森林調査に於ける統計数理的問題

松　下　嘉　米　男
林　　　知　己　夫
石　田　　正　次
藤　本　　　　熙
赤　池　弘　次
宇　沢　弘　文
植　松　俊　夫

(1953 年 1 月　受付)

Some Statistical Problems in Forest Survey

Kameo MATUSITA　Chikio HAYASHI
Masatugu ISIDA　Hiroshi FUJIMOTO
Hirotugu AKAIKE　Hirobumi UZAWA
Toshio UEMATSU

Chapter I

The errors in measurement of diameter and height

Assuming that the type of cross section of tree is ellipse, we showed the optmum method of measurement of diameter by calipers. Then we discussed the relation between the error of volume and the errors of diameter and height.

Chapter II

On the sampling units

We treaded a problem, how to make sampling units to estimate the total volume of stands in an area.

The optimum division of the area into sampling units was obtained considering the cost functions.

Chapter III

On the area sampling method

In this chapter we treated the sampling design of the forest susvey to estimate the total volume in Japan, using the longitudinal and latitudinal line for the sampling units.

Chapter IV

Variances of regression estimate

We calculated the variances of regression estimates of several types which are observed in forest survey.

The Institute of Statistical mathematics

緒　言

統計数理の対象として林学は非常に多くの興味ある問題を含んでいる．苗圃の問題，測樹，生長量，森林調査，森林施業等すべて新しい統計理論の実験分野をなしている．

統計数理研究所は林野庁計画課，林業試験所を協力して，統計数理的な立場から，林学に於ける諸問題の解決に当つているが，その中で妥当性を目ざす森林調査の方法論の中我々の立場からみて興味ある問題をここに取り上げて論ずることにする．

森林調査は林学に於ても又実際の森林経営の上からも，最も基礎的な分野であるが未だ多くの問題点を残している．理論家が手をつけていても，疊水練の域にあるものが多いのは遺憾なことである。ここに報告するものは林野庁計画課が企画した森林調査に於て生じた問題であるが，その多くは未だ完全な解決をみていない．今後も多方面からの検討を要するものである．しかし，実用上一応役に立つものもあり，又多くの批判者協同研究者を得るためにここに発表したわけである．

（統計数理研究所）

第 I 章

毎木調査に於ける誤差

§1.　ま え が き

毎木調査と云うものが林学に於て非常に身近かになりすぎるためか，現今でも相当あぶない方法がそのままなんの検討もなしに用いられている．そこには材積算定の精度上かなり問題となり得るものがみうけられるので，ここにその二三の事項をとり上げてみることにする．

毎木調査は計量的に森林を取り扱うための最も基礎となるものであつて，調査の項目は樹種，令級，胸高直径 (d)，樹高 (h) がその主なるものであるが，ここでは計量的に与えられる後二者について考えてみる．この二つの数量はそれ自体意味をもつほか更に

$$v=f(d,h)$$

として樹木の材積を算出しなければならないので，この三者間の種々の誤差が統計的問題となるわけである．又この函数型そのものも統計的対象として重大な問題を蔵している．つまり d と h だけの知識からどれほどの精度で v を推定できるかと云う測樹の問題が残るがここでは論じない．

§2.　胸高直径の測定

通常胸高直径は胸高断面積を求める目的で測定されているが，この目的のためには如何に胸高直径を測定するのが合理的であるかを考えてみよう．

一般に胸高直径の測定には実用的には次の三つの方法が行われている．つまり

a）最大径と最小径を測定して，その平均値を出す．

b）任意の方向の直径とこれに直交する方向の直径を測定して，その平均値を出す．

c）任意の一方向の直径を測定する．

今樹木の胸高断面が楕円であると考え，この三つの場合を検討することにする．

a）の場合

今楕圓の長軸を a，短軸を b とすればこの場合の胸高断面積 d は

$$d=\pi\left(\frac{a+b}{2}\right)^2$$

の式で求めていることになる

この式を變形すれば

$$d=\pi ab+\frac{1}{4}\pi(a-b)^2$$

となり常に $\frac{1}{4}\pi(a-b)^2$ だけ多めに見積つてしまうことになる．（楕圓の面積は πab であるから）

又 a と b の測定にそれぞれ ε_a, ε_b の測定誤差が生ずると考えれば,

$$d=\pi\left(\frac{a+b+\varepsilon_a+\varepsilon_b}{2}\right)^2$$

が胸高断面積となるここで今 ε_a, ε_b がまつたくの偶然誤差で

$$E(\varepsilon_a)=0\ ,\quad E(\varepsilon_b)=0\ ,\quad E(\varepsilon_a\cdot\varepsilon_b)=0$$

$$E(\varepsilon_a{}^2)=E(\varepsilon_b{}^2)=\sigma^2$$

と仮定すれば，平均的な意味では,

$$d=\pi ab+\frac{1}{4}\pi(a-b)^2+\frac{1}{2}\pi\sigma^2$$

となり常に $\frac{1}{4}\pi(a-b)^2+\frac{1}{2}\pi\sigma^2$ だけ多めに見積るわけである.

b） の場合

はじめに測つた任意の半径が長軸から θ だけ傾いていたとすればこの場合の胸高断面積 d は

$$d=\pi\left\{\frac{1}{2}\left(\frac{ab}{\sqrt{a^2\cos^2\theta+b^2\sin^2\theta}}+\varepsilon_1+\frac{ab}{\sqrt{a^2\sin^2\theta+b^2\cos^2\theta}}+\varepsilon_2\right)\right\}^2$$

ここで θ は 0 から $\pi/2$ までの間に等しく分布する確率変数と考えて θ と ε に関する平均値をとれば近似的に

$$E(d)=\pi ab\left\{1-\frac{1}{6}\frac{(a-b)^2}{a^2}\right\}+\frac{1}{2}\pi\sigma^2$$

となり　$-\dfrac{\pi ab}{6}\dfrac{(a-b)^2}{a^2}+\dfrac{1}{2}\pi\sigma^2$　だけの歪みを生ずる.

c） の場合

この場合は

$$d=\pi\left(\frac{ab}{\sqrt{a^2\sin^2\theta+b^2\cos^2\theta}}+\varepsilon\right)^2$$

であるから前の（b）と同様に d の θ と ε に関する平均値をとれば

$$E(d)=\pi ab+\pi\sigma^2$$

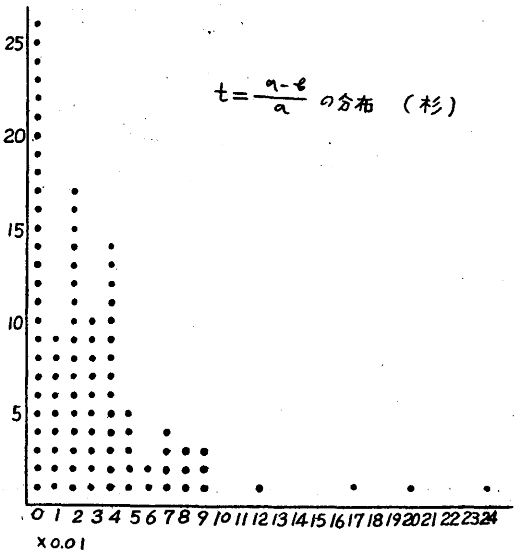
で σ^2 だけ常に多く見積つている.

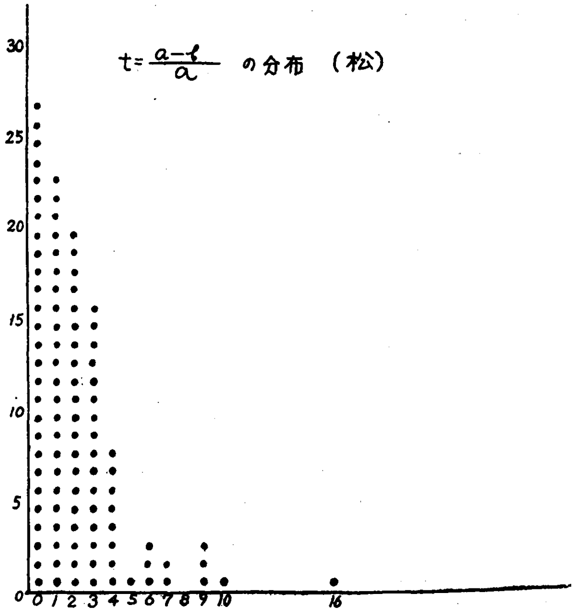
(a),(b),(c) の相対的な歪みを表にしてみると近似的に次のようになる.

	胸高断面積に及ぼす相対的な歪み	特　色
a） 長軸と短軸	$\frac{1}{4}t^2+\frac{1}{2}S^2$	常に正の歪み
b） 任意の直交二直径	$-\frac{1}{6}t^2+\frac{1}{2}S^2$	正とも負とも云えない
c） 任意の一直径	S^2	測定誤差だけで常に正の歪み

但　$t=\dfrac{a-b}{a}$,　$S^2=\dfrac{\pi\sigma^2}{\pi ab}$

この三つの方法を比較するためには t^2 と S^2 の大さを検討する必要がある. t の分布の一例には次のようなものがある.





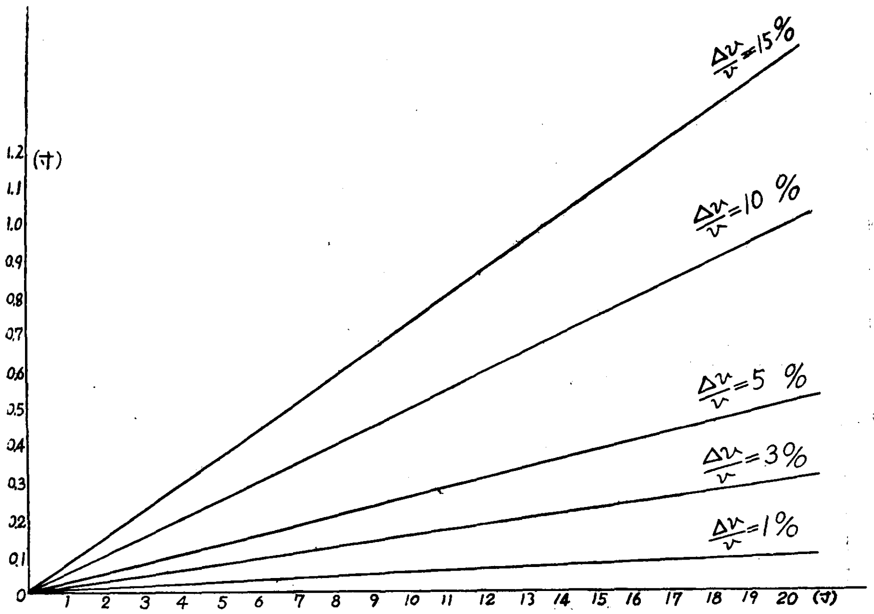
勿論これはごく一部の林分の調査結果であるから，この様子から結論を出すことは難かしいが，t の値はまず 0.02 から 0.04 を平均として J 型に近い分布をしていると考えて大過はあるまい．又 t と a の間の相関はみとめられない．

そこで a)，b)，c) のうち最もよい方法はどれかと云うに，もし測定誤差が非常に小さなものであれば，最も簡単な (c) の方法（任意の一方向）となるわけである．つまり偏平度の補正のつもりで二方向測定したことが，実はかえつて測定結果に歪みをもたらすことになつてしまう．極めて円に近い切断面をもつ樹木を測る場合にあつては t の項が無視出来るから，(b) の方法によつて二方向を測定し，測定誤差の項を二分の一とすることもあながち無意味ではないが（この場合はほかに目盛りのよみあやまりも発見出来る可能性がある）t の項は断面積に比例し，S の項は，その断面積に関係しない量であるから，相対的にみれば，太い樹木の場合は t の項が問題になり，細い樹木の場合は S の項が問題になると考えられる．

いづれにせよ，今まで唯無反省に2方向の測定が1方向よりも精度が高いと考えて来たことには問題があつて，むしろ1方向の方が，理論的にも実用上からも一応すぐれていると見るべきであろう．

以上の議論は，樹木が完全な楕円であるとの仮定した場合であつたが，これに対してはこの仮定は必ずしも満足されない，つまり樹木の断面積は非常に複雑であると云う反論も生ずるわけである．この反論は一応もつともなことではあるが，断面が楕円形に近い凸形であれば一方向ランダムの測定が常に有利である．又凹形の場合まで考えれば測定器具そのものを変えなければならぬ．

念のため胸高直径の測定誤差が蓄積に及ぼす割り合いを近似的にグラフに示してみると次のようになる．



§3. 樹高の測定

毎木調査を行う場合，樹高の測定は胸高直径の測定に比して一般に非常に粗雑に取り扱われがちであるので，ここに樹高の測定誤差が，どれくらい蓄積に影響するかを検討してみることにする．

まず立木幹材の材積は寺崎氏の式（林業試験報告第8号）を用いて求めるものとする．つまり

$$v=cg10^{ah-b/h}$$

但　v: 幹材積
　　g: 胸高，断面積
　　h: 樹高
　　a, b, c: 常数

に於て，h に Δh なる誤差がある場合，v にどれほどの狂いを生ずるかを計算する．

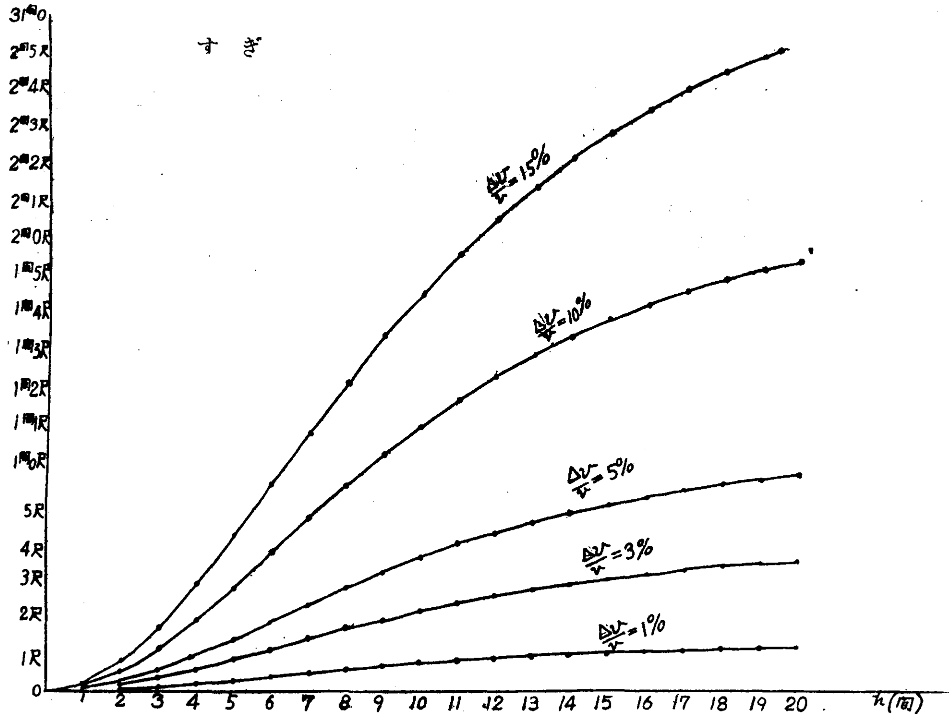
$$\frac{cg10^{a(h+\Delta h)-b/(h+\Delta h)}}{cg10^{ah-b/h}}=K$$

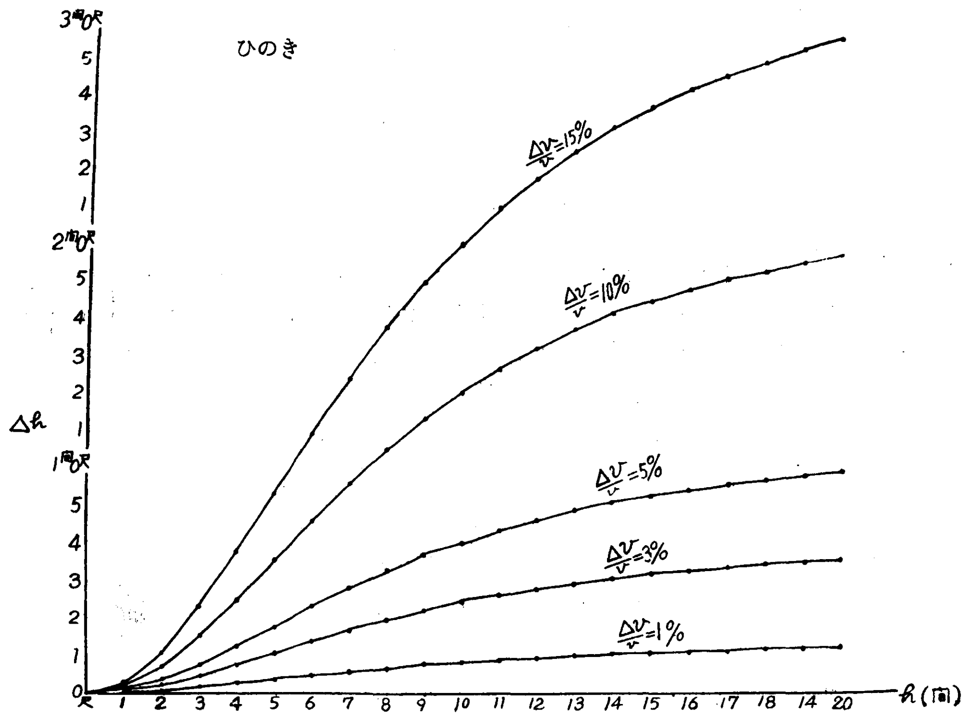
とおけば，近似的に次式を得る．

$$\frac{b}{h}\varepsilon^2-\left(ah+\frac{b}{h}\right)\varepsilon+\log K=0$$

但　ε : $\Delta h/h$（樹高の相対誤差）
　　K: $1+\Delta v/v$（Δv は材積の誤差）

この式からすぎ，v のきについてのグラフを作れば次のようになる．





このグラフから見られるように，我々がふだんに犯す樹高の誤差の材積えの影響はけして小さいものではない．

以上は単木の場合であるが更に林分の総蓄積と，樹高の誤差との関係を次に考えてみることにする．

材積の式

$$v=cg10^{ah-b/h}$$

に於て h に Δh なる誤差がある場合材蓄の誤差 Δv を近似的に計算てしみると

$$\Delta v=v\left(a+\frac{b}{h^2}\right)\Delta h\cdot\log_e 10$$

となるそこでこのような測定を一つの林分（本数 n）でしたときの総蓄積の誤差 ΔV_T を出せば

$$\Delta V_T=\sum_{i=1}^{n}\Delta v_i=\log_e 10\left\{a+\frac{b}{h_i^2}\right\}v_i\Delta h_i$$

となりこれが bias の first order となる．この bias の order は樹高の測定法（機差，個人差等）によつて定まるものであるから，樹高の測定にはこの辺の充分な検討が必要であろう．

実際の調査の場合は樹高を目測によつて求めることが多いが，非常になれているものであつても目測の精度は一般に信じられている値よりも遙かに大きいものであることに留意しなければならない．しかもその誤差は通常高い方を低くみる傾向があり．又測定を行う順序に関して樹木の大小がその測定値の誤差に關係してくるから，目測の誤差の扱いは一層困難をともなうものである．結局のところ樹高は何等かの方法で前述の條件を考えた実測を行わないかぎり，妥当な蓄積を求めることは出来ないと云いうるであろう．

§4. 樹高，胸高直径の平均値からの蓄積算出

樹高 h_i，胸高直径 d_i をもつ n 本の樹木を含む林分の総蓄積を求める場合，平均樹高 $\bar{h}$ と平均胸高直径 $\bar{d}$ を求めて，この二つの値から単木材積表により平均的な材積を求めこの値を n 倍して総蓄積とする方法が時々みうけられる．この場合寺崎氏の式を用いたときの誤差をここで論じてみる．つまり，

$$V_T=\sum_{i=1}^{n}v_i=\sum_{i=1}^{n}c\pi\left(\frac{d_i}{2}\right)^2 10^{ah_i-b/h_i}$$

の代りに簡便法として

$$V_T'=nc\pi\left(\frac{\bar{d}}{2}\right)^2 10^{a\bar{h}-b/\bar{h}}$$

を用いたときの誤差を考える．

そのために V_T と V_T' の比 r を作る．

$$r=\frac{V_T}{V_T'}=\frac{1}{n}\sum_{i=1}^{n}\frac{d_i^2}{\bar{d}^2}10^{a(h_i-\bar{h})-b(1/h_i-1/\bar{h})}$$

この式を展開すれば，近似的に

$$r=1+\frac{\sigma_d^2}{\bar{d}^2}+\frac{K}{n}\sum_{i=1}^{n}d_i^2(h_i-\bar{h})$$

但 σ_d^2: 胸直径の分散

K: $\dfrac{(a\bar{h}^2+b)\log_e 10}{\bar{d}^2\bar{h}^2}$

を得る．ここで h と d と間に一次的な関係つまり

$$(h_i-\bar{h})=\alpha(d_i-\bar{d})+\varepsilon_i$$

但 α: 常数

ε_i: 回帰直線のまわりの誤差

なる式が成立するとすれば，

$$r=1+\frac{\sigma_d^2}{\bar{d}^2}+K\left\{\alpha\mu_3+2\alpha\bar{d}\sigma_d^2+\frac{\sum\varepsilon_i(d_i-\bar{d})^2}{n}+\frac{2\bar{d}\sum(d_i-\bar{d})\varepsilon_i}{n}\right\}$$

となる．更にここで ε_i と d_i 及び d_i^2 とがそれぞれ独立であるとすれば，

$$r=1+\frac{\sigma_d^2}{\bar{d}^2}+\alpha K\{\mu_3+2\bar{d}\sigma_d^2\}$$

更に胸高直径の分布が対象であるとすれば μ_3 は0となるから，

$$r=1+\frac{\sigma_d^2}{\bar{d}^2}\left\{1+4.6\alpha\bar{d}\left(a+\frac{b}{\bar{h}^2}\right)\right\}$$

となり．この式によつて V_T を V_T' で代用した場合の誤差を近似的に計算することが出来る．

この式を用いて次に示すような杉の林分についての計算をしてみる．

す　ぎ

h ＼ d	1.5	2.5	3.5	4.5	5.5	6.5	7.5	
1.5	1							1
2.5	29	13						42
3.5	14	107	31					152
4.5		25	141	14				180
5.5		1	42	57	2			102
6.5				12	9		1	22
7.5					2	1		3
8.5							3	3
9.5							1	1
計	44	146	214	83	13	1	5	506

$\bar{h}$（間）	4.3676
σ_h^2	1.2058
$\bar{d}$（寸）	3.2984
σ_d^2	1.0385
$\sigma_d^2/\bar{d}^2$	0.0955
ρ	0.883
α	0.951
μ_3	0.7099

V_T	73.17石
V_T'	61.00石
誤差	+20%

寺崎氏によれば杉の常数は

$$a=0.0190\qquad b=1.379$$

であるから，r は約 1.22 で修正値は 74 石となる．はじめは約 20％ ほどくるつていたものをかなりの精度まで修正出来る．この近似は非常にきつい仮定のもとでの話であつて，常にこのような修正が出来るとは限らないことに留意しなければならぬ．いづれにせよ平均値から蓄積を求めれば約2割程度の誤差が生ずるのであつて，その大約の修正をほどこすにも上式のような比較的面倒な手段を用いなければならないのである．

次に樹高のみ平均値を用いた場合つまり

$$V_T'=c\pi10^{a\bar{h}-b/\bar{h}}\sum_{i=1}^{n}\left(\frac{d_i}{2}\right)^2$$

でV_T を代用した場合を考えれば近似的に

$$V_T'-V_T=c\cdot\log_e10\cdot10^{a\bar{h}-b/\bar{h}}\sum_{i=1}^{n}g_i(\bar{h}_i-\bar{h})$$

（但 g は胸高断面積）

で g と h との間に相関がなければこの誤差は 0 となる．しかし一般にこの間の相關はかなり高い．今大きく見積つてこの相変係数1をとすればこの間の誤差は

$$V_T'-V_T=c\cdot\log_e10\cdot10^{a\bar{h}-b/\bar{h}}\cdot\sigma_h\cdot\sigma_g$$

となる．

§5.　樹高の代りに回帰直線を代用する方法

この場合は

$$V_T = \sum_{i=1}^{n} v_i = \sum_{i=1}^{n} c\pi\left(\frac{d_i}{2}\right)^2 10^{ah_i - b_i/h_i}$$

の h_i の代りに h と d との間の回帰直線

$$h-\bar{h}=\alpha(d-\bar{d})$$

を代用する．つまり

$$V_T' = \sum_{i=1}^{n} c\pi\left(\frac{d_i}{2}\right)^2 10^{a\{\alpha(d_i-\bar{d})+\bar{h}\}-b/\{\alpha(d_i-\bar{d})+\bar{h}\}}$$

を以つて V_T の代用とするものである．この場合の誤差を次に計算してみる．

今回帰直線のまわりの誤差を ε_i とすれば

$$h_i=\alpha(d_i-\bar{d})+\bar{h}+\varepsilon_i$$

となるから，

$$V_T' \doteqdot \sum_{i=1}^{n} v_i 10^{-(a+b/h_i^2)\varepsilon_i}$$
$$= \sum_{i=1}^{n} v_i \left\{1 - i\left(a+\frac{b}{h_i^2}\right)\varepsilon_i \log_e 10\right\}$$

故にこの場合の

$$V_T - V_T' = \log_e 10 \sum_{i=1}^{n}\left(a+\frac{b}{h_i^2}\right)\varepsilon_i v_i$$

即ち，この方法が用いられるためには，つまり $V_T - V_T' = 0$ なるためには ε_i と v_i 及び v_i/h_i^2 間の相関が 0 でなければならない．しかしこの條件が成立することを望むことはむづかしいであろうから，この方法によつても bias をまぬかれることは出来ない．因みに樹高と胸高直径の相関図の一例をを後にあげてみる．この方法を用いる場合は回帰直線は大抵標本によつて計算されるので，それによる誤差も又考慮する必要がある．

§6.　Bitterlich の方法について

これはある林分の胸高断面積の総和を求めようとする方法である．次にこの方法の概略を述べてみる．

調査を実施しようとする林分を完全に覆う充分細かい格子網を考える．今林分の中に落ちた格子点の総数を N，又林分の樹幹の中に落ちた格子点総数を N_0 とすれば

$$R = \frac{N_0}{N}$$

は林分の全面積 S とその林分中の樹木の胸高断面積 S_0 との比

$$K = \frac{S_0}{S}$$

の近似値になる．そこで N のうちからランダムに n 個だけを抽出しそのうち樹幹の中に落ちる点の数 n_0 を数えれば

$$r = \frac{n_0}{n}$$

は R の標本値として K を推定することができるわけである．

胸高直径と樹高との相関図（桧）

h＼d	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	10.5	11.0	11.5	12.0	12.5	13.0	13.5
0.5																											
1.0		1	1																								
1.5		18	10	4	1																						
2.0		11	35	20	6	2																					
2.5		2	17	43	24	9	2																				
3.0		1	4	16	43	31	7	2																			
3.5				3	9	23	12	6	1	1																	
4.0					2	9	12	8	7	1	1																
4.5					1	5	15	17	10	4	3	2	1														
5.0						1	5	13	18	12	7	6	1	1	1												
5.5						2	5	16	26	33	24	19	7	3	1												
6.0								1	7	15	16	12	3	2	1												
6.5								1	2	6	10	11	9	4	2	2											
7.0							1	1	2	5	8	7	8	6	3	2	1	1	1	1				1			
7.5										1	3	4	5	5	2	2	2		1								
8.0											1	1	2	3	4	4	4	1									
8.5													1	2	2	2	2	2		1							
9.0															1	2	1	1	1	1	1						
9.5																	1			1	1	1					
10.0																				1	1		1				
10.5																											
11.0																											
11.5																				1							
12.0																											
12.5																											

相関係数	0.921		胸高直径(寸)	樹高(尺)
回帰直線	$h = 0.871d + 0.921$	平均	4.139	4.528
$2\sigma_h\sqrt{1-\rho^2}$	1.502	標準偏差	2.037	1.927

胸高直径と樹高との相関図（杉）

h＼d	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	10.5	11.0	11.5	12.0	12.5	13.0	13.5
0.5																											
1.0		1																									
1.5		5	8	2																							
2.0		4	16	16	2	1																					
2.5		1	8	12	7	5	1																				
3.0			2	9	17	10	4	1																			
3.5				3	14	13	7	3	1	1																	
4.0					4	16	15	10	4	2	2																
4.5					1	8	30	19	16	8	5	2	1														
5.0						2	12	31	25	22	17	12	5	3		1											
5.5							4	19	40	20	9	5	4	3	2	1	1										
6.0							1	4	20	25	17	7	4	4	2	1											
6.5				1			2	5	9	27	42	23	13	8	5	2											
7.0								1	1	4	9	15	16	17	5	2	1	1		1					1		
7.5									1	1	2	6	13	10	7	5	3	1									
8.0											3	2	6	5	6	2	2	1		1							
8.5											2	2	3	7	7	2	3	1									
9.0													2	2	3	2	2	1	2								
9.5													1		2	2	2	3	1								
10.0														1	1	1	2	2	2			1					
10.5															1			2	3	1							
11.0																				2	1						
11.5																											
12.0																										1	1
12.5																						1		1			
13.0																										1	
13.5																								1			
14.0																											1
14.5																											

			胸高直径（寸）	樹高（尺）
相関係数	0.891			
回帰直線	$h = 0.833d + 1.158$	平　均	4.973	5.551
$2\sigma_h\sqrt{1-\rho^2}$	1.776	標準偏差	1.972	1.955

しかし実際に n_0 を数えることが困難であるために次のような方法を考える．今仮にその林分の樹木の胸高直径が全部一率に t 倍であつたとする．（このとき各樹木は重なり合う場合もあつてよい。）そして抽出された n 個の格子点の各に立つて，その格子点が何木の樹幹（t 倍したもの）に含まれるかを数える．今 n 個の格子点のうちの i 番目の格子点が x_i 本の樹幹の中に含まれるとすれば

$$r = t^2 \sum_{i=1}^{n} x_i / n$$

となる．そこで x_i を数えるには調査員は t m だけはなれたとき巾 1 m の物体が常に画面一ぱいになるような視野を有する眼鏡をもつて抽出された i 格子点に立ち，各樹木を望んで，そのうち画面を完全に覆う樹木の本数を数えれば x_i が得られる．

この方法は毎木についての測定を行わずに胸高直径の総和が得られるので，一応注目されたわけであるが，この方法は次に述べるような非常な難点をもつている．つまり，

1） 樹木の影になるものが数えられない．
2） 樹木の胸高断面は完全な円でなければならぬ．
3） 林地の凸凹のはげしいところでは遠方の樹木の胸高直径が望めない．
4） 林分の面積を求める必要があり，面積の誤差はそのまま推定値の誤差に入つてくる．
5） 林分の周辺での bias を適当に補正しなければならぬ．
6） 格子点の確率抽出が事実上不可能である．つまりその結果の精度については何等客観的な保証は与えられない．
7） 胸高断面積は通常蓄積を求めるために測定されるわけであるが，この方法によつて求めた胸高断積総和から蓄積を出すことは非常な誤差をともなう．

等があげられる．以下この各々について述べてみる．

まず 1）は under estimate の原因となるものであるが，これをさけようとすれば，この危険の生ずる可能性のある樹立に対しては，結局一々その胸高直径を実際に測定しなければならない．

2）及び 3）はまづ平地に生えた通常の針葉樹にあつては問題がないが，斜立する樹木に対しては測定がむずかしい．斜面に於て斜立する樹木を測定する場合には更に面倒になつて，この方法では測定出来ない．傾面，下草等の障碍によつて調査が不能な場合も非常に多いと考えなければならない，又遠方の樹木の丁度胸高直径のところを望んで測定出来るかどうか疑問である．

4），5）は測量についての手間である．林分の面積を求めるのは無駄ではないが，5）の補正を行うために林分の周囲に一定巾の縁を付けるための測量に要する手間は相当なものである．地形によつては不可能な場合も生ずる．

次に 6）であるがこれがこの方法の最も大きい欠点である．この方法で，ランダムに標本格子点を抽出するためには，まずその林分の平面地図を作りこの地図を二つの座標軸によつて細分し，その格子点を等確率で抽出し，再び測量によつて抽出された格子点に到達するわけである．このとき問題になるのは樹幹の中に格子点が落ちた場合と測量の誤差である．前者は樹を切り倒さないかぎり bias を生ずるわけであつて，予め樹の中の格子点を除外して格子点を抽出し結果を出すためには，どんな太さの木がどこにあるかまで全林の様子が明確にならないかぎり（もしこれが出来るくらいなら調査はすんでいるはずである）不可能である．木を切り倒して調査を強行すれば次に測量の誤差の影響を考えなければならない．つまり実際の格子点は図上で抽出された格子点と一致せずその周囲にある散らばりをもつて分布することになるのでその時生ずる分散と bias の order を計算しなければならないが，これもその林分の樹木の立ち方によつて変つてくるので問題はむづかしい．この方法に追従する者の多くははこの格子点の抽出を次のように考えている．つまり格子点を充分細かにすれば結局標本格子点を抽出することは平面（林分）内の点をランダムに抽出するこ

とと同一視してよい，更に標本格子点は森林の中を任意にあるきまわつて，そのうちの適当な所を撰べばよいと（この場合樹幹中の格子点を除外したために生ずる bias は無視出来るものとしている）．この考え方は次のような理由によつて疑問の点を残す．

充分細かい格子点と云う前提に対して，歩巾が問題にならないかどうか，地形，地面の様子や個人差によつて場所によつて抽出確率が変らないかどうか．（この場合抽出は水平面について考えていると云うことが大切である）更にランダムに歩くと云つても，各格子点を独立に抽出するためにはどんな歩き方をすればよいか．これらの問題はすべて格子点の抽出の困難性を示すものであろう．これらの点を解決しない限り調査精度の客観的な保証は与えられない．

最後に 7) の問題であるが，回帰直線によつて胸高断面積の総和から蓄積を求め得るほど高くはない．つまりこの場合相関係数が高いと云う條件のほかに，回帰直線のまわりの分布の模様が問題になり，推定結果の精度を一般に求められない．

以上述べて来たように Bitterlich の方法はこれを実施する上に幾多の欠点を有しているので，これを実際の森林調査に用いることは無暴と云うことが出来よう．

増山氏は Bitterlich の方法のうち，前述の欠点 1) を除く方法を考えた．その方法の概略は次の通りである．

面積 S_i 周囲 l_i の任意の凸図形の周りに巾 α なる縁をつけ出来た図形の面積 u_i は次の式で表わすことが出来る．

$$u_i=S_i+\alpha l_i+\pi\alpha^2$$

ここで S_i l_i を各樹木の胸高断面積及びその周の長さと考えて，Bitterlich の方法と同様の格子点を作り，N 個の格子点中各 u_i 内に落ちた格子点数 N_0 を数えれば N_0/N は U/S（但し S はこの林分の前面積，$U=\sum_{i=1}^{M}u_i$，M はこの林分の総本数）の近似値となり

$$U=S_0+\alpha L+\pi\alpha^2 M$$

$$\text{但}\quad S_0=\sum_{i=1}^{M}s_i$$

$$L=\sum_{i=1}^{M}l_i$$

の式が成立する．N 個の格子点が n 個を抽出し，各点から半径 α（水平距離の意味で）内にある樹木の本数 x_j を数え $\sum_{j=1}^{n}x_j/n$ を作れば，この値は V/S の標本値となる．そこでこのような方法を α を変えて三度（これをα，β，γ とす）行えば，三つの式を作ることが出来るからこの式をといて，胸高断面積の総和，総本数，胸高断面の周囲の長さの総和の推定値を求めることが出来ることになる．

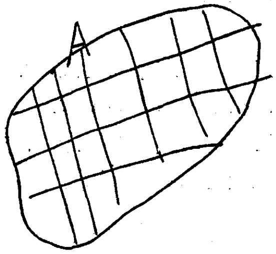
この方法によれば確かに樹木の胸高断面が凸形でありさえすればなんでもよいわけであるが，2) 以下の欠点はすべて又含んでいるわけである．しかも格子点から樹木までの水平距離を一本一本計るのであるから（この場合簡便なレインジファインダーでは精度の点で実用にならないし又，暗い林の中で樹木の胸高で調査者に一番近いところに――この点は調査理論から本質的なものである――ファインダーを合せることは出来ない．Bitterlich の方法の一点に立つただけで毎木に触れる必要がないと云う特色を失することになり，手数から云つても area sampling による毎木調査よりはるかに大がかりなものになつてしまう．（石田）

第 II 章

森林調査に於ける sampling unit の問題

森林調査に於て sampling により蓄積を推定せんとするとき，調査すべき林分をどう切つて

sampling unit をとれば一番よいかという事が一つの問題である．次にそれを論ずる．



この場合の調査は次の如きものである．

今 A を調査せんとする area とし，それを図の如く block に切り，各 block を sampling unit とす，その中の蓄積を $X_1,\cdots,X_N$ とす，但し N は block の数である．総蓄積は $V=\sum_{i=1}^{N} X_i$ である．

今 N けの block から n を等しい確率を与えて抽出する．それを $x_1,\cdots,x_n$ とし $\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$ とする時 $v=N\bar{x}$ により V を推定する．信頼度を n にとる時

$$P_r\left(|v-V|\leqq Nk\sqrt{\frac{N-n}{N-1}}\frac{\sigma}{\sqrt{n}}\right)\geqq\lambda=1-\frac{1}{k^2}$$

但し，σ^2 はこの時の母分散，即ち推定したい値 V が $v-Nk\sqrt{\frac{N-n}{N-1}}\frac{\sigma}{\sqrt{n}}$ と $v+Nk\sqrt{\frac{N-n}{N-1}}\frac{\sigma}{\sqrt{n}}$ の間にあるという事が確率 λ をもつていえるわけである．

この時の数度（相対精度）y は

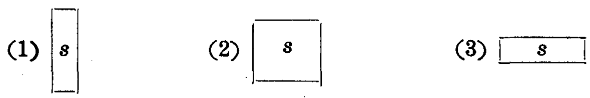
$$y=\frac{N\sqrt{\frac{N-n}{N-1}}\frac{\sigma}{\sqrt{n}}}{V}=\sqrt{\frac{N-n}{N-1}}\frac{1}{\sqrt{n}}\frac{\sigma}{\bar{X}}$$

但　$\bar{X}=\frac{1}{N}\sum_{i=1}^{N}X_i=\frac{V}{N}$　即母集団の平均で与えられる．

実際調査の場合は，先づ調査したい area を周測して，それに基づき地図を作成し，その上で block を切つて，それから sample spot を抽出し，次に当つた spot は確かな目標から測量して行つて，その spot を確認する．それから，その spot を周測して地図上の block を現地に確定する．それから，その中の毎木調査を行つて，かくして $x_1,\cdots,x_n$ の値を知り，従つて v の値を知る．

精度の式　　$y=\sqrt{\frac{N-n}{N-1}}\frac{1}{\sqrt{n}}\frac{\sigma}{\bar{X}}$

をみるに，これは $\sigma/\bar{X}$ に関係する，然るに，area を block に切る切方により，従つて母集団の作り方により $\sigma/\bar{X}$ は変つてくる．故に，切方により y がどう変るかをみるには，先づ $\sigma/\bar{X}$ が切方の函数としてどのように与えられるかを知らねばならぬ．所で切方は samplin unit にする block の形や面積に関係する．形は無数にとりうるわけだが，ここには，どの block もほぼ同じような形に，又ほぼ等面積になるように切るものとして考えてゆく事にする．そして形として代表的に



の3通りを考えてゆく．然るときは，形を指定すれば切方は Block の面積 s の函数となる．よつて切方 s と呼ぶ事にする．

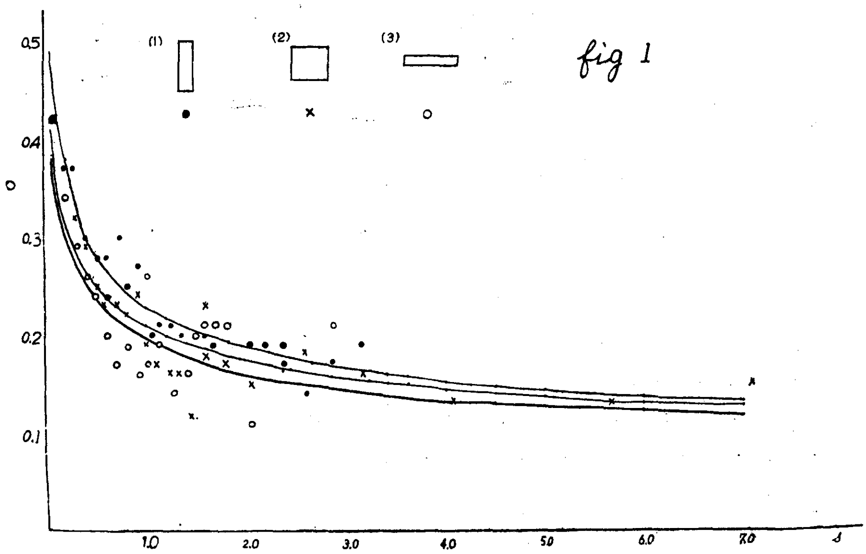
$Z=\sigma/\bar{X}$ は切方の函数．即ち形を指定すれば

Z は s の函数：$Z=Z(s)$

これがどのような函数形をもつかをこの前の山梨縣で調査からしらべた．その大体を述べる．先づ現地に於て，一つの林班をえらび，ここを全部毎木調査でしらべた．この林班は檜と唐松の混合林で檜が大部分であつた．広さは2町8反4畝，令級は5令級であつた．

毎木の際に地図上に一本一々の木の位置を目測で plot した．次に地図に対しいろいろな切方を試みて，それに対する Z を計算した．

但し木の位置を plot するのに目測によつたため，正確な位置とはいえない．故に一本一々の木を考えるよりもむしろ，area を先づ最小の単位 block に切つておき，その各々の中に入つている木の蓄積を計算しておき，次に area を sampling unit の block に切る時は，最小単位の block を集めるような切方をした．その方が計算の便からもよいと思える．こうしていろいろの切方 s に対する $Z(s)$ を計算して，それをグラフ用紙に点をとつて表すと fig. 1 の如くなつた．（図中●，×，○は夫々形が (1), (2). (3) の場合の点をあらわす．）これよりみれば $Z(s)$ は $a(s)^{-\lambda}$，形をなすと思はれる．勿論 $Z(s)$ の形は調査すべき林分の状態，広さ等に関係する．故に唯一箇所だけの調査から，上のように断定する事は出来ないわけだが，恐らく一般に上の形になるのではないかと思われる．以後 $Z(s)=as^{-\lambda}$ を仮定して論を進める事にする．



上の場合の点のちらばりを $Z=as^{-\lambda}$ であてはめて次の数値を得た．

(1) に対し　$Z=0.23102s^{-0.30112}$

(2) に対し　$Z=0.20951s^{-0.26928}$

(3) に対し　$Z=0.19809s^{-0.28455}$　（s は反）

fig. 1 よりみると (3) が一番下に表われているので $\boxed{}$ の形が一番よい事になるが，点は大体まざつているので，3つの形に大した差はない．もし同数の sample をとるものとすれば，$Z=\sigma/\bar{X}$ が精度を与えるわけだが，Z は s が増と，次第に減ずるから，s が大きい程よい事になる．然し，Z は s が一反歩位迄の時は急激にへるが，それ以後の減少は緩慢となり，s が2反5畝以後は殆ど減少しなくなる．故に cost を考えず，従つて回数の sample をとるとした場合は1反～2反位に切るのがよいという事になる．

それでは実際に調査を行う場合を考えて，cost を加味した時の一番よい切方はどうなるかを次に考えると，その場合，次の2通りの問題を考えればよいであろう．

(i) 調査に要する総 cost を与え，精度を最良にする切方

(ii) 推定に対して或精度を要求するとき，cost を最小にする切方

その場合，area を矩形と考え，sampling unit を正方形にとるものとす．実際には area は矩形でないが，必要なら area を区分けすれば各部は近似的に矩形とみてよいであろう，例えば fig. 2 のように

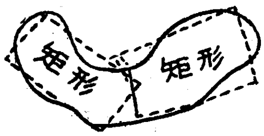
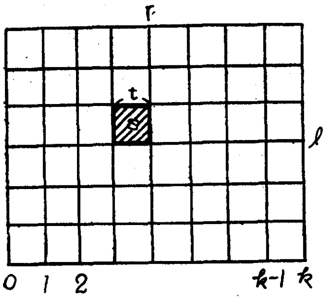


fig. 2

故に area が矩形の場合の問題がとければ充分である．この時は sampling unit を正方形としたが，前に sampling unit の形を3通り

考えた時の, $\sigma(s)/\bar{X}(s)$ は三つとも, 結果に影響する程の差のないことが明らかであるから, そのうちの正方形をとつて考えてみると, この場合, 切方は sampling unit, の一辺の長さ t で指定されるから, 「切方 t」とよぶ事にする.



面積A, 周圍ノ長サL, r≧ℓ

fig. 3

今 area を fig. 3 のようなものとし, これを N ヶの sampling unit に切るとす. 但し area の底辺がほぼ等分されるような切方 t のみを考える. (計算に当つては t が連続的に変るように扱う) よつて丁度 k 等分されるとすれば, $k=r/t$ (i), (ii) を考える場合に必要な式は, 精度及 cost の式である. 精度の式は, 切方 t, sample 数 n の時

$$y=y(t,n)=\sqrt{\frac{N-n}{N-1}}\frac{1}{\sqrt{n}}\frac{\sigma(s)}{\bar{X}(s)}$$

で与えられる. ここで $\dfrac{\sigma(s)}{\bar{X}(s)}as^{-\lambda}=at^{-2\lambda}$

又 $N\gg 1$ と考えてよいから $\dfrac{N-n}{N-1}\dfrac{1}{n}\fallingdotseq\dfrac{1}{n}-\dfrac{1}{N}$

$$\therefore \quad y=y(t,n)=\sqrt{\frac{1}{n}-\frac{1}{N}}\,at^{-2\lambda} \qquad \text{(a)}$$

次に cost の式を考えると

cost は次のようなものからなると考えてよいであろう.

(イ) area を周測するに要する cost

(ロ) sample spot 確認に要する cost (sample spot に到達迄)

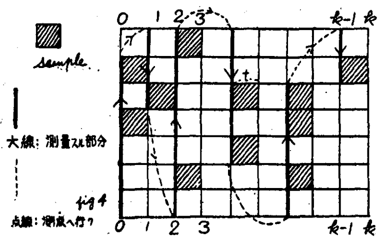
(ハ) sampling unit を周測して確定するに要する cost

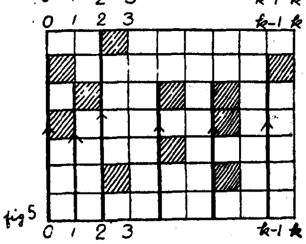
(ニ) 毎木に要する cost

今測量の cost は測る長さに比例し毎木の cost は面積に比例すると考え, 単位当りの測量及び毎木の cost を夫々 β, α とすれば, 切方 t で saple n けをとるとき, (イ)につき, $L\beta$, (ハ)につき $4n\beta t$, (ニ)につき $n\beta t^2$, 今(ロ)につき要する cost を D とすれば,

$$C(t,n)=n\alpha t^2+4n\beta t+D+L\beta$$

次に D がどう与えられるかを考うると sample spot の確認には, ある確かな目標から測つてゆくわけだが, 目じるしとしては周測の時の測点がよいと思われる. 今 area を切つた時 area の上辺及び下辺の分点に周測の時の測点があるものとすれば, sample が例えば図のようにおちた時図4の太線の部分を測ればよい. sample のいろいろなおち方に対する測るべき長さの平均を考える事にすれば, fig. 5 のように太線の部分を測ると考えても同じ事になる. 即ち各列に於て area の下辺から一番遠い sample までの距離を測るとして, そのいろいろの sample のおち方に対する平均値を考えそれが sample spot の確認のために測るべき距離 $\rho(t,n)$ を与えるとすればよい. 然れば $D=\beta\rho(t,n)$. $\rho(t,n)$ を計算すれば





$$\rho(t,n)=kl\left[1-\frac{k}{n+1}\left\{1-\left(1-\frac{1}{k}\right)^{n+1}\right\}\right]$$

となる.

$$\therefore \quad C(t,n)=n\alpha t^2+4n\beta t+\beta\rho(t,n)+L\beta \tag{b}$$

$\rho(t,n)$ は形が複雑で数値計算に不便なので，これを近似式でおきかえると，それには，t 従つて k をいろいろかえて（但し k の実際に考えうる範囲で）$\rho(t,n)$ グラフをかいて見て，グラフに一番近い直線をとつた．こうして実用的な近似式として

$$\rho(t,n) = 0.16\,ln+2l$$

が得られた．よつて

$$C(t,n) = n\alpha t^2+4n\beta t+\beta(0.16ln+2l)+L\beta \tag{b'}$$

この $y(t,n)$, $C(t,n)$ を用いて (i), (ii) が求められる．

（i） C を与えた総 cost とすれば問題は $C(t,n)\equiv C$ の下に $y(t,n)$を最小

（ii） ε を与えた精度とすれば問題は $y(t,n)\equiv\varepsilon$ の下に $C(t,n)$ を最小

これは微分を用いて計算をした結果は，最良の切方を与える t は次の式から求まることになる．

（i）の場合

$\overline{C}=C-L\beta$（周測以外の総 cost）とおくとき t は

$$(A\alpha-\overline{C}+2l\beta)(1-2\lambda)t^2+2A\beta(1-4\lambda)t-0.32A\lambda l\beta=0 \tag{c$_1$}$$

その時の sample 数は

$$n=\frac{C-2l\beta}{\alpha t^2+4\beta t+0.16l\beta} \tag{c$_2$}$$

その時の精度は

$$y=\sqrt{\frac{1}{n}-\frac{t^2}{A}}\,at^{-2\lambda} \tag{c$_3$}$$

から求められる．

（ii）の場合

t は

$$A\frac{\varepsilon^2}{a^2}\alpha(1-2\lambda)t^{4\lambda}-2A\frac{\varepsilon^2}{a^2}\beta(4\lambda-1)t^{4\lambda-1}-0.32\lambda Al\frac{\varepsilon^2}{a^2}\beta t^{4\lambda-2}-2\beta t-0.16l\beta=0 \tag{d$_1$}$$

その時の sample 数は

$$n=\frac{A}{Adt^{4\lambda}+t^2} \tag{d$_2$}$$

その時の area 周測以外の cost は

$$n\alpha t^2+4n\beta t+\beta(0.16n+2)l \tag{d$_3$}$$

から求められる．

α,β の数値は森林調査の実際の場合を考えて割出すべきものであるが，大体次の如く与えてよかろう．

α:

(イ) 樹高を時々正確にはかり．胸高直経 2 方向の場合：80 本を 2 人時

(ロ) 樹高目測，胸高直経 2 方向の場合：100 本を 2 人時

(ハ) 樹高目測，胸高直径 1 方向の場合：120 本を 2 人時

β:

(イ) 200 m 3 人時：山での測量として普通の場合

(ロ) 400 m 3 人時：山での測量容易の場合

この前の山梨の現地調査を参考にすれば，

α の(イ)，(ロ)，(ハ) を $\alpha_1,\alpha_2,\alpha_3$，$\beta$ の(イ)，(ロ)を β_1,β_2 とするに

$\alpha_1=0.00243$ 人時/m^2

$\alpha_2=0.00194$ 人時/m^2

$\alpha_3=0.00162$ 人時/m^2

$\beta_1=0.0150$ 人時/m

$\beta_2=0.0075$ 人時/m

なる数値を得る．

この α,β の値を用いれば，任意の area に対し，(i)，(ii) の場合につき (c) 式又は (d) 式から最もよい切方が計算出来る．今この前の山梨での調査の area につき，実際に計算してみた．結果その場合の数値は

$$\lambda=0.2846,\quad a=1.414,\quad A=28400\,\mathrm{m}^2,\quad l=119\,\mathrm{m},\quad r=239\,\mathrm{m}$$

α,β としては α の(ロ)，β の(イ)を用いる事にする．$\alpha_2=0.00194$ 人時/m^2，$\beta_1=0.015$ 人時/m.

（i） この時，与えた area の周測以外の総 Cost $\overline{C}$ と，その場合の最良の切方 s との関係は fig. 6 で与えられる．

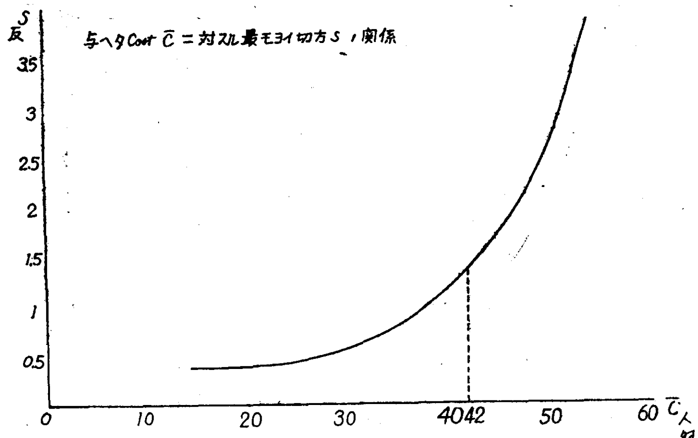


fig 6

α_2,β_1 に対し，$\overline{C}$ の値は 42 人時値位が標準と思はれるが，特にその時は

$s=1.4$ 反 $\quad n=8 \quad y=5.2\%$

（ii） この場合に与えた精度とその場合最良の切方 s との関係は fig. 7 で与えられる．

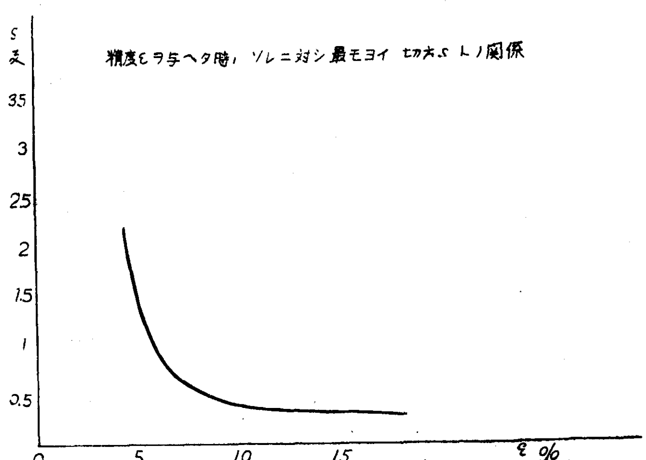


fig 7

特に $\varepsilon=5\%$ の時 $s=1.6$ 反　$n=7$　$C=45$ 人時

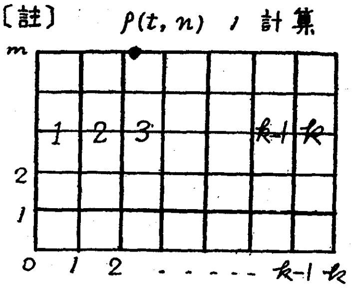
もし α, β として α の(イ), β の(ロ)を用いれば $\alpha_1=0.00243$ 人時/m², $\beta_1=0.015$ 人時/m この時は

(i) $\bar{C}=42$ 人時 に対し　$s=7$ 畝　$n=12$　$y=5.6\%$

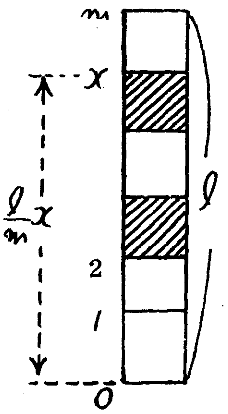
(ii) $\varepsilon=5\%$ に対し　$s=8$ 畝　$n=12$　$C=150$ 人時

以上よりみて，この場合の一番よい切方は1反～1反5畝位であると思はれる.

〔註〕



area を図の如く sampling unit に切りこれから n けの sample をとるのに，当つた sample が先づ k けの列（縦に細長い k けの矩形）に多項分布をなしておちるとす. 即ち，1, 2, $\cdots\cdots$, k の列に夫々 $n_1, n_2, \cdots\cdots n_k$ けおちる確率が $P_r(n_1, \cdots\cdots, n_k)=n!/n_1!\cdots\cdots n_1!(1/k)^n$ とす. 但し $\sum_{i=1}^{k} n_i=n$



このようなおち方をした時，一つの列に着目して，例えば三番目の列を考える. この列には m けの場所があつて，ここに n_i けの sample がおける. それが等しい確率でおるとすれば，一つのおち方の確率は $1/\binom{m}{n_i}$ である. 今下辺から一番遠い sample が下から x 番目を占める場合の数は $\binom{x-1}{n_i-1}$ 但し $n_i \leqq x$ ∴ この列で測るべき距離の平均値は

$$\sum_{x=n_i}^{m}\frac{l}{m}x\frac{\binom{x-1}{n_i-1}}{\binom{m}{n_i}}=\frac{n_i}{\binom{m}{n_i}}\frac{l}{m}\sum_{x=n_i}^{m}\frac{x(x-1)!}{n_i(n_i-1)!\,(x-n_i)!}$$

$$=\frac{ln_i}{m\binom{m}{n_i}}\sum_{x=n_i}^{m}\binom{x}{n_i}\quad ここで等式\ \sum_{x=n_i}^{m}\binom{x}{n_i}=\binom{m+1}{n_i+1}$$

を用うれば，

$$上式=\frac{l}{m}\frac{n_i}{\frac{m!}{n_i!\,(m-n_i)!}}\frac{(m+1)!}{(n_i+1)!\,(m-n_i)!}=\frac{l}{m}m+1\frac{n_i}{n_i+1}=\frac{n_i}{n_i+1}l\left(1+\frac{1}{m}\right)$$

他の矩形についても同様 ∴ 1, 2, $\cdots\cdots$, k 列中＝夫々 $n_1, \cdots\cdots, n_k$ 個おちるような sample のあらわれの時に測るべき距離の平均値に $\left(1+\frac{1}{m}\right)l\sum_{i=1}^{k}\frac{n_i}{n_i+1}$, $m \gg 1$ の時は $\fallingdotseq l\sum_{i=1}^{k}\frac{n_i}{n_i+1}$

∴ これを $P_r(n_1, \cdots\cdots, n_k)$ に関し平均値をとり，それを $\rho((t, n)$ とすれば

$$\rho(t,n)=l\sum_{i=1}^{k}E\left(\frac{n_i}{n_i+1}\right)$$

$$然るに\qquad E\left(\frac{n_i}{n_i+1}\right)=1-E\left(\frac{1}{n_i+1}\right)$$

$$且\qquad E\left(\frac{1}{n_i+1}\right)=\frac{k}{n+1}\left\{1+\left(1-\frac{1}{k}\right)^{n+1}\right\}\qquad (i=1, \cdots\cdots, k)$$

$$\therefore\qquad \rho(t,n)=lk\left[1-\frac{k}{n+1}\left\{1-\left(1-\frac{1}{k}\right)^{n+1}\right\}\right]$$

（植松，宇沢）

第 Ⅱ 章

森林資源調査に於ける area sampling

§1.　まえがき

一つの縣又は日本全国と云つたような比較的広い地域の森林資源の調査を sampling によつて実施する場合，既成の適当な sampling unit が見当らなかつたり，又あつてもその処理その他の理由によつて使用出来ないことが多い．このときは地形図を利用した area sampling の方法が都合がよいので次に5万分の1地形図を利用した全国森林資源調査の企画の概略を述べてみる．

§2.　調査単位と標本数（調査地点数）の決定

この調査は国土資源調査の一端として，日本全国森林の総蓄積を信頼度 95% で誤差5%以下の精度のもとに推定を行うとともに森林計画上基礎となる諸資料（樹種分布，生長量，土質，土地利用状況，施業の状態等）を得るために林野庁計画課が行つたものである．

この調査はその性格上多方面にわたる調査項目をもつので，特定項目についてだけの層別をさけ，一応地点抽出の便宜上，縣別のみの層化に留め，又，調査精度と要求する結果の性格の点から sub-sampling は行はず，日本全土から直接に調査単位を抽出する方針をたてた．調査単位としては行政区画等の既成の単位が，その大いさ，境界標識の不明確さ基礎資料蒐集の困難性等から不適当であつたので，新に日本全土を調査に適当な一定面積の單位に区分し，その単位を等確率で抽出する手段をとることにした．

調査地点数は上記の方針上

$$0.05 \geqq \frac{2\sigma}{\sqrt{n}\,\bar{X}}$$

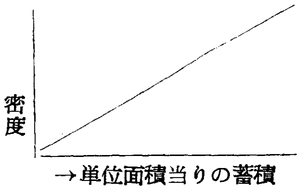
但　n：　調査地点数

$\bar{X}$：　各調査地点の蓄積の平均値

σ：　地点別蓄積の標準偏差

により定めることが出来る．

ここで変異係数 $\sigma/\bar{X}$ の値であるが単位面積当りの蓄積の分布は下に示すような0に偏したL分布をなすと考えられる．この場合の変異係数はまず1から1.5程度と思つていれば大過はないであろう．そこで今



$$\frac{\sigma}{\bar{X}} = 1.3$$

とすれば

$$0.05 \geqq 2\frac{1.3}{\sqrt{n}}$$

つまり

$$n \geqq 2704$$

となり，調査地点数 n は 3000 でまず充分である．

そこでこの 3000 の調査地点を各支庁，都，府，県別に次式によつてその面積に比例して，配分した．

$$k_i = \frac{a_i}{A} n$$

但　n：　総地点数（3000）

a_i：　各縣の全面積（地理調査所資料）

$A = \sum a_i$（日本の全面積）

k_i：　（縣の割当て地点数）

§3.　調査単位の面積の決定

調査単位の面積は大きいほど精度はよくなるわけであるが，第2章の sampling unit の項で述べるように，一般の林分ではいたずらに調査地点の面積を大きくすることは精度と手間を比較して損であつて大約1反から3反程度が望ましい．又土地利用の状況や地形等の土地の性格を示す調査項目を考えた場合には調査単位が大きくなるとその中が複雑化して分析上面白くない．更に調査班の構成，費用等の点からの限度もある．これらの理由を綜合して，我々の場合一調査単位の面積は 50 m×30 m，約1反5畝とした．

§4.　調査地点の抽出

我々の場合，日本全土を 30 m×50 m の小区画に区分して調査単位をつくり，そのうちから等確率で 3000 を抽出することを考えたのであるが現実にこれを直接行うことは不可能なので，次のような方法で抽出を行つた．

まず，日本全土を充分多くの等しい面積の区画に区分するような格子網を考え，この格子点の中からその 3000 を等確率で抽出する．更にその格子点を原点として，一定の方法でその点を含む 30 m×50 m の調査地点を撰定するのである．そして格子網は緯経度各 1′ 約 $1.8\times1.6\,\mathrm{km}^2$ によつて作ることにした．この大きさは我々の目的に対して充分な細かさをもつている．

ところが，この緯経度の網は緯度によつて変化するから，日本中どこでも目の大きさが同じとは云えない．しかしこのとき予め縣別に地点数を割り当てておいて（縣を一つの層と考えて）縣別に抽出を行えば実用上充分な精度で単位面積当り等しい確率で調査地点を定めることが出来る．

格子網として緯経度を利用するとすれば，最初に格子網をどう固定するかということを決めなければならない．勿論これはランダムに定めなければならないのであるが，我々は抽出の便宜上緯度については地理的な1分を，経度に対しては1分 10.4 秒の点を格子点とした．この決め方によつて調査理論上不都合を来すことはまずないと思われる．

§5.　現地調査の実際

地図上で抽出された調査原点を現地で確定して調査を実施するために，我々は次のような方法をとつた．

1.　調査原点位置を5万分の1地形図に図示し，その附近の基本図，航空写真等を利用して，その位置に到達する経路及び現地を確定する上に有効な標識を検討する．

2.　トラバース測量を利用するために調査原点を図示した地形図を平板にはり，現地附近で磁針偏角の決定を行う．

3.　調査原点附近で視野の出来るだけ広い地点に立ち三角点，嶺，峰，鞍部，その他の標識を利用して交会法によりこれを地図上に求める．

4.　この地点を起点として，調査原点を交会法及びトラバース法によつて確定する．

5.　ポケットコンパスによりの調査原点より真北に水平距離 50 m，真東に 30 m の区画測量を行つて，調査地点を定め，毎木測定，その他の調査を実施する．

以上のようにして定めた調査原点の位置の精度は ±100 m 以内であつた．この精度は我々の目的にとつて充分であつて，その手数もさほど面倒なものではない．

§6.　調査の精度

全国総蓄積の推定式は

$$v_T=\frac{A}{a}\sum_{i=1}^{n}v_i/n$$

但　A：日本の全面積

　　a：一調査地点の面積 $(30\times50\,\mathrm{m}^2)$

　　v_i：i 地点の蓄積

　　n：調査地点数(3000)

となりその分散 σ_{v_T} は

$$\frac{A^2}{a^2}\frac{\sigma_v{}^2}{n}$$

で与えられる．ここで $\sigma_v{}^2$ は調査単位間の蓄積の分散である．故に信頼度を 95% とすれば信頼巾 l は縣別による層化の効果，測量誤差を無視して

$$l=\pm 2\frac{A}{a}\frac{\sigma_v}{\sqrt{n}}$$

で計算出来る．

今 a を測定するのに地点で ε_i なる誤差があるとすれば，推定値 v_T の分散 σ_{v_T} は近似的に

$$\sigma_{v_T}^2=\frac{A^2}{a^2n^2}\left(\sigma_v{}^2+\frac{\sigma_a{}^2}{a^2}\bar{V}^2\right)$$

但 $\bar{V}$ は調査単位の母集団平均値

となる．

その他比率の推定の場合もまつたく同様にその推定値の分散を計算することが出来る．

（林，石田）

第 IV 章

森林調査に於ける回帰推定

§1. まえがき

各単位がもし，v_i の二つの標識をもつ大いさ N の母集団を考え，今 x_i の母集団平均

$$\bar{X}=\frac{1}{N}\sum_{i=1}^{N}x_i$$

が知られているとする．このとき N の中から等確率で n 個 $(x_1,v_1,x_2,v_i,\cdots,x_n,v_n)$ を抽出して x と v の間の一次回帰係数

$$b=\frac{\sum_{i=1}^{n}(x_i-\bar{X})(v-\bar{v})}{\sum_{i=1}^{n}(x_i-\bar{x})}$$

$$\text{但}\quad \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i\ ,\quad \bar{v}=\frac{1}{n}\sum_{i=1}^{n}v_i$$

を計算し

$$\bar{v}_L=\bar{v}+b(\bar{X}-\bar{x})$$

を以つて v の母集団平均値

$$\bar{V}=\frac{1}{N}\sum_{i=1}^{N}v_i$$

の推定値とするのがよく知られている回帰推定の方法である．

この方法は Watson (1937), Yates (1949) によつて林学に応用されている．Watson は一本の樹木についている葉の総面積を計算するために，x_i として一枚の葉の面積，v_i として一枚の葉の面積をとり，何枚かの葉の x_i と v_i を実測し，更に葉の全枚数 N とその全重量 $\sum_{i=1}^{N}x_i$ をしらべてこの方法を応用した．又 Yates は x_i として目測蓄積，v_i として実測蓄積をとり，この方法を利用して全林の目測値を一部の実測値で修正し全林総蓄積の推定を行つている．

この方法は比推定に比して次のような利点をもつている.

1) 比推定は非常に特別な場合のほかは bias estimatation であるが回帰推定は比較的実在性のある條件で unibias な推定を行うことが出来る.
2) 比推定の場合は有効な信頼巾を計算することが困難であるが, 回帰推定の場合は二三の條件でこれを行うことが出来る.
3) 比推定は唯一つの推定因子だけしか利用出来ないが, 回帰推定はその制限がない.

昭和 28 年度に林野庁計画課, 統計数理研究所が企画実施した森林調査はこの回帰推定法を利用したものである. つまり, 新森林法の実施に伴い, 森林計画の基礎資料として, 各森林区平均面積約 15000 町歩) を単位として, その総蓄積を推定し, 更に森林区を構成する各筆 (平均面積約1町歩) の蓄積を算出する調査を実施した.

調査方法はまず各筆について目測によりその蓄積を調査し, 更に N ケの全筆中から n ケの標本筆を等確率で抽出してその蓄積 v_i を調査し, 前に述べた方法によつて各森林区の総蓄積の推定を行つた. この場合回帰推定法の既成の理論では少々不充分な点があり, 二三の考察を行つたので, ここにそれを述べることにする.

§2. 最も理想的な場合

我々はまず n の標本によつて v と x との間の標本回帰直線

$$v_L = \bar{v} + b(x - \bar{x})$$

$$\text{但} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\bar{v} = \frac{1}{n}\sum_{i=1}^{n} v_i$$

$$b = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(v_i-\bar{v})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

を計算し x の代りに目測の平均筆蓄積

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

を代入して

$$\bar{v}_L = \bar{v} + b(\bar{X} - \bar{x})$$

を以つて, 筆の平均蓄積の推定値とし,

$$V_T = N\{\bar{v} + b(\bar{X} - \bar{x})\}$$

を以つて森林区の推定総蓄積 V_T としたわけである.

ここで x を固定した場合の v の平均値が x と一次的な関係があるとする. 即ち

$$v = \bar{Y} + \beta(x - \bar{x}) + e$$

$$\text{但} \quad \beta : b \text{ の母集団の値}$$

として, 更に e は平均値 0 で各 x について常に等しい moment をもつ確率変数とすると, N, n が充分大きければこの推定値は近似的に平均値

$$E(\bar{v}_L) = \bar{V}$$

分散

$$\sigma_L{}^2=\frac{\sigma_v{}^2(1-\rho^2)}{n}\left(1+\frac{1}{n}\right)$$

$$但\quad \sigma_v=\frac{1}{N}\sum_{i=1}^{N}(v_i-\overline{V})^2\ ,$$

ρ: v と x との間の相関係数

の Gauss 分布をなすことが証明出来る．(Annals of the I. S. M. vol. No. 1 1952) つまりこの場合は信頼度 95° で森林区の総蓄積は

$$N(\bar{v}_L\pm 2\sigma_L)$$

の間にあると云うことが出来るわけである．ここで仮定したことを云いなおしてみると次のようになる．

1) 目測の値 x_i と実測の値は直線的である．

2) 目測の誤差は x の如何によらず常に前の直線のまわりに同じ傾向で生ずる．

実際の調査の結果は調査員が非常によく訓練されているところではほぼ上の條件が満されていた．1) の條件はほぼ全部満されているが 2) の條件についてはしかし大半のものが以下に述べるような誤差の傾向をしめしている．

§3. 目測の誤差が x に比例している場合

目測の誤差はとかく蓄積の大きい筆に於いて大きくなりがちである．そして又蓄積の0の筆に於いては目測誤差は0である．つまり目測誤差の偏準偏差が x に比例している場合

$$\sigma_\varepsilon{}^2=\lambda x^2$$

である．

このときの推定値の分散を計算してみると近似的に

$$\sigma_L{}^2=\frac{\sigma_v{}^2(1-\rho^2)}{n}+\frac{\lambda}{n^2}\left(\overline{X}^2+2\overline{X}\frac{\mu_3}{\mu_2}-\frac{\mu_4}{\mu_2}\right)$$

但 μ_i は x の i 次の moment

となつて，一般に分散は大きくなるが sample 数が充分大きければ2次以上の moment も §2 の場合と同一の結果になる．しかしこの場合は n がどれほどになれば §2 と同様に取りあつかつてよいかは一般的に定めることは出来ないのであるが，我々の場合このような傾向をもつための λ の値は約0.1以下であるから変異係数の意味で $1/n^2$ の項は充分小さい．

比較的大きい筆から成つている森林区にあつては

$$\sigma_\varepsilon{}^2=\lambda(x+\alpha)^k$$

の型の場合があるこの推定値の分散は同様に

$$\sigma_L{}^2=\frac{\sigma_v{}^2(1-\rho^2)}{n}+\frac{\lambda}{n^2}\left\{(\overline{X}+\alpha)^k+k(\overline{X}+\alpha)^{k-1}\frac{\mu_3}{\mu_2}-\cdots\right\}$$

である．

§4. 目測蓄積の母集団平均がもとめられない場合

推定式

$$v_L=v+b(\overline{X}-\bar{x})$$

に於いて，N が非常に多い場合 $\overline{X}$ を求めることはかなりの手間を要する，そこではじめの n ケの sample とは独立に m ケの sample を抽出し $\overline{X}$ の代りに

$$\bar{x}_m = \frac{1}{m}\sum_{i=1}^{m} x_i$$

を使つた場合を考えてみるとこのときの推定値の分散は §2 と同じ條件のもとに近似計算で

$$\sigma_L^2 = \frac{\sigma_v^2(1-\rho^2)}{n}\left(1+\frac{1}{n}\right)+\beta^2\frac{\mu_2}{m}$$

となり m が n に比して充分大きければ §2 と同一の結果になり，各項の order がわかつている場合には $\bar{x}_m$ で $\bar{X}$ を代用することが出来る.

又一般に目測蓄積 a なる筆をこの方法で推定した場合，母集団の回帰直線からの分散は近似的に

$$\sigma_{aL}^2 = \frac{\sigma_v^2(1-\rho)^2}{n}\left\{1+\frac{(a-\bar{X})^2}{\mu_2}\right\}$$

となる．以上の分散の計算はすべて近似計算であるが，実用的にはこの程度の order で充分である。

§5. 回帰線が直線でない場合

同一森林区を多くの組の調査班が調査したり，又森林区の様相が地域的に非常にちがつている場合には回帰線が直線とならない場合がある．このときには異質的と思われる組に層別して調査を実施した．つまり

1) 調査班での層別
2) 目測蓄積での層別
3) 樹種，令級，林種による層別

等がそれである.

集計は各層別には回帰線が一次的であるとして，各層別に回帰推定を行い．その和をもつて総蓄積の推定値とした．この場合の分散は

$$\sigma_L^2 = \frac{1}{N^2}\sum_{i=1}^{k} N_i^2\sigma_{L}^2{}_i$$

但 k : 層の数
N_i : 各層の筆数
σ_{Li}^2: 各層の平均筆蓄積推定値の分散

である.

§6. 調査の企画

以上の考察によつて，我々は次のような調査の企画を行つた.

1) 層別 回帰線が直線であることが我々の條件であつたために，回帰線か直線にならない場合 §5 で述べた手段を講ずることが出来るようになるべく細部にわたる層別を行つた.
   つまり層別の項目は
   a) 目測蓄積
   b) 樹種
   c) 樹令
   (d) 樹種
   である.
2) 推定因子 推定の因子は前に述べた方法とまつたく同様に重相関係数を利用していくつでも取り上げることが出来るので，予備調査の結果まず次の因子を考えた.
   a) 目測蓄積 (x)
   b) 目測面積 (y)
   c) 令級 (z)

しかし最終的な結果として天然林，混交林等令級を確定出来ない林分があるので推定因子としては，x，y の二つを用いることにした．

3) 推定の方法 最小自重法により推定式

$$v_L=ax+by+c$$

の常数項 a，b，c を計算し，各の v と v_L との相関図を画く．

もしこのとき直線的な傾向がみられなければその要因をさがし §5 によつて，回帰直線を層別に分けて推定を行う．その必要を認めない場合には層別しないと考えて一つの推定式により推定を行う．

4) 信頼巾の計算 §2 の條件がみたされていないときには §3 等によつて別に分散の推定をする．

（林，石田）

§7. 結　論

以上のことを綜合してみると，結局のところ簡便法によつて蓄積を求めたのでは誤差の点で問題をのこすことがわかる．つまり蓄積を求めるためには樹高，胸高直径について念入りな毎木測定を行うことが必須であることになる．