstrap方法的漸近性，他指出[1]：当$R(X,F)$的分布弱收敛于某一极限时，$R(X^*,\hat{F}_n)$的Bootstrap分布以概率1弱收敛于同一极限。其它一些学者，如Singh[8]、Bickel和Freedman[9]、Freedman[10]、Beran[11]等人，都就具体应用问题讨论了Bootstrap方法的收敛性，得到了一些结果。但更一般的收敛性理论还有待于进一步的探讨。

2. 算法的稳健性问题：Beran[11]证明在一定条件下抽样分布的Bootstrap估计具有MinMax特性。另一方面由于Bootstrap方法是完全建立在观测子样$X=x=(x_1,\cdots,x_n)$的基础之上的。观测数据质量的好坏对Bootstrap精度的影响如何？当子样容量n小时（即小子样場合），Bootstrap方法的结果又怎样？这些问题均需要加以回答。

3. 算法的计算量问题：对于具体问题，要达到一定的精度要求，这需要多少计算量？特别在进行统计模拟计算时，N应如何选择？这个问题也值得研究。

其它还有一些问题，如Bootstrap方法与参数统计方法的联系等等。

目前，在我们的工作中，参数统计方法仍占着主导地位，人们习惯于“正态假设”等标准参数模型。Bootstrap方法的出现为解决实际问题开辟了一条新的途径。我们介绍

Bootstrap方法的目的，一是了解在应用统计学领域里的最新发展动态，更重要的是要把这种新方法运用飞行器试验学的具体问题中去，以期提高统计估值的精度和可信度。

最后需要指出的是，Bootstrap方法并不是在一切场合下都适用的。Beran[11]曾经具体给出了一个反例；Bickel和Freemdan亦讨论了Bootstrap估计失败的问题。这一问题应引起足够的重视。

附 表

(A) 总体中位数的估计

$E_*[R^*]$	无平滑Bootstrap		平滑 Bootstrap，z_i服从$[-\frac{d}{2},\frac{d}{2}]$上均匀分布				
	$\hat{F}_n$	$\hat{F}_{sym}$	$d=0$	$d=0.25$	$d=0.5$	$d=1$	$d=2$
平均值	1.01	1.00	1.00	1.01	1.00	0.99	0.93
标准偏差	0.31	0.33	0.32	0.32	0.32	0.30	0.26

(B) 标准偏差的估计

类别 / 方法	$\sigma_B(\hat{\rho})$	
	平均	标准偏差
无平滑 Bootstrap	0.206	0.066
正态平滑 Bootstrap	0.200	0.060
均匀平滑 Bootstrap	0.205	0.061

(C) 错分率的估计

分类 / 变量	$m=n=10$		$m=n=20$	
	均值	标准偏差	均值	标准偏差
R	0.062	0.143	0.028	0.103
$E_*[R^*]$	0.057	0.026	0.029	0.015
$\sqrt{V_{ar_*}[R^*]}$	0.131	0.016	0.097	0.010

(D) 非参数置信区间估计

方法	正态理论（真值）	百分值法	偏倚校正百分值法
置信区间	]0.23, 0.33[	[0.27, 0.67]	[0.22, 0.66]

参 考 文 献

[1] B.Efron, Bootstrap Methods: Another Look at the Jackknife, Ann. Statist. 7, 1—36, 1979.

[2] B.Efron Computers and the Theory of Statistics: Thinking the Unthinkble, SIAM Review 21, 460—480, 1979.

[3] B.Efron, Nonparametric Estimates of Standard Error: The Jackknife, the Bootstrap and Other Methods, Biometrika 68—3, 569—599, 1981.

[4] B.Efron, Censered Data and the Bootstrap, J.American Stasistical Association 76, 312—319, 1981.

[5] B.Efron, Computer Intensive Methods in Statistics, Some Recent Advances in Statistics, Academic Press, 1982.

[6] B.Efron, The Jackknife, the Bootstrap and Other Resampling Plans, CBMS—NSF, SIAM, Philadelphia, 1982.

[7] B.Efron, A Leisurely Look at the Bootstrap, the Jackknife and the Cross Validation, Amer. Statist. 37, 36—48, 1983.

[8] K.Sigh, On the Asymptotic Theory for the Bootstrap, Ann. Statist. 9, 1187—1195, 1981.

[9] P.J.Bickel & D.A.Freedman, Some Asymptotic Accuracy for the Bootstrap, Ann. Statist. 9, 1196—1217, 1981.

[10] D.A.Freedman, Bootstrapping Regression Models, Ann. Statist. 9,1218—1228,1981

[11] R.Beran, Estimated Sampling Distributions: the Bootstrap and Competitors,Ann. Statist. 10, 212—225, 1982.

[12] R.Beran, Jackknife Approximations to Bootstrap Estimates, Ann. Statist. 12, 101—118, 1984.

Bootstrap Method and Its Applications

Mao Ning

Abstract

A new statistical inference method, called Bootstrap method, was advanced by Prof. Efron in 1977. The method applies to a variety of estimation problems, so importance has been attached to by more and more statisticians. In the paper the basic idea of the method and its application examples are introduced. And the existing problems and the orienlation of the research are discussed as well.