新しいサンプリング調査法のデータ処理システムの開発

ODC 612

新しいサンプリング調査法のデータ処理システムの開発

Ⅰ 試験担当者

経営部経営第2科測定研究室長　粟屋仁志
主任研究室　西川 英
（現調査部海外林業調査科技術情報室長）
室員　天野正博
北海道支場経営部経営研究室長　真辺 昭

Ⅱ 試験目的

最近の森林資源の有効利用や公益的機能に対する認識の高まりにより，施業方法が多様化し，そのため複層構造の林分が次第に増加する傾向にあり，またわが国の森林蓄積の相当部分を占める天然林の多くは複層構造をなしている。 森林蓄積算定の基礎となる材積表は適用地域の平均的な値を示すもので個々の林分の正確な材積を求めるには標本木を伐倒して求めた実材積を用いる必要がある。

最近立木状態のままで比較的簡単かつ正確に上部直径を測定できる測樹器（デンドロメーターやテレレラスコープなど）が考案され，さらにこれらの器具を用いた新しい調査法として Grosenbaugh によって3－P（Probabilty Proportional to Prediction）サンプリングが提案されている。この方法は複層構造の蓄積推定に適しており，かつ正確な林分材積あるいは利用材積の推定が可能と考えられるが，わが国の林分に適した場合どのような問題点が生ずるか不明であった。この報告は現地適用試験によりこの問題点を解明し，従来の調査法に代る効率的な調査方法のシステムとデータ処理方法を開発することを目的と行なったものである。

Gro senbaugh が3－Pサンプリングの方法を提案して以来，その理論構成についていくつかの説が発表されているので，この報告では理論的面からの考案も含めている。

なお，この報告は主として真辺昭と天野正博が取りまとめたものである。

Ⅲ 試験の経過と得られた成果

1. はじめに

林分の蓄積調査では一般に全数調査は非能率的であり，また経済的にも引き合わない場合が多いので，標準地調査あるいは標本地調査（サンプリング）が用いられている。サンプリングには各種の方法が考えられ，例えば対象地域を固定半径プロット調査あるいはポイントサンプリンによって調べたり，一定本数ごとに標本木を抽出して必要な項目を測定し，最終的に材積表を用いて林分材積を求める方法などがある。

しかし調査木の大きさ，または価格の単木間の変動が大きいと，通常用いられている等確率抽出の方法は効率が悪い。それは総材積または総価格に対する寄与の少ない小径木を大径木と同じウェイトで調べることになるからである。また天然林や択伐林，間伐材ななどのように林分材積や利用材積を推定するための表が整備されていない森林では，これらの表を作成するだけでも大変な仕事量となる。

こうしたことからGro senbaughはもし推定しようとする変量と相関の高い（理想的には比例関係の成立する）補助変量が目測あるいは簡単な測定で単木ごとに求められるなら，この補助変数の大きさに比例する確率で標本木を抽出することによって調査効率を高めようとした。これが３－Ｐサンプリング（以下３ＰＳと略す）である。また３ＰＳは標本木の材積を表を用いることなく，デンドロメータのような高精度の測定器によって直接推定し，さらに得られたデンドロメータの複雑な測定値を電算機処理することによって，計算が自動的にかつ効率的に実行できるようになった。

ここではGro senbaugh・Loetcheらにしたがって，まず調査方法の概要を述べ，続いて我国での応用可能性を検討するとともに，我国の実情に合った形での３ＰＳシステムを提案する。

2. ３ＰＳの方法

Ｎ本からなる林木の母集団を考え，i 番目の木の推定しようとする変量の値を yi（$i=1$，……，Ｎ）とする。サンプリングの目的は，yi の母集団総合計 $\sum_{i=1}^{N} yi$ を求めることである。ここで yi に対応して i 番目の木に指定する補助変量の値を xi とする。xi は大きさの変量（size variable）と呼ばれ，yi と比例するかまたは高い相関をもっていることが望ましい。

母集団における xi の最大値 $xmax$ より大きい任意の乱数Ｌをとり，１から（Ｌ－１）までの範囲の乱数をＮ個用意する。

３ＰＳの抽出手続は，Ｎ本の木を任意の順に巡回して xi を求め，これを用意じた乱数の値と比較する。もし乱数が xi に等しいか，それより小さければこの木を標準木に選んで yi を実

測する。そのほかのときは次の木に移る。乱数は木が変るたびに新しいものを使用し，xi の値は記録しておく。

この方法によって n 本の木が選ばれたとすると，これは以下に示すLahiriの方法で大きさ１の標本を非復元で n 個抽出したのと同じになる。Lahiriの方法は，

1）１～Ｎの範囲から乱数を１つ選び，対応する木の番号 i をきめる。

2）１～Ｌの範囲で第２の乱数をとり，この値が上できめた i 番目の木の xi より小さいか等しいときその木の標本とする。

3）そのほかの時は i 番目の木を除いた残りの木で，大きさ１の標本が得られるまで以上の手続きを繰り返す。

この方法で大きさ変量 xi に比例した確率で標本木が選ばれることは次のようにしてわかる。

ステップ1）ではどの木の抽出確率も最初の抽出では $1/N$ に等しい。サンプリングでは非復元だから，特定の木が２回目の抽出で出てくる条件付確率は $1/(N-1)$ に等しい。したがって i 番目の木が２回目の抽出でえられる確率は，$\overline{\sum_{i \neq j}}\ 1/N \cdot \frac{1}{N-1} = 1/N$ である。同様に３回目の抽出で i 番目の木が選ばれる確率も $1/N$ になる。

次に1)の操作で i 番目の木が取り出されたとき，その xi が与えられた乱数より大きくなる確率は xi/L に等しい。なぜなら，同等な出現確率をもつ１～Ｌまでの乱数のうち，xi に等しいかそれより小さいものの割合は xi/L だからである。したがって1)，2）を通じて，特定の木 i が yi の測定のために選び出される確率は xi/NL となる。また１つの試行（２つの乱数をとり出して参照する手続き）で抽出がおこなわれないで終る確率は

$$q = \sum_{i=1}^{N} (1/N)(1 - xi/L) = 1 - \Sigma\, xi/NL$$

だから，$pi = xi/NL$ とおくと，大きさ１の標本の抽出が i 番目の木の選出によって終結する機会は

$$pi + q\,pi + q^2 pi + \cdots\cdots + q^{N-1}\, pi \propto pi/(1-q) = xi/\overline{\Sigma} xi$$

すなわち，母集団の各立木はそれぞれの xi に正確に比例する確率で抽出される。

この方法母集団の各立木が全部調べられるまで続けると xi に正確に比例する確率で，独立な大大きさ１の標本があるランダムな個数 n 個だけ抽出される。（非復元で）

不偏な推定量

３ＰＳによる林分材積を誘導するため，確率変数 αi をつぎのように定義する。

$$\alpha i = \begin{cases} 1 : \text{確率 } xi/L \text{ で標本が抽出されたとき。} \\ 0 : \text{標本が抽出されなかったとき。} \end{cases}$$

i番目の木にLahiriの方法が適用されたとすると

$$E\left(\alpha_i \frac{y_i}{x_i}\right) = \frac{y_i}{L} \quad \text{ゆえに}$$

$$y_i = LE\left(\alpha_i y_i / x_i\right)$$

すべてのiについてたしあげると，各サンプリングは独立に行われるので

$$\sum_{i=1}^{N} y_i = \sum_{i=1}^{N} LE\left(\alpha_i y_i / x_i\right)$$

$$= L \sum_{i=1}^{n} \left(y_i / x_i\right) = Y_u \qquad (1)$$

となる。（Grosonbaugh はこれを３PFIRSTと呼んでいる。）なお３PSはY_uを調整しない推定量（unadjusted estimator）ともいう。

期待標本サイズと分散

３PSの標本サイズnは確率変数であって，一意的には定まらない。推定精度と調査費用の見積りにはnが直接の関係をもつので，nの期待値とその分散を知っておかなければならない。

さきのα_iを使うと

$$n = \sum_{i=1}^{N} \alpha_i$$

また$E(\alpha_i) = x_i / L$ より

$$E(n) = E\left(\sum_{i=1}^{N} \alpha_i\right) = \sum_{i=1}^{N} \left[E(\alpha_i)\right] \qquad (2)$$

$$= \sum_{i=1}^{N} x_i / L$$

この結果から標本の大きさの期待値はLの選び方と，母集団における大きさ変数の総計$\sum_{i=1}^{N} x_i$によって決定されることがわかる。

つぎに標本サイズnの分散は，α_i間の独立性によって，α_1，α_2……α_Nの分散の和に等しくなる。すなわち，$E(\alpha_i^2) = E(\alpha_i) = x_i / L$ に注目すると

$$V(n) = V\left(\sum_{i=1}^{N} \alpha_i^2\right) = V\left(\sum_{i=1}^{N} \alpha_i\right)$$

$$= \sum_{i=1}^{N} \left[E(\alpha_i^2) - \left(E(\alpha_i)\right)^2\right]$$

$$= \sum_{i=1}^{N} x_i / L - \sum_{i=1}^{N} x_i^2 / L^2 \qquad (3)$$

(3)式を変形すると

$$V(n) = \sum_{i=1}^{N} x_i / L - \left\{\left(\sum_{i=1}^{N} x_i\right)^2 \sum_{i=1}^{N} x_i\right\} / \left\{L^2 \left(\sum_{i=1}^{N} x_i\right)^2\right\}$$

ここで$ne = E(n)$とおいて

$$6x^2 = \left(\sum_{i=1}^{N} x_i^2 - N\bar{x}^2\right) / N$$

を用いると

$$V(n) = ne - ne^2 / N - ne^2\, 6x^2 / N\bar{x}^2$$

$$= ne - ne^2 \left(1 + 6x^2 / \bar{x}^2\right) / N \qquad (4)$$

$$= ne - ne^2 \left\{1 + \left(C \cdot V \cdot (x)\right)^2\right\} / N \qquad (5)$$

となって，期待標本サイズの分散式は期待標本サイズそのもので表わせる。Grosenbaughは(5)を３PFIFTHとよんでいる。

偏りのある推定量

３PSの設計では必要とする推定精度に合わせて期待標本サイズneをきめ，これから

$$L = \sum_{i=1}^{N} x_i / ne$$

としてLを求める。したがって(1)は

$$Y_u = \sum_{i=1}^{n} y_i L / x_i = \left(\sum_{i=1}^{n} y_i / x_i\right)\left(\sum_{i=1}^{N} x_i\right) / ne \qquad (6)$$

と書ける。

偏りのある推定量は，これに現実にえられた標本サイズnと期待標本サイズneの比をかけて補正したものである。すなわち

$$Y_a = (Y_u)\, ne / n = \left(\sum_{i=1}^{n} y_i / x_i\right)\left(\sum_{i=1}^{N} x_i\right) / n \qquad (7)$$

(6)と(7)の違いは分母がneからnに変わっていることであるが，その他に補助変数の母集団総計$\sum_{i=1}^{N} x_i$が必要なことに注意しなければならない。このように，期待値の代りに現実のnを使い，また標本から得られる追加情報$\sum_{i=1}^{N} x_i$を利用することで精度が向上するのである。Gro senbaugh はY_uを調整された推定量（adjusted estimator）とよび３PS EVENTHと表わしている。次に偏りの大きさであるが，SchreuderらはY_uの期待性を

$$E(Y_a) = \sum_{i=1}^{N} y_i\, p_i\, X / L \qquad (8)$$

と表わしている。ここで $X = \sum_{i=1}^{N} x_i$，$p_i = \sum_{j=1}^{N} (p_{ij}) / j$ で p_{ij} は大きさjの標本に立木iが含まれる確率である。したがってY_aの偏りは

$$E(Y_a) - \sum_{i=1}^{N} y_i = \frac{X}{L} \sum_{i=1}^{N} Y_i \left(p_i - L/X\right) \qquad (9)$$

となる。p_iは定数とはなりえないから，明らかにY_aは偏りをもつ。しかしp_iの計算には，可能なすべての組合せの同時出現確率を求めなければならないので，(9)式による偏りの評価は事実上不可能である。

Y_uの分散と分散推定量

Y_uの分散はY_uが不偏なことから，$Y = \sum_{i=1}^{N} y_i$ とおいて

$$V(Y_u) = E(Y_u)^2 - (Y)^2$$

期待値の計算で用いた確率変数α_iを使い，$\pi_i = x_i / L$ とすると

$$V(Y_u) = E\left(\sum_{i=1}^{N} \alpha_i\, y_i / \pi_i\right)^2 - Y^2$$

$$= E\left(\sum_{i=1}^{N} \alpha_i^2\, y_i^2 / \pi_i^2 + \sum_{i \neq j}^{N} \alpha_i\, \alpha_j\, y_i\, y_j / \pi_i\, \pi_j\right) - Y^2$$

ここで$E(\alpha_i^2) = E(\alpha_i)$であり，かつ$\alpha_i$と$\alpha_j$は独立だから

$$E(\alpha_i \alpha_j) = E(\alpha_i)E(\alpha_j)$$

したがって

$$V(Y_u) = \sum_{i=1}^{N} \pi_i y_i^2/\pi_i^2 + \sum_{i=1}^{N} \pi_i \pi_j y_i y_j/\pi_i \pi_j - Y^2$$

$$= \sum_{i=1}^{N} y_i^2/\pi_i + \sum_{i=1}^{N} y_i y_j - \left(\sum_{i=1}^{N} y_i^2 + \sum_{i=1}^{N} y_i y_j\right)$$

$$= \sum_{i=1}^{N} y_i^2/\pi_i - \sum_{i=1}^{N} y_i^2$$

$$= \sum_{i=1}^{N} y_i^2 (1-\pi_i)/\pi_i$$

これに，$\pi_i = x_i/L$を代入すると

$$V(Y_u) = \sum_{i=1}^{N} y_i^2 L/x_i - \sum_{i=1}^{N} y_i^2 \qquad (10)$$

となる。

$V(Y_u)$の不偏な標本推定量は

$$v(Y_u) = \sum_{i=1}^{n} y_i^2 (1-\pi_i)/\pi_i^2$$

$$= \sum_{i=1}^{n} y_i^2 L^2/x_i^2 - \sum_{i=1}^{n} y_i^2 L/x_i \qquad (11)$$

である。 これが不偏なことは

$$E[v(Y_u)] = E\left[\sum_{i=1}^{N} \alpha_i y_i^2 (1-\pi_i)/\pi_i^2\right]$$

$$= \sum_{i=1}^{N} \left[E(\alpha_i) Y_i^2 (1-\pi_i)/\pi_i^2\right]$$

$$= \sum_{i=1}^{N} \pi_i y_i^2 (1-\pi_i)/\pi_i^2 = V(Y_u)$$

で証明される。

Y_aの分散と分散推定量

期待値のところで述べたのと同じ理由で，偏りのある推定量については正確な分散も求められない。このため幾つかの仮定のもとに近似式が導かれている。

確率変数の分散は条件付分散の期待値と条件期待値の分散の和に等しい（Raj）という定理からY_aの分散は

$$\gamma ar(Y_a) = \underset{n}{E}[\gamma ar(Y_a \mid n = n^*)] + \underset{n}{Var}[E(Y_a \mid n = n^*)] \qquad (12)$$

と表わすことができる。ここで$E_n[Var(Y_a \mid n = n^*)]$まず標本サイズを特定の値$n^*$に固定して$Y_a$の分散を計算し，そのあと$n$のとりうるすべての値（$n = 0, 1, 2, \cdots\cdots, N$）にわたってそれらの経過を平均するという演算を示している。

$\underset{n}{Var}[E(Y_a \mid n = n^*)]$も，標本サイズを$n^*$に固定して求めた推定値の，$n$の異なるすべ

ての標本間の分散である。

一定サイズn^*の繰返し標本で，立木iをその中に含む標本の出現確率p^*がx_iに比例すれば，Y_aのタイプの推定量は不偏である。すなわち$E(Y_a) = Y$。3PSでは，異なる大きさのすべての標本を考えたときの無条件確率はx_iに比例するが，一定サイズn^*の標本への（条件付）抽出確率は一般にx_iに比例しない。しかしここでは$p_i^* \propto x_i$と仮定して$Var(Y_a)$の近似を行なう。そうすると，$E(Y_a \mid n = n^*) = Y$はn^*をどのように選んでも一定になるから，（12式）の右辺の第2項はゼロになる。

次に⒀の右辺第1項の〔 〕内を評価するため，もう一つ復元抽出の仮定を設ける。

これによって

$$Var(Y_a) = \underset{n}{E}\left\{\left[\sum_{i=1}^{N} z_i (y_i/z_i - Y^2\right]/n\right\}$$

ここで$Z_i = x_i/\sum_{i=1}^{N} x_i$，$Y = \sum_{i=1}^{N} y_i$

$n = n_e + \varepsilon$として$1/n$をTaylor 展開すると

$$1/n = (1 - \varepsilon/n_e + \varepsilon^2/n_e^2)/n_e$$

したがって，$E(\varepsilon) = 0$により

$$Var(Y_a) \fallingdotseq V_1 = \left[\sum_{i=1}^{N} z_i (Y_i/z_i - Y)^2/n_e\right]\left[1 + V(n)/n_e^2\right] \qquad (13)$$

$V(n)$はnの分散で(4)または(5)で与えられる。

標本サイズの変動を無視してV_1の右辺の第2項の〔 〕の中を1とすると

$$V_2 = \left[\sum_{i=1}^{N} z_i (Y_i/z_i - Y)^2/n_e\right] \qquad (14)$$

が得られる。これは復元抽出，固定標本サイズで大きさに比例する抽出確率を用いる場合の分散で，Grosenbaughの3PEIGHTHはこのn_eをnでおきかえたものである。

同じ仮定のもとでの3PEIGHTHに対応する標本推定量は

$$v(Y_a) = v_1 = \sum_{i=1}^{N} (y_i X/x_i - Y_a)^2/n(n-1) \qquad (15)$$

でGrosenbaughはこれを3PNINTHと呼んでいる。Schreuderらは抽出率が高いとき，非復元抽出による精度の向上を説明するため，有限補正に類似した係数$(N-n_e)/N$を使って，近似式

$$V_3 = (V_1)(N-n_e)/N \qquad (16)$$

と対応する推定量

$$v_2 = (v_1)(N-n)/N \qquad (17)$$

を示している。

$Var(Y_a)$のもう1つの近似方法として，

$$Ya = (Yu)(ne/n)$$

から出発してTaylor展開を利用するものがある。この場合の母分散は，

$$Var(Ya) \doteqdot V_4 = \sum_{i=1}^{N} zi\,(yi/zi - Y)^2\,(1 - xi/L)/ne \qquad (18)$$

これを(14)と比較すると，$(1 - xi/L)$ はおおまかな有限補正に相当し，非復元抽出での利得を表わしている。V_4 に相当する推定量は，

$$V_3 = \sum_{i=1}^{n} (yi\,X/xi - Ya)(1 - xi/L)/n(n-1) \qquad (19)$$

である。

(18)式から，さらに n の変動性を考慮に入れた近似

$$V_5 = (V_4)\,[\,1 + V(n)/ne^2\,] \qquad (20)$$

を考えることもできる。

ここに示したYaの母分散とその推定量はいずれも標本サイズを固定した復元抽出の場合の式がもとになっている。しかし復元抽出の仮定は抽出率がごく小さいときしか許されず，また固定標本サイズのときの抽出確率がxiに比例するという仮定も３ＰＳでは成立しない。

さて以上述べたように３ＰＳには不偏な推定量と偏りのある推定量の２つがあって，前者については正確な分散公式とその不偏推定量が知られているが，偏りのある推定量ではそのような一般式を導くことはできない。一方推定精度の点では後者の推定量がはるかに優れているので実用上はいくつかの仮定のもとに分散の近似公式を求めている。次節ではモンテカルロ実験によってこれらの仮定の影響を検討するとともに，近似度を向上させるために用いた補正係数の有効性を確めることにする。

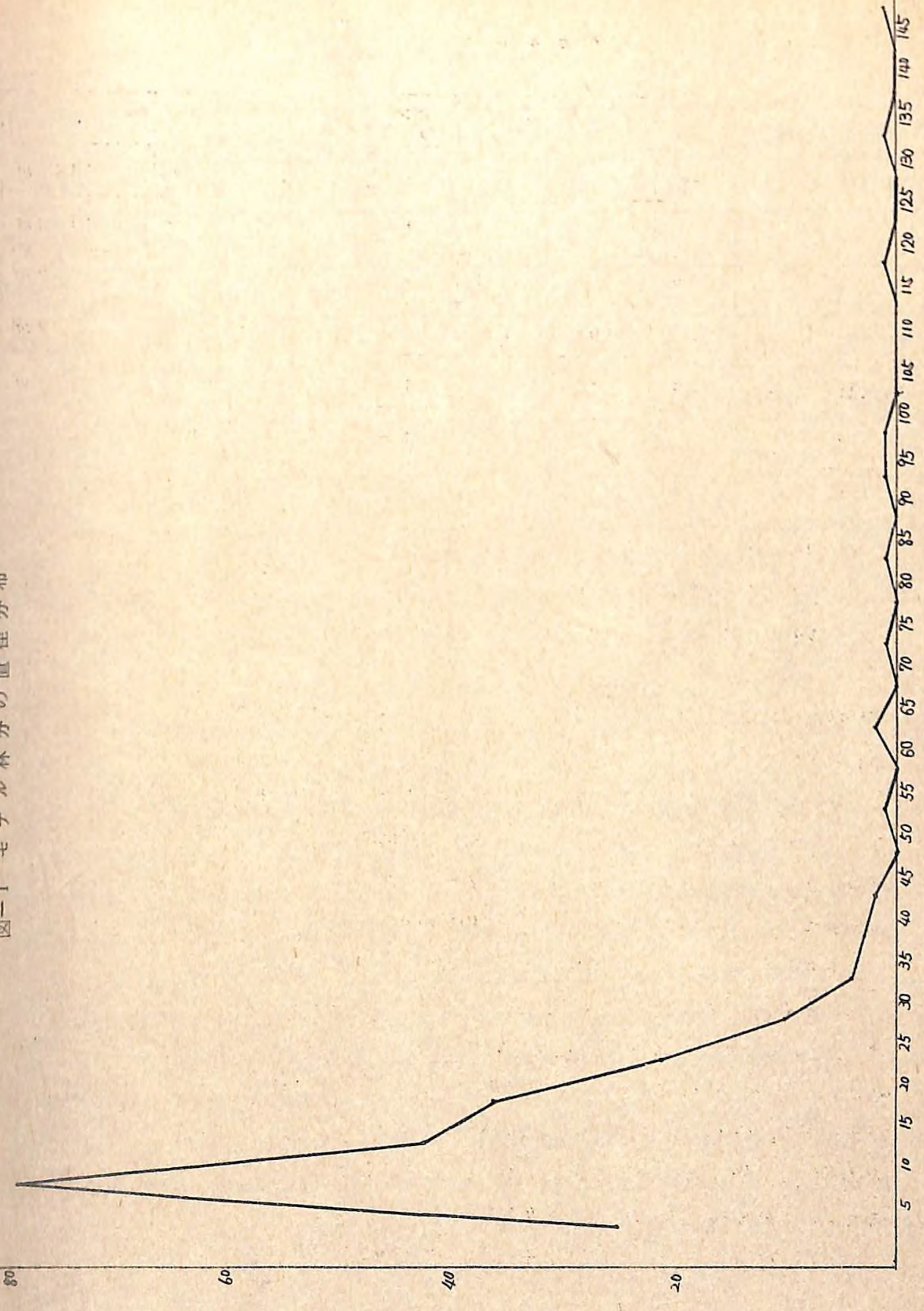
3. モンテカルロ実験による３ＰＳの検討

いままであげてきた３ＰＳの幾つかの問題点を検討するため，モンテカルロ実験を行なった。これは電算機内にモデル林分を設定し，その林分に対して繰り返し３ＰＳを試みることであり，

〔表－１〕　モデル林分の構造

樹種	ブナ，ナラ，ヒメシャラ	胸高直径の平均	16.91 cm
面積	0.22 ha	〃 範囲	4 cm 〜 147 cm
本数	235 本	林分材積の平均	0.49 m³
林分材積	115.5 m³	〃 範囲	0.03 m³ 〜 17.3 m³

※　林分材積は各単木を２ｍ毎に玉伐りすると想定し，その部位の直径をテレレラスコープで読み取り，末口二乗法によって計算した。



図—１　モデル林分の直径分布

〔表－2〕　モンテカルロ実験の結果

補助変数	ne	$\bar{n}$	$V(n)$	サンプル数の範囲	$\bar{Y}u$	$\bar{Y}a$	$\bar{Y}u$ の範囲	$\bar{Y}a$ の範囲	bu/Y	ba/Y	vu/Vu	va/Va	Vu/Va
目測材積	2.3	3.0	2.97	1 ～ 8	159.7 m³	114.2 m³	31.2 ～ 386.9 m³	71.7 ～ 163.9 m³	0.38	−0.01	1.21	0.89	26.41
〃	3.3	4.2	3.13	1 ～ 9	147.9 m³	115.2 m³	37.6 ～ 320.0 m³	77.3 ～ 169.4 m³	0.28	−0.01	1.37	1.15	19.38
〃	4.6	5.5	3.42	1 ～ 11	136.9 m³	115.2 m³	45.7 ～ 278.9 m³	75.0 ～ 174.3 m³	0.18	−0.003	1.17	1.33	9.22
胸高直径	9.0	10.8	5.06	6 ～ 16	148.1 m³	123.2 m³	23.7 ～ 293.1 m³	32.4 ～ 219.1 m³	0.28	0.07	1.31	1.90	2.37
胸高直径の二乗	3.3	4.0	3.49	1 ～ 9	141.7 m³	117.1 m³	15.8 ～ 333.1 m³	50.2 ～ 213.8 m³	0.23	0.01	1.03	1.03	7.0
〃	4.2	5.1	4.16	1 ～ 11	114.1 m³	115.5 m³	12.6 ～ 328.2 m³	48.7 ～ 179.3 m³	0.25	0.01	0.93	1.43	10.61
〃	5.6	6.1	3.37	2 ～ 12	127.5 m³	116.1 m³	42.0 ～ 248.7 m³	58.3 ～ 162.3 m³	0.10	0.005	1.07	1.31	5.94

※　ne＝期待サンプル数　$\bar{n}$＝実際に抽出されたサンプル数の平均　$V(n)$＝抽出されたサンプル数の分散　$\bar{Y}u$＝調整されない推定値 Yu の平均

$\bar{Y}a$＝調整された推定値 Ya の平均　Y＝林分材積の実測値　$ba = est.\{bias(Ya)\} = (\sum_{j=1}^{K} Yaj / K - Y)$　K＝モンテカルロ実験の繰り返し数（200）

$Va = est.\{Var(Ya)\} = \sum_{j=1}^{K}(Yaj - Ya)^2/(K-1)$　$va = \{E(v(Ya))\} = \sum_{j=1}^{K}〔v_j(Ya)〕/K$

実験的な平均値や分散を計算することによって３ＰＳの推定式の精度を確かめるとともに，通常行っているランダムサンプリングとの比較も行なってみた。

モデル林分は富士山麓の天然広葉樹林を対象とした。その林分構造を〔表－１〕に，直径分布を〔図－１〕に示す。このモデル林分に対し，補助変数や乱数の最大値Ｌをいろいろ変えて，各々２００回づつ３ＰＳを試みた結果が〔表－２〕である。なお補助変数として目測材積，胸高直径および胸高直径の二乗をとった。利用材積との相関係数はそれぞれ０.９８７，０.９０６，０.９５４とかなり高いので，どれも補助変数としての資格は充分あると考えられる。

この表で調整しない推定値 Yu と調整された推定値 Ya をみると，Ya の方が情報量が多いだけに推定精度もかなり高い。Vu/Va あるいは Yu と Ya の範囲をみると推定値のバラツキも Ya の方がはるかに小さい。このことから(7)式の ne/n という補正係数は十分その効果を発揮していると云える。分散の推定式については $v(Yu)$，$v(Ya)$ ともに近似の程度はそれほどよくない。なお vu/Vu，va/Va はそのほとんどが１より大きいが，これは期待サンプル数が確率変数であるため，３ＰＳの推定パラメータを θ としたとき

$$Var(\theta) = \underset{n}{E}\{Var(\theta/n)\} + Var\{E(\theta/n)\} \qquad (21)$$

となるので，どうしても３ＰＳの分散推定式はそのパラメータを過大評価してしまうことによる。

実測のために抽出されたサンプル数の $\bar{n}$ を期待サンプル数 ne と比較すると，実際に抽出されたサンプル数の方が，どの場合も推定式を上回る結果となった。これはモンテカルロ実験を行なう際に，サンプル数が０となったときは再度サンプリングをやり直したことも若干影響を与えていると思われるが，それ以外に，モデル林分の性質上こうなったのか，推定式自身に問題があるのか今後検討を要する。またサンプル数の範囲をみるとかなりのバラツキがあるが，サンプル数の少ないとき，Yu の推定精度はかなり落ちるが Ya ではサンプル数が多いときに比してそれ程推定精度は落ちていない。これもさきほどの ne/n という補正係数が有効に作用しているためと考えられる。なお３ＰＳにおける問題点の１つとして，実際に野外調査を行ったとき，サンプル数が０になると林分材積の推定ができないという欠点があげられている。今回の３ＰＳでサンプル数が０となった割合を〔表－３〕に示す。このようにサンプル数が０となる危険性は，期待サンプル数 ue が確率変数である以上常につきまとうので，乱数の最大値Ｌを設定する際こうしたことも考慮する必要があろう。

〔表－３〕　３ＰＳにおいてサンプル数が０となった回数

補助変数	ne	サンプル数が０の回数	
		１回目	２回目
目測材積	2.3	11／200	9／200
〃	3.3	2／200	3／200
〃	4.6	2／200	2／200

※　３ＰＳは２００回づつ試行した。

補助変数については目測材積がもっともよいが，胸高直径の二乗のときもそれに劣らない精度で林分材積を推定している。胸高直径を補助変数した場合は前２者に比して精度が悪い。このことから補助変数としての良し悪しは林分材積との相関に比例することが確認できる。

つぎに他のサンプリング方法と３ＰＳを比較するために，同じモデル林分に対しランダムサンプリングを行った。その結果は〔表－４〕のようになったが，この林分にたいしては明らかに３ＰＳの方が精度も効率も優れている。この理由の一つには，対象林分が多数の小径木の中

〔表－４〕　任意抽出の結果

サンプル数	推定値の平均	標準偏差	推定値の範囲
5	178.1㎥	342.0	1.7 ～ 1255.5㎥
10	174.8㎥	258.0	6.1 ～ 633.3㎥
15	175.7㎥	222.4	9.7 ～ 522.0㎥
20	175.1㎥	203.4	20.1 ～ 396.4㎥
25	174.2㎥	199.1	19.4 ～ 456.9㎥
30	173.9㎥	195.0	20.1 ～ 449.1㎥
35	172.8㎥	188.8	38.9 ～ 401.0㎥
40	172.2㎥	185.5	44.9 ～ 339.2㎥
45	165.9㎥	179.3	42.1 ～ 337.9㎥
50	165.4㎥	178.1	68.5 ～ 343.9㎥
100	148.9㎥	155.5	74.3 ～ 225.1㎥

※　各サンプル数について２００回づつランダムサンプリングを行なった結果である。

に大径木が点在するという林分構造であり，それに対し単木を単位とした任意抽出をすると，大径木が標本木として抽出されたときはどうしても過大評価になってしまうことが考えられる。したがってこのような林分に対してはプロットサンプリングを行えばもう少し精度の良い結果が得られると思われる。ただしこの場合でも天然林ではかなりの本数の材積測定が必要であると考えられるので，僅か５本程度のサンプル木を選ぶだけで効率よく林分材積の推定ができる３ＰＳの方が有利であろう。

4.　３ＰＳ手順とそのデータ処理

試験的に我国の森林に３ＰＳを適用した経験から，Grosenbaugh の３ＰＳを簡易化した形での３ＰＳ手順とそのデータ処理方法について提案する。なおここでは林分材積の推定を目的とし，補助変数として目測材積の使用を前提としている。実際に３ＰＳを試みるときは，随時他のものをこの両者にあてても，手順や処理形態は同じである。提案するシステムの概略は以下のようである。

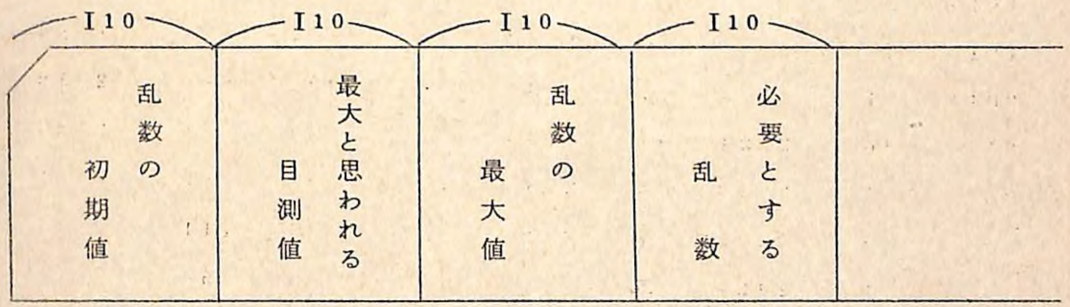
1）乱数表の作成
2）テレレラスコープによる現地調査
3）データ集計
4）電算機によるデータ処理

なお必要とする器具はテレレラスコープと携帯用の電卓である。

○乱数表の作成

任意抽出や系統的抽出のようにあらかじめ標本数を決めてサンプリングを行なうのとは異なり，３ＰＳでは実際に現地調査をしなければ標本数が掴めない。しかし調査計画はある程度立てる必要があるため，３ＰＳ理論の中でふれた(2)式を用いて期待サンプル数を推定する

〔図－２〕　乱数表作成のための入力様式例



※　乱数の初期値は１～３２７６７の間の任意の整数
※　必要とする乱数は想定される立木本数より大きな数とする。

〔図－3〕 乱 数 表

*****ランスウ ヒヨウ*****

ランスウ ノカズ ＝ 250 サイダイスイテイチ ＝ 147　　ランスウ ノ サイダイチ ＝ 200

1= 8	2= −1	3= 71	4= 0	5= 126	6= −1	7= 20	8= 35	9= 1	0= 98
11= 78	12= 20	13= 31	14= −1	15= 14	16= 129	17= −1	18= 32	19= −1	10= 33
21= −1	22= 48	23= 72	24= 2	25= 41	26= 34	27= −1	28= 62	29= −1	20= 13
31= 111	32= 115	33= 124	34= −1	35= 14	36= 116	37= 127	38= −1	39= 25	30= 90
41= 54	42= 105	43= 97	44= 74	45= 8	46= −1	47= 52	48= 95	49= 70	40= −1
51= 75	52= 11	53= 66	54= −1	55= 46	56= 59	57= 137	58= −1	59= 63	50= −1
61= 24	62= 79	63= 24	64= 74	65= 10	66= 26	67= 93	68= 63	69= −1	60= 25
71= 81	72= 30	73= 147	74= 8	75= −1	76= 28	77= 121	78= 143	79= 2	70= −1
81= 72	82= 2	83= 64	84= −1	85= 33	86= −1	87= 56	88= 120	89= 139	80= −1
91= 71	92= 0	93= 92	94= 59	95= 136	96= −1	97= 59	98= 134	99= −1	90= 52
101= 98	102= 77	103= 18	104= 5	105= 57	106= 123	107= −1	108= 9	109= 13	100= 115
111= 124	112= −1	113= 13	114= 97	115= 74	116= 8	117= −1	118= 47	119= 64	110= −1
121= 33	122= −1	123= 43	124= 45	125= 55	126= 113	127= 120	128= 140	129= −1	120= 75
131= 11	132= 50	133= 83	134= 35	135= 0	136= 104	137= 94	138= 65	139= −1	130= 36
141= 7	142= 118	143= 134	144= −1	145= 54	146= 104	147= 95	148= 68	149= −1	140= 60
151= 140	152= −1	153= 76	154= 16	155= −1	156= 35	157= −1	158= 74	159= 10	150= 40
161= 30	162= 141	163= −1	164= 82	165= 31	166= −1	167= 25	168= 90	169= 55	160= 114
171= 122	172= 145	173= 4	174= 22	175= 55	176= 115	177= 125	178= −1	179= 17	170= −1
181= 63	182= −1	183= 19	184= 17	185= −1	186= 75	187= 11	188= 53	189= 101	180= 86
191= 44	192= 49	193= 81	194= 29	195= 133	196= −1	197= 47	198= 68	199= −1	190= 64
201= −1	202= 27	203= 114	204= 123	205= 147	206= 7	207= −1	208= 8	209= −1	200= 60
211= 140	212= −1	213= 75	214= 13	215= 94	216= 65	217= −1	218= 40	219= 29	210= 136
221= −1	222= 60	223= 140	224= −1	225= 77	226= 18	227= 13	228= 101	229= 84	220= 38
231= 15	232= −1	233= 13	234= 111	235= 115	236= 124	237= −1	238= 14	239= 116	230= 127
241= −1	242= 25	243= 90	244= 54	245= 105	246= 97	247= 74	248= 8	249= −1	240= 52

チユウイ −1 ノ ランスウ ニ タイオウスル サンプルボク ハ ソクテイシナイ

必要がある。したがって調査担当者はまずおおまかな目安として調査対象林分の材積Xを推定するとともに，最大と思われる木の目測材積Xmax とおおざっぱな立木本数Nmaxを知っておく必要がある。つぎに望ましいと思われる期待サンプル数neを決めると，乱数の最大値Lは

$$L = X/_{ne}$$

で与えられる。但しL＞Xmax でなければならない。このようにしてLが決まれば，現地調査に必要な乱数表は〔図－2〕，〔図－3〕のような入出力様式で得られる。

o テレレラスコープによる現地調査

Grosenbaugh はデンドロメータによる材積測定を行なっているが，我国は地形が急峻であるため，器械の設定，移動に手間のかかるデンドロメータよりはテレレラスコープの方が実用的であると判断した。

さて，現地調査としては，まず調査対象林分の各立木をある順序で目測を行ない，その目測値が乱数より大きいか等しければその木をテレレラスコープで実測し，再び次の木の目測に移る。テレレラスコープによる利用材積の測定の仕方は，立木を一定長の丸太に区切りその上部直径をテレレラスコープで読み取り算出するが，測すすべき上部直径の高さをテレレラスコープの仰角に換算するために簡単な電卓がいる。

このようにして現地調査を行なうが，そのとき野帳の記載要領を〔図－4〕に示す。

〔図－4〕　野帖の記載要領

目測推定値野帖

番号	樹種	乱数	目測値	備考	番号	樹種	乱数	目測値	備考
1	ブナ	3	4.5		11				
2					12				
3					13				
4					14				
5					15				
6					16				
7					17				

実測値野帖

番号		乱数		目測値		総材積		備考	
高さ	上部直径	材積	品質	備考	高さ	上部直径	材積	品質	備考
2.0					22.0				
4.0					24.0				
6.0					26.0				
8.0					28.0				
10.0					30.0				

o データの集計

3PSによって得られたデータを電算機で処理する前に入力データ量を少量化するため，各測定された丸太別のデータを単木毎の利用材積として集計する。末口二乗法を例とすると，D_{ij}がi番目の標本木の j 番目の丸太でK番丸太までとれるとし，ℓを丸太の長さとすれば，

$$i\text{番目の丸太の利用材積} = \left\{ \sum_{j=1}^{K} D_{ij}^2 \right\} \times \ell$$

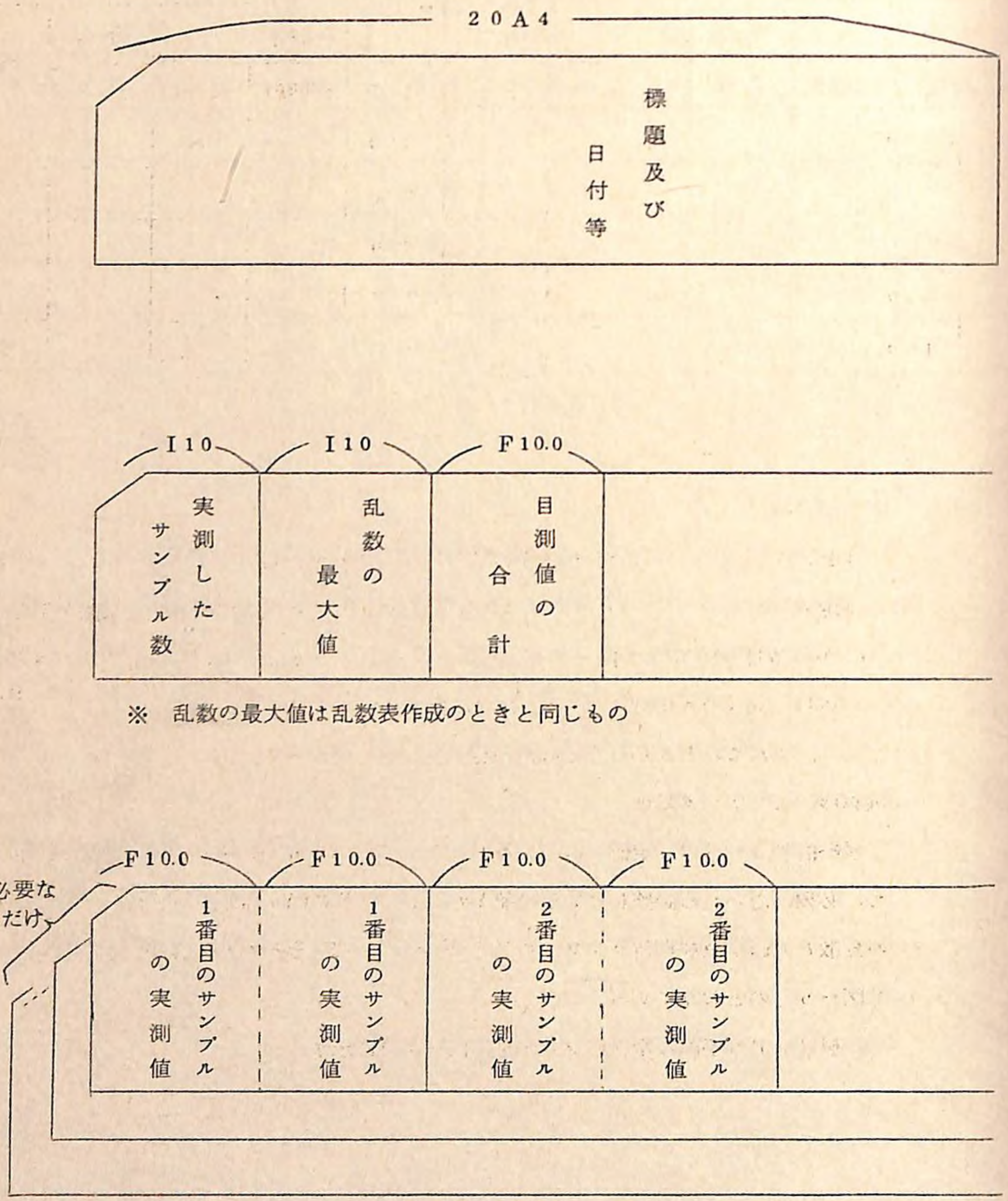
となる。またこれとは別に全目測値の合計Xを計算しておく。

o 電算機によるデータ処理

最終的に3PSの推定値を計算するわけだが，Grosenbaugh のとりあげた推定値のうち，必要と思われる調整しない推定値Y_u，その分散$v(Y_u)$，調整された推定値Y_a，その分散$v(Y_a)$及び期待サンプル数$E(n)$を計算する。その入力と出力の様式は〔図－5〕，〔図－6〕のとおりである。

最後に3PS手順の全体のフローを〔図－7〕に示す。

〔図－5〕 3PSのデータ処理のための入力様式



※ 乱数の最大値は乱数表作成のときと同じもの

※ 1枚のカードに4組のデータをパンチする。

〔図－6〕 3PSデータ処理の出力様式

3－P SAMPLING TEST DATA 1978. 7. 20

サンプル ノ カズ ＝ 8 ランスウ ノ サイダイチ ＝ 50 モクソク ノ ゴウケイ ＝ 229.70

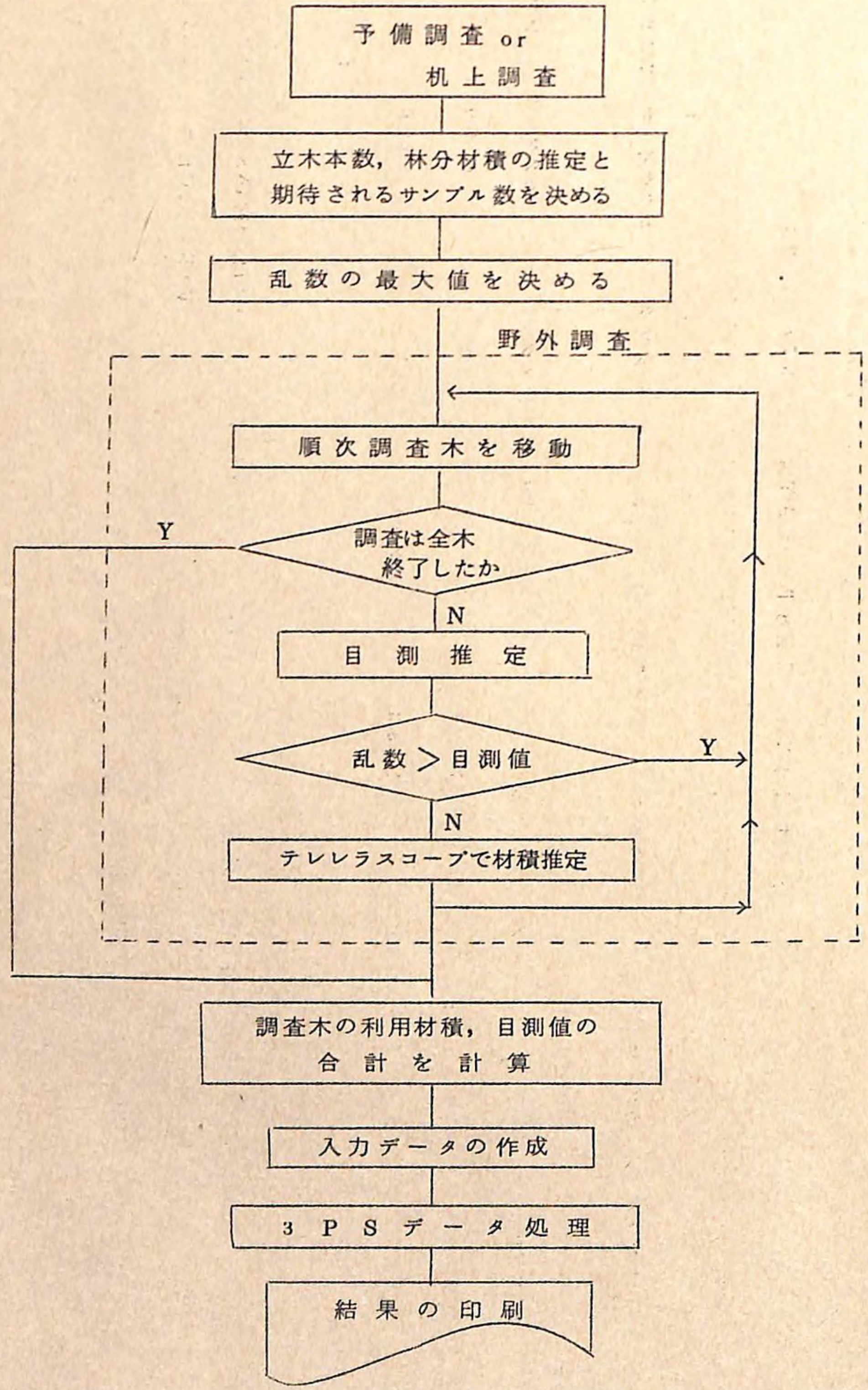
モクソクチ	ジツソクチ	モクソクチ	ジツソクチ	モクソクチ	ジツソクチ	モクソクチ	ジツソクチ
10.00	3.36	30.00	16.11	0.70	0.27	10.00	4.55
0.14	0.06	30.00	17.33	24.00	12.11	10.00	6.12

UNADJUSTED ESTIMATE MEANS ＝ 192.331 ADJASTED ESTIMATE MEANS ＝ 110.446

VARIANCE ＝ 3188.375 VARIANCE ＝ 59.391

ESTIMATED SAMPLE NO. ＝ 4.594

〔図－7〕　３ＰＳの流れ図



5. まとめ

３ＰＳの理論及びその推定式の検討を行なったが，ここで改めて３ＰＳの特徴をあげると，つぎの３点に要約される。

1）補助変数に比例した標本抽出

2）高精度の器械による材積測定

3）電算機によるデータ処理

1）については今回目測材積と胸高直径を用いたが，どちらも測定が容易であり，とくに目測は林業において慣習的に多用されており，それなりの技術を有した者も多いので，2)の上部直径の測定可能な器械に結びつければ，林業の特性をうまくいかしたサンプリングが可能であろう。また2)は材積表を必要としないことを意味しており，そういった点から，我国の森林調査でも有効に活用できる分野がある。3）についてはこの報告でも一応電算機を利用したが３ＰＳで使われる程度の乱数や推定式の算出には廉価な電卓でも十分計算可能である。

なおここで新しい森林調査法として３ＰＳをとりあげてきたが，森林調査にあたって全てに万能であるような調査法というのは存在せず，対象とする森林の構造や調査目的の相異に応じて，調査方法も柔軟に変えていくことが必要であり，そのためにも今後種々の調査方法を森林に適用していく努力が必要であろう。

引　用　文　献

1）Hans T. Schreudrr ：３－P Sampling and Some Alternatives，I. Forest Science，vol 14，PP 429－453，1969

2）L. R. Grosenbaugh：Three－Pee Sampling Theory and program THRP′ for computer generation of selection criteia，U. S. For. Sew. Res. Paper PSW－21，1－53，1965

3）G. P. Patil，E. C. Pielou，W. E. Waters Ed.：Statistical Ecology vol. 2，Pennsylvania State Univ. Press，1971

4）Des Raj ：Sampling Theory，McGraw－Hill，1968

5）鈴木太七・久田直弘：Grosenbaugh の３Ｐ法について，８８回日林講 PP 95－96