1つの特性値に過ぎないものとなる。

大きさN（偶数）の母集団で$u_i=-1$の個数が$N_{-1}=N/2$個の場合、$u_i=1$の個数も$N_+=N/2$個となり、μやσ_jは値がなくなる。一方この集団から大きさn（奇数）の標本を抽出してmを計算する。この場合$T\neq 0$で、mは${}_NC_n$通りできて期待値が求まる。

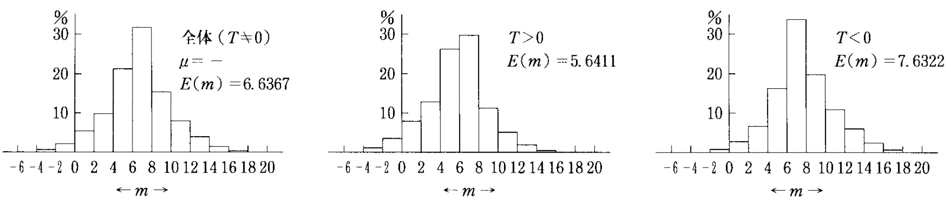
—— mの分布の検証例8 ——

$N=18$として、$z=\{(x_i, u_i)\}_N$の内容を次のとおりとする。$u_i=-1$の個数$N_{-1}=9$に注意を要する。$z=\{(3.01\ 1), (4.15, -1), (4.22, -1), (4.84, 1), (5.09, 1), (5.83, -1), (6.07, 1), (6.46, -1), (6.61, 1), (6.96, 1), (7.11, -1), (7.45, 1), (7.79, -1), (8.09, 1), (8.14, -1), (8.72, -1), (9.37, 1), (9.55, -1)\}$とおく。これを母集団として、大きさ5の標本を等確率、非復元抽出する。nが奇数であるため$T=0$はなく、標本はすべて有効である。mは常に計算され、その度数分布は表や図で示すことができる。標本は全部で${}_{18}C_5=8568$通りあり、そのうち$T>0$のものは4284通り、$T<0$のものも同数ある。x_iの単純平均（u_iが関与しない。$n=1$の場合に相当）を$\bar{x}$とすると、$\bar{x}=6.6367$であるが、全体の$E(m)$はこれに一致している。（表7，図13）以上で標本平均mに関する検証を終える。μは存在するが$E(m)$は存在しない、あるいはμは存在しないが$E(m)$は存在する、など特異な現象が見られた。

表7　mの度数分布表

m	全体（$T\neq 0$）		$T>0$		$T<0$	
以上～未満	実度数	比率%	実度数	比率%	実度数	比率%
−6～−4	5	0.06	5	0.12	0	0
−4～−2	52	0.61	48	1.12	4	0.09
−2～0	182	2.12	151	3.52	31	0.72
0～2	458	5.35	341	7.96	117	2.73
2～4	837	9.77	552	12.89	285	6.65
4～6	1820	21.24	1122	26.19	698	16.29
6～8	2715	31.69	1271	29.67	1444	33.71
8～10	1324	15.45	478	11.16	846	19.75
10～12	683	7.97	217	5.07	466	10.88
12～14	338	3.94	80	1.87	258	6.02
14～16	121	1.41	17	0.40	104	2.43
16～18	30	0.35	2	0.04	28	0.65
18～20	3	0.04	0	0	3	0.07
計	8568	100.00	4284	100.00	4284	100.00

図13（表7に対応）　**図13a**　**図13b**



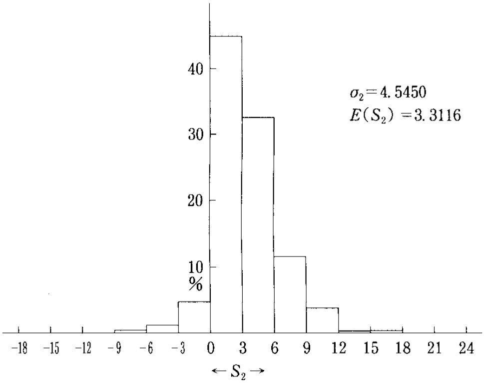
S_j の期待値

S_j の内容は㉛による。母集団における㊻に対応する。その分母は $T=\sum U_i$ であるので、S_j の期待値を求める方法は、m の期待値を求める方法と同じである。結果は一層複雑で、例えば S_2 については、σ_2 が存在するとして、

表 8　S_2 の度数分布表

S_2 以上～未満	実度数	比率%
−15～−12	1	0.03
−12～−9	2	0.05
−9～−6	15	0.39
−6～−3	45	1.16
−3～0	184	4.75
0～3	1744	44.99
3～6	1261	32.53
6～9	447	11.53
9～12	146	3.77
12～15	13	0.33
15～18	14	0.36
18～21	4	0.10
計	3876	100.00

図14　（表 8 に対応）



$$E(S_2)=\sigma_2(1+\beta) \tag{94}$$

の形となるが、β には最低次の項だけを見ても母集団の 4 次の積率が含まれる。ここではその具体的な式は示さず、検証例を示すことによって S_2 の確率分布を概観する。S_2 は分散を表わす確率変数であるが、変域が負に亙ることがあるなど、ユニークな点に注意を向けたい。

—— S_2 の分布の検証例 1 ——

前号の 6 頁の 8 で設定した大きさ19の母集団

表 9　S_2' の度数分布表

S_2'	全体 （$T\neq0$）		$T>0$		$T<0$	
以上～未満	実度数	比率%	実度数	比率%	実度数	比率%
−15～−12	4	0.11	3	0.09	1	6.25
−12～−9	9	0.26	8	0.23	1	6.25
−9～−6	25	0.71	25	0.71	0	0
−6～−3	78	2.22	75	2.14	3	18.75
−3～0	277	7.88	273	7.80	4	25.00
0～3	1305	37.12	1298	37.09	7	43.75
3～6	1058	30.09	1058	30.23	0	0
6～9	456	12.97	456	13.03	0	0
9～12	191	5.43	191	5.46	0	0
12～15	63	1.79	63	1.80	0	0
15～18	38	1.08	38	1.09	0	0
18～21	12	0.34	12	0.34	0	0
計	3516	100.00	3500	100.00	16	100.00

から、等確率、非復元で抽出した大きさ$n=4$の標本に基づく標本分散S_2（㊿における$j=2$の場合）の分布を、度数分布の表と図で示すこととする。この例は$u_i=-1$の個数が$N_{-1}=1$で、無効標本はない。常に$T>0$である。S_2の実現値${}_{19}C_4=3876$個のうち負に亙るものが247個（6.37%）ある。（表8，図14）

—— S_2の分布の検証例2 ——

前記のmの分布の検証例6と同じ母集団$N=19$, $u_i=-1$の個数$N_{-1}=3$と標本規模$n=4$により、S_2の分布状況をみることとする。S_2に対して無効標本があり、これを除いたS_2をS_2'と書く。標本は全部で${}_{19}C_4=3876$通り、そのうちS_2に対する無効標本が360通りある。$E(S_2')$はこの無効標本を除いた3516通りについての期待値である。（表9，図15）

図15 （表9に対応）

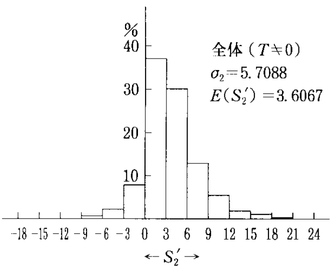


図15a

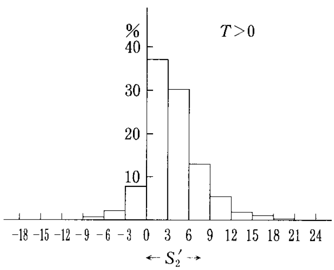
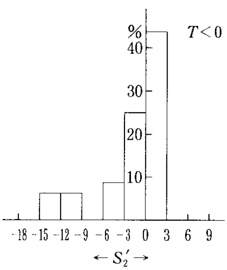


図15b



—— S_2の分布の検証例3 ——

前記のmの分布の検証例7と同じ母集団$N=19$, $u_i=-1$の個数$N_{-1}=9$と標本規模$n=4$により、S_2の分布状況をみることとする。S_2に対して無効標本があり、これを除いたS_2をS_2'と書く。負の値がたくさん実現している。標本は全部で${}_{19}C_4=3876$通り、そのうちS_2に対する無効標本が1620通りある。$E(S_2')$はこの無効標本を除いた2256通りについての期待値である。母集団の分散$\sigma_2=-7.4002$

表10 S_2'の度数分布表

S_2'	全体（$T\neq0$）		$T>0$		$T<0$	
以上～未満	実度数	比率%	実度数	比率%	実度数	比率%
−27～−24	2	0.09	0	0	2	0.21
−24～−21	6	0.27	2	0.16	4	0.41
−21～−18	14	0.62	2	0.16	12	1.24
−18～−15	27	1.20	7	0.54	20	2.07
−15～−12	49	2.17	13	1.01	36	3.73
−12～−9	76	3.37	23	1.78	53	5.49
−9～−6	115	5.10	46	3.57	69	7.14
−6～−3	194	8.60	75	5.81	119	12.32
−3～0	324	14.36	158	12.25	166	17.18
0～3	662	29.34	319	24.73	343	35.51
3～6	362	16.05	246	19.07	116	12.01
6～9	180	7.98	154	11.94	26	2.69
9～12	131	5.81	131	10.16	0	0
12～15	48	2.13	48	3.72	0	0
15～18	48	2.13	48	3.72	0	0
18～21	18	0.80	18	1.40	0	0
計	2256	100.00	1290	100.00	966	100.00

図16　（表10に対応）

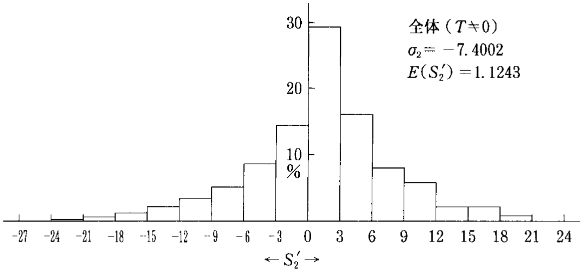


図16a

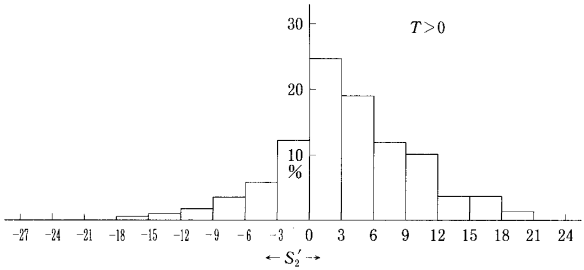
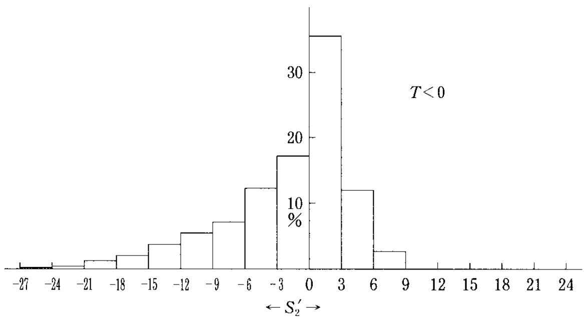


図16b



に対して $E(S_2') = 1.1243$ で、両者に大差が生じている。（表10，図16）

—— S_2 の分布の検証例 4 ——

前記の m の分布の検証例 8 と同じ母集団 $N = 18$, $u_i = -1$ の個数 $N_{-1} = 9$ と標本規模 $n = 5$ により、S_2 の分布状況をみることとする。この例では σ_2 は存在しない。しかし標本分散 S_2 は常に存在し、無効標本もない。S_2 の期待値は $E(S_2) = -7.2609$ と負になっている。S_2 の度数分布の表や図をみても、負の側に広く分布していることがわかる。（表11，図17）

表11　S_2 の度数分布表

S_2	全体 ($T \neq 0$)		$T>0$		$T<0$	
以上～未満	実度数	比率%	実度数	比率%	実度数	比率%
−150～−140	1	0.01	0	0	1	0.02
−140～−130	3	0.04	1	0.02	2	0.05
−130～−120	4	0.05	1	0.02	3	0.07
−120～−110	8	0.09	3	0.07	5	0.12
−110～−100	10	0.12	3	0.07	7	0.16
−100～−90	24	0.28	11	0.26	13	0.30
−90～−80	36	0.42	13	0.30	23	0.54
−80～−70	54	0.63	23	0.54	31	0.72
−70～−60	75	0.88	34	0.79	41	0.96
−60～−50	123	1.44	60	1.40	63	1.47
−50～−40	195	2.28	93	2.17	102	2.38
−40～−30	302	3.52	152	3.55	150	3.50
−30～−20	524	6.12	269	6.28	255	5.95
−20～−10	982	11.46	537	12.54	445	10.39
−10～0	2327	27.16	1276	29.79	1051	24.53
0～10	3561	41.56	1655	38.63	1906	44.49
10～20	329	3.84	143	3.34	186	4.34
20～30	10	0.12	10	0.23	0	0
計	8568	100.00	4284	100.00	4284	100.00

図17　(表11に対応)

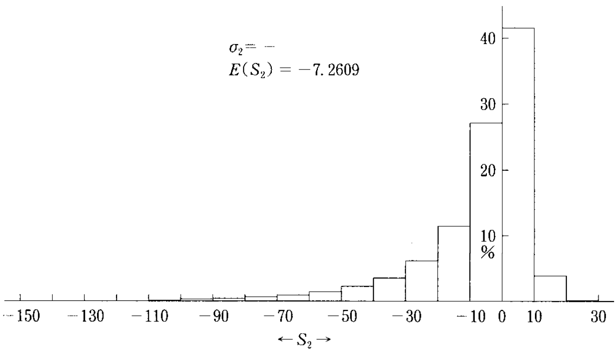


図17a

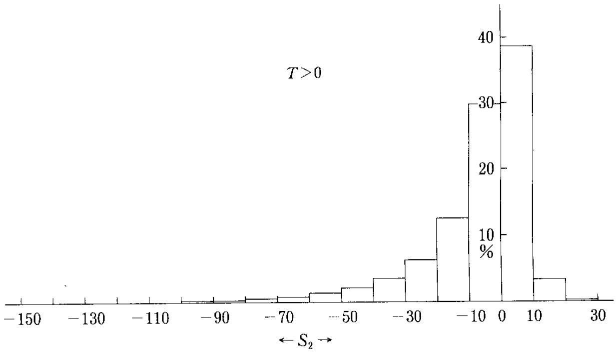
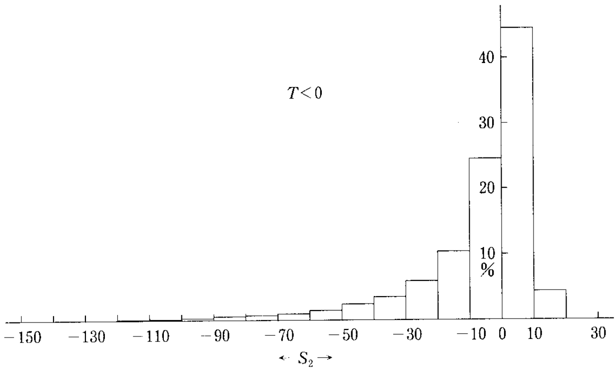


図17b



標本の大きさと m の期待値等の関係

これまで適宜大きさ N の有限母集団を模型的に設定し、適宜大きさ n の標本を等確率、非復元で抽出し、m や S_2 について期待値や標本分布、およびそれぞれに対する母集団特性値との関係などを検証してきた。ここで、それらを整理して一望できるようにしたのが表12と図18である。基本模型は $N=19$, $N_{-1}=0$, $N_{+1}=19$, $z=\{(x_i,\ u_i)\}_N=\{$ (3.01　1), (4.15, 1), (4.22, 1), (4.84, 1), (5.09, 1), (5.83, 1), (6.07, 1), (6.46, 1), (6.61, 1), (6.96, 1), (7.11, 1), (7.45, 1), (7.79, 1), (8.09, 1), (8.14, 1), (8.72, 1), (9.37, 1), (9.55, 1), (11.34, 1) $\}$ で、u_i はすべて 1 、従来型である。表12の母集団 2 ～ 7 は、この基本模型から N_{-1} を少しずつ増やして母集団を変化させたものである。

標本の大きさ n は 3 以上 N 以下の奇数としてある。$n=1$ は u_i に関係せず、x_i の単純平均と同じであるため取り上げてない。n が偶数の場合は、n と N_{-1} の関係により期待値が存在しない場合があり、表示してない。表中の μ は㊺で与えられる母集団平均、m は⑹⓪, ⑹⑨で与えられる標本平均、$E(m)$ は m の期待値、$B(m)$ は m の μ に対する偏りで $E(m-\mu)$ に等しく、$V(m)$ は m の分散で $E\{m-E(m)\}^2$ に等しく $MSE(m)$ は m の平均平方誤差で $E(m-\mu)^2$ に等しい。なお、S_2 等についても表12に準じた表ができるが、省略する。

表12　標本の大きさと m の期待値等の関係

母集団	$N=19$	n	3	5	7	9	11	13	15	17	19
1	$N_1=0$, $N_{+1}=19$, $\mu=6.8842$	$E(m)$	6.8842	6.8842	6.8842	6.8842	6.8842	6.8842	6.8842	6.8842	6.8842
		$B(m)$	0	0	0	0	0	0	0	0	0
		$V(m)$	1.2436	0.6529	0.3997	0.2591	0.1696	0.1076	0.0622	0.0274	0
		$MSE(m)$	1.2436	0.6529	0.3997	0.2591	0.1696	0.1076	0.0622	0.0274	0
2	$N_1=1$, $N_{+1}=18$, $\mu=7.0082$	$E(m)$	7.1185	7.0403	7.0248	7.0180	7.0144	7.0120	7.0104	7.0092	7.0082
		$B(m)$	0.1102	0.0322	0.0165	0.0098	0.0061	0.0038	0.0021	0.0009	0
		$V(m)$	2.5433	0.9391	0.5388	0.3400	0.2194	0.1379	0.0792	0.0348	0
		$MSE(m)$	2.5554	0.9401	0.5391	0.3401	0.2194	0.1380	0.0792	0.0348	0
3	$N_1=3$, $N_{+1}=16$, $\mu=6.8185$	$E(m)$	6.8116	6.7874	6.7937	6.8076	6.8123	6.8149	6.8165	6.8176	6.8185
		$B(m)$	−0.0069	−0.0311	−0.0248	−0.0109	−0.0062	−0.0036	−0.0020	−0.0008	0
		$V(m)$	4.7694	3.1604	1.5788	0.7146	0.4238	0.2556	0.1430	0.0617	0
		$MSE(m)$	4.7694	3.1613	1.5794	0.7148	0.4238	0.2556	0.1430	0.0617	0
4	$N_1=5$, $N_{+1}=14$, $\mu=7.4756$	$E(m)$	7.1973	7.4547	7.6326	7.7071	7.6686	7.5576	7.5160	7.4918	7.4756
		$B(m)$	−0.2783	−0.0209	0.1570	0.2316	0.1930	0.0820	0.0404	0.0162	0
		$V(m)$	7.2449	7.4921	6.2095	4.1924	2.1528	0.7520	0.3640	0.1459	0
		$MSE(m)$	7.3224	7.4925	6.2342	4.2461	2.1901	0.7587	0.3656	0.1459	0
5	$N_1=7$, $N_{+1}=12$, $\mu=6.1480$	$E(m)$	6.7639	6.6276	6.4735	6.3017	6.1176	5.9433	5.8520	6.0740	6.1480
		$B(m)$	0.6159	0.4796	0.3255	0.1537	−0.0303	−0.2047	−0.2960	−0.0740	0
		$V(m)$	8.7354	10.8419	11.4250	10.8019	9.1232	6.5293	3.3604	0.6141	0
		$MSE(m)$	9.1148	11.0719	11.5310	10.8526	9.1241	6.5713	3.4481	0.6195	0
6	$N_1=9$, $N_{+1}=10$, $\mu=0.26$	$E(m)$	6.8409	6.7861	6.7143	6.6166	6.4753	6.2533	5.8538	4.9215	0.26
		$B(m)$	6.5809	6.5261	6.4544	6.3566	6.2153	5.9933	5.5938	4.6615	0
		$V(m)$	9.9004	14.0514	17.4981	20.5837	23.5186	26.4689	29.5147	31.7438	0
		$MSE(m)$	53.2088	56.6410	59.1568	60.9896	62.1486	62.3889	60.8050	53.4732	0
7	$N=18$, $N_1=N_{+1}=9$, μ なし	$E(m)$	6.6367	6.6367	6.6367	6.6367	6.6367	6.6367	6.6367	6.6367	
		$V(m)$	9.2401	15.4345	22.5031	30.9699	41.5668	55.6184	76.1063	112.3919	

表の数値は、横方向に系列をなしている。母集団が N の値と N_1 の値ごとに異なるため、縦方向には系列をなしていない。

図18a（母集団 1 に対応）

$N=19,\ N_1=0,\ N_{+1}=19$
$\mu=6.8842$

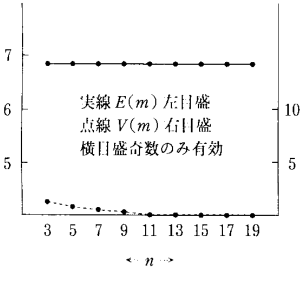


図18b（母集団 2 に対応）

$N=19,\ N_1=1,\ N_{+1}=18$
$\mu=7.0082$

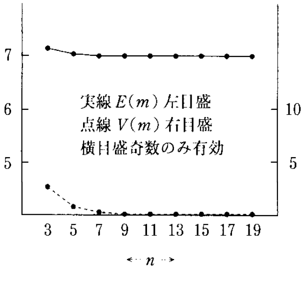


図18c（母集団 3 に対応）

$N=19,\ N_1=3,\ N_{+1}=16$
$\mu=6.8185$

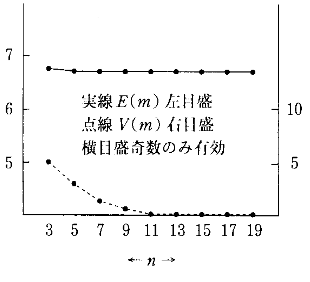


図18d（母集団 4 に対応）

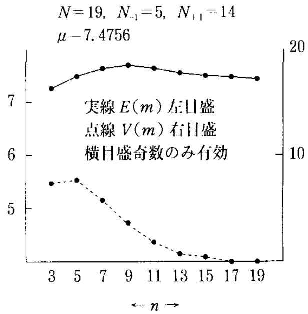


図18e（母集団 5 に対応）

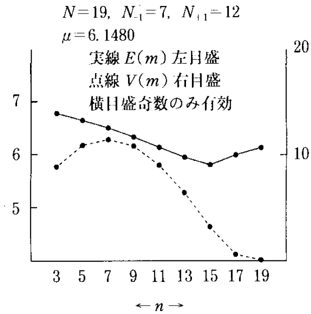


図18f（母集団 6 に対応）

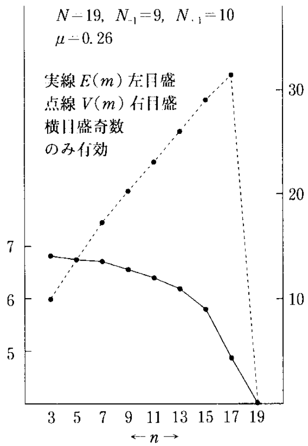
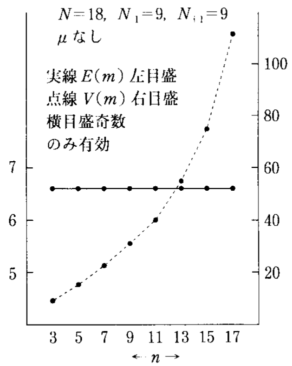


図18g（母集団 7 に対応）



以下に表12の各母集団について説明する。

母集団 1 ： $N=19$, $N_{-1}=0$, $N_{+1}=19$ 。上記の基本模型そのもので、$u_i=-1$ はない。常に $E(m)=\mu$, $B(m)=0$, $V(m)=MSE(m)$ である。標本の大きさ n の増大と共に $V(m)$, $MSE(m)$ が単調に減少するのは既知の法則の通りである。（図18a）

母集団 2 ： $N=19$, $N_{-1}=1$, $N_{+1}=18$ 。母集団 1 において 6 番目の単位を（5.83, 1）から（5.83, −1）に変更したもの。m に僅かに偏りが出ている。変更が少ないため、標本の大きさと m の特性値の関係は従来型と似ている。分散などが全体として大きくなっている。（図18b）

母集団 3 ： $N=19$, $N_{-1}=3$, $N_{+1}=16$ 。母集団 2 において11番目の単位を（7.11, 1）から（7.11, −1）に、15番目の単位を（8.14, 1）から（8.14, −1）に変更したもの。$n=5$ における $E(m)$ に僅かの凹みが見られる。分散は水準が更に高まっているが、n の増大と共に単調に減少している。（図18c）

母集団 4 ： $N=19$, $N_{-1}=5$, $N_{+1}=14$ 。母集団 3 において 3 番目の単位を（4.22, 1）から（4.22, −1）に、8 番目の単位を（6.46, 1）から（6.46, −1）に変更したもの。期待値、分散など、n の増大と共にはっきりした起伏

が生じている点に注目を要する。(図18d)

母集団5：$N=19$, $N_{-1}=7$, $N_{+1}=12$。母集団4において16番目の単位を(8.72, 1)から(8.72, −1)に、18番目の単位を(9.55, 1)から(9.55, −1)に変更したもの。特に分散の変化が興味深い。標本の大きさが3, 5, 7と増大するにつれて分散は増大し、その後減少している。(図18e)

母集団6：$N=19$, $N_{-1}=9$, $N_{+1}=10$。母集団5において12番目の単位を(7.45, 1)から(7.45, −1)に、13番目の単位を(7.79, 1)から(7.79, −1)に変更したもの。特に分散は標本の大きさの増大と共にNの手前まで増大し続け、$n=N$で急に0になるという奇態を呈している。(図18f)

母集団7：$N=18$, $N_{-1}=9$, $N_{+1}=9$。母集団1の19番目の単位を取り除き、$N=18$とし、$u_i=-1$の個数と位置を母集団6と同じにしたもの。μはない。ここではmの期待値は一定だが、分散は標本が大きくなるほど大きくなっている。(図18g)

期待値や分散の特徴

新統計集団では、負の構成単位の存在によって、従来の集団に比べ様々な目新しい現象がみられる。以下にその幾つかを掲げる。

1) 標本平均mの分布範囲が拡大することがある。従来の集団では、x_iの最小値と最大値の間に分布するが、新しいmはこの範囲をはみ出すことがある。分子と分母の相関の状況に関係する。

2) 標本分散S_2の分布範囲が拡大することがある。従来の集団では完全平方和の分散は0以上だが、新しいS_2は負になることがある。

3) Nが偶数で、$N_{-1}=N/2$のときはμやσ_2は存在しない。分母が0になるためである。

4) nが偶数で、$2N_{-1}$以下のときは、mやS_2の期待値は存在しない。統計量の分母が0になることがあるためである。

5) mをμ(存在する場合。以下同じ)に関連づけると、mはμに対して偏りをもつ。S_2がσ_2に対して偏りをもつのは、従来と同じ。

6) $n=1$においては、mはx_iの平均に、S_2はx_iの分散に等しくなる。いずれもu_iに関係しない。

7) nが奇数で増大すると、mの分散が増大することがある。これはnの増大と共に分母の変動も増大し、分子の変動がそれを吸収しきれない場合に起きる。従来と反対の現象であることに注意を要する。

8) $n=N$においては、$m=\mu$, $S_2=\sigma_2$となる。一致性の一種であり、7)と矛盾しない。

おわりに

本稿の内容に対して、読者は腑に落ちないかも知れない。「x_iがすべて正でもその平均が負になることがあるとか、分散が負になることがあるとか、これが何の役に立つのか」と。

筆者は本稿の内容を、本来そうであるのに、これまでその半分しか見ていなかったのだと考えている。自然数だけしか知らなかった昔の人にとって、負の数は俄かに信じ難いものであったに違いない。また、「分散が負になるとすれば、標準偏差はどうなるのか」との疑問に対しては、「虚数」とするのが既存理論との接続上よいと筆者は考えている。このほかにも決め方はあるが、どう決めるかは写像の選択の問題である。こうして、分散は負に、標準偏差は虚に値域を広げる。数平面上でいえば、分散は横の座標軸を原点から左方に、標準偏差は原点を角(かど)に90度曲がって上方または下方に伸びていくことになる。

また、「本稿における母集団と標本の関係の研究は、統計的推測にどう役立つのか」との疑

問に対しては、前号で詳述した写像の概念が解決のよい助けとなるであろう。

このことに関連して、あえて重ねて述べたいのは、前述のとおり、統計量の分布の範囲が、従来の集団より広がること、例えば、正のデータによる平均が負になったり、分散が負になったり、標本平均の分散が標本規模の拡大と共に大きくなったりするなどの例がみられることである。これは設定した例においてそうなったということであって、すべての場合にそうなるわけではない。より適切な言い方をすれば、新統計集団においては、これらの状況が実に多様であり、従来の統計集団をその一部として丸々包含しているということである。例えば、標本の大きさと標本平均の分散との関係をみると、従来 σ^2/n などの形が知られていて、n の増大と共に小さくなっていくと理解されているが、これは、新統計集団の中では単なる一例に過ぎなく、前述のように n の増大に伴って却って大きくなる場合や、n^{-1} よりも高い負のオーダーで、従来よりも急速に小さくなっていく場合もあり、その態様は無限に多い。この千変万化の態様は、比推定論の理解によって完全に納得できるものである。我々は、この無限の多様性の中から目的に合わせて最適なものを選択するという立場である。前述の n の増大と共に分散が大きくなっていく現象を無益とせず、n の増大と共に分散が小さくなっていく現象と等価同等同意義に捕えようという態度である。これは、新統計集団において正の構成単位と負の構成単位を同等に扱い、分散の正も負も対等に考えることと共通した理念によるものである。より大切な点は、我々がその多様な状態の中のどれに効用を見出すかということにある。

この際、負の分散の効用を1つ紹介しておく。正の分散を差推定する場合、分散の推定値が負になることがあるのは以前から知られている。標本抽出が副標本（部分標本。2段標本ではない）方式で行われ、全標本による統計量の分散を副標本の分散から合成するとき、副標本レベルので負の分散が生じた場合、そのまま用いると、全体の分散が適正に求められる。負の分散を嫌って無視すると、全体の分散が不当に過大なる。

本題に関し、研究分野はなお広い。連続分布やパラメトリックな分野は、とりわけ研究のし甲斐があろう。

［正誤］前号8頁右欄下6行目 ${}_{19}P_4=93042$ は ${}_{19}P_4=93024$ の誤り。続く3行も同じ。

参考文献

(1) Funatsu, Y. (1982) A METHOD OF DERIVING VALID APPROXIMATE EXPRESSIONS FOR BIAS IN RATIO ESTIMATION, Journal of Statistical Planning and Inference 6 (1982) 215-225, North-Holland Publishing Company.

(2) Funatsu, Y. (1982) A NOTE ON KOOP'S PROCEDURE TO OBTAIN THE BIAS OF THE RATIO ESTIMATE, Sankyā, The Indian Journal of Statistics, vol.44, Series B, Pt.2, pp.219-222

(3) 船津好明　「負の分散について」ESTRELA　2000年9月号(No.78), 34-45頁

(4) 船津好明　「新統計集団(1)」明星大学経済学研究紀要第32巻第1•2号、1-9頁、平成13年3月、明星大学