研究ノート

プログラムの適否判定についての統計学的一考察

石　川　　浩*

A Statistical Approach to the Evaluation of the Validity of a Computer Program

By
Hiroshi Ishikawa

Associate Professor, School of Business, Kagawa University
2—1 Saiwai-cho, Takamatsu-shi 760 JAPAN

1. はじめに

近年，工学の分野におけるコンピューターの導入は著しいものがあり，高度に複雑な科学計算を内包したコンピューター・プログラムが数多く開発されるに至っている。一般にある与えられた課題に対して，その解決へ向けての技術的なアプローチの手法は必ずしも単一のものではなく，したがって開発されるプログラムも各社各様のものとなるであろうことは容易に推察されるところである。それゆえ，ある社のプログラムが技術的に妥当なものであるか否かを，何らかの基準で経済的かつ簡便に判別することができうるとするならば好ましい限りである。もちろん，そのプログラムの論理を逐一たどって行けば妥当性の評価は厳密な観点からなしえようが，多大の労力を要して経済的でなく，かつ短時日の処理という点で難点が大きい。

本稿は上述の問題に対して，通産省の委託を受けて高圧ガス保安協会の内部に設置されたA-ADHOC委員会における，筆者の分担課題についての報告を基にして，問題の所在とその解決への手がかりを明らかにしようとするものである。[1] すなわち，ある課題に対し

---

* 香川大学商業短期大学部（〒760高松市幸町２—１）

て叡智を集結して作成された標準プログラムが一方に存在するとき，他方において全く独立にある社で作成されたプログラム（これを以後においては検定プログラムと称することとする）があるものとして，その適否を「有限母集団における比率の推定手法」を用いて統計的に判定することができるか否かという点について論じるものである。

2．問題の所在と明確化

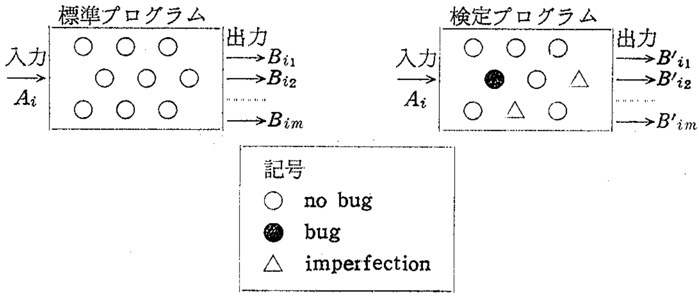

図1 プログラムの入出力状態

さて，標準プログラムは言うまでもなく何らの bug をも有しないとみなせうるものであって，図1に示すように，第 i 番目のケースの入力 A_i に対して出力 $B_{ij}(j=1,2,\cdots,m)$ をもつ。一方，同じ入力条件に対して検定プログラムは内部に保有するかも知れない bug や理論適用等における不備（imperfection）等によって標準プログラムとは異なった部分を含む出力 $B'_{ij}(j=1,2,\cdots,m)$をもつであろうと予想される。もちろん，

$$B'_{ij} \equiv B_{ij} \quad \textit{for all i's and j's}$$

であれば理想的に最良のプログラムと判定されるが，実際問題としては完全に同一のプログラムでなければこのようなことは起こりえないであろうから，各々の出力 B_{ij} に対してある許容誤差レベルを定めて，検定プログラムの対応する出力 B'_{ij} がその許容範囲内にあればそのパスは合格と考え，その他の場合は不合格とみなすことによって，検定プログラムが全体としてどの程度の不合格率をもてば否と判定するかを考える必要があろう。

与えられた検定プログラムの構造を詳細にトレースすれば，もちろん，そのプログラムの適否はほぼ完全に判定されるものと考えられようが，経費と効率の点から難点がある。そこで何らかの統計的手法を援用することとなるが，この際には，予め種々の例題を準備しておき，第 k 番目の例題$(k=1,2,\cdots,N)$について入力条件 $A_i^{(k)}$ を種々に変え $(i=1,2,\cdots,l)$，その各々に対して標準プログラムの出力 $B_{ij}^{(k)}$ $(j=1,2,\cdots,m)$と対応する検定プロ

グラムの出力 $B'_{ij}{}^{(k)}(j=1,2,\cdots,m)$ を比較する必要が生じる。さらにはN個の例題すべてについて比較検討を行うことはやはり経費と効率の点から難点を生じることが容易に推察されるから，N個の中からランダムにある少数のサンプル例題を n 個選び出し，その結果について適否判定を行うという処理方法を取ることが考えられる。

以上のような場合には以下のように解決されなければならない様々な問題点が存在する。

(i) 何例程度の例題を用意しておくべきか（Nをいくらにするか)。またプログラムの bug や imperfection に敏感な例題の内容はどのようなものとしておくべきか（検定プログラムの構造は未知であるという点に留意して)。

(ii) ランダムに少数個の例題を選んで（n 個のサンプルに基づいて)，母集団（N個全体の）特性を推測する場合の信頼度の問題。

(iii) 選ばれた例題に対する入力条件 $A_i(i=1,2,\cdots,l)$)は幾種類程度選択するか（l はいくらとすべきか)。また A_i をどう散らしたらプログラムの欠陥に最も sensitive となるのか。

(iv) 1つの入力 A_i に対して考える出力 B_{ij} として何を選ぶか。例えば，構造解析用のプログラムということを例にとれば，応力，ひずみあるいは固有振動数などが B_{ij} として考えられるが，これらは必ずしも独立なものではなく，考える B_{ij} の組合せによっては非常に相関のある場合が生じよう。あるいは逆に理論的には非常な相関があるべきにもかかわらず，内包する bug や imperfection によって一見独立なように見えるという矛盾点を以て適否判定の check に利用しうるか。

(v) 選ばれた B_{ij} に対して，適否判定の許容誤差レベルをどのように与えるべきか。検定プログラムの対応する出力 B'_{ij} との比をとり $|(B'_{ij}-B_{ij})/B_{ij}|$ と normalize した形で異なる種類（異なる j）の出力の許容誤差を一律に与えることが可能か。あるいは出力毎に許容誤差レベルを変えるべきかについて技術的観点からの判断はどうか。

(Vi) ある例題のある入力条件 $A_i{}^{(k)}$ に対する検定プログラムの各々の出力 $B'_{ij}{}^{(k)}$ の合否を積重ねて，全体としての合否の情報に基づいて最終的なプログラムの適否判定をいかに行うべきか。

上述のようにプログラムの適否判定の問題は非常に広範囲の内容を含み，問題の完全な解決のためには少なからぬ時日を要するように思われる。本稿の課題はプログラムの適否判定に際して，有限母集団のあるクラスに入るものの比率を推定する手法が応用できるか

否かを明らかにすることにあるが，上述のすべての問題点を包括することはあまりに大部にわたることを鑑み，本稿においては問題解決へ向けての基礎的な手がかりを与えることを主眼とした。

3. 有限母集団における比率の推定手法

総数 N 個の例題からランダムに n 個の例題標本を選ぶ場合を考える。ある例題に対して，入力条件 $A_i(i=1,2,\cdots,l)$ のそれぞれに対して複数個の出力 $B_{ij}(j=1,2,\cdots,m)$ が対応する場合が一般であるが，はじめに述べた問題点ならびに基礎的な観点からの検討ということを考えて１入力（$l=1$），１出力（$m=1$）の場合を例にとって話を進める。すなわちこの場合にはある例題 (k) に関して，１入力 A_k に対する標準プログラムの出力 B_k ならびに検定プログラムの出力 B'_k のみを考えればよい。

さて，用意されている全例題（N 個の母集団）に対してもしも標準プログラムを用いれば解 $B_k(k=1,2,\cdots,N)$ が得られ，検定プログラムを用いれば対応する解 $B'_k(k=1,2,\cdots,N)$ が得られるものとしよう。B_k に対する B'_k の差異は元々検定プログラム中に含まれる確定的な bug や imperfection に基づくものと考えられるので，差異が生じたとした場合，その差異の変動が統計的な性質を本質的に有するものなのか否かという点については議論の余地があるであろうが，ここでは例題がうまく設定されており，B_k と B'_k との差異は例題毎に統計的な振舞いをするものと仮定しよう。k 番目の例題に関しての許容誤差レベルが相対誤差の形で $\varDelta_k$ と与えられているものとして，その例題に対する検定プログラムの合否を

$$\left.\begin{aligned} \left|\frac{B'_k-B_k}{B_k}\right| &\leq \varDelta_k \longrightarrow \text{合格} \\ \left|\frac{B'_k-B_k}{B_k}\right| &> \varDelta_k \longrightarrow \text{不合格} \end{aligned}\right\} \qquad (1)$$

と判定する。もちろん $\varDelta_k$ については別の観点からの技術的な考察に基づいて基本的資料が与えられているものと考える。場合によっては $\varDelta_k\equiv\varDelta\ (k=1,2,\cdots,N)$ ということも起こりうるであろう。

3･1 ２クラス分類における個数の分布

判定基準式（１）に基づいて N 個の全例題に対して検定プログラムの合否を判別したとすれば，その結果は合格のクラス G(Good) と不合格のクラス B(Bad) という２つのクラス

に分類することができる。このようにして検定プログラムの２クラス分類の母集団が概念的に構成される。この母集団においてクラスＢ（不合格)に属するものの比率を p とする。

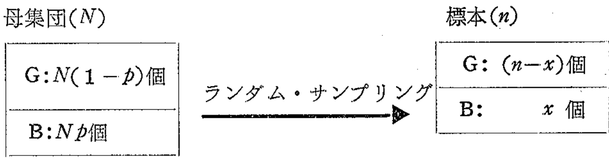

図２　単純ランダム・サンプリング

全例題について検定プログラムを走らせることは不経済であるから，任意に選んだ n 個の例題のみを用いて各例題に対する合否判定を式(１)に基づいて実際に行い，それらの結果をＧとＢの２クラスに分類する。この関係は丁度図２に示した有限母集団からの単純ランダム・サンプリングの関係で表すことができる。すなわちここでの問題は N 個の母集団におけるクラスＢ（不合格）に属する成員の比率 p が未知であるから，これを n 個のランダム・サンプルにおけるクラスＢの成員の比率を基にして推定し，以て検定プログラムの適否を判定しようとするものである。

はじめに，寸法 n のランダム・サンプルにおけるクラスＢに属するものの個数 X の分布を求めよう。

事象〔$X=x$〕というのは，母集団(N)から選ばれた n 個のサンプル中で，クラスＧに属するものの個数が（$n-x$）個，またクラスＢに属するものの個数が x 個であるという事象である。この事象の確率は以下のようにして求められる。

(i)　N 個のものから n 個を選び出す相異なる場合の数: $\binom{N}{n}={}_NC_n$

(ii)　母集団でＧに属する $N(1-p)$ 個のものから，標本中でＧに属する $(n-x)$ 個のものを選び出す場合の数: $\binom{N(1-p)}{n-x}={}_{N(1-p)}C_{n-x}$

(iii)　母集団でＢに属する Np 個のものの中から，標本中に x 個のものを選び出す場合の数: $\binom{Np}{x}={}_{Np}C_x$

(ii)の場合の各々に対して，(iii)の場合のいずれが対応してもそれは事象〔$X=x$〕を与え，かつ各々の場合はいずれも同様に確からしく生じるから，結局 X の確率関数は

$$f_X(x)=P[X=x]=\frac{\binom{Np}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}} \qquad (2)$$

ただし，x は $0\leqq x\leqq \min(n,Np)$ の整数

ここに，min(n, Np)とはnとNpとのうちいずれか小さい方を表す。式(2)の確率関数をもつ確率変数Xは超幾何分布に従うことがわかる。[2] Xの特性量は以下の如くである。

Xの平均： $$E[X]=\sum_{x=0}^{\min(n,Np)} x f_X(x)=np \tag{3}$$

Xの分散： $$V[X]=E[(X-E[X])^2]$$
$$=\left(\frac{N-n}{N-1}\right)npq \tag{4}$$

ただし，$q=1-p$ (5)

Xの変動係数：

$$\delta_X=\sqrt{\left(\frac{N-n}{N-1}\right)\left(\frac{q}{np}\right)} \tag{6}$$

3·2 2クラスの場合の比率の推定

寸法nのランダム・サンプル中ではクラスBに属するものの個数がXであるから，母集団におけるクラスBの比率pを

$$Y=X/n \tag{7}$$

で推定しようとするのは自然の試みである。Xが確率変数であるから，当然に比率pの推定量Yもまた確率変数となる。

Yの確率関数を$f_Y(y)$とすれば，

$$f_Y(y)=P[Y=y]=P[X/n=y]=P[X=ny]$$

$$=f_X(ny)=\frac{\binom{Np}{ny}\binom{N(1-p)}{n(1-y)}}{\binom{N}{n}} \tag{8}$$

Yの平均：$$E[Y]=E\left[\frac{X}{n}\right]=p \tag{9}$$

すなわち，比率の推定量Yは不偏推定量である。

また，Yの分散は，

$$V[Y]=V\left[\frac{X}{n}\right]=\frac{1}{n^2}V[X]$$

$$=\left(\frac{N-n}{N-1}\right)\left(\frac{pq}{n}\right) \tag{10}$$

Yの変動係数:

$$\delta_Y=\frac{1}{\sqrt{n}}\sqrt{\frac{q}{p}}\sqrt{\left(\frac{N-n}{N-1}\right)} \tag{11}$$

これはXの変動係数と全く同じである。

式(11)において$n=1$とすると,

$$\delta_Y(n=1)=\sqrt{q/p} \tag{12}$$

表1　$\delta_Y(n=1)$とpとの関係 $(q=1-p)$

p	0.001	0.005	0.01	0.05	0.1	0.2	0.5	0.8	0.9
$\sqrt{\frac{q}{p}}$	31.6	14.1	9.9	4.4	3.0	2.0	1.0	0.5	0.3

表1に$n=1$の場合の$\delta_Y(n=1)$と母集団比率pとの関係を示すが，明らかに，サンプルの大きさ(n)を固定すれば，クラスBに入るものの全体に対する比率pの推定量Yの変動係数δ_Yは，母集団においてクラスBの占める割合が大きくなるにつれて減少することがわかる。$p=5$％以下では変動係数の値が大きい。換言すれば，母集団において占める割合が小さいようなものの比率を正確に推定するためには大きなサンプルを必要とすると言えよう。

3・3　信頼限界

はじめに母集団においてクラスB（不合格）に属するものの総数$A=Np$についての推論を考えよう。いま，サンプル中においてn個のうちのa個がクラスB（不合格）に属することが観測されたとして，Aの信頼区間を求めてみる。

n個のサンプル中でクラスBに属するものの個数がa以下となる確率がある与えられた小さなα_U（例えば0.025）となるようなAの値$\hat{A}_U$を計算すれば，これが信頼区間の上限を与える[3]。サンプル中のBに属するものの個数Xは式(2)で表される確率関数をもつから，求める$\hat{A}_U$の値は次式を満足する。

$$\sum_{x=0}^{a} f_X(x|\hat{A}_U, N-\hat{A}_U)=\sum_{x=0}^{a}\frac{\binom{\hat{A}_U}{x}\binom{N-\hat{A}_U}{n-x}}{\binom{N}{n}}=\alpha_U \tag{13}$$

ここに，$f_X(x|\hat{A}_U, N-\hat{A}_U)$は$Np=\hat{A}_U$を与えたときの$X$の確率関数である。さて，$\alpha_U$を与えたとき，式(13)を満たす$\hat{A}_U$は一般には整数とはならないから，実際には式(13)の左辺がα_Uより小さいか，あるいは等しいような$\hat{A}_U$のうち，もっとも小さな整数をとり，これを$\hat{A}_U$とすればよい。

同様にして，Aの信頼区間の下限$\hat{A}_L$は，

$$\sum_{x=a}^{\min(n,\hat{A}_L)} f_X(x|\hat{A}_L, N-\hat{A}_L)=\sum_{x=a}^{\min(n,\hat{A}_L)} \frac{\binom{\hat{A}_L}{x}\binom{N-\hat{A}_L}{n-x}}{\binom{N}{n}} \leqq \alpha_L \tag{14}$$

を満たす最大の整数として求めることができる。ここに，α_Lは下限界を与える信頼レベルである。

以上の$\hat{A}_L, \hat{A}_U$を用いて母集団比率pの信頼限界が以下のように求まる。

$$\left.\begin{array}{ll} p\text{の下側信頼限界：} & p_L=\hat{A}_L/N \\ p\text{の上側信頼限界：} & p_U=\hat{A}_U/N \end{array}\right\} \tag{15}$$

3・4　比率推定における標本寸法nと標本抽出率$\frac{n}{N}$

母集団の各成員はそれぞれクラスG（合格）および，クラスB（不合格）のいずれかに分類されるものとし，クラスBに入る成員の比率pの推定量を式(7)のようにYとすればYの分布はすでに述べた式(8)から

$$f_Y(y)=\frac{\binom{Np}{ny}\binom{N(1-p)}{n(1-y)}}{\binom{N}{n}} \tag{16}$$

ただし，$0\leqq y\leqq \min\left(1, \frac{Np}{n}\right)$

であり，かつyは$\frac{1}{n}$毎の離散的な値をとる。

$$E[Y]=p \tag{17}$$

$$V[Y]=\left(\frac{N-n}{N-1}\right)\left(\frac{pq}{n}\right)=\sigma_Y{}^2 \tag{18}$$

$$\text{ただし，}q=1-p \tag{19}$$

またσ_Yは標準偏差であって，

$$\sigma_Y=\sqrt{\frac{N-n}{N-1}}\quad\sqrt{\frac{pq}{n}} \tag{20}$$

Yの変動係数 $$\delta_Y=\sqrt{\left(\frac{N-n}{N-1}\right)\left(\frac{q}{np}\right)} \tag{21}$$

ここで，推定量Yの誤差限界をdとし，また現実の誤差がdより大きいという結果になる確率をα以下に抑えるものと考えれば，必要な標本寸法nは

$$P[\,|Y-p|>d\,]\leqq\alpha \tag{22}$$

を満足する最小の整数として求めることができる。ただし式（22）は母数pを含むから，これが既知の場合はよいが，未知であれば，例えば標本による推定値を第1近似として用いることになろう。

ところでYは式（16）のように超幾何分布に従う確率変数であるが，ある条件の下で正規近似が満足されるものとすれば，そのような場合には，平均p，分散$\sigma_Y{}^2$の正規分布が適用されることになる。したがって式(22)を満足する関係式はきわめて簡単に

$$d\geqq k\sigma_Y\left(=k\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{pq}{n}}\right) \tag{23}$$

ただし。kは標準正規分布の$\left(1-\frac{\alpha}{2}\right)$%点

となる。それゆえ，

$$n\geqq\frac{\left(\frac{k^2pq}{d^2}\right)}{1+\frac{1}{N}\left(\frac{k^2pq}{d^2}-1\right)} \tag{24}$$

を満足する最小の整数が求める標本寸法nを与える。なお，$\frac{n}{N}$の値を標本抽出率と呼んでいる。

4. 二，三の解析結果

図3～図5に超幾何分布の確率関数$f_X(x)$を図示した。いずれも一例として母集団の大きさ$N=100$の場合を選び，標本寸法nを$n=5,10,20$とし，pをパラメータとして描いたものである。pもしくはnが小さい場合には著しく左右非対称の分布となるが，pもしくはnが大きくなるにつれて非対称度は小さくなる。pが1.0に近くなっても分布形の非対称度は大きくなる。なお，超幾何分布は本来離散的な分布であるから確率関数は整数値のxの所のみで値をもつ棒グラフのように表すべきであるが，ここでは繁雑さを避けるた

め折れ線グラフとして表示した。

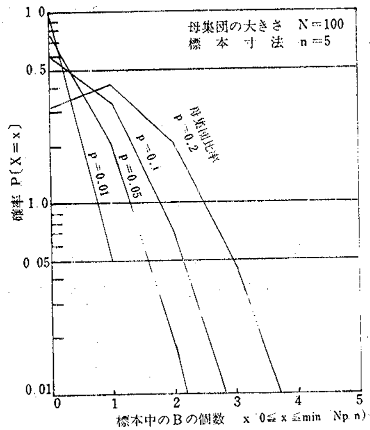

図3　超幾何分布の確率関数($N=100; n=5$)

$$P\{X=x\}=\binom{Np}{x}\binom{N(1-p)}{n-x}\Big/\binom{N}{n}$$

ただし，$0\leqq x\leqq \min(n, Np)$

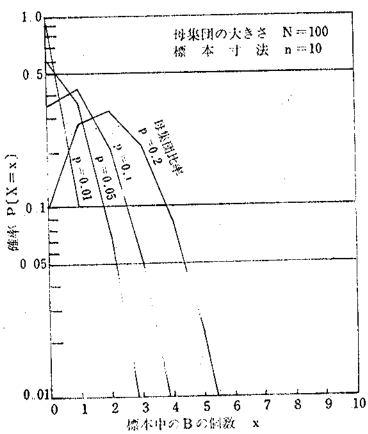

図4　超幾何分布の確率関数

($N=100; n=10$)

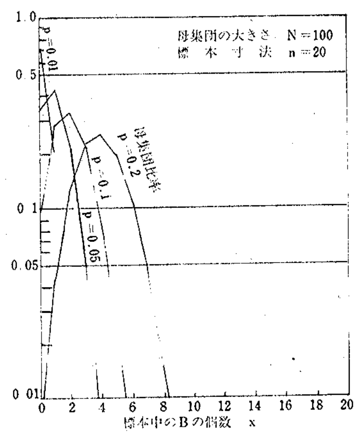

図5　超幾何分布の確率関数

($N=100; n=20$)

図6（$N=50$）および図7（$N=100$）は母集団比率 p の推定量 Y の変動係数 δ_Y が p および標本寸法 n によってどのように変化するかを調べたものである。$N=50$ および 100は，プログラムの適否判定のために現実的観点から準備されるであろう例題の総数を便宜的にその程度のものとした場合ということから選んだ。明らかに標本寸法 n が小さな場合には変動係数の値が大きく，また p が小さいような母集団の比率を推定する場合にも変動係数の値は大きい。

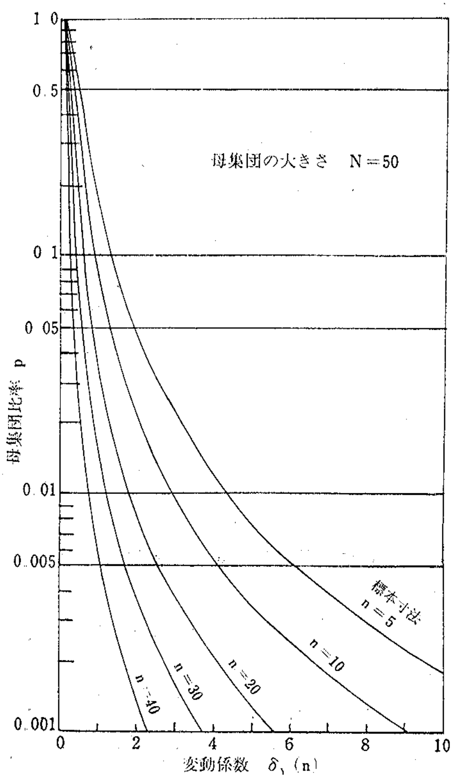

図6 母集団比率の推定量 Y の変動係数 δ_Y に及ぼす母集団比率 p および標本寸法 n の影響（$N=50$）

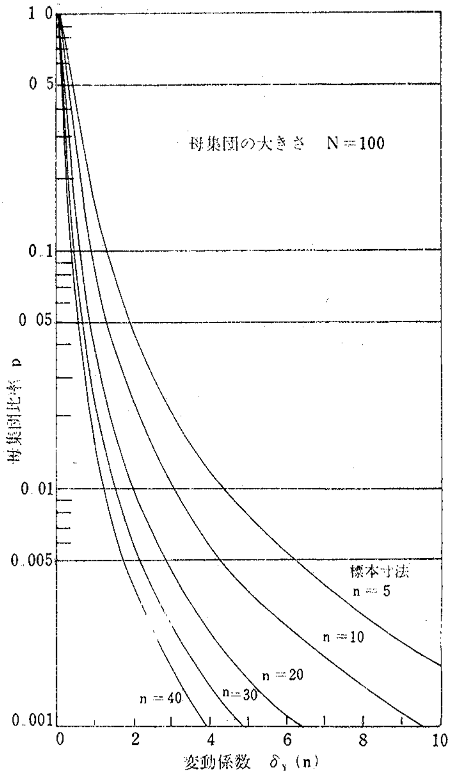

図7　母集団比率の推定量 Y の変動係数 δ_Y ($N=100$)

図8 ($N=50$) および図9 ($N=100$) は一定の変動係数 δ_Y を与えるための p と n との関係を示したものである。それゆえ，これらの図は一定の変動係数 δ_Y を与えるためには，比率 p が異なればどの程度の標本抽出が必要かということを示すものとも解釈される。

例えば $N=100$ の場合（図9），母集団中に占める割合が $p=10$ %程度のものを変動係数 $\delta_Y=0.3$ に抑えて比率推定する場合には全体のうちの約半分程度を抽き取って調べる必要があるが，$p=50$%程度のものであれば，全体の約1割程度を調べるだけでよいと考えられる。

もちろん小さな p のものを精度よく推定しようとすれば標本抽出率が100%，すなわちほぼ全数を調査し尽さなければならないということになるが，例えば図9において，$p=$

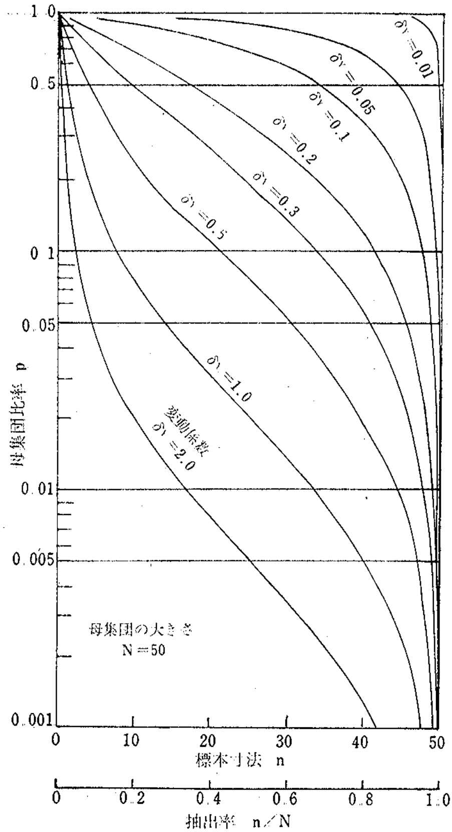

図８　変動係数 δ_Y を一定とした場合の p と n の関係
（標本抽出率 n/N の挙動）（N=50）

1％のものをδ_Y=2.0程度で収めたとすれば（抽出率20％），$\sigma_Y=0.01\times2.0=0.02$ であり，またp=10％のものを同じく δ_Y=2.0 に収めるとすれば（抽出率２～３％），$\sigma_Y=0.1\times2.0=0.2$であって絶対的な$\sigma_Y$ の値ははるかに異なるので，この辺の事情もよく勘案しておく必要があろう。

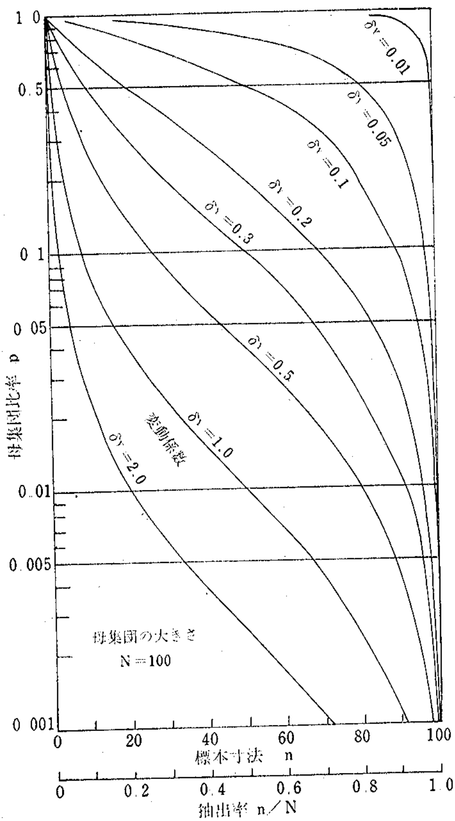

図9 変動係数 δ_Y を一定にした場合の p と n の関係 ($N=100$)

図10($\delta_Y=0.01$), 図11($\delta_Y=0.1$) および図12($\delta_Y=1.0$) は同じ変動係数 δ_Y を与える p と n との関係を母集団の大きさ N をパラメータとして表示したものである。δ_Y が大きくなるほど，標本抽出率は小さくてすむことがわかる。

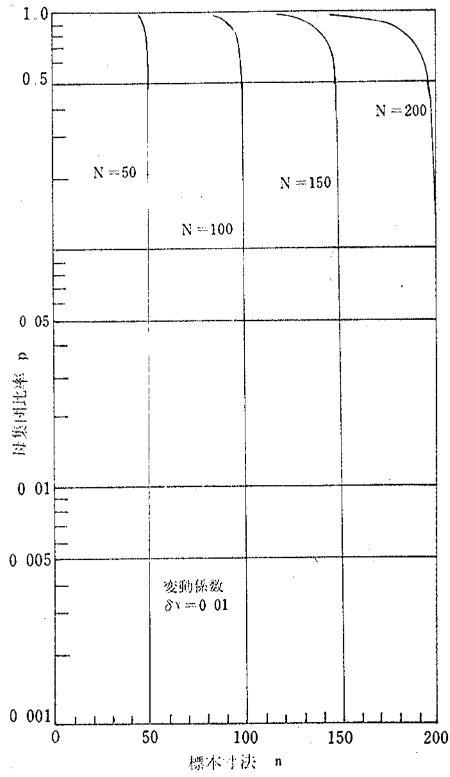

図10　等変動係数 δ_Y の下での母集団の大きさ N と標本寸法 n の関係（$\delta_Y=0.01$）

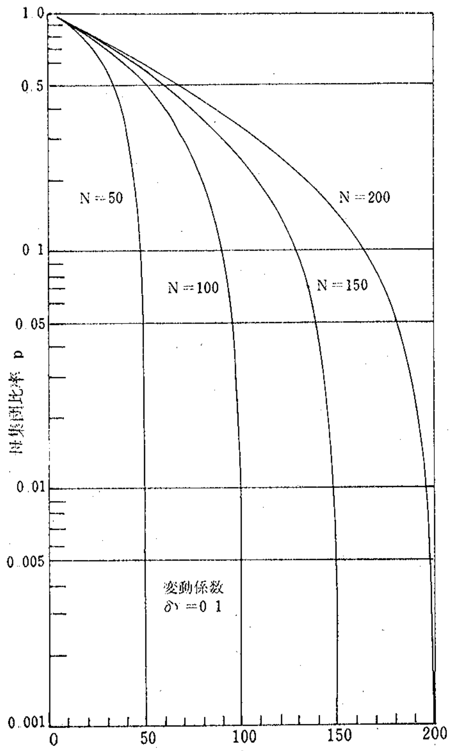

図11 等変動係数 δ_Y の下での N と n の関係(($\delta_Y = 0.1$)

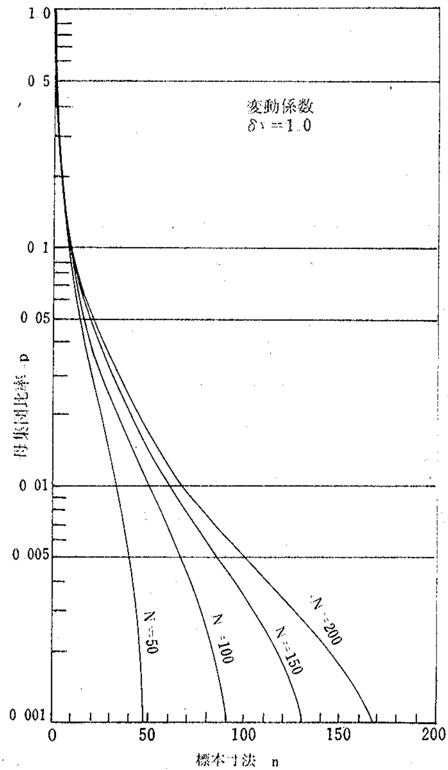

図12 等変動係数 δ_Y の下での N と n の関係 ($\delta_Y=1.0$)

5. プログラムの適否判定

N 個の例題のうちからランダムに n 個を選んで，ある入力条件の下で検定プログラムを走らせ出力 B'_k $(k=1,2,\cdots,n)$ を観測する。これを同じ入力条件に対応した標準プログラムの出力 $B_k(k=1,2,\cdots,n)$ と比較し，許容誤差 $\varDelta_k(k=1,2,\cdots,n)$ を用いて式 (1) の判定条件によって各々の k に対して合否を判定し，n 個のサンプル中で不合格となったものの個数 a を計数する。

技術的な判断に基づいて，許容誤差 $\varDelta_k$ は全例題 N 個の中で不合格になったものの比率

p がある値 p_0 以下の場合にはその検定プログラムを認定するような基準で定められているものとしよう。Δ_k と p_0 とはもちろん密接に結びついていて，Δ_k を厳しくすれば p_0 が大きくなり，逆に Δ_k の制約が緩やかな場合には p_0 は小さな値としてもよいことが予想される。

さて，n 個のサンプル中で a 個が不合格（クラスBに属する）となったことが観測されたとしたとき，先に第3・3節に述べたように母集団の不合格率 p の上下限信頼限界を式 (13) ～式 (15) を用いて p_U および p_L として求めることができる。それゆえ，これらの値をプログラムの適否判定の許容不合格率 p_0 と比較することによって，プログラムの適否判定を行うことができる。p が不合格率であることを考え，例えば最も保守的な立場として上側信頼限界 p_U を用いて

$p_U \leqq p_0$　　ならば　合格と認定

その他の場合　　不合格と認定

とすることが考えられる。p_U の計算に用いる危険率 α_U の値をいくらにすべきかは技術的な判断を要する。なお，N 個の全例題中何個位のサンプルを選べばよいのか，またその時の適否判定の精度はどうか等については，前節における議論を参照すればよい。

6. おわりに

プログラムの適否判定に関連して，有限母集団の比率の推定手法を応用する可能性について基礎的な観点から考察した結果について論じた。はじめに述べたようにプログラムの適否判定には非常に様々の問題点があり，それらが解決されて初めてこの問題が解決を見るわけであるが，基本的には統計的手法の応用の可能性が示唆されたものと考えられる。

同一入力条件に対する複数個の出力 B_{ij} ($j \geqq 2$) を考える場合には，例えばクラスター・サンプリングの考え方等が援用できそうに思えるが，これにいては今後の課題として考えて行きたい。

参考文献

〔1〕 石川浩,「有限母集団における比率推定手法を応用したプログラムの適否判定について」,「耐震設計用プログラムに係る認定方法の検討」(高圧ガス保安協会),昭和56年3月，pp.350～361.

〔2〕 木村等，石川浩,「確率・統計学入門」,昭和55年3月，新日本印刷所，p.156.

〔3〕 鈴木ら共訳・コクラン原著,「サンプリングの理論と方法(1)」,昭和47年6月，東京図書(株)，p.63.