

寄稿

MS-Excelを用いたサイズ度数分布から
年齢組成を推定する方法の検討相澤 康¹, 滝口直之¹Consideration of the Methods for Estimating the Age-composition from the
Length Frequency Data with MS-ExcelYasushi AIZAWA¹ and Naoyuki TAKIGUCHI¹

はじめに

年齢形質が得られていない生物種の年齢別サイズ、個体数等を知るには、体長等の度数分布を正規分布に分解する手法が用いられている。小型計算機による多くのプログラムがあり、赤嶺 (1995)、山川 (1997a) によりレビューされ、水産研究現場で活用されている。しかし、これらの多くは BASIC プログラムであり、MS-Windows ソフトが普及した近年の環境では使い易いとは言えない。また、BASIC プログラムを使用する場合は、ある程度プログラムを習熟する必要があり、計算過程が確認できない欠点がある (東海, 1997)。

このような状況の中、MS-Windows ソフト上で使い易いプログラムの提供が望まれ、表計算ソフトにより、網目の選択性曲線や年齢組成を推定する方法が紹介されている (東海, 1997, 山川, 1997b, 1997c, Yamakawa and Matsumiya, 1997, 藤森・東海, 1999, 相澤・倉田, 1999)。

本報では、表計算ソフトの MS-Excel ((株) マイクロソフト社) を用いて度数分布を正規分布に分解する手法として、Hasselblad (1966) の方法とアドインソフトである Solver による方法を紹介し、その実用性について検討したので報告する。

材料および方法

Hasselblad の方法

Hasselblad の方法 (以下、Hasselblad 法という) は、最尤法によるパラメータの推定方法であり、赤嶺 (1995) に詳述されている。表計算ソフトになじみやすくするため、以下の式により計算した。記号を示す。

 f_{ij} : 年齢 i , 階級 j の個体数 N : 正規分布 K_i : 年齢 i の個体数 (推定するパラメータ) P_i : 年齢 i の年齢組成 (推定するパラメータ) μ_i : 年齢 i の平均体長 (推定するパラメータ) σ_i : 年齢 i の標準偏差 (推定するパラメータ) $g_{.j}$: 階級 j の混合正規分布 $C_{.j}$: 階級 j の度数分布 (実測値) $F_{.j}$: 階級 j の組成 (実測値) $x_{.j}$: 階級値 h : 階級幅 m : 年齢数 n : 階級数 i : 年齢を表わす添字 j : 階級を表わす添字年齢 i の体長組成は正規分布と仮定しているので、

$$f_{ij}^{old} = [N(\mu_i, \sigma_i, x_{.j}) \cdot P_i]^{old} \cdot h \quad (1)$$

となる。また、

$$g_{.j} = \sum_{i=1}^m f_{ij} \quad (2)$$

であり、ただし、

$$\sum_{j=1}^n g_{.j} = 1 \quad (3)$$

である。階級 j で $g_{.j} = F_{.j}$ になるように、 f_{ij} を比例配分する。ここでは簡単に記述するために Ψ とおくと、

$$\psi_{ij} = (f_{ij}/g_{.j})^{old} \cdot F_{.j} \quad (4)$$

となり、このとき

1999年5月31日受付, 1999年10月12日受理

¹ 神奈川県水産総合研究所Kanagawa Prefectural Fisheries Research Institute, Youroushi, Misaki,
Miura, Kanagawa 238-0237, Japan

$$F_j = \sum_{i=1}^m \Psi_{ij} \quad (5)$$

であることに注意する. 次に $P \rightarrow \mu \rightarrow \sigma$ の順に最適化をしていく.

$$P_i^{new} = \sum_{j=1}^n \Psi_{ij} \quad (6)$$

$$\mu_i^{new} = \sum_{j=1}^n [(\Psi_{ij}/P_i^{new}) \cdot x_{ij}] \quad (7)$$

$$(\sigma_i^{new})^2 = \sum_{j=1}^n [(x_{ij} - \mu_i^{new})^2 \cdot (\Psi_{ij}/P_i^{new})] \quad (8)$$

(7), (8) 式は通常の度数分布から標本平均と標準偏差を求める計算式である. 反復計算はMS-Excelの反復計算機能を用いた. すなわち, 初期値をワークシートに入力し1回の計算を行った後, 入力セルから出力セルを参照することで新しい解を代入して, 反復計算を行った.

ワークシートの構成, 計算方法の設定および計算の手順は付録に示す. 精度の設定はMS-Excelの既定値とした.

Solverによる計算法

SolverとはMS-Excelにアドインされた最適解を探索するソフトである.

Solverによる計算法は最小二乗法, x^2 最小化法および最尤法を適用した. Hasselblad法も最尤法であるので, ここでは区別するために対数尤度最大化法という. 以下の式をそれぞれ目的関数とし, Solverにより最適解を探索した.

$$d^2 = \sum_{j=1}^n (F_j - g_j)^2 \rightarrow \text{Min} \quad (9)$$

$$x^2 = \sum_{j=1}^n [(F_j - g_j)^2 / g_j] \rightarrow \text{Min} \quad (10)$$

$$\text{MLH} = \sum_{j=1}^n [(F_j \cdot \ln(g_j))] \rightarrow \text{Max} \quad (11)$$

混合正規分布の和が1になるように (3) 式を, また, 各パラメータが負に収束しないように

$$P_i, \mu_i, \sigma_i \geq 0 \quad (12)$$

の制約条件を設定した.

ワークシートの構成, パラメータ・オプション設定および操作の手順は付録に示す. 精度, 公差および収束は, 原則として既定値としたが, 対数尤度最大化法についてはHasselblad法と精度, 計算時間を比較するため, 精度を既定値 10^{-5} の他に 10^{-4} と 10^{-6} に設定した.

表1. 例題のパラメータ

AGE (i)	1	2	3	4	5
Parameters					
K	8000	4000	2000	1000	500
P	0.52	0.26	0.13	0.06	0.03
μ	11	19	26	32	37
σ Sample 1	1	1	1	1	1
Sample 2	2	2	2	2	2
Sample 3	3	3	3	3	3
Sample 4	1.00	1.30	1.55	1.75	1.90
Sample 5	2.00	2.60	3.10	3.50	3.80

表2. 例題の体長度数分布

BL(cm)	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
0.5	0	0	2	0	0
1.5	0	0	7	0	0
2.5	0	0	19	0	0
3.5	0	1	47	0	1
4.5	0	8	102	0	8
5.5	0	36	198	0	36
6.5	0	127	345	0	127
7.5	7	345	539	7	345
8.5	140	731	753	140	731
9.5	1036	1205	943	1036	1205
10.5	2817	1547	1059	2817	1550
11.5	2817	1547	1072	2817	1556
12.5	1036	1209	990	1036	1232
13.5	140	749	851	140	796
14.5	7	409	712	10	483
15.5	4	300	615	33	376
16.5	70	402	576	193	425
17.5	518	610	576	631	534
18.5	1408	775	584	1140	618
19.5	1408	776	569	1140	631
20.5	518	611	526	632	574
21.5	70	397	464	201	478
22.5	5	259	406	73	387
23.5	35	246	363	143	329
24.5	259	320	340	322	306
25.5	704	392	326	489	302
26.5	704	392	310	490	298
27.5	259	317	288	331	284
28.5	36	226	260	171	260
29.5	19	178	233	122	232
30.5	130	183	209	166	206
31.5	352	205	193	221	185
32.5	352	203	178	225	167
33.5	130	172	163	177	152
34.5	26	137	146	126	137
35.5	66	118	128	108	120
36.5	176	113	110	110	103
37.5	176	101	91	103	85
38.5	65	76	72	77	69
39.5	9	46	53	44	54
40.5	0	22	36	19	40
41.5	0	8	23	6	29
42.5	0	2	12	2	20
43.5	0	1	6	0	13
44.5	0	0	3	0	8
45.5	0	0	1	0	4
46.5	0	0	0	0	2
47.5	0	0	0	0	1
48.5	0	0	0	0	1
49.5	0	0	0	0	0
50.5	0	0	0	0	0
	15499	15502	15499	15498	15500

計算例題

今回は, 標準偏差を異にする5つの計算例題を用意した. 条件を表1, データを表2に示す. この例題は以下の特徴を備えている.

- 1) 正規分布を合計し作成したデータで、混合正規分布 g は小数点以下を四捨五入した。
- 2) 平均体長 μ は対数関数的に単調増加、個体数 K は指数関数的に単調減少する。
- 3) 標準偏差 σ がすべての年齢で同一の3例題と、 μ の平方根に比例する2例題がある。 σ の取り方でモードの明瞭・不明瞭ができるが、すべての例題で1, 2才はモードがはっきりしている。
- 4) 初期値を表3に示す。計算方法の特性を把握するため、あえて条件の悪い初期値を与えた。

また、この他に実際に現場から得られたマダイ尾叉長さ数分布についても適用した(表4)。

表3. 例題計算の初期値

AGE (i)	1	2	3	4	5
Parameters					
K	3000	3000	3000	3000	3000
P	0.2	0.2	0.2	0.2	0.2
μ	10	20	25	30	35
σ	1	1	1	1	1

- 1) このデータは、1993年から1997年の間、神奈川県沿岸の市場において水揚げされたマダイを調査員が測定したものである(相澤・倉田, 1999)。
- 2) 神奈川県沿岸のマダイの成長式は既に今井(1997)により報告されており、10才魚は66.37 cmと推定されている。これ以上の年齢のマダイは全体の0.5%以下であるので、計算の上では大きな支障とはならないものと判断し、10才以上を一括した。
- 3) 神奈川県では、完全に漁獲加入する年齢は2才以上である(今井, 1996)。よって、0, 1才魚は漁獲の選択性を受けていると予想される。
- 4) 各年齢のモードは不明瞭で、ヒストグラムは右に長く裾をひく。今回のように正規分布の分解に供するには、質の悪いデータと思われる。

計算の初期値は表5に示した。 μ は成長式(今井, 1996)から初期値を求めた。

結果

例題の推定結果を表6に示す。Solverでは、得られた最適

表4. 市場調査から得たマダイの尾叉長さ数分布(1993-1997年)

FL(cm)	Number	FL(cm)	Number	FL(cm)	Number	FL(cm)	Number
1.5	0	21.5	919	41.5	377	61.5	35
2.5	0	22.5	1044	42.5	332	62.5	37
3.5	0	23.5	977	43.5	300	63.5	30
4.5	1	24.5	976	44.5	278	64.5	19
5.5	5	25.5	919	45.5	266	65.5	31
6.5	1	26.5	1000	46.5	217	66.5	18
7.5	1	27.5	1048	47.5	193	67.5	25
8.5	1	28.5	1078	48.5	189	68.5	19
9.5	3	29.5	1041	49.5	190	69.5	12
10.5	1	30.5	933	50.5	159	70.5	14
11.5	6	31.5	919	51.5	137	71.5	6
12.5	28	32.5	824	52.5	120	72.5	9
13.5	113	33.5	823	53.5	127	73.5	5
14.5	131	34.5	734	54.5	95	74.5	6
15.5	161	35.5	687	55.5	94	75.5	2
16.5	214	36.5	633	56.5	82	76.5	2
17.5	249	37.5	541	57.5	66	77.5	3
18.5	310	38.5	533	58.5	67	78.5	3
19.5	468	39.5	477	59.5	59	79.5	1
20.5	652	40.5	482	60.5	53	80.5	2

表5. マダイデータ計算の初期値

Hasselblad法は初期値A, SolverにはAをHasselblad法で10回反復した初期値Bを用いた。

A) For Hasselblad's method

AGE (i)	0	1	2	3	4	5	6	7	8	9	10
Parameters											
K	500	5000	5000	5000	2000	2000	1000	500	200	200	200
P	0.023	0.231	0.231	0.231	0.093	0.093	0.046	0.023	0.009	0.009	0.009
μ	13.99	21.98	27.61	33.76	39.56	44.44	48.84	53.05	56.66	59.66	66.37
σ	1.5	1.5	1.5	1.5	2	2	2	2	2.5	2.5	2.5

B) For Solver, given from solution of Hasselblad's method, iterated 10 times

AGE (i)	0	1	2	3	4	5	6	7	8	9	10
Parameters											
K	893.0	5335.5	5779.5	4692.2	2538.1	1480.9	773.7	458.9	214.9	227.1	219.1
P	0.039	0.236	0.256	0.207	0.112	0.065	0.034	0.020	0.010	0.010	0.010
μ	15.90	22.25	27.87	33.40	38.99	44.11	49.06	53.26	56.80	59.85	67.08
σ	2.19	2.17	2.15	2.23	2.02	2.15	1.96	2.05	2.41	2.64	4.81

表6. 例題の計算結果

Hasselblad's method									
	AGE (i)	1	2	3	4	5	d^2	χ^2	MLH
Parameters									
Sample1	K	8000	3999	1999	1000	500	1.649E-08	4.176E-05	-2.659
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	37.00			
	σ	1.00	1.00	1.00	1.00	1.00			
Sample2	K	7999	4002	2006	982	514	4.740E-08	5.730E-05	-3.221
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	31.98	36.94			
	σ	2.00	2.00	2.00	1.97	2.03			
Sample3	K	7996	4037	1508	1096	861	3.367E-07	9.049E-05	-3.392
	P	0.52	0.26	0.10	0.07	0.06			
	μ	11.00	19.01	25.42	29.96	35.56			
	σ	2.99	2.99	2.66	3.01	3.34			
Sample4	K	8000	4001	1999	1002	497	9.738E-09	4.037E-05	-2.800
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	37.01			
	σ	1.00	1.30	1.55	1.75	1.88			
Sample5	K	7986	4146	1365	937	1066	1.107E-07	8.560E-05	-3.250
	P	0.52	0.27	0.09	0.06	0.07			
	μ	11.00	19.07	25.67	29.44	34.71			
	σ	2.00	2.65	2.67	3.36	4.30			

Least-squares (Solver)									
	AGE (i)	1	2	3	4	5	d^2	χ^2	MLH
Parameters									
Sample1	K	8000	3999	1999	1000	500	3.887E-09	4.185E-05	-2.659
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	37.00			
	σ	1.00	1.00	1.00	1.00	1.00			
Sample2	K	8002	4000	1999	997	504	1.197E-08	6.010E-05	-3.221
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	31.99	36.98			
	σ	2.00	2.00	2.00	2.00	2.01			
Sample3	K	7984	4100	1459	973	983	1.046E-07	1.816E-04	-3.392
	P	0.52	0.26	0.09	0.06	0.06			
	μ	10.99	19.05	25.59	29.77	34.98			
	σ	3.00	3.04	2.70	3.21	3.71			
Sample4	K	8000	4000	1999	998	500	5.197E-09	3.940E-05	-2.800
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	36.99			
	σ	1.00	1.30	1.55	1.75	1.90			
Sample5	K	7996	4076	1491	1058	879	3.803E-08	8.375E-05	-3.250
	P	0.52	0.26	0.10	0.07	0.06			
	μ	11.00	19.04	25.71	29.86	35.27			
	σ	2.00	2.62	2.82	3.69	4.18			

χ^2 minimum (Solver)									
	AGE (i)	1	2	3	4	5	d^2	χ^2	MLH
Parameters									
Sample1	K	8000	3999	1999	1000	500	1.051E-08	4.116E-05	-2.659
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	37.00			
	σ	1.00	1.00	1.00	1.00	1.00			
Sample2	K	7999	4001	2006	976	520	3.812E-08	5.623E-05	-3.221
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.01	31.98	36.92			
	σ	2.00	2.00	2.00	1.96	2.05			
Sample3	K	7999	3949	2031	1256	263	3.065E-08	3.907E-05	-3.392
	P	0.52	0.25	0.13	0.08	0.02			
	μ	11.00	18.96	25.92	32.59	38.13			
	σ	3.00	2.98	3.07	3.41	2.66			
Sample4	K	8000	4000	2000	996	502	6.759E-09	3.928E-05	-2.800
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	36.99			
	σ	1.00	1.30	1.55	1.75	1.90			
Sample5	K	7994	4048	1734	1363	361	3.235E-08	7.624E-05	-3.250
	P	0.52	0.26	0.11	0.09	0.02			
	μ	11.00	19.02	25.87	31.66	37.60			
	σ	2.00	2.62	2.97	3.99	3.73			

MLH (Solver)									
	AGE (i)	1	2	3	4	5	d^2	χ^2	MLH
Parameters									
Sample1	K	8000	3999	1999	1000	500	1.946E-08	4.183E-05	-2.659
	P	0.52	0.26	0.13	0.06	0.03			
	μ	11.00	19.00	26.00	32.00	37.00			
	σ	1.00	1.00	1.00	1.00	1.00			
Sample2	K	7989	4078	1626	882	927	1.490E-05	1.506E-03	-3.222
	P	0.52	0.26	0.10	0.06	0.06			
	μ	11.00	19.05	25.71	30.21	35.11			
	σ	2.00	2.05	1.67	2.22	2.80			
Sample3	K	8188	3949	1396	948	1019	4.890E-06	3.762E-04	-3.392
	P	0.53	0.25	0.09	0.06	0.07			
	μ	11.10	19.26	25.41	29.77	35.23			
	σ	3.05	2.89	2.26	2.31	3.31			
Sample4	K	7998	4008	1860	716	916	2.812E-05	3.238E-03	-2.802
	P	0.52	0.26	0.12	0.05	0.06			
	μ	11.00	19.00	25.87	30.85	34.93			
	σ	1.00	1.30	1.45	2.12	2.80			
Sample5	K	7982	4308	1204	980	1025	3.860E-06	6.037E-04	-3.250
	P	0.51	0.28	0.08	0.06	0.07			
	μ	10.99	19.20	25.52	29.71	35.31			
	σ	2.00	2.71	2.13	2.41	3.83			

解を初期値として再度操作を繰り返すと精度の高い解が得られた。これら5例題についての推定結果は、操作を5回繰り返して得た解である。各年齢のモードがはっきりしている例題1では、例題条件と同一の推定結果が得られた。また、全ての例題でモードのはっきりしている1, 2才は適正な解が得られたが、例題3と5についてはモードのはっきりしない3~5才は例題条件とかけ離れた解で精度も悪かった。

マダイ尾叉長データの適用例では、初期値Aを用いると、Hasselblad法と最小二乗法は収束したが、対数尤度最大化法では収束せず、 χ^2 最小化法ではパラメータが0や明らかに異常な解に収束してしまった。大域的に解を探索するために、解の探索方法を共役傾斜法、微分係数を中央に設定しても、同じ結果となったので、Solverには初期値AをHasselblad法で10回反復して得た初期値Bを用いた。結果を表7と図1に示した。マダイの適用例においてCPUにPentium 100 MHzの場合の計算時間および反復回数を表8に示した。最尤法であるHasselblad法と対数尤度最大化法による収束状況を図2に示す。Solverでは収束するまでの計算時間と回数は精度の設定により変化する、それぞれ2~6

分、16~57回であった。

考 察

本報は赤嶺 (1982, 1984, 1985, 1995) および Akamine (1987) の "Polymodalな度数分布を正規分布に分解する方法" について、MS-Excelによる計算法を提案するものである。

年齢組成Pを年齢別個体数Kに、体長組成Fを体長度数分布Cに置き換えて計算する手法も考えられる。この場合、Hasselblad法では計算が可能で同じ解を得ているが、Solverでは適当な初期値を与えないとKがほとんど変化せず、収束しなかったり明らかに異常な解に収束してしまった。PをKに置き換えて計算する場合には事前にHasselblad法で得た解を初期値とする必要があった。

計算例題では、 σ が大きく、モードのはっきりしない例題3~5では推定精度が低かったが、条件のよい初期値を用いれば精度の高い解が得られると思われる。

マダイの適用例における計算時間は、Solverは2~6分、Hasselblad法は1000回で10分であった。Akamine (1987) および赤嶺 (1995) はHasselblad法では反復回数200回で

表7. マダイデータの計算結果

Hasselblad's method														
AGE (i)	0	1	2	3	4	5	6	7	8	9	10	d ²	x ²	MLH
Parameters														
K	1906	3820	6371	4976	2570	1293	556	461	258	177	226	2.554E-05	2.385E-02	-3.647
P	0.08	0.17	0.28	0.22	0.11	0.06	0.02	0.02	0.01	0.01	0.01			
μ	18.3	22.5	27.8	33.6	39.6	45.1	49.6	53.4	57.9	61.4	64.9			
σ	3.15	1.88	2.61	2.82	2.50	2.26	1.61	1.90	2.29	4.81	6.36			

Least-squares (Solver)														
AGE (i)	0	1	2	3	4	5	6	7	8	9	10	d ²	x ²	MLH
Parameters														
K	1195	4664	6322	4832	2426	1394	712	398	237	208	225	2.497E-05	2.181E-01	-3.647
P	0.05	0.21	0.28	0.21	0.11	0.06	0.03	0.02	0.01	0.01	0.01			
μ	16.7	22.3	27.9	33.6	39.4	44.5	49.1	53.2	56.8	59.9	67.0			
σ	2.46	2.08	2.63	2.83	2.59	2.83	2.51	2.15	2.40	2.67	4.79			

x ² minimum (Solver)														
AGE (i)	0	1	2	3	4	5	6	7	8	9	10	d ²	x ²	MLH
Parameters														
K	1051	5523	5611	4628	2501	1401	773	449	222	214	239	6.533E-05	9.756E-03	-3.649
P	0.05	0.24	0.25	0.20	0.11	0.06	0.03	0.02	0.01	0.01	0.01			
μ	16.1	22.6	28.1	33.4	39.0	44.1	49.0	53.3	56.8	59.9	67.1			
σ	3.25	2.54	2.61	2.74	2.41	2.27	1.97	2.05	2.41	2.67	4.97			

MLH (Solver)														
AGE (i)	0	1	2	3	4	5	6	7	8	9	10	d ²	x ²	MLH
Parameters														
K	1000	5341	5782	4706	2493	1400	772	453	220	223	222	3.181E-05	1.037E-01	-3.647
P	0.04	0.24	0.26	0.21	0.11	0.06	0.03	0.02	0.01	0.01	0.01			
μ	16.2	22.5	28.1	33.5	39.0	44.1	49.0	53.3	56.8	59.9	67.1			
σ	2.48	2.30	2.48	2.76	2.46	2.29	1.97	2.05	2.40	2.69	4.85			

十分であると述べており、この場合は2分となるので、いずれの方法を用いても計算時間においては実用上問題ないと考えられる。最尤法のHasselblad法と対数尤度最大化法では、精度についてはオプションで設定するので優劣を比較することはできない。ただし、今回の設定では、対数尤度最大化法は反復回数が少なくとも1回あたりの計算時間が長いので、同じ計算時間内ではHasselblad法が精度が高い結果となった。また、Solverを用いた場合の安定性は最小二乗法>x²最小化法>対数尤度最大化法で、条件の悪いデータでは後2者については初期値をHasselblad法で求める必要があった。赤嶺(1984, 1985, 1995)およびAkamine(1987)は解の探索方法と目的関数について、

- 1) ニュートン法は解近傍での収束が速く精度が高いが、初期値が解から離れていると振動・発散を起こす。
 - 2) Hasselblad法は反復回数を多くする必要があるが、計算量が少ないので計算時間は問題にはならない。また、安定的に解が得られ、パラメータの多いモデルに向いている。
 - 3) 目的関数を最小二乗法にすると条件の悪いデータでも安定的に解が得られる。
- と述べており、今回は解の探索方法が準ニュートン法で違いがあるものの、同様の結果が得られた。

Akamine(1987)、赤嶺(1995)は、これらに加えて

- 4) 最尤法は理論的には最良であるが、飛び離れていたり、

- 裾を長くひく異常なデータを含む場合には適用できない。
 - 5) 最小二乗法系の手法では、x²最小化法が最尤法に近い解を得られる。
 - 6) 最小二乗法系の手法は一部のデータを削除するロバスト的推定が容易である。
- としている。

今回の結果から、最尤法を適用する場合には計算時間、安定性および精度を考慮して、Hasselblad法を用いることが望ましい。ただ、理論的には最良であっても、現場で得られるデータは必ずしも条件が良くなく、 $P_i > P_{i+1}$ 、 $\sigma_i < \sigma_{i+1}$ 等の条件設定や一部のデータを削除する場合も考えられるので、安定的に解が得られ、しかも条件設定と操作が容易なSolverによる最小二乗法が実用的である。また、赤嶺(私信)はHasselblad法においてもPが一定な場合等の特殊な例についても解説しており、今後表計算ソフトによる計算法を検討していきたい。

本報ではMS-Excelを用いて、混合正規分布のパラメータ推定法を紹介した。両方法とも作図機能を併用することで、計算過程が確認しやすい。また、目的関数の微分、2階微分が必要ないため、アルゴリズムの導出が容易で、応用範囲が広く、実用的と考えられる。

ただし、初期値の決定や得られた解の評価等は使用者が行う必要がある。どの方法を使用するか、また、制限条件

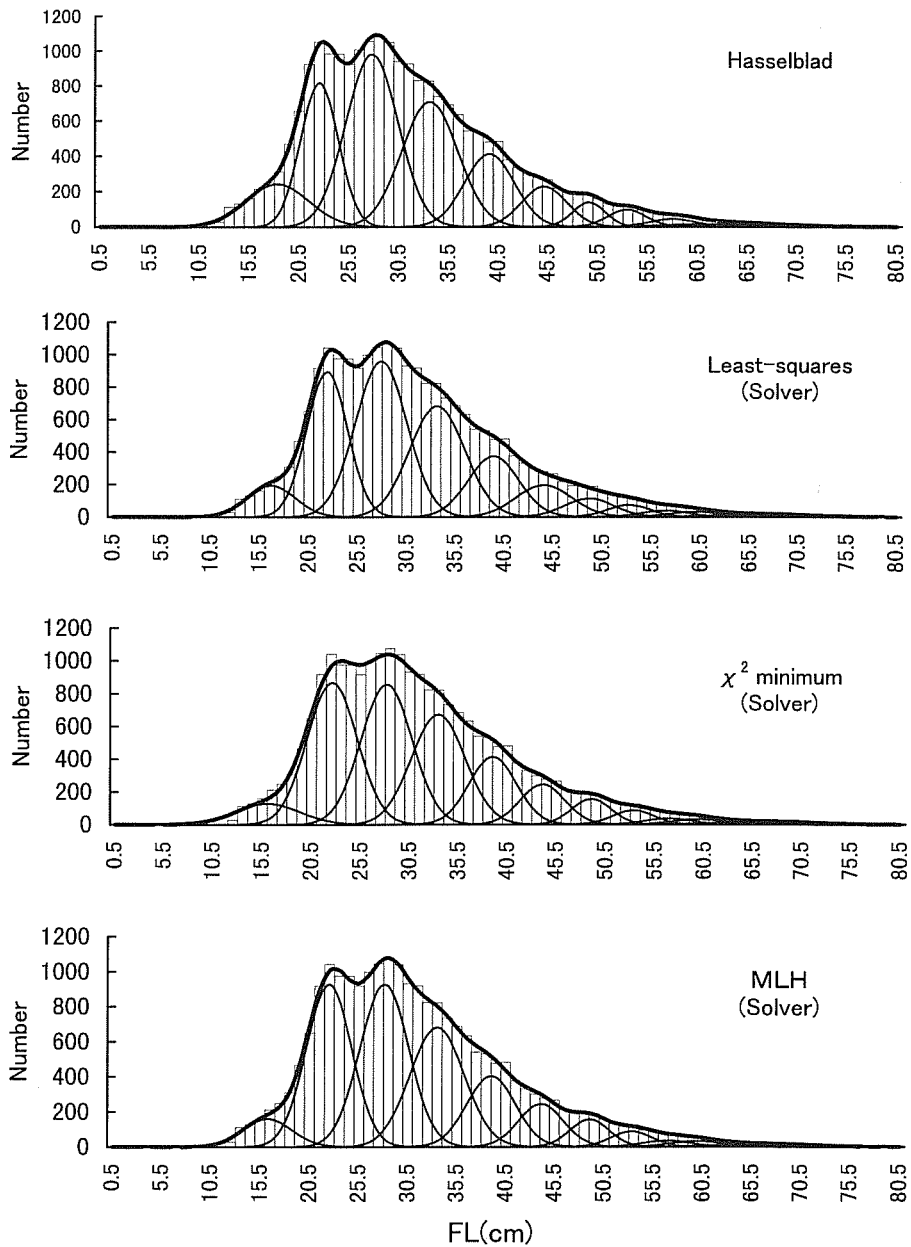


図1. マダイデータの計算結果. ヒストグラムは尾叉長の度数分布の実測値実線は正規分布を示す.

表8. 各方法の反復計算回数と計算時間. 精度はSolverのオプションで設定する.

Methods	Number of iteration	Time (min)	Precision
Hasselblad's method	1000	10	—
Least-Squares(Solver)	51	5	10^{-5}
χ^2 minimum (Solver)	40	4	10^{-5}
MLH (Solver)	16	2	10^{-4}
ditto	19	2	10^{-5}
ditto	57	6	10^{-6}

をどのように設定するかは、データの質で選択すればよいだろう。

謝 辞

本研究を進めるにあたり、貴重なご助言と文献、そして励ましを賜った中央水産研究所 赤嶺 達郎博士並びに三重県水産技術センター山川 卓博士に深謝の意を表します。中村良成主任研究員をはじめとする神奈川県水産総合研究所栽培技術部の皆様からはご指導とご協力を、資料整理には西元裕子さんと古知美華さんにご協力をいただいた。漁業

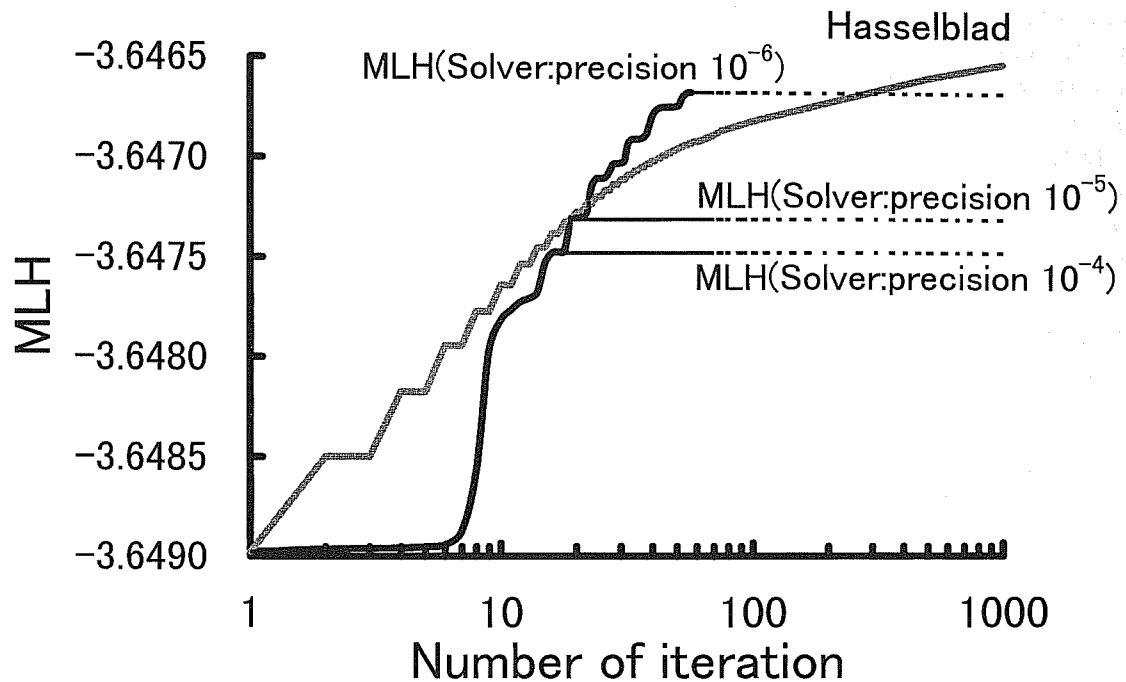


図2. Hasselblad法と対数尤度最大化法 (Solver) による対数尤度の収束状況。

協同組合関係者と市場関係者の皆様には市場調査で貴重なマダイの測定に便宜を図っていただいた、心よりお礼申し上げます。

引用文献

- 相澤 康・倉田健太郎 (1999) 神奈川県横須賀市大楠漁業協同組合で水揚げされたマダイの年齢組成の推定方法について、神水総研研報, 4, 19-26.
- 赤嶺達郎 (1982) Polymodalな度数分布を正規分布へ分解するBASICプログラム、日本研報告, 33, 163-166.
- 赤嶺達郎 (1984) Marquadt法によるPolymodalな度数分布を正規分布へ分解するBASICプログラム、日本研報告, 34, 53-60.
- 赤嶺達郎 (1985) Polymodalな度数分布を正規分布へ分解するBASICプログラムの検討、日本研報告, 35, 129-159.
- Akamine, T. (1987) Comparison of algorithms of several methods for estimating parameters of a mixture of normal distributions. Bull. Jap. Sea Reg. Fish. Res. Lab., 37, 259-277.
- 赤嶺達郎 (1995) 水産資源学における成長式に関する数理的研究。中央水研研報, 7, 189-263.
- Hasselblad, V. (1966) Estimation of parameters for a mixture of normal distributions. Technometrics, 8, 431-444.
- 今井利為 (1996) 神奈川県におけるマダイ種苗放流効果の推定。栽培技術研究, 25, 59-74.
- 真子 渺・松宮義晴 (1977) 銘柄組成による年齢組成推定法。西海区水研研報, 50, 1-8.
- 東海 正 (1997) MS-Excelのソルバーによる曳網の網目選択性Logistic式パラメータの最尤推定。水産海洋研究, 61, 288-298.
- 藤森康澄・東海 正 (1999) 石田の方法と北原の方法によるMS-Excelを用いた刺網の網目選択性曲線の推定。水産海洋研究, 63, 14-25.
- 山川 卓 (1997a) 体長組成法、「水産動物の成長解析」(赤嶺達郎・麦谷泰雄編), 水産学シリーズ 115, 恒星社厚生閣, 東京, pp. 39-51.

京, pp. 39-51.

山川 卓 (1997b) 複数体長組成データの解析によるイセエビの成長と齢別組成および加入の推定。水産海洋研究, 61, 23-32.

山川 卓 (1997c) イセエビの資源評価と漁業管理。三重県水技研報, 7, 1-96.

Yamakawa, T. and Y. Matsumiya (1997) Simultaneous analysis of multiple length frequency data sets when the growth rates fluctuate between years. Fisheries Sci., 63, 708-714.

付 録

1. Hasselblad 法

(1) ワークシートの構成 (図A)

① B22:F24

パラメータを入力するセル。

② A25:A75, B83:A133, A141:A191, A199:A249

階級値 (x_{ij}) を入力するセル。

③ B25:F75

年齢別体長組成 (f_{ij}) を計算するセル。(1) 式。

パラメータの入力セルを参照して

“=NORMDIST (x_{ij} , μ_i , σ_i) · h” を入力する。

④ G24:G75

混合正規分布。G25:G75は(2)式, G24は(3)式。

⑤ H25:H75

I25:I75の体長度数分布 (C_{ij}) を体長組成に変換するセル。

⑥ I25:I75

体長度数分布 (実測値) を入力するセル。

⑦ J24:J75, K24:K75, L24:L75

χ^2 最小化法, 最小二乗法および対数尤度最大化法の目的関数のセル, J24:L24は(9)~(11)式の左辺で誤差の平方和(d^2), χ^2 値(χ^2)および対数尤度(MLH), J25:L75が右辺各項.

⑧ B80:F80, B83:F133

年齢組成(P_i)の最適解を計算するセル, B80:F80は(6)式, B83:F133は(4)式.

⑨ B138:F139, B141:F191

平均体長(μ_i)の最適解を計算するセル, B138:F138は P^{new} を参照, B139:F139は(7)式左辺, B141:F191は右辺各項.

⑩ B196:B249

標準偏差(σ_i)の最適解を計算するセル, B196:F197は P^{new} と μ^{new} を参照, B198:F198は(8)式左辺の平方根, B199:F249は右辺各項.

⑪ B16:F18

計算結果を出力するセル, B22:F24を参照, PのセルB22:F22には総個体数(I24)を乗じて, Kに変換.

⑫ 混合正規分布と体長組成(実測値)を示すグラフ

細実線は正規分布③ B25:F75を, 太実線は混合正規分布④ G25:G75を, ヒストグラムは体長組成(実測値)⑤ H25:H75を参照.

⑬ 各パラメータの計算結果を示すグラフ

それぞれB16:F18を参照.

(2) 操作の手順

①パラメータの初期値入力

グラフと目的関数の値を見ながらB22:F24に適当な初期値を入力する. G24($\sum P = \sum g$)=1に注意する. B196:F198にパラメータ new が出力される.

②反復計算機能への切り替え(図B)

MS-Excelの[ツール]-[オプション]-[計算方法]のダイアログボックスを開き, 手動, 再計算に設定する.

③パラメータ new の入力

B22:F24にB196:F198を参照する数式(すなわち, B22に“=B196”, B23に“=B197”...F24に“=F198”)を入力する. 省力化のため, B22に“=B196”を入力し, B22:F24にコピーする.(入力中は計算しない.)

④再計算

f・9キーを押して計算する. B16:F18に結果が出力される.

2. Solverによる方法

(1) ワークシートの構成

Hasselblad法のワークシートを流用できる. ただし, Solverのみ使用の場合は行77以下は不要.

(2) 操作の手順

①初期値の入力

Hasselblad法に同じ.

②Solverの起動とパラメータの設定(図C, D)

MS-Excelの[ツール]-[ソルバー]のダイアログボックスを開き, パラメータを設定する.

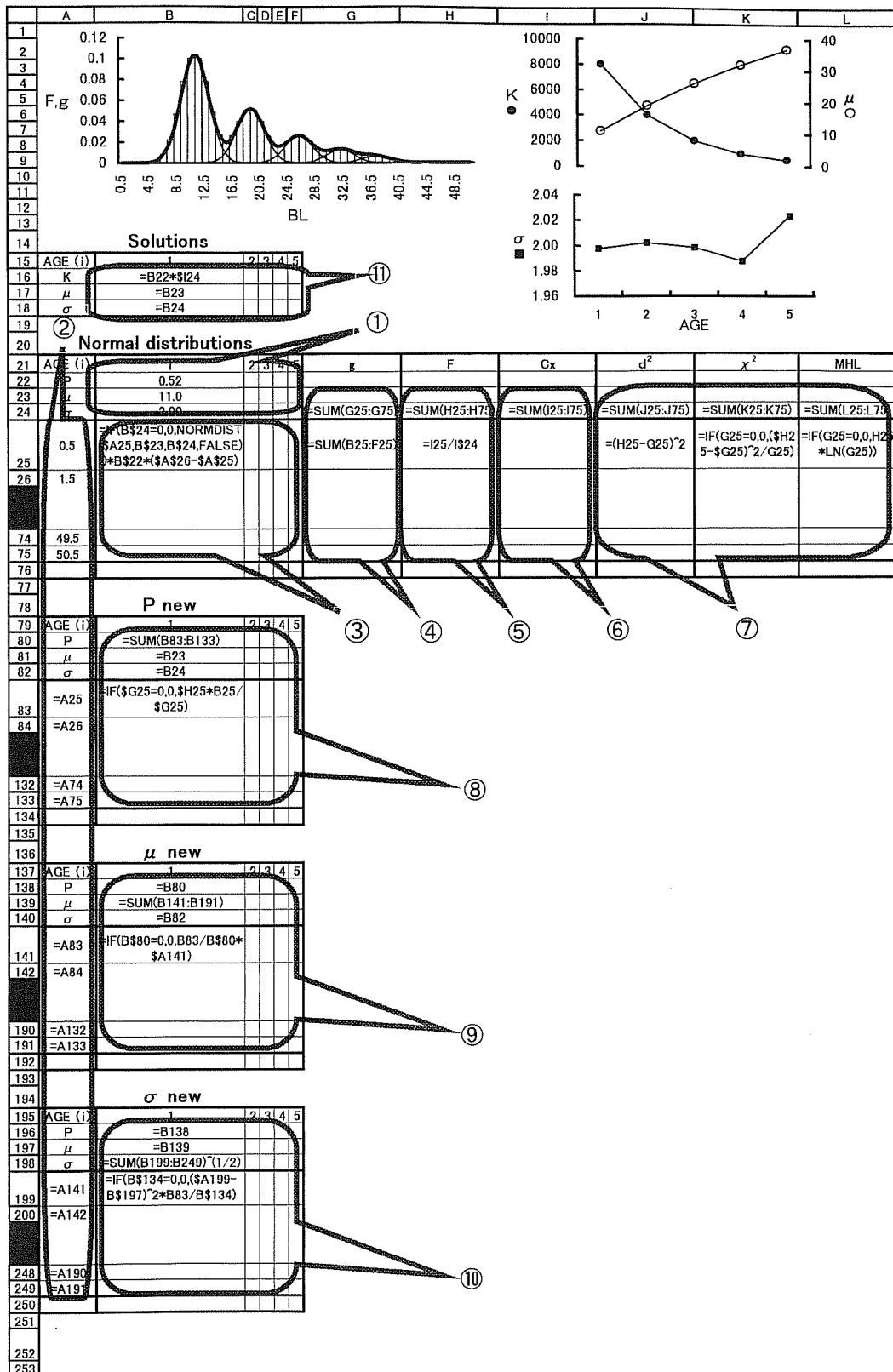
目的セルは誤差の平方和(J24)と χ^2 値(K24)は最小値, 対数尤度(L24)は最大値に, 変化させるセルはP, μ , σ のパラメータ(B22:F26)を, 制約条件は(3)式と(12)式を設定する. また, オプション設定で解の探索方法他を設定する.

3. その他

(1) ERRが生じないように, ③と⑦~⑩は“IF”関数を使用する.

(2) 絶対参照“\$*\$*”により入力作業の省力化をする.(例えばB25に正規分布の数式を入力し, B25:F75にコピーすることで, ③の範囲を一括して入力する.)

(3) $i > 5$, $j > 50$ の場合は, それぞれ列行を挿入して, ワークシートを作り直す.



図A. ワークシートの構成

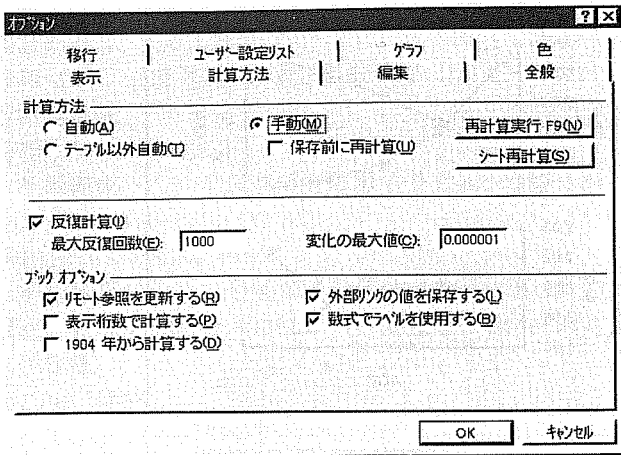


図 B. MS-Excel のオプションダイアログボックス

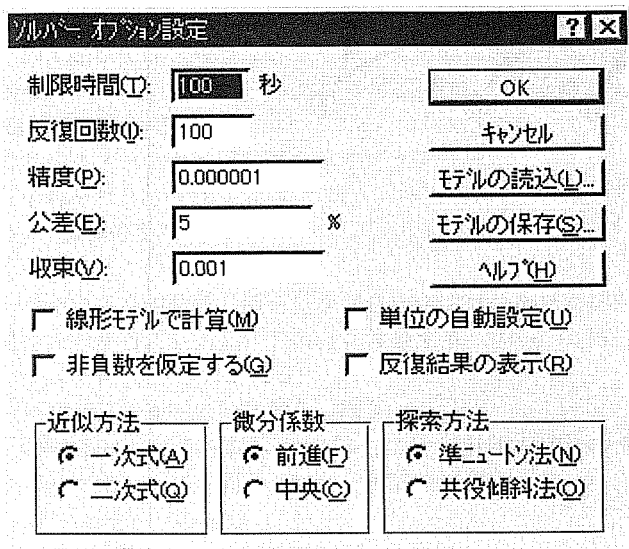
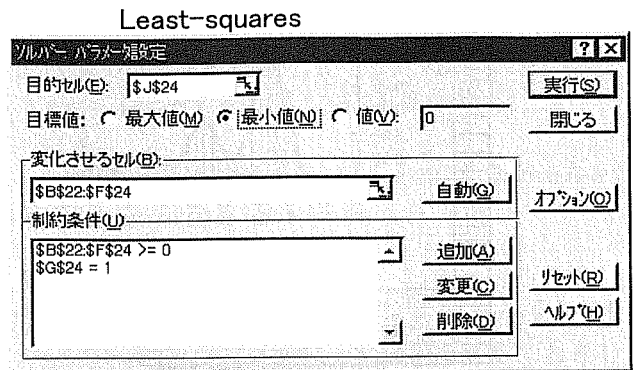
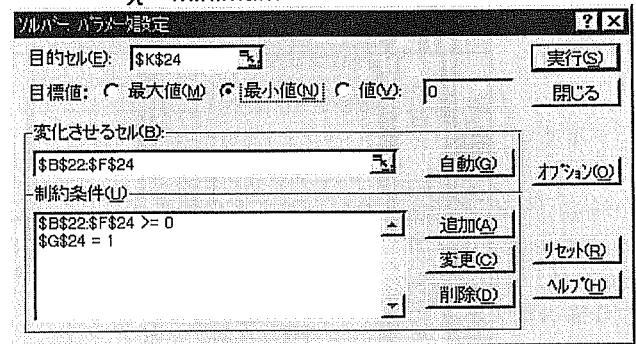


図 D. Solver のオプションダイアログボックス



X^2 minimum



MHL

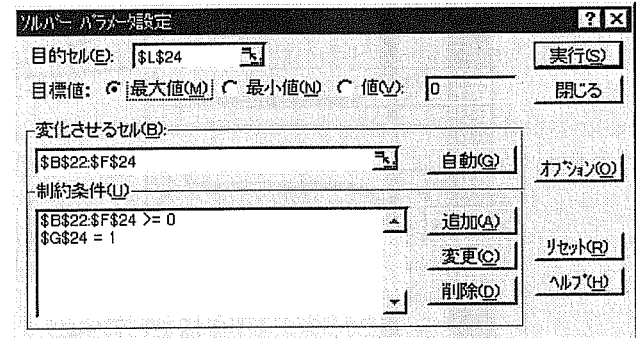


図 C. Solver の設定ダイアログボックス