8

実験データの採り方（I）

1. 標本調査法と単純ランダムサンプリング

1.1 サンプリング

本章の標本調査法（methods of sampling survey）は，対象母集団からの標本抽出の方式，抽出標本による推測方法や推定精度等について理論的に明らかにすることを目的とする。

すなわち，実際面における技術と経済および数理統計の理論によって，客観的合理性をもつ標本抽出方式の選定，推定値の算出法や，それらの精度や信頼性に関して考察される。これに対して，旧来からの全数調査法（センサス，sensus）があり，数年ごとに実施される人口調査は1つの例である。全数調査は経済・労力等に困難があるばかりでなく，その精度に関しても論議が多い。

実際，ある工場における原材料から，半製品や製品までの品質管理，工場実験，工程解析，出荷検査において，また試験研究や開発研究を推進するにも，合理的な標本抽出方式の選定が必要である。

1.2 母集団，サンプル，サンプリング

まず，対象とする統計的集団特性を知るために，その一部を抜きとって観測または測定が必要になる。この対象となる統計集団を母集団といい，この母集団から一部を抜き取って測定するものを標本（サンプル，sample）という。またこのようにサンプルを抽出することを標本抽出（サンプリング，sampling）という。サンプリングによって推定を行う目的は，形式的に

（i） 母集団のある特性値の平均，総計または母集団内の構成率を推定する。

（ii） 母集団のある特性の分布型，すなわちバラツキ具合や偏りの程度を知る。

の2つに区分される。ただし，上の（i）と（ii）を同時に知ることを目的とする場合も多い。

ここで重要な注意は，我々が知る対象は常に母集団であって，個々の観測値ではないということである。すなわち，母集団特性を推定するためにサンプルを抽出するのであって，サンプル値を得ることが最終目的ではない。したがって，サンプルとして必要な条件は，目的とする対象母集団を，最も有効かつ経済的なサンプリングによって，偏り（bias）がなく，最大限に精度よく推定できることである。

母集団の大きさは数式計算の上で，サンプルの単位によって，母集団が有限個の要素から成り立つ場合と，無限個から成り立つ場合に依存する。前者を有限母集団（finite population），後者を無限母集団（infinite population）という。工場などの少量生産では有限母集団が多いが，ロットの大きさによって実用上から統計的に無限母集団と考えて支障ない場合も多く，理論的に無限母集団として論じる方が一般的で，そのとり扱いも容易である。実際，無限母集団のとり扱いは有限としたときより必要サンプル数は若干増すが，推定上の安全性を考慮して推定されることになる。

1.3 信頼性，精度，確度

これまで誤差（error）という術語を，その意味内容を明確にせずサンプリング誤差や測定誤差などの上で使用してきた。この誤差の内容は大別して次の3種類に分類される。(a) 信頼性（reliability），(b) 精度（precision），(c) 確度（accuracy）。

(a) 信頼性とは，サンプリングや測定が常に決められたとおり実施され，充分信頼できるか否かということである。すなわち，操作上の間違い，例えば担当者が無作為抽出法を不注意に偏りのある有意抽出で行っているような場合は，規定どおりのサンプリングや測定が行われず，サンプリング誤差や測定誤差を検討したり，試験法の精度，正確さを論じても意味がない。データに関する信頼性はすべての論議や分析に先行して考える必要がある。

(b) 精度とは，同じサンプリング法で同一の母集団から繰り返しサンプリングしたとき，または同じ測定法で同一サンプルを測定したとき，データの平均値に対するバラツキをいう。したがって，このようなバラツキは管理状態になければならない。通常，精度を示すためには次のような尺度を用いる。

（i） 標準偏差 σ　（ii） 分散 σ^2　（iii） 変動係数 $C \cdot V = \sigma/\mu$　（iv） 範囲または平均範囲　（v） 信頼区間

例えば，正規母集団からのサンプリングを考えるとき，ある信頼係数（1-α）を定めて，n を標本数，β を推定の精度とし，（v）の信頼区間の尺度で

（b-1）母標準偏差 σ が未知の場合

$$\pm\beta=\pm t(n-1,\alpha)\ s/(n-1)^{1/2}$$

ここに，s は標本標準偏差，$t(n-1,\alpha)$ は危険率 α で自由度 $n-1$ の t 分布における棄却限界である。

（b-2）母標準偏差 σ が既知の場合

$$\pm\beta=\pm u(\alpha)\ \sigma/n^{1/2}$$

ここに，$u(\alpha)$ は危険率 α で正規分布における両側棄却限界である。

（c）確度とは，同一母集団から同じサンプリング法で繰り返しサンプリングしたとき，または同一サンプリングを同じ測定法で繰り返し測定したときに，その標本平均が $n\to\infty$ のときに確率収束する値と母集団平均値の偏りをいい，これを確度，または裏の意味で偏り（bias）という。確度は信頼性と精度が考察された後に，はじめて論議される。

1.4 サンプリング誤差と測定誤差

これまで，サンプリングの誤差と測定誤差を別々に定義づけてきたが，実際にデータを得る際には，この2つの誤差が相伴って同時に起きる。すなわち，観測者が得るデータは，サンプリングして測定するという2つ過程を経ているので，この2つの誤差が加算されている。したがって，工場等における工程の管理や解析等では，この2つの誤差を分離して考察する必要がある。

2つの精度（標準偏差）を σ_S と σ_M とし，2つの誤差が互いに独立とすれば分散の加法性から

$$\sigma=(\sigma_S{}^2+\sigma_M{}^2)^{1/2} \tag{8.1}$$

が成立する。したがって，この2つの誤差を分離するには，この式を利用すれば算出できる。しかし，破壊試験などでは，この2つの誤差が完全に交絡するので，これらの誤差の分離はできない。また，2つの誤差が完全に独立でなく，その間の相関性が極めて弱いと考えられるならば，上式を実用的に用いてもよい。2つの誤差に明らかな相関性があり，この母相関係数を ρ とすると，精度は次式で示される。

$$\sigma=(\sigma_S{}^2+\sigma_M{}^2+2\rho\sigma_S\sigma_M)^{1/2} \tag{8.2}$$

サンプリングと測定の2つ誤差に分離されるときは，各大きさの違いにより精度を上げる重み（weight）が異なってくる。すなわち，データの精度をある一定値以下に抑えたいときに，σ_S または σ_M の一方のみを小さくしてもあまり意味がない。

サンプリング誤差と測定誤差を考えるとき，必ず，この2つを分離する必要がある。データはこれらが一緒になっているので，σ_S または σ_M の一方が求められると，他方が得られる。通常は σ_M が統計的または技術的に得られやすい。また，σ_M は均一なサンプルが得られるときには，これを繰り返し測定すると，そのデータのバラツキから推定される。

ただし，破壊試験では，測定するごとにサンプルが異なるので，サンプリング誤差と測定誤差を分離できない。このようなときは，できるだけ均一な資料を作成して，そのデータのバラツキによって測定の精度 σ_M を定義する。

1.5 誤差の加法性

いま，ある母集団からの観測特性値が互いに独立な正規分布 $N(\mu,\ \sigma^2)$ にしたがうとする。この母集団から抽出される標本を代表する確率変数 X の構造式は次式で示される。

$$X=\mu+E_S+E_M \tag{8.3}$$

ここに，E_S はサンプリング誤差，E_M は測定誤差を表し，両者は互いに独立とする。さて，このような関係式では，母集団から抽出される1個のデータの分散 σ^2 は，E_S と E_M が互いに独立であることから

$$\sigma^2=\sigma_S{}^2+\sigma_M{}^2$$

となる。また，$X_1, X_2, \cdots, X_n$ を互いに独立とし，$a_1, a_2, \cdots, a_n$ を定数とするとき，ある1つの確率変数 Y が，

$$Y=a_1X_1+a_2X_2+\cdots+a_nX_n$$

のように定式化できれば，Y の分散は

$$V(Y)=a_1{}^2V(X_1)+a_2{}^2V(X_2)+\cdots+a_n{}^2V(X_n)$$

である。したがって，n 個の測定値の平均 $\bar{x}$ の分散は

$$V(\bar{x})=V(x)/n=\sigma^2/n=(\sigma_S{}^2+\sigma_M{}^2)/n$$

となる。

1.6 単純ランダムサンプリング法（単純無作為抽出法）

（1） 単純ランダムサンプリング

大きさ N の母集団から，n 個の標本を抽出する際に，この母集団

から標本を無作為 (random) に抽出し，母集団の平均値を推定したり，この推定精度を算出したりする方法をランダムサンプリング法 (random sampling) という。この際，N 個から n 個を抽出する方法は ${}_NC_n$ 通りあり，どの1通りも等確率で選出されるサンプリング法を単純ランダムサンプリング (simple random sampling) という。このとき，母集団と標本による平均値と分散は次式で示される。

$$\text{母平均}\ \mu=\sum_{i=1}^{N} X_i/N,\ \text{標本平均}\ \bar{x}=\sum_{i=1}^{n} x_i/n,$$

$$\text{母分散}\ \sigma^2=\sum_{i=1}^{N}(X_i-\mu)^2/n,\ \text{不偏分散}\ u^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1).$$

ここに，$X_1, \cdots, X_N$ は母集団における値，$x_1, \cdots, x_n$ は標本値を意味する。この方法が適用されるのは，母集団について（i）統計的にも技術的にも全く予備知識を有しないとき，（ii）後記の層別，集落，副次などの方法によって母集団に詳細な条件区分を付けた後で，最終的にその抽出範囲の枠内 (frame) で無作為化を必要とするときに基礎的な方法となる。したがって，母集団に関し何らかの予備知識があれば，その知識をできる限り利用し，経済的に推定精度の高いサンプリング法を採用するのがよい。

［**例題 8.1**］ かます入りの澱粉 500 俵の中から，無作為に 16 俵を抜き出し重量をはかった。この 16 俵の平均は 27 kg，標準偏差は 3 kg であった。したがって，500 俵の総重量は $27\times500=13500$kg と推定された。また，16 俵の平均値の標準偏差は，各俵の標準偏差の $1/\sqrt{16}=1/4$ であるから，$\hat{\sigma}(\bar{x})=0.75$kg になる。したがって，500 俵全重量の推定値として 13.5ton，その標準偏差は 0.35ton である。

（2） 無作為化

ランダムサンプリングの原則は，母集団や層，群などの対象集団中のサンプル単位の1つ1つが全く平等の機会で抜き取られることで，無作為化の方法を具体的に定めなければ，得られたデータに偏りが入り，推測に致命的失敗が起きる。実際には，乱数表やサイコロ等を用いて，実際に適した方法を工夫して定める必要がある。

（3） ランダムサンプリングの基礎数理

（i） 無限母集団の場合

いま，平均μ，分散σ^2の無限母集団から無作為標本$x_1, x_2, \cdots, x_n$を得たとする。このとき，標本平均$\bar{x}$は母数μの1つの不偏推定量$E(\bar{x})=\mu$で，標本平均の分散は$V(\bar{x})=\sigma^2/n$である。不偏分散を

$$u^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1)$$

とすると，この統計量は母分散σ^2の不偏推定量，すなわち$E(u^2)=\sigma^2$である。

（ii） 有限母集団の場合

有限母集団の平均と分散を次のように定義する。

$$\mu=E(x)=\sum_{i=1}^{N}x_i/N \tag{8.4}$$

$$\sigma^2=E\{(x-\mu)^2\}=\sum_{i=1}^{N}(x_i-\mu)^2/N \tag{8.5}$$

また，同一母集団から元に戻さず標本x_iとx_kを無作為に抽出するとき，これらの共分散は次のようになる。

$$E\{(x_i-\mu)(x_k-\mu)\}=\sum_{i=1}^{N}\sum_{k=1}^{N}(x_i-\mu)(x_k-\mu)/N(N-1), \quad i\neq k \tag{8.6}$$

ここに，

$$\sum_{i=1}^{N}(X_i-\mu)=(X_1-\mu)+(X_2-\mu)+\cdots+(X_N-\mu)=0 \tag{8.7}$$

であるから，この式を2乗して

$$\sum_{i=1}^{N}(X_i-\mu)^2+\sum_{i=1}^{N}\sum_{k=1}^{N}(X_i-\mu)(X_k-\mu)=0, \quad i\neq k \tag{8.8}$$

を得る。この式に式（8.4）を代入すると，

$$N\cdot\sigma^2+\sum_{i=1}^{N}\sum_{k=1}^{N}(X_i-\mu)(X_k-\mu)=0, \quad i\neq k \tag{8.9}$$

となる。この関係を（8.6）に代入して

$$E\{(X_i-\mu)(X_k-\mu)\}=-\sigma^2/(N-1) \tag{8.10}$$

を得る。

いま，同一有限母集団から元に戻さずn個の標本を抽出したとき，その標本平均$\bar{x}$の期待値と分散を考える。標本平均は無限母集団の場合と同様に

$$E(\bar{x})=\mu \tag{8.11}$$

が成立する。すなわち，標本平均は母平均μの不偏推定量で，標本

分散は

$$\begin{aligned}\sigma^2(\bar{x}) &= E\{(\bar{x}-\mu)^2\} \\ &= E[\{(x_1-\mu)+(x_2-\mu)+\cdots+(x_n-\mu)\}^2]/n^2 \\ &= [\sum_{i=1}^{n} E\{(x_1-\mu)^2\} + \sum_{i=1}^{n}\sum_{\substack{k=1 \\ i\neq k}}^{n} E\{(x_i-\mu)(x_k-\mu)\}]/n^2 \\ &= n\sigma^2\{1-(n-1)/(N-1)\}/n^2 \\ &= (\sigma^2/n)\{1-(n-1)/(N-1)\} \qquad (8.12)\end{aligned}$$

すなわち，有限母集団からの標本抽出の場合には，無限母集団の場合の $\sigma^2(\bar{x})=\sigma^2/n$ に更に係数 $\{1-(n-1)/(N-1)\}<1$ が付き，これを有限修正（finite multiplier）という。有限母集団からの標本抽出は，無限母集団からの標本抽出に比し，標本平均の分散がこの修正分だけ小さくなる。また，N が大きくなるほど，無限母集団の場合の σ^2/n に近付く。

[例題 8.2] 大型錠剤の 1,000 錠ずつの小分けの計数検査に，ある精度の重量管理をしたい。いま，1 ロットの大きさを 16 万錠とし，実験的にそのロットから 100 錠を無作為抽出して秤量し，100 回反復して次表 8－1 を得た。この錠剤 1,000 錠の重量とその重量変動を求め，重量によって錠数が管理できる精度を考察する。

[解] 表 8－1 から，$x=619.26$mg，標本分散 $u^2=50.84$，標本標準偏差 $u=7.13$mg を得る。u^2 を母分散と見なし，有限修正を無視して標本平均の標準偏差 $\hat{\sigma}(\bar{x})=0.713$mg が得られる。したがって，1,000 錠の平均重量は $y=619.26$g で，1,000 錠の重量の標準偏差は $1000\,\hat{\sigma}(\bar{x})=0.713\times1,000=0.713$g となる。すなわち，99% の信頼区間は，(617.12, 621.40) と得られ，この信頼区間の内側と外側とで計数作業の正否が検出される。

図 8－1 観測データ

1	627	617	616	621	619	615	627	615	625	625
2	620	630	619	618	620	581	613	613	628	612
3	610	625	622	615	623	615	613	620	614	625
4	620	619	618	624	616	626	613	616	617	630
5	624	635	617	635	612	620	614	617	615	629
6	619	614	621	619	615	619	622	625	615	621
7	621	623	623	621	620	616	617	612	609	611
8	615	627	616	619	620	615	623	614	618	627
9	610	617	614	616	614	619	618	615	627	627
10	632	628	625	619	642	615	617	623	611	614

2. 層別サンプリング法

2.1 層別サンプリング法とは

単純サンダムサンプリングでは，母集団が種々の要因を含み分散が大きいと，平均値の推定精度が悪くなり，サンプルを多くとる必要がある。ここでは母集団をいくつかの層に分けて，その各層からランダムにサンプリングする。このようなサンプリング法を層別サンプリング（stratified sampling）という。この方法では層別の仕方をできる限り層内分散 σ_w^2 を小さく均一にして，層間分散 σ_b^2 をできる限り大きく層別するとその精度は同じサンプル数のランダムサンプリング法に対して非常によくなる。

2.2 層別サンプリングの数理

いま，母集団の平均と分散を各 μ と σ^2，有限母集団の大きさを N とする。さらに，この母集団は L 個の層に層別されて，各層 i の大きさは N_i, $i=1, 2, \cdots, L$ とする。この母集団において各層 i から n_i 個のサンプルを無作為に抽出し，母集団全体から合計 n 個のサンプルを抽出するとする。このとき，$N=\sum_{i=1}^{L} N_i$, $n=\sum_{i=1}^{L} n_i$ である。また，各層における特性値を x_{ik}, $k=1, 2, \cdots, N_i$, とし，母平均と母分散を各 μ_i と σ_i^2, $N_i/N=P_i$ とおくと

$$\mu=\sum_{i=1}^{L} P_i \mu_i \tag{8.13}$$

$$\mu_i=\sum_{k=1}^{Ni} x_{ik}/N_i \tag{8.14}$$

$$\sigma^2=\sum_{i=1}^{L} P_i \sigma_i^2+\sum_{i=1}^{L} P_i(\mu-\mu_i)^2=\sum_{i=1}^{L}\sum_{k=1}^{N_i}(x_{ik}-\mu)^2/N \tag{8.15}$$

である。ここに $\sigma_i^2=\sum_{k=1}^{N_i}(x_{ik}-\mu_i)^2/N_i$ で上式の第1項を層内分散 σ_w^2, 第2項を層間分散 σ_b^2 という。すなわち，$\sigma_w^2=\sum_{i=1}^{L} P_i \sigma_i^2$, $\sigma_b^2=\sum_{i=1}^{L} P_i(\mu-\mu_i)^2$ であるから，

$$\sigma^2=\sigma_w^2+\sigma_b^2$$

を得る。

さて，層別サンプリング法による標本平均およびその分散は

$$\bar{x}=\sum_{i=1}^{L} P_i \bar{x}_i \tag{8.16}$$

で示される。ここに

$$\bar{x}_i=\sum_{k=1}^{n_i} x_{ik}/n_i \tag{8.17}$$

である。また

$$\sigma^2(\bar{x})=\sum_{i=1}^{L} (N_i-n_i)/(N_i-1)\}\ \{\sigma_i^2/n_i\}\ P_i^2 \tag{8.18}$$

$$=\sum_{i=1}^{L} \{(1/n_i)-(1/N_i)\}P_i^2\,\sigma_i^2, \quad (N_i \gg 1) \tag{8.19}$$

$$=\sum_{i=1}^{L} P_i^2\cdot\sigma_i^2/n_i=\{\sum_{i=1}^{L} N_i^2\sigma_i^2/n_i\}/N^2,$$

$$(n_i/N_i<0.1 \text{ のとき}) \tag{8.20}$$

である。母集団の総計を推定するには，上式で $\bar{x}_i$ を各 N_i 倍，および $\sigma^2(\bar{x})$ の式中で N_i^2 倍して加算すればよい。すなわち

$$Y=\sum_{i=1}^{L} N_i\bar{x}_i=\sum_{i=1}^{L} \{(N_i/n_i)(\sum_{k=1}^{n_i} x_{ik})\} \tag{8.21}$$

$$\sigma_Y{}^2=\sum_{i=1}^{L} N_i^2\{(N_i-n_i)/(N_i-1)\}\,(\sigma_i^2/n_i) \tag{8.22}$$

$$=\sum_{i=1}^{L} N_i^2\cdot\{(1/n_i)-(1/N_i)\}\sigma_i^2 \quad (N_i \gg 1 \text{ のとき}) \tag{8.23}$$

$$=\sum_{i=1}^{L} N_i^2\cdot\sigma_i^2/n_i, \quad (n_i/N_i<0.1 \text{ のとき}) \tag{8.24}$$

となる。この $\sigma_Y{}^2$ は層内分散 σ_i^2 だけで，近似的に決まり，層間分散 $\sigma_b{}^2$ には関係しない。すなわち，層間に大きな分散をもたせ，層内分散をできるだけ小さくなるように層別しサンプリングすると，相当に精度のよい推定ができる。この際，各層から抜き取るサンプルの大きさ n_i を定めるのは，調査に要する経費を少なくし，かつ精度をよくするために種々の工夫がなされる。例えば，比例サンプリング法，ネイマンサンプリング法，デミング (Deming) サンプリング法などがある。通常は $n_i/N_i<0.1$，かつ N_i＝一定の簡単な場合が多い。

2.3 比例サンプリング

層別サンプリングにおいて，各層の大きさ N_i に比例して，一定の比率 a で標本 n_i 個を各層から抽出する方法を比例サンプリング法

(size proportional sampling) という。この方法は各層内のバラツキ σ_i^2 が大約一定のときに役立ち，推定精度が最もよくなる。

$$n_1/N_1=n_2/N_2=\cdots=n_i/N_i=\cdots=n_L/N_L=a$$

$$n_i=nP_i=aN_i \tag{8.25}$$

とすると，式 (8.16) と (8.17) から

$$\bar{x}_i=\sum_{i=1}^{L}\sum_{k=1}^{n_i}x_{ik}/n \tag{8.26}$$

$$=\sum_{i=1}^{L}\sum_{k=1}^{n_i}x_{ik}/aN, \quad (N_i=N/L=\text{一定}) \tag{8.27}$$

ここに，$a=n/N$ である。層別サンプリング法では，平均値の推定は N_1 が一定でなくとも，得られたすべてデータを層に関係なく加えればよいので，計算は簡単になる。同様に，このサンプリング法ならすべてのサンプルを一緒にして測定したり，あるいは混合資料としてもよい。このとき，標本平均の分散は

$$\sigma^2(\bar{x})=\sum_{i=1}^{L}\{(N-n)/(NP_i-1)(\sigma_i^2/n)P_i^2\} \tag{8.28}$$

$$=\{(1/n)-(1/N)\}\sum_{i=1}^{L}P_i\sigma_i^2, \quad (N_i\gg 1) \tag{8.29}$$

$$=\{(1-a)/n\}\sum_{i=1}^{L}P_i\sigma_i^2 \tag{8.30}$$

$$=\sum_{i=1}^{L}P_i\sigma_i^2/n=\sigma_w^2/n \quad (n/N=a<0.1) \tag{8.31}$$

$$=\sum_{i=1}^{L}\sigma_i^2/nL \quad (n/N<0.1,\ N_i=N/L=\text{一定}) \tag{8.32}$$

となる。ただし，P_i が一定，すなわち N_1 が一定ならば，$\sigma_w^2=\sum_{i=1}^{L}\sigma_i^2/L$ である。工場では，式 (8.32) の条件が近似的に成立することが多い。このとき，X の平均値の推定精度を $\sigma^2(\bar{x},\ P)$ で，またランダムサンプリングの推定精度を $\sigma^2(\bar{x},\ R)$ で示すと，式 (8.15) から

$$\sigma^2(\bar{x},\ R)=\{\sum_{i=1}^{L}P_i\sigma_i^2+\sum_{i=1}^{L}P_i(\mu-\mu_i)^2\}/n=(\sigma_w^2+\sigma_b^2)/n$$

であり，式 (8.31) の結果とあわせて

$$\sigma^2(\bar{x},\ R)-\sigma^2(\bar{x},\ P)=\sigma_b^2/n\geqq 0$$

を得る。すなわち，実用的には層別効果がなく $\mu_i=\mu$，すなわち $\sigma_b^2=0$ のときにランダムサンプリングの精度に等しくなる。他方，層

別が非常にうまく行われ，各層内で均一と見なせるときは近似的に $\sigma_w^2=0$，すなわち $\sigma^2(\bar{x}, P)=0$ であり，サンプルの平均値はそのまま母平均になる。したがって，与えられた母集団での特性値 x_{ij} をできるだけ各層内で均一になるよう層別をすると σ_w^2 は小さくなる。すなわち，σ_b^2 は大きくなり精度のよい母平均の推定が行える。一般に，どのように層別しても $\sigma_w^2>0$ であるから，技術的または経済的に容易に層別が行えるなら，積極的に層別した方が有利である。

［**例題 8.3**］ 層別比例推定とその精度

大口取り扱いのアンプル業者は，その配下に幾軒かの下請け工場を有し，まとまった量の受注には，これらの工場からのアンプルをとり混ぜ，大ロットとして納入することが多い。いま，表8－2のように製作所ごとに層別された1納入ロットについて，そのアンプル口径の平均を推定する。このロットからサンプル数 $n=100$ とする層別比例サンプリングによって，ロット平均およびその推定精度について考察する。N_1, N_2, N_3, N_4 に対するサンプル数は各々40本，30本，20本，10本で，これらの測定値より表8－3の計算がなされた。

これらの表から，ロット中のアンプルの口径の平均値は，次のように推定される。

$$\bar{x}=(3.24\times0.4)+(2.71\times0.3)+(3.53\times0.2)+(3.18\times0.1)$$
$$=3.133(\mathrm{mm}).$$

また，この推定精度は式（8.30）を用いて

$$\sigma^2(\bar{x}, P)=\{(1-a)/n\}\sum_{i=1}^{L}P_i\sigma_i^2$$
$$=(0.4\times0.23^2+0.3\times0.18^2+0.2\times0.30^2$$
$$+0.1\times0.32^2)\times999/100000$$
$$=0.00999\times0.05912=0.000590$$

表8－2 ロットの層別

層	アンプル数
N_1	40000
N_2	30000
N_3	20000
N_4	10000

表8－3 計算結果

サンプル数		$\bar{x}_i$	$\hat{\sigma}_i$	$P_i=N_i/N$	a
		mm	mm		
n_1	40	3.24	0.23	0.4	
n_2	30	2.71	0.18	0.3	1/10000
n_3	20	3.53	0.30	0.2	
n_4	10	3.18	0.32	0.1	

を得，$\sigma(\bar{x}, P)=0.0243$(mm) となる。ここで，$\bar{x}$ の推定精度が極めて高いことに着目したい。すなわち，母平均の推定精度は理論的に層内分散 σ_w^2 のみに影響される。したがって，ロットの品質平均の推定精度をよくするには，そのロット間変動に大きく影響する要因（製作所）に関して層別し，その層内変動をできる限り小さくしておくのが効果的である。

2.4 ネイマンサンプリング

各層内のバラツキ σ_i の大きさに比例して，各層内からの標本の抽出率を決定する方法をネイマンサンプリング（Neymann sampling）という。

いま，層の数を L 個とすると，各層からの抽出率は

$$n_1/N_1 \sim \sigma_1,\ n_2/N_2 \sim \sigma_2,\ \cdots,\ n_L/N_L \sim \sigma_L$$

であり，また各層からの標本数は

$$n_i = n(N_i\sigma_i/\sum_{i=1}^{L} N_i\sigma_i),\quad i=1,\ 2,\ \cdots,\ L \tag{8.33}$$

となる。これは $\sum_{i=1}^{L} n_i = n$ の条件下で $\sigma(\bar{x}, P)$ を最小にする。

このとき母平均値の推定値と精度は次式で示される。

$$\bar{x} = \sum_{i=1}^{L} (P_i/n_i)\sum_{k=1}^{n_i} x_{ik} = \sum_{i=1}^{L} P_i\bar{x}_i \tag{8.34}$$

$$\begin{aligned}\sigma^2(\bar{x}, N) &= (\sum_{i=1}^{L} P_i\sigma_i)^2/n - \sum_{i=1}^{L} P_i\{(n_i-1)/n_i\}\sigma_i^2\{N_i/(N_i-1)\}/N \\ &= (\sum_{i=1}^{L} P_i\sigma_i)^2/n \end{aligned} \tag{8.35}$$

このサンプリング法は，層内分散の大きい層からは抜取り比を大きくする。$\sigma(\bar{x}, N)$ は $\sigma(\bar{x}, P)$ より小さくなって平均の推定精度はよいが，標本平均の計算には，各層の重みつき平均値を用いなければならない。ネイマンサンプリングで σ_i が一定のときは，通常の比例サンプリングになる。よく管理された工場では σ_i があまり違わないので，比例サンプリング法を用いることが多い。

［**例題 8.3**］ ネイマンサンプリングの実施例

同一ロットの顆粒を用いて 4 種の打錠機による錠剤重量の変動は，従来の生産実績から各々 σ_i $(i=1,\ 2,\ 3,\ 4)$ として知られてい

表8－4　生産量とサンプル数

機種	重量変動(mg)	当日の生産量
A	$\sigma_1=1.35$	$N_1=20$万錠
B	$\sigma_2=1.70$	$N_2=40$万錠
C	$\sigma_3=1.50$	$N_3=60$万錠
D	$\sigma_4=1.75$	$N_4=60$万錠

表8－5　層別の標本平均と標本標準偏差

機種	サンプル数	$\bar{x}_i$ (mg)	u_i (mg)
A	28	201.63	1.43
B	70	202.08	1.86
C	93	205.02	1.51
D	100	204.98	1.73

る（表8－4）。標本数を一定として，最も精度よくロットの平均重量を知る方法について考察する。このときの生産量を N_i（$i=1, 2, 3, 4$）として表8－4に与えている。ただし，サンプル数は300錠とする。

［解］表8－4のようにいくつかの層別ができ，しかも大約の σ_i（$i=1, 2, 3, 4$）が知れているとき，一定の標本数で最も推定精度のよい抽出法は，ネイマンサンプリング法である。したがって

$$n_i=n(N_i\sigma_i/\sum_{i=1}^{L}\sigma_i),\quad i=1,\ 2,\ 3,\ 4,\quad n=300$$

の式を用いて，各層からの抜取り個数は，次のように算出される。

$$n_1=300\times2\times1.35/29.0=810/29=27.9=28.$$

同様に，$n_2=70$, $n_3=93$, $n_4=109$ を得る。また，それぞれの P_i は

$$P_1=N_i/N=0.11,\ P_2=0.22,\ P_3=0.33,\ P_4=0.34$$

である。このように各層からランダムサンプリングし秤量した結果が表8－5である。

表8－5の結果から推定される平均重量は

$$\bar{x}=\sum_{i=1}^{L}P_ix_i=203.99(\mathrm{mg})$$

であり，またこの平均値の推定精度は次のように得られる。

$$\sigma^2(\bar{x},\ N)=(\sum_{i=1}^{L}P_i\sigma_i)^2/n=1.653/300=0.0091=(0.096)^2$$

2.5　デミングサンプリング

ネイマンサンプリングの考え方に，さらに各層からの抽出の経費をも考慮し，各抽出率，すなわち各層からの抽出個数 n_i を決定するのがデミングサンプリング（Deming sampling）である。いま，第 i 層での1個当りの測定経費を C_i とすると，総経費 $C=\sum_{i=1}^{L}C_in_i$ を一定におさえ，母平均の推定精度を最大にするには，次のように n_i

を決定すればよい。

$$n_1/N_1 \sim \sigma_1/C_1^{1/2},\ n_2/N_2 \sim \sigma_2/C_2^{1/2},\ \cdots,\ n_L/N_L \sim \sigma_L/C_L^{1/2}$$

または

$$n_i = (C/C_i)(N_i\sigma_i C_i^{1/2}/\sum_{i=1}^{L} N_i\sigma_i C_i^{1/2}),\quad i=1,\ 2,\ \cdots,\ L \quad (8.36)$$

である。このとき母平均の推定と推定精度は (8.16), (8.17) および (8.18) ～ (8.21) 式で示される。C_i が一定ならば，ネイマンサンプリングと同じになる。

9

実験データの採り方（II）

1. 集落サンプリング

集落サンプリング法（cluster sampling）は母集団をいくつかに群別し，その群をランダムにサンプリングし，その群内は全部調べる方法である。社会調査などでは町や村等を調査区域にとるので，この群を集落（cluster）と呼んでいる。この際には集落間の相違をできるだけ小さくなるように，すなわち σ_b をできるだけ小さくとり，したがって集落内 σ_w をできるだけ大きくとると有効である。極端に $\sigma_b=0$ ならば，1つの集落サンプリングをすればよいわけである。例えば，底が平らで壁の垂直なタンクに入った液体が，水平の層には全く均一で，深さとともに濃縮している場合には，表面から底まで垂直に集落をつくりサンプリングするのが典型的な集落サンプリングで，この集落を均一に混合して測定すれば，そのデータはそのままタンク全体の平均値を示す。

いま，M を全集落の数，m をサンプリングした集落の数，N_i を第 i 集落に含まれるサンプル単位数，N を全サンプル単位数，$n=\sum_{i=1}^{m} N_i$ をサンプリングした全サンプル単位数，また $Y_i=\sum_{k=1}^{N_i} x_{ik}$ とすると，母平均の推定には，

$$\bar{x}=(M/Nm)\sum_{i=1}^{m} Y_i=(M/Nm)\sum_{i=1}^{m}\sum_{k=1}^{N_i} x_{ik} \tag{9.1}$$

$$=\sum_{i=1}^{m}\sum_{k=1}^{N_i} x_{ik}/\bar{N}m=\sum_{i=1}^{m}\sum_{k=1}^{N_i} x_{ik}/n, \quad (N_i=\bar{N}=N/M) \tag{9.2}$$

$$=\sum_{i=1}^{m}\sum_{k=1}^{N_i} x_{ik}/\sum_{i=1}^{m} N_i=\sum_{i=1}^{m}\sum_{k=1}^{N_i} x_{ik}/n \tag{9.3}$$

を用いる。総計の推定には

$$Y=(M/m)\sum_{i=1}^{m}\sum_{k=1}^{N_i}x_{ik} \tag{9.4}$$

$$=(N/n)\sum_{i=1}^{m}\sum_{k=1}^{N_i}x_{ik}, \tag{9.5}$$

を用いる。この分散は，ランダムサンプリングの式で，集落の大きさを考慮に入れ，σ^2 のかわりに $\sigma_b{}^2$ を用いればよい。ここに

$$\sigma_b{}^2=\sum_{i=1}^{M}(\sum_{k=1}^{N_i}x_{ik}-U)^2/M,\quad U=\sum_{i=1}^{M}Y_i/M$$

である。このとき，

$$V(\bar{x})=(M/N)^2\{(M-m)/(M-1)\}\sigma_b{}^2/m \tag{9.6}$$

$$=(M/N)^2\sigma_b{}^2/m,\quad (m/M<0.1) \tag{9.7}$$

N_i が一定で，$N_i=\bar{N}$ であれば

$$V(\bar{x})=\{(M-m)/(M-1)\}\sigma_{b'}{}^2/m \tag{9.8}$$

$$\fallingdotseq \sigma_{b'}{}^2/m,\quad (m/M<0.1) \tag{9.9}$$

を得る。ただし

$$\sigma_{b'}{}^2=\sum_{i=1}^{M}(\mu_i-\mu)^2/M,\quad \mu_i=\sum_{k=1}^{N_i}x_{ik}/N_i,\quad \mu=\sum_{k=1}^{M}\mu_k/M$$

である。総計の推定精度は上の式を N^2 倍すればよい。

次に，ランダムサンプリング法と集落サンプリング法における平均値の推定精度を比較する。N_i が一定で$\bar{N}$のとき，集落サンプリング法では

$$V_C(\bar{x})=\{(M-m)/(M-1)\}\sigma_{b'}{}^2/m$$

$$\fallingdotseq\{(M-m)/M\}\sigma_{b'}{}^2/m, \tag{9.10}$$

であり，ランダムサンプリング法で同じサンプル数 $m\bar{N}$ をとれば

$$V_R(\bar{x})=\{(N-m\bar{N})/(N-1)\}\sigma^2/m\bar{N}$$

$$\fallingdotseq\{(N-m\bar{N})/N\}\sigma^2/m\bar{N}$$

$$=\{(M-m)/M\}\sigma^2/m\bar{N} \tag{9.11}$$

を得る。

このことから，相対効率（relative efficiency, R. E.）を次のように定義し，この値を 1 と比較し，大小によって，どれほど精度の上で優れているかが知れる。

$$R.E.=V_C(\bar{x})/V_R(\bar{x})\fallingdotseq\bar{N}\sigma_{b'}{}^2/\sigma^2=\{\sigma^2+\bar{N}\sigma_{b'}{}^2-(\sigma_{b'}{}^2+\sigma_w{}^2)\}/\sigma^2$$

$$=1+\{(\bar{N}-1)/\sigma^2\}\{\sigma_{b'}{}^2-\sigma_w{}^2/(N-1)\} \tag{9.12}$$

上式右辺の第 2 項について

$$\delta=\{\sigma_{b'}{}^2-\sigma_w{}^2/(\bar{N}-1)\}/\sigma^2 \tag{9.13}$$

とおくと，相対効率は次のように簡単に示される。

$$R.E.=1+(\overline{N}-1)\delta \tag{9.14}$$

ここで δ を集落の級内相関係数という。したがって，$\delta>0$ であれば集落サンプリング法の方の精度が低く，$\delta=0$ であれば両者の精度は等しく，$\delta<0$ であれば集落サンプリング法の方が精度は高い。また，$\sigma^2=\sigma_{b'}^2+\sigma_w^2$ の関係を式（9.12）に代入すると

$$\begin{aligned}\delta&=\{\sigma^2-\overline{N}\sigma_w^2/(\overline{N}-1)\}\sigma^2=1-\{\overline{N}/(\overline{N}-1)\}(\sigma_w^2/\sigma^2)^2\\&=\{-\sigma^2/(\overline{N}-1)+\overline{N}\sigma_{b'}^2/(\overline{N}-1)\}/\sigma^2\\&=-1/(\overline{N}-1)+\overline{N}(\sigma_{b'}^2/\sigma^2)^2/(\overline{N}-1)\end{aligned}$$

を得る。したがって，

$$-1/(\overline{N}-1)\leqq\delta\leqq1$$

である。ここで，$\delta=1$ は各集落内のすべての計量単位が等しく，集落内が完全に等質な $\sigma_w=0$ のときである。また，$\delta=-1/(\overline{N}-1)$ は $\sigma_{b'}=0$ のとき成立し，各集落内が最大限に異質なときを意味する。また，$\sigma_{b'}$ が一定なら各集落の大きさ $\overline{N}$ が大きいほど総平均の精度は悪くなる。実際には，集落が大きくなれば σ_b が小さくなりがちで，$\sigma_{b'}$ が 2 乗で影響するので集落化と層別を考えねばならない。

［**例題 9.1**］ 集落サンプリング法の実施例

1 台に 16 対の杵と臼をもつロータリー型打錠機について，錠剤の重量変動は 16 対の杵と臼の相対的位置にも生じる。いま，顆粒 1 ロットを何も調整せず打錠したところ，毎回の重量は表 9－1 のように測定された。ここで，毎回得られる錠剤を 1 つの集落と考え，各杵間変動を集落内変動とし，集落間変動は時間的に全くランダムに同一分布をすると考えるとき，でき上がった錠剤 1 ロットの平均重量と推定精度は集落サンプリングで推定される。ここに，この打錠機の性能を毎分 20 回転で，ロットの大きさを 19,200 錠とする。

表 9－1

測定	杵番号															
回数	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	154	153	155	154	156	152	154	155	155	153	155	156	155	153	157	156
2	153	154	156	155	154	154	155	155	156	154	154	155	155	154	156	154
3	154	152	157	156	155	153	154	155	154	155	154	155	156	155	156	155
4	155	152	155	155	153	154	155	154	155	155	154	154	154	155	156	156
5	153	153	154	154	154	155	155	155	156	154	155	155	155	154	155	155
6	154	151	155	156	155	153	155	156	155	155	155	156	154	154	156	156

単位：mg

上の記号に対応すると $M=1200$, $m=6$, $N=19,200$, $N_i=16$, $n=96$ で，錠剤の測定重量から 150mg 引いた平均値の推定値は

$$\bar{x}=(M/Nm)\sum_{i=1}^{m}\sum_{k=1}^{N_i}x_{ik}=\{1200/(19200\times6)\}\times443=4.61(\mathrm{mg})$$

である。したがって，錠剤の平均重量は 154.61mg と推定される。また，総重量の推定値は，$154.61\times19200=2968512$mg である。

精度については

$$V(\bar{x})=\{(M-m)/(M-1)\}(\sigma_b^2/m)(M/N)^2$$
$$=(1200/19200)^2\{(1200-6)/(1200-1)\}\sigma_b^2/6$$

であり，σ_b^2 を推定する必要がある。

$$W=\sum_{i=1}^{m}Y_i/m=73.8$$

とおくと，

$$\sigma_b^2=\sum_{i=1}^{m}(Y_i-W)^2/(m-1)=16.84/(6-1)=3.37$$

が得られる。このことから

$$V(\bar{x})=0.0022=(0.047)^2$$

を得る。したがって，このロットの平均重量は 154.61mg と推定され，この 95% の信頼区間は $154.61\pm(0.047\times2)$mg，すなわち (154.516, 154.704)mg である。

2. 多段サンプリング法

2.1 多段サンプリング法

多数の箱詰めの豆球，ボルトや瓶詰の化学薬品などでは，まず箱や瓶などをランダムに選定し，次に，その箱などの中から個々のサンプルを抽出する。サンプルを抜きとる際に，このように抽出単位を段階的に下げてサンプリングする方法を多段サンプリング (multi-stage sampling) または副次サンプリング (subsampling) という。このようなサンプリング法では，まず第 1 段階として箱，束，瓶などを単位としてサンプリングする。これを 1 次サンプリング単位 (primary sampling unit) という。その後，第 2 段階として得られた 1 次サンプリング単位からランダムサンプリングして 2 次サンプリング単位 (secondary sampling unit) を選定する。このような 2 段階でサンプリングする方法を 2 段サンプリング (two-stage sampling) という。例えば，大箱入りの消しゴム，タバコ，

石炭などをサンプリングする際に，まず第1段階として大箱をサンプリングし，さらにその大箱から小箱をサンプリングし，次に小箱の中から数個の消しゴムなどを抜きとる3段サンプリング法も考えられ，さらに4段，5段サンプリング法も同様に考えられる。前節までのランダムサンプリング法，層別サンプリング法，集落サンプリング法はいずれも1段サンプリング法（one-stage sampling）である。

2.2 多段サンプリングの数理

1段階目の抜取り比を a_1，第 i 段階の抜取り比を a_i とすると，K 段サンプリングでの全抜取り比 a は，各段階ごとに比例サンプリングすると $a=a_1\times a_2\times\cdots\times a_K$ となる。

いま，大きさ N の有限母集団が，M 個の1次単位からなり，各 N_i 個の2次単位を含んでいるとする（$i=1, 2, \cdots, M$）。第 i 番目の1次単位の中の第 k 番目の2次単位の特性値を x_{ik} とし，その1次単位での平均を μ_i とすると，母集団については

$$\mu=\sum_{i=1}^{M}\sum_{k=1}^{N_i}x_{ik}/N=\sum_{i=1}^{M}N_i\mu_i/N=\sum_{i=1}^{M}P_i\mu_i \tag{9.15}$$

$$\sigma^2=\sum_{i=1}^{M}\sum_{k=1}^{N_i}(x_{ik}-\mu)^2/N=\sum_{i=1}^{M}\sum_{k=1}^{N_i}(x_{ik}-\mu_i)^2/N+\sum_{i=1}^{M}N_i(\mu_i-\mu)^2/N$$

$$=\sum_{i=1}^{M}P_i\sigma_i^2+\sum_{i=1}^{M}P_i(\mu_i-\mu)^2=\sigma_w^2+\sigma_b^2 \tag{9.16}$$

が成立する。ここに，

$$P_i=N_i/N,\quad \sigma_i^2=\sum_{k=1}^{N_i}(x_{ik}-\mu_i)^2/N_i$$

である。この母集団から m 個の1次単位をとり，その各1次単位から n_i 個のサンプルをとると，総計の推定は次式で行われる。

$$Y=(M/m)\sum_{i=1}^{m}(N_i/n_i)\sum_{k=1}^{n_i}x_{ik} \tag{9.17}$$

いま，1次抜取り比を $a_1=m/M$，2次抜取り比を $a_{2i}=n_i/N_i$ とすると，上式は

$$Y=(1/a_1)\sum_{i=1}^{m}(1/a_{2i})\sum_{k=1}^{n_i}x_{ik} \tag{9.18}$$

となる。第2段階サンプリングで，$a_{21}=a_{22}=\cdots=a_{2m}(=a_2)$ なら

$$Y=\sum_{i=1}^{m}\sum_{k=1}^{n_i}x_{ik}/a_1a_2 \tag{9.19}$$

$$=(N/n)\sum_{i=1}^{m}\sum_{k=1}^{n_i}x_{ik} \quad (N_i=\text{一定}) \tag{9.20}$$

である。また，サンプルからの母平均の推定は次式で計算される。

$$\bar{x}=(M/Nm)\sum_{i=1}^{m}(N_i/n_i)\sum_{k=1}^{N_i}x_{ik} \tag{9.21}$$

$$=\sum_{i=1}^{m}\sum_{k=1}^{n_i}x_{ik}/n \quad (a_2 \text{ と } N_i \text{ が一定}). \tag{9.22}$$

実際には，a_2 を一定にすることが多いので，式 (9.22) が適用できる。この分散は次のように得られる。

$$V(\bar{x})=(1/N^2)\{M^2[(M-m)/(M-1)]\sigma_Y^2/m$$
$$+(M/m)\sum_{i=1}^{M}N_i^2[(N_i-n_i)/(N_i-1)]\sigma_i^2/n_i\} \tag{9.23}$$

$$\fallingdotseq(1/N^2)\left\{m(1-a_1)\sigma_Y^2/a_1^2+(1-a_2)\sum_{i=1}^{M}n_i\sigma_i^2/a_1a_2^2\right\}$$
$$(M\gg1,\ N_i\gg1), \tag{9.24}$$

$$\fallingdotseq(1-a_1)\sigma_b^2/m+(1-a_2)\sigma_w^2/m\bar{n}$$
$$(N_i=\bar{N},\ \text{すなわち } n_i=\bar{n} \text{ 一定}), \tag{9.25}$$

$$\fallingdotseq\sigma_b^2/m+\sigma_w^2/m\bar{n}=\sigma_b^2/m+\sigma_w^2/n$$
$$(m/M<0.1,\ n_i/N_i=a_2<0.1), \tag{9.26}$$

ここに，

$$\sigma_Y^2=\sum_{i=1}^{M}(Y_i-\bar{Y})^2/N$$

$$Y_i=\sum_{k=1}^{N_i}x_{ik},\quad \bar{Y}=\sum_{i=1}^{M}Y_i/M$$

$$\sigma_i^2=\sum_{k=1}^{N_i}(x_{ik}-\mu_i)^2/N_i$$

とある。とくに，N_i が一定のときは

$$\sigma_b^2=\sum_{i=1}^{M}(\mu_i-\mu)^2/M,\quad \sigma_w^2=\sum_{i=1}^{M}\sum_{k=1}^{N_i}(x_{ik}-\mu_i)^2/N$$

である。また，総計の推定量 Y の分散はこれまでと同様に $\bar{x}$ の分散を N^2 倍すればよい。

［**例題 9.2**］ ある副原料は，木箱の中に密封された紙函 6 個が入って梱包され，通常 3 ないし 10 の木箱が 1 度に入荷される。そして 3 箱に 1 箱をランダムに選び，各箱から 2 函をとり，各函の上部と下

部から2検体をとる。この2検体を混合し含水率を測定して，入荷ロットの平均品質が推定される。いま，14箱入荷して5箱をランダムに選び，各箱より紙函2函を測定して次の含水率を得た。このとき，木箱を1次サンプリング単位，紙函を2次サンプリング単位とすると，これは2段サンプリングと考えられる。この場合のロットの平均品質とその精度を計算することにする。

表9-2　測定された含水率

抽出箱No.	4	7	9	10	13
函1	3.1	2.7	4.6	3.5	3.4
函2	2.9	3.0	3.8	3.6	3.2

上の記号に対応させると，$M=14$, $N_i=\overline{N}=6$, $N=84$, $m=5$, $n_i=\overline{n}=2$, $n=m\overline{n}=10$ で，また $a_1=m/M=5/14$, $a_2=n/N=1/3$ である。そこで，母平均の推定は次式で与えられる。

$$\overline{x}=(M/Nm)\sum_{i=1}^{m}(N_i/n_i)\sum_{k=1}^{n_i}x_{ik}=\{14/(84\times5)\}(6/2)\times33.8=3.38$$

この推定精度は次式で得られる。

$$\hat{\sigma}^2(\overline{x})=(1-a_1)\hat{\sigma}_b^2/m+(1-a_2)\hat{\sigma}_w^2/m\overline{n}$$
$$=(1-5/14)\hat{\sigma}_b^2/5+(1-1/3)\hat{\sigma}_w^2/10$$

ここに，

$$\hat{\sigma}_w^2=\{(\overline{N}-1)/\overline{N}\}\{1/m(\overline{n}-1)\sum_{i=1}^{m}\sum_{k=1}^{n_i}(x_{ik}-\overline{x}_i)^2\}$$

かつ

$$\hat{\sigma}_b^2=\{(M-1)/M\}\{\sum_{i=1}^{m}(\overline{x}_i-\overline{x})^2/(m-1)-[(\overline{N}-\overline{n})/\overline{N}]\hat{\sigma}_w^2/\overline{n}\}$$

$$=\sum_{i=1}^{m}(\overline{x}_i-\overline{x})^2/(m-1)-\hat{\sigma}_w^2/\overline{n}\quad(M\gg1,\ a_2<1/10\text{のとき})$$

であるから，

$$\hat{\sigma}_w^2=(1/5)\times0.2050=0.0410$$

を得る。同様にして $\hat{\sigma}_b^2=0.2336$ が得られる。したがって，式 (9.24) を用いて

$$\hat{\sigma}^2(\overline{x})=(9/14)(0.2336/5)+(2/3)(0.0410/10)=0.03276$$
$$\hat{\sigma}(\overline{x})=0.181$$

を得る。すなわち，ロットの平均含水率の推定値は3.38%で，この95%の信頼区間は（3.01%，3.75%）である。