予 測 の 統 計 的 方 法

1. はじめに

予測の目的に統計的方法が役に立つかと問われれば，統計学者としての私の立場からは，確かに役立つと答えたいところである．また後に述べるように，統計的解析の目的は，たとえ表面的には予測を目的としていない場合でも，実は何らかの意味で予測をねらったものであることが多い．しかしながら，それが直接に予測に役立つかといわれれば，率直にいって，統計的方法が具体的な予測の問題に直ちに有効である場合は多くないであろうといわざるを得ない．

というのは，統計的方法というものは，かなり概念を広く解釈するとしても，要するに数字データの形式的処理法であると見なさざるを得ない．しかしながら，ORの場に現われるような具体的な予測ないし予知の問題においては，形式的な数学的処理に適さないような，理論的，経験的知識にもとづく判断，あるいは数字に表わし難い質的情報の利用が，重要な意味をもってくるからである．この「特集」で，他の方々によって扱われるような，技術，経済，天気，災害等の予測ないし予報の問題の中で，おそらく経済予測の場合を除けば，統計的方法が用いられる場面はあまり多くないであろう．またマクロ・エコノメトリック・モデルによる計測と予測というような複雑精緻な方法が用いられる経済予測の場合でも，実はそれは多くの前提条件，あるいは将来の世界的国内的政治社会状況についての一連の仮定の上に立っているのであって，純粋に将来の経済の状態を「当てる」という意味での予測の観点からは，むしろその前提条件の変化のほうが重要な意味をもつことが多い．モデルを用いる計測の目的は，単純な予測よりも，いくつかの政策目標ととり得べき政策の間の整合性をチェックしたり，ある種の状況の変化が経済にどのような影響を及ぼすであろうかを調べたりすることにあるというべきである．

しかし，もちろん私は統計的方法が予測の目的に役立たないというつもりではない．統計的方法は予測のための判断の根拠，あるいは資料を提供するという点では非常に有効である．私があらかじめ断っておきたいと思うのは，何らかの形式的な統計的方法によって予測の最終的な結論を出すことが可能な場合はあまりないし，またそのことを目標として主観的判断の要素までをふくめた「予測の形式的一般理論」などを作ろうとしても，得るところは少ないであろうということなのである．

2. ベイズ予測

以上のことを前提にした上で，予測と関連した統計的方法について述べよう．ここでその内容は2つに分けられる．1つは直接に予測を目ざした方法であり，第2は直接には予測を目的としないけれども，実は予測ということを念頭においた方法である．

まず前者から述べよう．それは形式的にはつぎのような形で定義できる．いまデータ $X_1, \cdots, X_n$ にもとづいて，ある量 Y を予測したいものとしよう．ここでデータから Y が正確に予測できる

場合，たとえば天文観測の結果にもとづいて日蝕の時・所を計算するような場合は，統計的な問題は生じないから，ここでは除くことにしよう．すなわち $X_1,\cdots,X_n$，Y の関係には何らかの不確実性がともなうとする．このことを $X_1,\cdots,X_n$，Y がある同時確率分布に従うという形で表わすことにする．ここでさらに2つの場合に分けられる．

1つはその同時分布が完全にわかっている場合，すなわちたとえば同時確率密度関数が，

$$f(x_1,\cdots,x_n,\ y)$$

という形で与えられている場合である．

このときには，データ $X_1=x_1,\cdots,X_n=x_n$ が観測されたとすれば，このときの Y の条件付確率密度関数が，ベイズの定理により，

$$f(y|x_1,\cdots,x_n)=\frac{f(x_1,\cdots,x_n,\ y)}{\int f(x_1,\cdots,x_n,\ y)dy}$$

で与えられるから，Y についての判断はすべてこの条件付分布を用いて行なうことができる．このような場合は，せまい意味の統計的予測とは区別して確率予測とよぶことにしよう．

本来の統計的予測の問題は，$X_1,\cdots,X_n,Y$ の同時分布が未知母数 θ をふくむ場合である(θ はベクトルであってもよい)．このとき同時確率密度関数を，

$$f(x_1,\cdots,x_n,y;\theta)$$

と表わそう．

ここでさらに2つの場合がある．1つは θ について何らかの確率分布，いわゆる事前分布が仮定される場合であり，もう1つはそれを仮定しない場合である．いわゆるベイジアンの立場に立つ人はつねに事前分布を用いるべきことを主張しているので，前者をベイジアン予測問題とよぼう．

ベイジアン予測問題においては，θ の事前分布の密度関数を $\pi(\theta)$ とすると，$X_1,\cdots,X_n,Y$ の同時密度関数は，

$$\tilde{f}(x_1,\cdots,x_n,y)=\int f(x_1,\cdots,x_n,y;\theta)\pi(\theta)d\theta$$

と表わされるから，問題は確率予測の場合に帰着する．そうして Y の条件付分布すなわち事後分布は，

$$\tilde{f}(y|x_1,\cdots,x_n)=\frac{\int f(x_1,\cdots,x_n,y;\theta)\pi(\theta)d\theta}{\iint f(x_1,\cdots,x_n,y;\theta)\pi(\theta)d\theta dy}$$

で与えられる．この場合，問題は確率分布の計算だけである．

例題　$X_1,\cdots,X_n,Y$ について，つぎのような構造を仮定する．

$$X_i=\alpha+\beta z_i+u_i \qquad (i=1,2,\cdots,n)$$
$$Y=\alpha+\beta z_0+u_0$$

ここで z は X および Y に影響する説明変数で，$z_1,\cdots,z_n$ はデータにおけるその値，z_0 は Y について想定される値である．さらに $u_1,\cdots,u_n,u_0$ は互いに独立に平均0，分散 σ^2 の正規分布に従うと仮定する．

そうすると X_i，Y の同時密度は，

$$f(x_1,\cdots,x_n,y;\alpha,\beta,\sigma)$$
$$=\frac{1}{(\sqrt{2\pi}\sigma)^{n+1}}\exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\alpha-\beta z_i)^2-\frac{1}{2\sigma^2}(y-\alpha-\beta z_0)^2\right\}$$

となる．そこで，α,β および σ について，特定の範囲に入るという事前の情報がないとすると，事前密度は，

$$\pi(\alpha,\beta,\sigma)=1/\sigma$$

となるものと見なしてもよい．(このような π は α，β,σ の全範囲で積分すると無限大になってしまうから確率分布にはならない．しかしベイジアンのやり方ではこういう事前密度をも許すのがふつうである．)

このとき，

$$\iiint f(x_1,\cdots,x_n,y;\alpha,\beta,\sigma)\,\pi(\alpha,\beta,\sigma)d\alpha d\beta d\sigma$$
$$\propto\iiint\frac{1}{\sigma^{n+2}}\exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\alpha-\beta z_i)^2-\frac{1}{2\sigma^2}(y-\alpha-\beta z_0)^2\right\}d\alpha d\beta d\sigma$$

となるが，この積分を計算するために，ここで，

$$\hat{\beta}=\sum(z_i-\bar{z})(x_i-\bar{x})/\sum(z_i-\bar{z})^2$$
$$\hat{\alpha}=\bar{x}-\hat{\beta}\bar{z}$$

$$\hat{\sigma}^2=\sum(x_i-\hat{\alpha}-\hat{\beta}z_i)^2/(n-2)$$

とおけば，若干の計算の後，

$$\iiint f(x_1,\cdots,x_n,y;\alpha,\beta,\sigma)\frac{1}{\sigma}d\alpha d\beta d\sigma$$

$$\propto\int\frac{1}{\sigma^n}\exp\Bigg[-\frac{1}{2\sigma^2}(n-2)\hat{\sigma}^2 -\frac{1}{2\sigma^2\left\{1+\frac{1}{n}+\frac{(z_0-\bar{z})^2}{\sum(z_i-\bar{z})^2}\right\}}(y-\hat{\alpha}-\hat{\beta}z_0)^2\Bigg]d\sigma$$

$$\propto\left[1+\frac{(y-\hat{\alpha}-\hat{\beta}z_0)^2}{\hat{\sigma}^2(n-2)\left\{1+\frac{1}{n}+\frac{(z_0-\bar{z})^2}{\sum(z_i-\bar{z})^2}\right\}}\right]^{-\frac{n-1}{2}}$$

となる．すなわち $x_1,\cdots,x_n$ いいかえれば $\hat{\alpha},\hat{\beta},\hat{\sigma}$ が与えられたとき，条件付で，

$$\frac{Y-\hat{\alpha}-\hat{\beta}z_0}{\sqrt{1+\frac{1}{n}+\frac{(z_0-\bar{z})^2}{\sum(z_i-\bar{z})^2}}\,\hat{\sigma}}$$

が自由度 $n-2$ の t 分布に従うことがわかる．

ベイジアンの予測理論は，最近一部の統計学者達によって熱心に推進されているが，その基本的問題点は，事前分布の想定に関して，主観的ないし恣意的な要素が入るということである．ベイジアンの立場に立つ人々は主観確率論の観点から，むしろ主観性を積極的に肯定しているが，しかし科学的方法としての統計的予測という観点からはやはり主観的な要素を導入することには問題があるといわねばならない．

しかしながら，さらに立ち入っていえば，予測の問題においては，ベイジアン予測における主観的要素の影響は，母数推測の場合より小さいことに注意しなければならない．なぜならば $X_1,\cdots,X_n$ と Y とが独立な場合，Y の条件付分布は母数が与えられたときの Y の分布と，母数の事後分布とが合成されたものと考えることができるが，n がある程度大きければ，事後分布の変動は小さくなって，Y の条件付分布は大部分母数が与えられたときの分布で決定され，しかも n が大きければ母数の事後分布そのものが事前分布によって影響されることが少なくなるからである．

そうして $X_1,\cdots,X_n,Y$ の同時分布が複雑な構造をもっているとき，ベイジアン予測の方法はとにかく Y の条件付分布を計算する手続きを与えるから，それが有効に用いられる場合も少なくないと思われる．

私自身は主観確率論に立つベイジアンの立場はとらないが，予測の問題においてはベイジアンの方法を一概に否定することはできないと思っている．

3. 非ベイズ予測

非ベイジアンの予測理論においては，母数推測の場合に対応して，いくつかの予測形式が区別される．いろいろな予測形式とそれについての理論について，その概要だけを後の章でのべるが，くわしくは私の著書「統計的予測論」(培風館)を参照していただきたい．ここではその基本的な点だけをのべよう．

最も基本的な予測形式は点予測および区間予測である．点予測とは $X_1,\cdots,X_n$ から，Y の値についての1つの予測量

$$\hat{Y}=\hat{y}(X_1,\cdots,X_n)$$

を計算するものである．そうして，

$$E_\theta(Y-\hat{Y})=0 \qquad \forall\,\theta$$

のとき，$\hat{Y}$ は不偏予測量であるとよばれる．そうして不偏予測量の中で誤差の分散

$$V_\theta(Y-\hat{Y})=E_\theta(Y-\hat{Y})^2$$

を最小にするものを最小分散不偏予測量とよび，最も望ましい予測量であると考える．

ところでもし $X_1,\cdots,X_n$ と Y とが独立ならば，$\hat{Y}$ と Y も独立になるから，$E_\theta(Y)=\mu(\theta)$ と表わせば，後の章でのべるように Y の最小分散不偏予測量を求めることは $\mu(\theta)$ の最小分散不偏推定量を求めることと同等になる．したがって問題は不偏推定論に帰着する．

たとえば先の例題において $E(Y)=\alpha+\beta z_0$ であるから，線形モデルについての推定論でよく知られているように最小分散不偏予測量は，

$$\bar{Y}=\hat{\alpha}+\hat{\beta}z_0 \qquad (\hat{\alpha},\hat{\beta}\text{ は最小2乗推定量})$$

となる．

他方，区間予測とは $X_1 \cdots X_n$ から２つの量

$$\underline{Y}=\underline{y}(X_1,\cdots,X_n),\ \bar{Y}=\bar{y}(X_1,\cdots,X_n)$$

を計算して　$\underline{Y}<Y<\bar{Y}$ となるであろうという形で予測する方式である．そうして，

$$P_\theta(\underline{Y}<Y<\bar{Y})\geqq 1-\alpha \quad \forall\theta$$

となるとき $[\underline{Y},\bar{Y}]$ を信頼係数 $1-\alpha$ の予測区間という．ただし上記の確率は $\underline{Y}$, Y, $\bar{Y}$ の同時分布について計算されるものであるから，$\underline{Y}$, $\bar{Y}$ が与えられたときに Y が区間 $[\underline{Y},\bar{Y}]$ にふくまれる確率を表わすわけではない．

前記の例題において，

$$P_\theta\left\{\frac{|Y-\hat{\alpha}-\hat{\beta}z_0|}{\sqrt{1+\dfrac{1}{n}+\dfrac{(z_0-\bar{z})^2}{\sum(z_i-\bar{z})^2}}\ \hat{\sigma}}<t_\alpha\right\}=1-\alpha$$

ただし t_α は自由度 $n-2$ の t 分布の両側の点となるから，

$$\bar{Y}(\underline{Y})=\hat{\alpha}+\hat{\beta}z_0\underset{(-)}{+}t_\alpha\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(z_0-\bar{z})^2}{\sum(z_i-\bar{z})^2}}$$

とすれば，信頼係数 $1-\alpha$ の予測区間が得られる．

信頼係数一定の予測区間の中では，その平均長さ　$E_\theta(\bar{Y}-\underline{Y})$ が小さいものが望ましいと考えられる．上記の例題の t 区間は（ある種の正則条件を満たす）予測区間の中で最短であることが示される．

予測区間の中には分布形を仮定しないノンパラメトリック区間もあることに注意しておこう．すなわち $X_1,\cdots,X_n$, Y が互いに独立に同じ連続分布に従うとき，$X_1,\cdots,X_n$ を大きさの順に並べた順序統計量を　$X_{(1)}<X_{(2)}<\cdots\cdots<X_{(n)}$　とすると，

$$\Pr\{X_{(i)}<Y<x_{(i+1)}\}=1/(n+1)$$

となるから，

$$\Pr\{X_{(i)}<Y<X_{(n-j+1)}\}=1-(i+j)/(n+1)$$

となる．したがって $\alpha=(i+j)/(n+1)$ となるように i, j を選べば，$X_{(i)}<Y<X_{(n-j+1)}$ が信頼係数 $1-\alpha$ の予測区間となる．

このような方法は n が小さいときには，適当な α の値が得られるような i, j の組が存在しないという欠点があるが，n が大きいときには，最短予測区間とほとんど同じものが得られるという点で，正確な分布型を仮定した場合と比べて効率が落ちないことに注意しよう．

なぜならば，いまかりに Y の分布が，対称な一山型の密度関数をもつとすれば，その分布の片側 $\alpha/2$%点（両側 α%点）を $\theta+\xi_\alpha$（θ は分布の中央値）とすれば，分布がわかっているときの最短予測区間は，

$$\theta-\xi_\alpha<Y<\theta+\xi_\alpha$$

になる．その長さは $2\xi_\alpha$ となるから，どんな予測区間もその期待値がこれより短くなることはないのに対して，n が大きければ　$i=\alpha(n+1)/2$ とするとき，

$$E(X_{(n-i+1)}-X_{(i)})\doteqdot 2\xi_\alpha$$

となるので，予測区間 $X_{(i)}<Y<X_{(n-i+1)}$ はほぼ最短予測区間に一致することになるからである．

いま $X_1,\cdots,X_n$, Y が平均 μ, 分散 σ^2 の正規分布に従うとしよう．$\alpha=0.05$ とするとき，$n=39$ とすると，ちょうど，

$$X_{(1)}<Y<X_{(39)}$$

$$(\text{すなわち}\ \min X_i<Y<\max X_i)$$

が95%予測区間を与える．このときの区間の平均の長さは統計数値表を用いて計算すると，

$$E(X_{(39)}-X_{(1)})=2\times 2.15\sigma=4.30\sigma$$

となる．

他方，この場合の最短予測区間である t 区間の平均の長さは，

$$2\,t_\alpha\sqrt{1+\frac{1}{n}}E(S),\qquad S^2=\sum(X_i-\bar{X})^2/(n-1)$$

で与えられる．ここで $n=39$ を代入すると，再び数表から，

$$t_{.05}(38)=2.024,\qquad E(S)=0.9935\sigma$$

となることがわかるから，t 区間の平均長さは，

$$2\times 2.024\times\sqrt{40/39}\times 0.9935\sigma=4.07\sigma$$

となる．したがってノンパラメトリック区間は t 区間に対し，平均的に約1.06倍の長さになるにすぎないから，実用上は大きな差はないといっても

よいであろう．

予測区間に対して，片側予測限界も応用の場では重要な意味をもつことがある．すなわちそれは，

$$P_\theta\{Y>\bar{Y}\}\leqq\alpha \quad \text{あるいは} \quad P_\theta\{Y<\underline{Y}\}\leqq\alpha \quad \forall\theta$$

となるような限界 $\bar{Y}$，あるいは $\underline{Y}$ を求める問題である．これは数学的には予測区間を求めるのと同じような方法で求められるが，実際的にはやや異なる意味をもつであろう．

予測区間，あるいは予測限界を計算することは理論上，計算上複雑な問題をふくんでいない場合でも，実際には行なわれないことが多い．しかしながら，点予測値を計算するだけでは，予測の精度について誤った印象を与えることが多いから，できる限り予測区間の計算を行なうことが必要であると思う．とくにモデルの「適合度」や母数の「標準誤差」で見る限り，データに対する「あてはまり」がよいように見えるモデルにおいても，予測区間を計算すると，その幅が非常に大きくなってしまうことがあることに注意を要する．

4. 推測と予測

つぎに直接には予測を目的としないけれども，実は予測ということを念頭においた方法を考えよう．そもそも統計的データ解析の目的は，母数の推定や検定というような形式的レベルから離れて，本来の目的にもどって考えれば，構造分析，予測，制御の3つに分けられる．ここで構造分布とは，データの背後にある客観的対象の「真の構造」をつきとめようとすることであり，予測とはまだ観測されていない量について判断することであり，制御とは今後実現すべき量を望ましい水準にもたらすために適当な手段をとることであるということができよう．ここで抽象的に考えれば，まず正しく構造を把握し，それにもとづいて正しく予測し，それから適切に制御を行なうというのが本来論理的に正しい手続きであると思われるかも知れない．しかしながら現実にはそれが不可能であって，真の構造がわからないままに，経験的に得られた手がかりを用いて予測を行なったり，あるいは対象の構造が「ブラック・ボックス」のままで制御をしたりしなければならないことも多い．たとえばマクロ・エコノメトリック・モデルを用いる予測において，多くの方程式のうち，あるものは経済理論的な意味のはっきりした消費関数，投資関数等であるが，いくつかは単なる経験的統計関係式である場合がふつうである．予測のためにはこのような式を用いることはさけられない．

実際，現実のデータ解析においては，データについて多数のモデルをあてはめてみて，その上でいろいろな意味で適当と思われるものを選ぶのがふつうである．このようなやり方はふつうの統計的推測の論理から見れば，その正当性を根拠づけることは困難である．最後に選ばれたモデルが「正しい」モデルであるという保証はないし，また最初から考えられたモデルの中に正しいものがふくまれているとも限らないであろう．したがって「構造分析」の観点からは，このような手続きは疑わしいということになる．しかしながら「予測」の観点からすれば，問題になっている変量の分布を最もよく「近似」できるモデルを選ぶという考え方でこのような手続きを把え直すことができる．

その点が最も明確なのは回帰分析の場合である．データ $X_1,\cdots,X_n$ の変動を，いくつかの説明変数 $Z_1,\cdots,Z_p$ を用いて，

$$X_i=\beta_0+\beta_1Z_{1i}+\cdots+\beta_pZ_{pi}+u_i \quad (i=1,\cdots,n)$$

という回帰モデルを用いて説明しようとする場合，現実には説明変数として選ばれる可能性のある変数は多数あるのがふつうである．実際にはその中から比較的少数のものを選ばねばならないが，そのためには数多くのケースをテストしなければならない．この場合，回帰モデルを真の「因果構造」を表わすモデルと考えることは不可能であって，このような分析の目的は，このようなモ

デルを前提にして計算された係数の推定値 $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_p$ を用いて，Z の将来の値に対応する値 Y を，

$$\hat{Y}=\hat{\beta}_0+\hat{\beta}_1Z_1^*+\cdots\cdots+\hat{\beta}_pZ_p^*$$

という形で予測すること，そのために最も適当なモデルを求めることであると考えねばならない．このような観点から，最近このいわゆる「変数選択」の問題について，いろいろな基準が提案されている．これについては，この雑誌の昨年5月号(第23巻第5号)の特集「回帰分析」に筆者をふくむ何人かの人々による紹介があるので，参考にしていただきたい．

ここではより簡単な場合について説明しよう．いまある作物について k 種の品種があったとしよう．このおのおのについてそれぞれ n 回ずつ実験を行なって，単位面積当りの収量にして，

$$X_{ij} \qquad i=1,\cdots,k;\ j=1,\cdots,n$$

という値が得られたとしよう．このとき，

$$X_{ij}=\mu_i+u_{ij}$$

というモデルを想定し，u_{ij} は互いに平均0, 分散 σ^2 の正規分布に従うものと仮定すれば，品種の間に収量の差がないか否かを見るには，仮説

$$\mu_i\equiv\mu \qquad i=1,\cdots,k$$

を検定すればよい．そのためには統計量

$$F=\frac{n\sum(\bar{X}_i-\bar{\bar{X}})^2/(k-1)}{\sum\sum(X_{ij}-\bar{X}_i)^2/k(n-1)}$$

を計算して，自由度 $(k-1,\ k(n-1))$ の F 分布の上側パーセント点と比較すればよいというのが，ふつうの数理統計学の教科書に書いてあることである．

しかしこのような検定を行なうことの意味はどこにあるのであろうか．もし k 品種が，とにかくすべて異なる品種であるならば(同じ品種に異なるラベルをつけただけということはないとすれば)，とにかくその平均収量が完全に等しいということはあり得ないであろう．したがって仮説が正しくないことはいわば最初から自明である．そうすると仮説検定を行なうことはそもそも無意味ではなかろうか．

この場合，仮説検定の一つの解釈として，つぎのように考えることができる，つまりわれわれが検定しようとしているのは．絶対的に μ_i がすべて等しいかどうかではなく，それらを等しいものとして扱ってよいかどうかである．つまりデータに照らしてそれが許されないかどうかを見ようとしているのである．さらにくわしくいえば，われわれが知りたいのは，単なる「母数」μ_i ではなく，今後の各品種の収量 $Y_i,\ i=1,\cdots,k$ の平均あるいは期待値としての $\mu_i=E(Y_i)$ である．そうしてそれを推定するのに，すべての品種をいっしょにして，

$$\hat{\mu}_i{}^0=\bar{\bar{X}} \qquad i=1,\cdots,k$$

とするのがよいのか，それとも，別個に考えて，

$$\hat{\mu}_i=\bar{X}_i \qquad i=1,\cdots,k$$

としなければならないのかを知ることが問題である．そうすると前者は一般に偏りをもつけれども，分散は後者より小さくなる．したがって偏りが小さければ，前者のほうがよい推定値を与えるであろう．仮説検定はこのどちらを選ぶべきかを決めるための手続きであると解釈できる．これが北川敏男氏の「推測過程論」の考え方である．

より具体的に2つの推定量の平均2乗誤差の k 品種についての和を取ると，

$$\sum E(\hat{\mu}_i{}^0-\mu_i)^2=\sum E(\bar{\bar{X}}-\mu_i)^2=kV(\bar{\bar{X}})$$
$$+\sum(\mu_i-\bar{\mu})^2=\sigma^2/n+\sum(\mu_i-\bar{\mu})^2$$
$$\sum E(\hat{\mu}_i-\mu_i)^2=\sum E(\bar{X}_i-\mu_i)^2$$
$$=\sum V(X_i)=k\sigma^2/n$$

となるから，

$$\lambda=n\sum(\mu_i-\bar{\mu})^2/(k-1)\sigma^2$$

とおけば，$\lambda>1$ ならば $\hat{\mu}_i$ のほうがよい推定量になり，$\lambda<1$ ならば $\hat{\mu}_i{}^0$ のほうがよい推定量ということになる．$(k-1)\lambda$ が F 統計量の非心母数の値になっているから，その値を基準にして選択を行なうのは合理的である．

しかしながら，ふつうに用いられる5%点などは F 値の限界としては大きすぎる．というのは，

そうすると $\lambda>1$ でも仮説が棄てられないで，$\hat{\mu}_i^0$ のほうが選ばれる可能性が大きくなってしまうからである．より適当な F の限界値の定め方にはいろいろな考え方があるが，大まかにいえば $F\fallingdotseq 2$ を限界とするのがほぼ適当である．

ところでさらに問題を上記のように考えるならば，μ_i の推定値として $\bar{X}_i$ $\bar{\bar{X}}$ のどちらかを取るのではなく，F の値に応じて適当な比重で結合したものを用いたらよいのではないかと考えられる．実際 C. Stein は $k\geqq 4$ のとき，平均2乗誤差の和がつねに $k\sigma^2/n$ より小さくなるようにすることができることを示した． ここで Stein の作ったものとやや異なる形を示すと，

$$\hat{\mu}_i^*=\bar{\bar{X}}+\alpha(\bar{X}_i-\bar{\bar{X}})$$

$$\left(\text{ただし，}\ \alpha^{-1}=1+\frac{f}{f+2}\,\frac{k-3}{k-1}\,\frac{1}{F},\ f=k(n-1)\right)$$

とすれば，つねに，

$$\sum E(\hat{\mu}_i^*-\mu_i)^2<k\sigma^2/n=\sum E(\hat{\mu}_i-\mu_i)^2$$

となることが証明されている．（この形については竹内啓「計量経済学の研究」（東洋経済新報社）第7章，Stein 推定量一般については，篠崎信雄「Stein 推定量について」日本数学会1977年秋季総合分科会特別講演予稿　参照.）

上記のような「平凡な」統計的推測の問題も，予測の観点を導入すると，いろいろ興味ある問題をふくんでいることがわかるであろう．

5. むすび

最初に述べたように，具体的な予測の問題において，形式的な統計的方法によって最終的な結論が得られることは少ないとしても，予測のために統計的方法が果たすべき役割は大きい．そうして統計的方法が予測のためにより有効となるためには，予測のための統計的方法にもっと関心が向けられることが望ましいし，また統計的推測の手法を予測の観点から見直すことも必要である．すでにこのような方向についていくつかの成果は得られているが，なお未解決の問題も多く残されている．この分野の問題について多くの理論家，実際家の関心が今後より多く寄せられることを期待したい．

たけうち・けい　1933年生
1956年　東京大学経済学部卒
現　在　東京大学経済学部教授　経済学博士

交換雑誌をご利用下さい

下記の雑誌が交換によって学会事務局に来ておりますので，どうぞご利用下さい．カッコ内は交換先です．

Cahiers du Centre d'Etudes de Recherche Operationnelle (Institut de Statistique de l'Universite Libre de Bruxelles, Belgique)

Transactions of the Royal Society of Canada (Royal Society of Canada, Canada)

数学的実践与认訳（中国科学院数学研究所）

Aplikace Mathematiky (Ceskoslovenska Akademie ved Matematicky Ustav Csav, Czechoslovakia)

Ekonomicko Matematicky Obzor (Czechoslovak Academy of Sciences, Czechoslovakia)

Mathematical Proceedings of the Cambridge PhilosophicalS ociety (Cambridge Philosophical Society, England)

Steelresearch (The Corporate Laboratories of the British Steel Corp., England)

International Journal of Production Research (University of Birmingham, England)

Journal of Royal Statistical Society (Royal Statistical Society, England)

Revue Française d'Automatique, Informatique, Recherche Operationnelle (A. F. C. E. T., France)

Mathematische Operationsforschung und Statistik (Universitäts-Bibliothek der Humboldt-Universität, Germany)

Indagations Mathematicae (Library Mathematisch Centrum, The Netherlands)

Hong Kong Productivity News (The Hong Kong Productivity Centre, Hong Kong)

Mathematical Chronicle (University of Auckland, New Zealand)

Dissertationes Mathematicae (Polskiej Akademii Nauk, Poland)