鳥大農研報（Bull. Fac. Agric., Tottori Univ.）45 135～143（1992）

不偏分散の小標本特性に関する研究

笠原浩三*・今井鐳蔵**

平成4年6月30日受付

On Small Sample Properties of the Unbiased Variance

Kozo KASAHARA* and Raizo IMAI**

This research attempts to clarify the small sample properties of the unbiased variance using the Monte Carlo Experiments. The unbiased variance consists with the population variance on the infinity samples. But the consistency of the variance is not confirmed with small samples. The result of this research suggensts the following points: 1) Unbiased variance estimators have only a little bias on small samples; 2) Variance of the unbiased variance of small samples is very large, and the size of the variance can not be reduced by the repetition of experiments many times. Thus, we developed a bias function:

$f(n)=114.6073-0.3066n+0.0017n^2$.

Furthermore, we modified the t-distribution curve using this bias function. The t-distribution table was adjusted for a test of a statistical hypothesis. The adjusted t-distribution table allows us to estimate directly the significance level with the bias of the "unbiased variance".

緒　　論

平均μ，分散δ^2からなる母集団から抽出したn個の標本の平均値$\bar{x}$の期待値は母集団平均μに等しいことはよく知られており，無意識にこの性質を利用していることが多い。しかし標本分散推定量は一般に母集団分散に等しくはならず，それは母集団分散値よりもやや小さい値として得られる。すなわちやや小さ目の偏りをもった値であるため，偏差の二乗和をnで割らないで(n－1)で割って偏りのない値として使用している。この値がいわゆる不偏分散といわれるものである。

このように一般に母分散が未知である場合には，標本から得られる情報を生かして推計されたいわゆる不偏分散推定量を母分散の偏りのない推定量として用いている

*　鳥取大学農学部農林総合科学科情報科学講座

*　*Department of Agricultural Information Science, Faculty of Agriculture, Tottori University*

**鳥取大学農学部農林総合科学科経営管理学講座

***Department of Farm Business Manegement, Faculty of Agriculture, Tottori University*

が，その根底には不偏分散$u^2(=\Sigma(x-\bar{x})^2/(n-1))$は母分散$\sigma^2$に一致することを理論的に証明できることに依拠していることにほかならない。しかしここに現実的に１つの問題が指摘される。不偏分散が母分散に理論的に一致するといっても，それは抽出回数を無大限に近づけた場合に得られる現象であって，現実的に得られるデータはそれ程回数多く抽出可能なものではない。さらに問題となるのは抽出される標本数が実際にはそれ程多くはなく，小標本に限られているという点である。抽出標本数が少なく，しかも抽出回数が有限回，場合によっては現実問題としてたったの１度きりということもあり得るが，そのような場合には不偏分散推定量が母分散に一致するということはなんら保証されることにはならないのである[2]。現実にもし抽出標本数が少ないいわゆる小標本において，不偏分散推定量が母分散に必ずしも一致しないということが明らかにされるならば，それに伴って従来の統計学の標本理論に修正が迫られることになる。例えば，不偏分散を用いたｔ分布，ｆ分布曲線などに小標本の偏りに応じた補正を施すことが必要になるであろう。

かくして本論文では，小標本から得られる不偏分散推定量の不偏性，有効性，一致性などについて実験的に明らかにした上で，体系的な小標本特性に関する考察を試みるものである。さらに，不偏分散推定量の偏りの大きさを定量的に明らかにした上で，統計的仮説検定のためのｔ分布表について若干の修正を試みることとする。

分析方法と疑似乱数

小標本特性を解明する方法としては，数学的証明法と実験的に事実関係を確認するモンテ・カルロ法があるが，数学的証明法では極く簡単な例を除いては現実には極めて難解な数式展開となり，一般には接近が困難である。そのため本稿では，後者のモンテ・カルロ法[4,5]により不偏分散の小標本特性の解明を試みることとする。

かくして，モンテ・カルロ実験の対象とする母集団を設定することが必要になるが，ここでは予め母平均値，母分散値が分っていることが必要なので次のように人工的に母集団を設定し，母平均，母分散を計測しておくこととする。(統計学通論[1]，p. 219の矩形型乱数表右下隅の100個分を利用した)

48,25,43,68,30,　92,45,88,78,21,　50,86,53,39,22,
26,15,53,80,09,　57,48,51,00,73,　49,85,83,36,26,
60,58,77,24,24,　20,78,71,36,56,　79,25,46,44,39,
53,39,91,56,53,　82,39,43,44,08,　45,15,41,45,80,
25,21,14,50,08,　59,29,74,68,40,　48,17,57,68,60,
54,89,95,72,48,　30,62,42,92,63,　37,85,60,79,95,
50,96,71,75,86,　05,94,07,62,47,

これによって，この母集団の平均値μは51.84，分散σ^2として610.4544，を得ることができる。

このように予め母集団の平均，分散が既知の人工的な母集団を設定することによって，抽出される標本に基づいて得られる標本分散推定量との比較が可能になり，推定量特性が分析できることとなる。

また一致性などの一般的特性の確認のため一部に無限母集団を使用したが，この無限母集団は０から99までの整数で構成される一様分布として設定したものである。このように無限母集団を設定することによって，無限回の抽出の結果母集団の各要素の抽出率は等しくなり，無限母集団でありながら母平均，母分散を計算することが可能になるのである。かくしてこのようにして設定した無限母集団の平均値μは49.5となり，分散σ^2は833.25として得ることができる。

また，モンテ・カルロ実験の場合，実験結果に直接影響を与えるいま１つの条件として，使用する乱数特性問題をあげることができる。一般に使用可能な公認乱数の数は有限個でありしかもそれ程多い数ではない[7]。ここでは使用上便利なパーソナル・コンピュタ（FM—R70 HD)による疑似乱数を用いることとする。パソコンの疑似乱数についてその乱数性を正確にチェックすることは極めて難しい問題であるが[4]，最も重要な性質である周期性と一様分布性について吟味した結果を整理すると次のとおりである。

周期性について

初めの６桁が再び現われるまでの乱数発生回数は
918,899～918,904個の間

初めの７桁が再び現われるまでの乱数発生回数は
1,577,480～1,577,486個の間

初めの８桁が再び現われるまでの乱数発生回数は
8,388,608～8,388,615個の間

であるが，いずれもそれ以降の乱数は全く別の乱数となっており（したがって，これらの桁の数値が現われたのは全くの偶然によるものであり，疑似乱数の一回転を意味するものではない)，少なくても発生される8,388,615個の範囲においては循環していることは認められなかった。

一様分布性について

０から99までの一様乱数について発生させた1,200,000個についてその発生度数を整理し，相対誤差を

第１表　パソコンによる疑似乱数の一様分布性

区　間	発生度数	相対度数	相対誤差
$x \leq 10$	119566	0.099638	0.000362
$10 < x \leq 20$	119927	0.099939	0.000061
$20 < x \leq 30$	120239	0.100199	0.000199
$30 < x \leq 40$	119939	0.099949	0.000051
$40 < x \leq 50$	119415	0.099512	0.000488
$50 < x \leq 60$	120177	0.100147	0.000147
$60 < x \leq 70$	120259	0.100215	0.000215
$70 < x \leq 80$	119804	0.099836	0.000164
$80 < x \leq 90$	120219	0.100182	0.000182
$90 < x \leq 100$	120455	0.100379	0.000379
合　計	1200000	1.000000	—

注）FM—R70HDによる疑似乱数

求めると第１表のようになる。相対誤差のレンジを求めると4.37×10^{-4}となり，0.05%以下の誤差となる。

以上の吟味を踏まえて，本稿ではパソコンによる疑似乱数を用いることとする。

不偏分散の小標本特性の吟味

不偏性（unbiasedness）

推定量の望ましい特性を吟味する指標として不偏性（unbiasedness），有効性（efficiency），一致性（consistency）の３指標があるが[5,6,8]，不偏性とは推定量の期待値が真の値に一致する性質のことである。すなわち，ここでの問題に置きかえると，反復推計された不偏分散推定量の平均値が母分散に一致するかどうかということである。もちろん反復回数が無限回の場合には不偏分散は母分散に一致することが証明される。

いま母集団からn個の標本（x_1, x_2, ……, x_n）を抽出すると，その標本分散S^2の期待値は次のようになる。

$$E(S^2)=E\{\Sigma(x_i-\bar{x})^2/n\}$$
$$=E\{\Sigma x_i^2/n-\bar{x}^2\}$$
$$=\Sigma(x_i^2)/n-E(\bar{x}^2) \quad (1)$$

ここで，

$$E(x_i^2)=\sigma^2(x_i)+E(x_i)^2=\sigma^2+\mu^2 \quad (2)$$
$$E(\bar{x}^2)=\sigma^2(\bar{x})+E(\bar{x})^2=\sigma^2/n+\mu^2 \quad (3)$$

であるから，(1)式は次のようになる。

$$E(S^2)=n(\sigma^2+\mu^2)/n-(\sigma^2/n+\mu^2)$$
$$=\sigma^2(n-1)/n \quad (4)$$

したがって，

$$E(n/(n-1)\times S^2)=\sigma^2 \quad (5)$$

となり，かくして，不偏分散

$$u^2=\Sigma(x_i-\bar{x})^2/(n-1) \quad (6)$$

の期待値は次のように母分散に一致することとなる。

$$E(u^2)=\sigma^2 \quad (7)$$

しかしここで大切なことは，これは不偏分散の期待値であって，それは抽出回数を無限に続けたときに平均的な値として期待されることを意味しているものである。少数回または，たった１度きりの抽出結果では(7)式の成立をなんら保証するものではないことに注意することが大切である。少数回または，１度きりの抽出結果となると，不偏分散推定量の期待値よりも，むしろ不偏分散推定量そのもののバラツキの状態が重要な要素となってくるものと思われる。

かくしてとくに小標本の特性に焦点を当てて考察を行うこととする。以下に不偏分散推定量に関するモンテ・カルロ実験結果について集約してみよう。第２表はこのうち不偏性に関する特性の一部をまとめたものである。このモンテ・カルロ実験は，先に設定した100個の有限母集団から非復元抽出，復元抽出の両方法により反復10,000回くり返し行った結果の標本分散値と不偏分散値

第２表　不偏分散推定量の不偏性

イ）非復元抽出

標本数	標本分散	相対誤差	不偏分散	相対誤差
5	495.792	18.78	619.740	1.52
10	556.714	8.80	618.571	1.32
20	585.581	4.07	616.402	0.97
30	596.453	2.29	617.020	1.07
50	604.225	1.02	616.556	0.99
80	608.991	0.24	616.699	1.02

ロ）復元抽出

標本数	標本分散	相対誤差	不偏分散	相対誤差
5	490.685	19.620	613.356	0.475
10	549.615	9.966	610.684	0.037
20	579.519	5.068	610.020	0.071
30	590.326	3.297	610.682	0.073
50	599.829	1.741	612.071	0.264
80	603.035	1.215	610.668	0.035
100	605.004	0.893	611.115	0.108

注）母集団特性：平均＝50.45，分散＝610.454
抽出方法：100個の有限母集団からの抽出，反復回数10,000回
非復元抽出の標本数100個の場合は，標本分散＝母分散となり，さらに不偏分散＝母分散×(n−1)/nとなる。

のそれぞれの算術平均値である。

これによるとまず，復元抽出，非復元抽出とも不偏性において，不偏分散推定量は標本分散推定量より明らかに優っていることを認めることができる。ただし，非復元抽出の標本数80個の場合は標本分散は不偏分散より母分散により近い値を示しているが，これは有限母集団を100個として作成したことによるものである。すなわち，標本数を次第に増やして丁度100個にして両分散値を求めると，標本分散は母分散に一致し，不偏分散は母分散の$(n-1)/n$倍となることは明白である。つまり，有限母集団においては，ある抽出標本数からは標本分散は不偏分散より優ることとなるのである。

いずれにしても，有限母集団による攪乱影響を除けば不偏分散は標本分散より不偏性において優ることは明らかである。しかしここで１つ特徴的な現象を指摘したい。それは不偏分散推定量の標本数と相対誤差の関係についてである。標本数が極めて小さい場合は相対誤差がやや大きくなるものの，相対誤差は標本数によって余り影響を受けないように思われることである。

この点については以下の有効性との関係が出てくるので改めて考察の対象とする。

有効性（efficiency）

有効性とは推定量の分布が高度に集中していることである。すなわち，いかに不偏性に優れていても逆に分散が大きければ総体として望ましい推定量とはいい難くなる。つまり望ましい推定量とは，真の値θの周辺部分に集中して分布していて，偏りもないことが必要なのである。

さて，モンテ・カルロ実験の結果をこの有効性に則して集約すると，第３表のようになる。すなわち，不偏分散推定量のバラツキ（不偏分散推定量の標準偏差で表している）を標本数および抽出反復数との関係において整理したものである。これによると，標本数を増やしていくと不偏分散のバラツキはやや低下する傾向が見られるものの，抽出反復数を増やしていくことによっては余り影響を受けないようである。ただし，抽出反復数が極めて少ない10回または100回程度の場合は，推定結果が不安定で一定の傾向が攪乱されているようである。

すなわち，不偏分散推定量は抽出標本数を増やしていくことによって有効性を高めていくことができるが，抽出反復数を増やしても有効性を高める効果が認め難いことである。このことは現実問題では，第(7)式が抽出回数を無限大にした場合に期待できる結果であるとしても，小標本では不偏分散推定量のバラツキが大きく不偏性を期待することはかなり難しいことを意味するものである。

ここに我々は小標本において２つの点で注意を払うことが必要になってくる。すなわち，１つは小標本で生じる偏りと，いま１つは小標本の範囲においては有効性が劣ることである。我々は小標本における不偏分散の利用にあたっては，不偏分散推定量には偏りの面とバラツキの面の二重のバイアスを含むことに注意しなければならないといえよう。

一致性（consistency）

一致推定量とは標本数を無限に大きくした場合に目標とする値に完全に一致する推定量のことである。すなわち，母集団の未知母数θの推定量Tについて，任意の数εに対して

$$\mathrm{Lim}\ \mathrm{Pr}\ \{\,|\,T-\theta\,|\geqq\varepsilon\}=0 \tag{8}$$

第３表　不偏推定量の有効性

標本数	抽出反復数	不偏分散の平均値	不偏分散のバラツキ
5	10	498.020	426.225
	100	635.656	361.866
	1000	627.167	367.739
	10000	613.356	354.428
10	10	667.993	222.714
	100	613.432	204.695
	1000	618.457	215.576
	10000	610.684	224.617
20	10	564.692	199.285
	100	580.783	139.321
	1000	607.927	150.783
	10000	610.020	151.001
30	10	507.746	90.529
	100	616.797	136.265
	1000	609.322	131.530
	10000	610.682	122.546
40	10	615.042	102.122
	100	615.758	94.143
	1000	609.409	107.456
	10000	612.408	105.639
50	10	587.636	73.938
	100	605.553	102.920
	1000	612.806	94.026
	10000	612.071	93.938

注）母集団特性：平均＝50.5，分散＝610.454
　　抽出方法：100個の有限母集団からの復元抽出
　　不偏分散のバラツキは不偏分散推定量の標準偏差である。

第４表　不偏分散推定量の一致性

イ）抽出標本数が10個の場合

抽出反復数	不偏分散の平均値	不偏分散の相対誤差	不偏分散のバラツキ
5	774.011	7.1094	170.922
10	766.178	8.0495	232.250
50	750.388	9.9444	282.163
100	805.185	3.2481	256.074
500	808.620	2.9559	267.055
1000	824.736	1.0218	267.147
5000	832.555	0.0834	267.668
10000	833.219	0.0037	267.164
20000	833.452	0.0243	267.149
30000	832.836	0.0497	266.755
50000	831.998	0.1502	265.428

ロ）抽出標本数が40個の場合

抽出反復数	不偏分散の平均値	不偏分散の相対誤差	不偏分散のバラツキ
5	879.049	5.4964	148.293
10	819.597	1.6386	137.891
50	835.328	0.2494	127.781
100	815.781	2.0963	122.580
500	822.874	1.2453	119.990
1000	824.130	0.6145	124.397
5000	834.780	0.1836	122.993
10000	835.117	0.2240	121.368
20000	834.432	0.1418	121.481
30000	834.279	0.1235	121.782
50000	834.276	0.1232	121.665

ハ）反復回数を100回にして標本数を大きくした場合

抽出標本数	不偏分散の平均値	不偏分散の相対誤差	不偏分散のバラツキ
100	834.463	0.1456	67.198
200	839.579	0.7595	55.794
300	826.822	0.7714	47.037
500	830.314	0.3523	34.094
1000	834.374	0.1345	25.350
1200	832.903	0.0417	21.413
1400	834.204	0.1144	20.279
1600	834.996	0.2094	15.985
1800	834.274	0.1229	19.615
2000	834.987	0.2085	16.939

注）母集団特性：平均＝45.5，分散＝833.250
抽出方法：一様分布無限母集団からの復元抽出
不偏分散のバラツキは不偏分散推定量の標準偏差である。

が成り立つとき，推定量Ｔを母数θの一致推定量というものである。先の(7)式は標本数を大きくしていった場合にも保証されなければならないのである。

すなわち，一致性は標本数が大きい場合の大標本特性といえる。そのため本稿で対象としている小標本の範囲においてはこの一致性を期待することは困難となる。

表４表イ），ロ）は無限母集団を設定して，小標本の範囲内（標本数10と40の場合に限定して）において抽出反復数の増加の関係で不偏分散推定量の一致性を垣間見たものである。さらにハ）はいわゆる大標本まで拡張して，抽出標本数の増加に対する不偏分散の一致性について整理したものである。

これによると，小標本の範囲においても不偏分散の一致性の傾向を確認することができるが，大標本まで拡張すると不偏分散の一致性および有効性を明確に確認することができる。またこれらのことから，小標本においても抽出反復数を大きくしていくことによって一致性を高めることが可能であることである。しかしながら，抽出反復数を相当程度増していっても不偏分散のバラツキはほとんど縮小しないが標本数を増やすことによっては可能であることを改めて確認することができる。

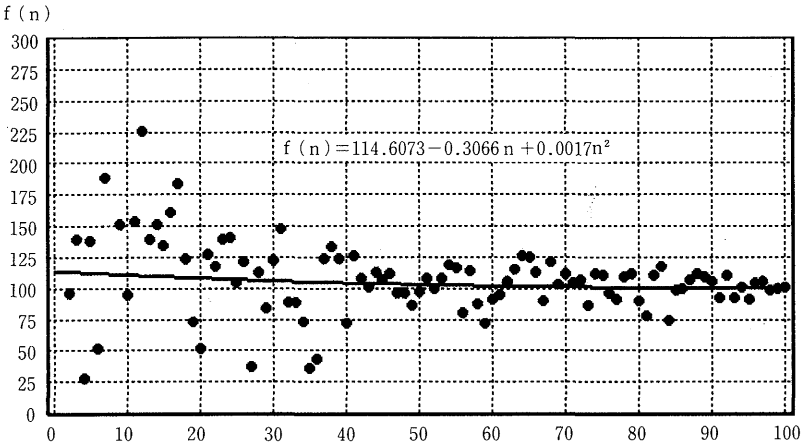
小標本における標本数と不偏分散推定値の偏りの関係については以下体系的に検討する。

小標本における不偏分散推定量の偏り関数

以上の考察の結果から，小標本においては不偏分散推定量は母分散に必ずしも一致しておらずなんらかの偏りが認められることである。さらに抽出標本数が少ない場合は勿論のこと標本数が大きい場合においても不偏分散推定量のバラツキは相当大きなものであることが確認された。しかし，無限母集団を設定の上，抽出標本数を大きくしていくと，不偏分散推定量は確実に母分散に近づけることができ，一致性のあることを確認することができた。

ここではさらに先の部分考察を総括する意味で不偏分散の小標本特性を標本数との関係において体系的に明らかにしてみたい。第１図は抽出標本数を２個から100個まで変化させ，各標本数毎に不偏分散量を2000回くり返し推計したものの算術平均値をプロットしたものである。推計式はこのプロット点を対象に回帰分析により推計したもので，不偏分散推定量の偏り関数と称すべきものである。その偏り関数は次のとうりである。

不偏分散推定量の偏り関数（100個の有限母集団）



第1図　不偏分散推定量のバラツキと偏り関数
(100個の有限母集団から抽出されたn個の標本から推計された不偏分散推定量2,000回分の平均値，100が母分散に一致する値)

$$f(n)=114.6073-0.3066n+0.0017n^2$$
$$(.5975)\quad(.3492)$$

決定係数$R^2=0.0125$

推計結果は，括弧内の通常のt-検定値，および決定係数R^2の大きさから判断して統計的仮説検定を必ずしもパスできるものではないが，不偏分散には標本数に対応した一定の偏りのあることは確かである。それは真の値すなわち母分散よりやや大きい方に偏っているようである。

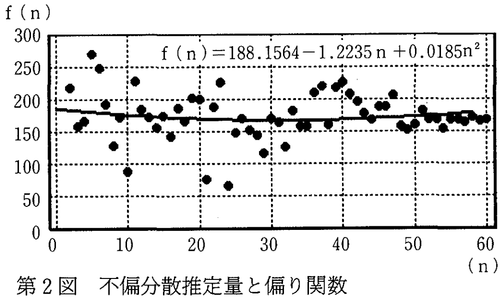
さらに60個の有限母集団を対象にした不偏分散推定量のプロット図は第2図に示すとうりであり，推計した偏り関数は次のとうりである。

不偏分散推定量の偏り関数（60個の有限母集団）

$$f(n)=188.1564-1.2235n+0.0185n^2$$
$$(1.5975)\quad(.9961)$$

決定係数$R^2=0.0202$

同様にして各nに対応する不偏分散推定量の偏り関数



第2図　不偏分散推定量と偏り関数
(60個の有限母集団を対象にした場合)

を推計することができるが，これらのことから，不偏分散推定量は母分散よりやや大きい値に偏っていることが認められるものである。

なおここで注意しておきたいことは，回帰分析の推計に使用したプロットされた数値である。この数値はいわゆる100個の有限母集団からくり返し推計された不偏分散推定量2,000回分の平均値である。したがって，もしくり返し回数を変えた場合にはその平均値も変ってくることになる。一般にはくり返し回数を大きくするとプロット点は回帰式の回りに集中してくることが予想される。よってここで推計された偏り関数のt-値，決定係数の値にはそれ程決定的な意味を持っていないことに注意することが必要である。

検定のためのt分布統計量の偏り修正

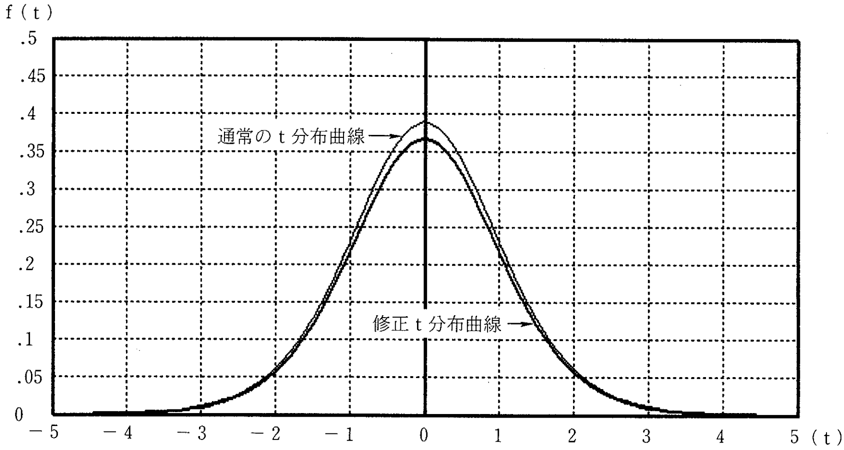
母平均μ，母分散σ^2の母集団からn個の標本を抽出するときの標本平均の分散は母分散σ^2の$1/n$に等しいことはよく知られている統計理論の最も基本的なものである[1,9]。したがって，次のように定義される統計量zは基準型の正規分布をすることとなる。

$$Z=(x-\mu)/(\sigma/n^{1/2}) \tag{9}$$

しかしながら，実際に統計的仮説検定などを行う場合には母分散σ^2は未知であることが多い。その場合にはσ^2の代りに標本から計算可能な第(6)式に示される不偏推定量u^2を利用する。その統計量をtとすると，

$$t=(x-\mu)/(u^2/n)^{1/2} \tag{10}$$

となり，tは自由度n-1のt分布に従うこととなる。



第3図　不偏分散推定量の偏りを修正したt分布曲線（n＝10の場合）

さてここで，小標本においてu^2に偏りがあるとすれば統計量tもその偏りの影響を受けることとなる。事前に不偏推定量の偏りが分かっている場合にはそれを修正して用いることの方が優ることはいうまでもない。いま不偏分散推定量の偏り率をλとして，偏りを含む不偏分散推定量をu^{2*}とすると，

$$u^{2*}=\lambda\times u^2 \tag{11}$$

となり，(10)式のtは次のようにt^*として表される。

$$t^*=\{(x-\mu)/(u^2/n)^{1/2}\}\times(1/\lambda^{1/2}) \tag{12}$$

すなわち，通常のt統計量を$(1/\lambda^{1/2})$倍することが必要となる。

いま自由度をnとすると，通常のt分布の密度関数は次の式で与えられる。

$$f(t)=[\Gamma((n+1)/2)/\{(n\pi)^{1/2}\Gamma(n/2)\}]\times(1+t^2/n)^{-(n+1)/2} \tag{13}$$

ここで，Γはガンマ関数を表す。さらに偏り率λは，今回のモンテ・カルロ実験結果から小標本の偏り関数として得られている。母集団規模100個で，抽出標本数がn個の場合のλは

$$\lambda\times 100=114.6073-0.3066n+0.0017n^2 \tag{14}$$

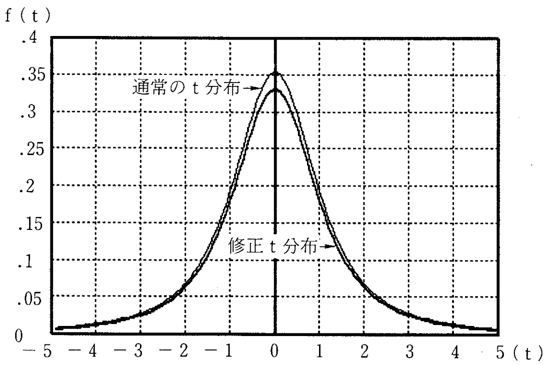
となるから，(12)式のt^*の実現値は(13)，(14)式から容易に求めることができる。

いまnが10の場合について考えてみると，(13)式の$\Gamma((n+1)/2)$，$\Gamma(n/2)$はそれぞれ52.34277，24となるから第1項目は次のような定数となり，

$$\Gamma((n+1)/2)/\{(n\pi)^{1/2}\Gamma(n/2)\}=0.38910$$

さらに(14)式は$\lambda=1.11711$なる定数として得られるからt^*は次のようになる。

$$t^*=0.38910\times 0.94613\times(1+t^2/10)^{-5.5} \tag{15}$$



第4図　修正t分布曲線（n＝2の場合）

かくして(15)式からtとt^*の対応する点を求めて容易にグラフにすることができる。

$t^*(0.0)=0.36814$

$t^*(1.0)=0.36814\times(1.1)^{-5.5}=0.21795$

$t^*(1.5)=0.36814\times(1.225)^{-5.5}=0.12057$

$t^*(2.0)=0.36814\times(1.4)^{-5.5}=0.05785$

第3図はこのようにして描いた修正t分布曲線である。また第4図は同様な手順によりnが2の場合の修正t分布曲線を描いたものである。

さて，第3図，4図にはともに比較の意味で通常のt分布曲線も示してあるが，両曲線の関係は相互に縦方向に伸縮した関係にある。ただし，通常のt分布曲線は確率密度関数を示すのに対し，ここでの修正t分布曲線は密度関数ではなく，単にtに対応する$1/\lambda^{1/2}$倍のf(t)ということになる。もし修正t分布曲線も密度関数として描くならば，縦軸の目盛りを$1/\lambda^{1/2}$倍するか，もしくは修正t

分布曲線を縦軸方向に$1/\lambda^{1/2}$倍に引き伸ばさなければならない。

すなわち，修正 t 分布曲線は通常のt分布密度関数を縦方向に$\lambda^{1/2}$倍縮めた関係にある。したがって，通常のt分布密度関数内の面積は丁度1.0になるのに対して，修正 t 分布曲線内の面積はその$1/\lambda^{1/2}$倍になっている。しかし両曲線とも任意の t の値に対して，

$\Pr\{|t| \geqq t_0\}$ なるt_0は同じ値を与える。

かくして，統計的仮説 t 検定においては，一定の帰無仮説の下で得られる t 統計量t_0が棄却域の中に落ちるか否かの判定には従来の t 分布表を利用しても良いこととなる。しかしこの場合には，通常の不偏分散値u^2に対して修正済不偏分散値λu^2を用いるか，または不偏分散値に偏りのないことが必要である。

したがって，母分散の代りに通常の不偏分散値u^2をそのまま使用してt検定を行う場合には，予め t 分布表の方に$\lambda^{1/2}$倍する修正を施しておくのが便利である。第５表はそのために修正したt分布表である。この表によれば，小標本における t 検定も従来と全く同じ要領で，不偏分散の偏りの存在を意識せず行うことができる。なお，表中の数値は偏りのない不偏分散を用いた時の，

$\Pr\{|t| \geqq t_0\} \rightarrow t_0$，を与えるものである。

総　　　括

本稿は母集団分散の不偏推定量として用いられる不偏分散の小標本特性をモンテ・カルロ実験によって解明したものである。通常標本から母分散を推計する場合不偏分散u^2を用いるが，これは不偏分散推定量の期待値が母分散に一致するという理論的根拠に基づいている。すなわち，抽出回数を無限回くり返した場合に得られる最終的平均値が一致するということである。しかしながら，現実には抽出回数は数回程度に限定されていたり，場合によっては１度きりということもある。このような場合には不偏分散の期待値が母分散に一致するという理論的根拠はなんら保証されるものにはならない。かくして現実的な意味において不偏分散に関する小標本特性を解明することが必要になるのである。

100個の有限母集団を人工的に作成し，モンテ・カルロ実験を行った結果，不偏分散の小標本に関する基本的な特性を明らかにすることができた。それらを総括すると以下のようである。

１）小標本においては，不偏分散推定量は確かに標本分散推定量より不偏性において優っている。しかし，なお偏りの存在も認められた。

２）しかし，不偏分散推定量の有効性はかなり劣る。すなわち，バラツキ（不偏分散推定量の標準偏差）は相当大きく，しかもそのバラツキの大きさは抽出回数を増やしても必ずしも小さくすることができない。

３）また，一致性については，大標本特性として抽出標本数を増やすことによって相当程度目標値に近づけることが可能であるが，小標本の範囲においても不偏分散の一致性の傾向を確認することができる。

第５表　統計的仮説検定のための修正 t 分布表

n	0.25	0.10	0.05	0.025	0.01	0.005
1	2.58108	6.75023	13.58428	27.2113	68.0571	136.120
2	1.71218	3.11772	4.59404	6.62547	10.5968	15.0429
3	1.51694	3.50947	3.39355	4.45347	6.22824	4.94757
4	1.43170	2.27022	2.95668	3.72236	4.90306	5.96107
5	1.38359	2.14308	2.73400	3.36448	4.28840	5.07671
6	1.35251	2.06409	2.59912	3.15338	3.93804	4.58535
7	1.33065	2.00992	2.50853	3.01414	3.71251	4.27456
8	1.31415	1.97022	2.44331	2.91533	3.55519	4.06070
9	1.30131	1.93985	2.39393	2.84135	3.43904	3.90456
10	1.29084	1.91570	2.35496	2.78376	3.34975	3.78531
11	1.28209	1.89585	2.32349	2.73741	3.27865	3.69120
12	1.27465	1.87923	2.29730	2.69923	3.22062	3.61486
13	1.26808	1.86499	2.27519	2.66716	3.17235	3.55169
14	1.26238	1.85270	2.25610	2.63983	3.13128	3.49828
15	1.25734	1.84182	2.23950	2.61606	3.09601	3.45250

注）数値は$\Pr\{|t| \geqq t_0\} \rightarrow t_0$を与える。

4）小標本においては不偏分散推定量の偏りが少なからず認められることから抽出標本数と不偏分散の偏りの関係を体系的に分析し，その関係を偏り関数として計測した。

5）計測した偏り関数を利用して修正 t 分布曲線を描き，実際に統計的仮説検定を行う場合に都合のよい修正t分布表を作成した。これによって，通常と全く同じ手順で，不偏分散の偏りの存在を意識せず検定が可能となるであろう。

参　考　文　献

1）北川敏男・稲葉三男共著 ：統計学通論，共立出版株式会社，東京（1985）pp. 116-122

2）北川敏男：統計情報論 2，共立出版株式社，東京（1987），pp. 191-196

3）小林竜一：応用統計学，共立出版株式会社，東京（1984）pp. 8-20

4）宮武修・脇本和昌：乱数とモンテカルロ法，森北出版株式会社，東京（1988）pp. 36-54

5）森棟公夫：経済モデルの推定と検定，共立出版株式会社，東京（1985）pp. 71-83

6）鍋谷清治：数理統計学，共立出版株式会社，東京（1987）pp. 47-123

7）坂本慶行・石黒真木夫・北川源四郎：情報量統計学，共立出版株式会社，東京（1983）pp. 217-222

8）ウォナコット，T.H.・ウォナコット，R.J.：統計学序説，国府田恒男・田中一盛・細田雄三共訳，培風館，東京（1990）pp. 117-128

9）鷲尾泰俊：推定と検定，共立出版株式会社，東京（1978）pp. 84-98