

# Kaplan-Meierの方法と数量化理論I類を組み合わせた複合的寿命推定法

上田 徹・原田 要之助

## 1. ま え が き

電子交換機などの大規模かつ複雑なシステムの寿命（故障時間）はシステム構成要素の数、冗長度や故障率ばかりでなく、システムの設置場所、保守員の技術力、予備パッケージの配備体制などの質的条件に依存すると思われる。従来行なわれてきた故障率にもとづく寿命推定法は故障率の推定精度に依存するが、構成要素の故障率推定精度は必ずしも高くない。そのような場合に質的データ（環境条件、属性）を利用した寿命推定法の適用が考えられる。質的データを使う手法としてよく知られている林の数量化理論第I類（ダミー変数法）[1]の利用が考えられるが、この方法では故障時間のわかっているデータしか使えない。これに対して信頼度関数のノンパラメトリックな推定法である Kaplan-Meier の方法[2]では属性データを用いないかわりに、故障時間データばかりでなく稼働中のシステムの年令も用いる。本稿ではこれら2手法の長所を組み合わせた寿命推定法（複合的寿命推定法）を提案する。

数量化理論第I類では各属性をアイテムと呼び、アイテムは複数のカテゴリに分けられる。カテゴリの数やカテゴリ間の境界は必ずしも明確でなく、あいまいさを含んでいる。そこであいまい

うえだ とおる, はらだ ようのすけ

日本電信電話公社 武蔵野電気通信研究所

さを考慮した推定法も提案する。

これら新しい手法の有効性のある疾患の発病年令データおよび交換機の故障時間データを用いて検証する。

## 2. 既存手法の概要

### 2.1 電子交換機の信頼度設計法[3]

電子交換機のシステム不稼働原因で最も影響が大きいものは2重化されている中央制御装置（以下CCと呼ぶ）の同時故障である。そこで両CCの停止確率推定法を以下に示す。

1CCの状態として、

G：正常

D：故障中で自動診断未了（相手CCもその後、故障したため自動診断ができず、手動診断中の状態も含む）

M：自動診断では原因が発見できず手動診断中

R：パッケージ差し替え中

の4状態を考える。2CCの状態を、各CCの状態I, Jの組合せ[I, J]で表わすと図1に示すような状態遷移が考えられる。両CCとも状態Gでない（図1の状態4, 5, 6, 7の）ときシステムは停止する。なお時間  $\Delta t$  のあいだに他の状態に遷移する確率は、

(図1の矢印線上の記号)  $\times \Delta t$

で与えられるものとする。このとき、定常状態

( $t \rightarrow \infty$ ) においてシステムが停止状態にある確率  $h$  は、

$$h \doteq f^2 \left\{ \frac{2}{r^2} + \left( \frac{1}{r} + \frac{2}{\nu m} \right) \left( \frac{1}{a} + \frac{q}{m} \right) + \left( \frac{q}{m \cdot r} \right) \right\} \quad (1)$$

で与えられる.[3]

わが国の代表的な電子交換機である D10 形交換機は多数の部品を使用しているが、方式設計時には十分な故障データがなく、導入当初の故障率推定も安全側に評価した値を用いざるを得ない.[4] しかし、運用実績を積みにつれ、故障の発生間隔分布や修理時間分布に依存せず、観測データのみから推定するノンパラメトリックな手法、たとえば Kaplan-Meier の方法の適用が可能となってくる。

さらに電子交換機の運用実績によればシステムの停止は必ずしもハードに帰因せず、保守者の不慣れや雷などの自然現象によるシステム停止のあることもわかり、これらの要因も考慮する必要がある。保守者の習熟度を客観的に調査することは困難であるが、交換機の規模が大きければ保守者が多く、比較的短時間で復旧する可能性がある。また地域的な訓練体制や熱心さの相違があるかもしれない。これらのことから比較的手しやすい属性として交換機規模、地域性(所属通信局)などを用いて故障にいたるまでの期間(寿命)を数量化理論第 I 類により推定することが考えられる。

## 2.2 Kaplan-Meier の方法

現時点までに得られた故障システム  $F_h$  の故障までの稼働期間データと正常システム  $G_i$  の連続稼働期間(年令)データを年令順に並べた図 2 に示すようなデータがあるものとする。このとき、Kaplan-Meier の方法では期間  $t_h$  よりも長く稼働する確率  $P(t_h)$  の推定量  $\hat{P}(t_h)$  は、

$$\hat{P}(t_0) = 1, \quad (2)$$

$$\hat{P}(t_1) = 1 - f_1/N \quad (3)$$

$$\hat{P}(t_h)/\hat{P}(t_{h-1}) = 1 - f_h / \left\{ N - \sum_{j=1}^{h-1} (f_j + n_j) \right\} \quad (4)$$

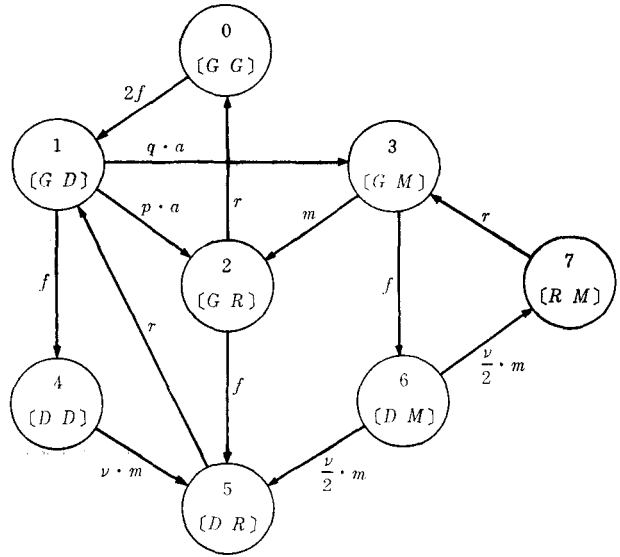


図 1 2重化されたCCの状態遷移

$f$ : CC故障率  $a$ : 自動診断終了率  $p$ : 自動診断適中率  
 $p+q=1$   $m$ : 手動診断終了率 ( $\nu=2$ : 並列修理)  
 $\nu=1$ : 直列修理  
 $r$ : パッケージ取替終了率

で与えられる。ここで、

$f_h$ : ちょうど稼働期間  $t_h$  で故障した標本数、

$n_h$ : 稼働期間  $t_h$  で未故障の標本数、

$N$ : 全標本数

である。

## 2.3 林の数量化理論第 I 類

システム属性が既知で、ダミー変数

$$d_H(jk) = \begin{cases} 1 : \text{システム } H \text{ がアイテム } j \text{ のカ} \\ \quad \text{テゴリ } h \text{ の属性をもつとき,} \\ 0 : \text{その他} \end{cases} \quad (5)$$

が与えられているものとする。ただし、 $H$  は図 2  $F_h, G_h$  のように並べられ、稼働期間(以下、「年令」と呼ぶ)  $t_h$  で故障した  $f_h$  個のシステムがあるときには、その  $i$  番目のシステムを  $(h, i)$  と表わすことにする。(  $i=1, 2, \dots, f_h$  )

故障時の年令(以下、「寿命」と呼ぶ)の推定量  $t_H^{(1)}$  は次の方法で得られる。

[方法 1]

$$t_H^{(1)} = \sum_j \sum_k d_H(jk) x_{jk}^{(1)} \quad (6)$$

ただし、 $x_{jk}^{(1)}$  は寿命のわかっているシステム

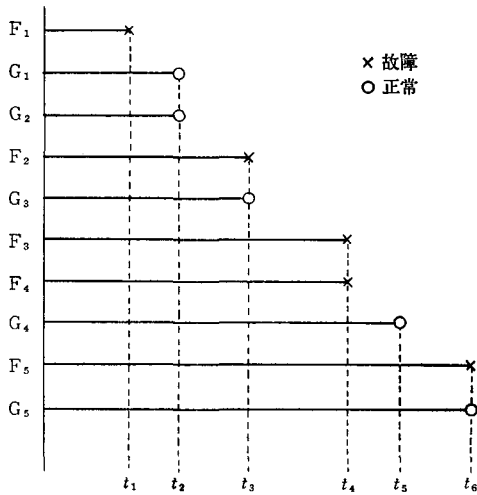


図 2 ある故障データ例

$H=(h, i)$ の属性データ  $d_{(h, i)}(jk)$ を用いて,

$$Q^2 = \sum_h \sum_{j=1}^{f_h} (t_h - t_{(h, i)}^{(1)})^2$$

$$= \sum_h \sum_{i=1}^{f_h} \{t_h - \sum_j \sum_k d_{(h, i)}(jk) x_{jk}^{(1)}\}^2 \quad (7)$$

を最小にするように決められる。

### 3. 新しい寿命推定法

#### 3.1 複合的寿命推定法

数量化理論第I類で用いられる評価尺度, すなわち式(7)の自然な拡張として,

$$Q_w^2 = \sum_h \sum_{i=1}^{f_h} w_h (t_h - t_{(h, i)}^{(1)})^2 \quad (8)$$

を最小にするように式(6)の  $x_{jk}$  を求めることが考えられる。ここで重み  $w_h$

$$\sum_h \sum_{i=1}^{f_h} w_h = \sum_h f_h \cdot w_h = 1 \quad (9)$$

となるように決める。ところで  $t_h$  なる値をとる標本数は  $f_h$  であったが, Kaplan-Meier の方法により, 母集団での

$$q(t_h) \triangleq \text{Prob}(t_{h-1} < \text{寿命} \leq t_h) \quad (10)$$

$$\text{の推定量 } \hat{q}(t_h) = \hat{P}(t_{h-1}) - \hat{P}(t_h) \quad (11)$$

を求めることができるので,  $t_h$  は  $\hat{q}(t_h)$  により重みづけられたほうが母集団での比率を反映できる。そこで,

$$f_h \cdot w_h = \hat{q}(t_h)$$

すなわち,

$$w_h = \hat{q}(t_h) / f_h \quad (12)$$

とする。以上から数量化理論第I類と Kaplan-Meier の方法を組み合わせた次の複合的寿命推定法が得られる。

[方法2]

$$t_H^{(2)} = \sum_j \sum_k d_H(jk) x_{jk}^{(2)} \quad (13)$$

ただし,  $x_{jk}^{(2)}$  は式(12)で与えられる  $w_h$  および属性データ  $d_{(h, i)}(jk)$  を用いて,

$$Q_w^2 = \sum_h \sum_{i=1}^{f_h} w_h (t_h - t_{(h, i)}^{(2)})^2$$

$$= \sum_h \sum_{i=1}^{f_h} w_h \{t_h - \sum_j \sum_k d_{(h, i)}(jk) x_{jk}^{(2)}\}^2 \quad (14)$$

を最小にするように決められる。

式(14)は自乗誤差  $(t_h - t_{(h, i)}^{(2)})^2$  を重み  $w_h$  で評価することを意味しているが, 式(14)の代りに,

$$Q_v^2 = \sum_h \sum_{i=1}^{f_h} \{v_h (t_h - t_{(h, i)}^{(2)})\}^2 \quad (15)$$

すなわち, 誤差自身に重みをつけることが考えられる。この場合にも式(9), (12)の考え方は適用でき,

$$v_h = \hat{q}(t_h) / f_h \quad (16)$$

で与え得る。式(14), (15)の相違は式(12)で与えられる重み  $w_h$  を用いて自乗誤差  $(t_h - t_{(h, i)}^{(2)})^2$  に対して  $w_h$  倍するか,  $w_h^2$  倍するかの差である。2種類の重み  $w_1, w_2$  に対して,

$$w_1/w_2 > 1 \text{ ならば } (w_1/w_2)^2 > w_1/w_2$$

なので, 式(15)による評価は多数派にさらに重みを加えることになり, 自乗誤差が式(14)による場合よりも改善できると思われる, 第2の複合的寿命推定法が得られる。

[方法3]

$$t_H^{(3)} = \sum_j \sum_k d_H(jk) x_{jk}^{(3)} \quad (17)$$

ただし,  $x_{jk}^{(3)}$  は式(16)で与えられる  $v_h$  および属性データ  $d_{(h, i)}(jk)$  を用いて,

$$Q_v^2 = \sum_h \sum_{i=1}^{f_h} v_h^2 \{t_h - \sum_j \sum_k d_{(h, i)}(jk) x_{jk}^{(3)}\}^2 \quad (18)$$

を最小にするように決められる

[方法1]～[方法3]から得られるシステムHの寿命推定量 $t_h^{(1)}$ は現時点のシステムHの年令 $t_h$ よりも小さくなる場合がある。こうした場合を避けるため現稼動システムHに仮想寿命 $\tilde{t}_H$ を与え、試行錯誤的に最適な寿命を推定する次のような方法が考えられる。

[方法4] 現在、稼動中のシステムの集合をG、現時点までの故障システムの集合をFとする。

(i) [方法2]により集合G内のシステム*i*の寿命推定値 $t_i^{(2)}$ および $Q_w^2$ を求める。 $(i=1, 2, \dots, \sum_h n_h)$

(ii) システム*i*( $i \in G$ )の現時点での年令 $t_i$ と寿命推定値 $t_i^{(2)}$ とを比較する。

(iii) すべての*i*( $i \in G$ )について、
$$t_i \leq t_i^{(2)} \quad (19)$$

ならば[方法2]が有効として打ち切る。

(iv) 式(18)が成り立たないG内のシステムが存在するならばその集合をΓとする。Γ内システム数はγである。

(v) Γ内のシステムの現年令と推定寿命との差が*m*番目(初期値*m*=1)に大きいシステム*a*に対して寿命 $\tilde{t}_n^{(m)}$ を仮定する。ただし、*n*は*n*回目のくりかえしを意味し、

$$\tilde{t}_n^{(m)} = t_a + n \quad (20)$$

である。また*n*は適当なところで打ち切ることとし、たとえば $t_a/2$ 以下とすることが考えられる。

(vi) システム*a*の仮想寿命 $\tilde{t}_n^{(m)}$ およびすでに得られた $\tilde{t}^{(1)}, \dots, \tilde{t}^{(m-1)}$ も含めて年令 $t_h$ での故障システム数 $f_h^{(n)}$ および $\hat{q}(t_h^{(n)})$ を求め、式(12)から新しい重み $w_h^{(n)}$ を求める。式(14)の右辺において $w_h$ の代わりに $w_h^{(n)}$ を使って得られる $Q_w^2$ の値を $Q_n^2$ とする。ただし $Q_0^2 \triangleq Q_w^2$ である。

(vii)  $A \triangleq Q_{n-1}^2 / \sum_h (f_h^{(n)} \cdot w_h^{(n)}) > B \triangleq Q_n^2 / \sum_h f_h^{(n-1)} \cdot w_h^{(n-1)} \quad (21)$

ならば*n*に1加算し、(v)にもどる。

$A \leq B$ ならば、

$$\tilde{t}^{(m)} = t_a + (n-1) \quad (22)$$

とし、*m*に1加算して $m \leq \gamma$ ならば(v)にもどり、 $m > \gamma$ ならば(viii)にゆく。ただし、 $n=1$ ならば、

$$\tilde{t}_a^{(m)} = t_a + 1 \quad (23)$$

とする。

(viii) 以上から得られた年令 $\tilde{t}^{(1)}, \dots, \tilde{t}^{(\gamma)}$ を最終の推定寿命とする。

[方法4]については全システムの属性データを必要とすること、計算時間が長いことなどの欠点があり、本稿では考え方を示すにとどめ、[方法ℓ]、ℓ=1, 2, 3についてのみ精度を検討する。

### 3.2 分類のあいまいさの克服法

前節3.1で述べたどの方法も式(5)で与えられるダミー変数の値を必要とするが、各アイテムの分類は一意でなかったり、おおよそこのあたりといったあいまいさをもつ場合が多いと考えられる。そこで式(5)の代りに、

$$d_H^*(jk) : \begin{cases} = a_H(jk) : \text{システムHがアイテム} \\ \quad j \text{のカテゴリ} k \text{の属性を確率} \\ \quad a_H(jk) \text{でもつとき,} \\ 0 : \text{その他,} \end{cases} \quad (24)$$

$$\text{ただし, } \sum_k a_H(jk) = 1 \quad (25)$$

と一般化した形が考えられる。具体的な $a_H(jk)$ の与え方としては、

- (i) アイテムの特徴を考慮してあらかじめ与える方法
  - (ii) 試行錯誤的にいろいろな値を与える方法
  - (iii) AICなどの評価尺度にもとづき、 $a_H(jk)$ も推定対象として評価する方法
- などが考えられる。

## 4. 複合的寿命推定法の評価

### 4.1 循環器系疾患の発病年令推定問題

電子交換機の故障までの時間の推定に先立ち、分析データとして公知のものを用いて[方法1]～[方法3]の推定精度を比較する。

文献[5]の循環器系疾患データを用いて発病年

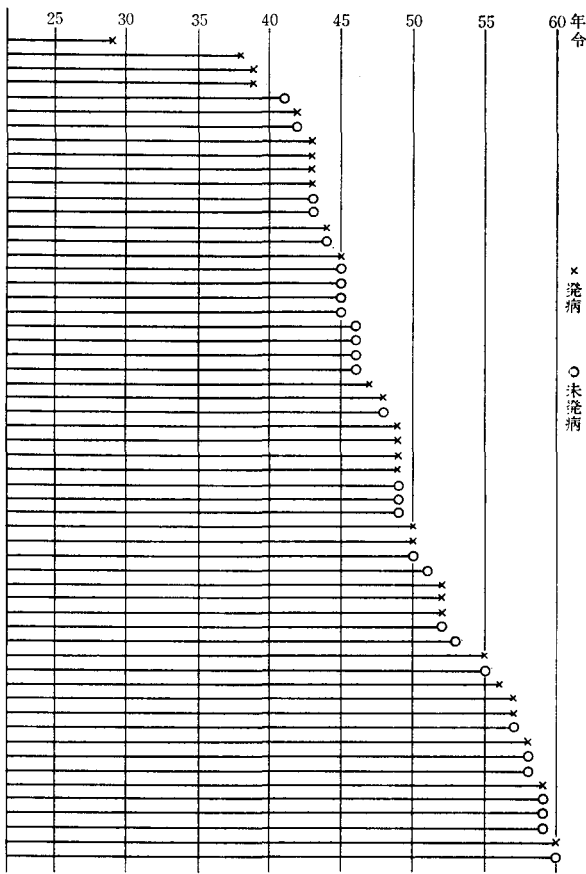


図 3 発病時の年齢あるいは時点  $T$  での年齢

表 2 各方法の推定誤差

	絶対推定誤差の平均		推定誤差の分散	
	(A)	(B)	(A)	(B)
方法 1	0.894	8.181	1.287	94.67
方法 2	0.904	7.740	1.380	87.18
方法 3	0.911	7.399	1.527	80.02

(A): 時点  $T$  までに発病したデータ

(B): 時点  $T$  での未発病データ

命を推定する問題に節 3 の方法を適用してみる。半数のデータを推定誤差評価用にもちいることとし、初回の検査（属性データ収集）後、7 年経過した時点  $T$  で、それまでに発病した人のデータ、すなわち、初回検査後の発症年数が 6 以下のデータが集められたものとする。

図 3 に発病年の年齢あるいは時点  $T$  での年齢を若いもの順に示す。

[方法 2], [方法 3] をもちいる場合の重み  $w_h, v_h$  を求める過程および導出結果を表 1 に示す。これらデータおよび属性データにもとづいて各方法を適用し、得られた推定誤差を表 2 に示す。

[方法 1] は時点  $T$  までの発病年齢の自乗推定誤差を最小にするものなので、当然、時

表 1 重みの導出過程（循環器系疾患の場合）

$t_h$	$f_h$	$n_h$	$\frac{\hat{P}(t_h)}{\hat{P}(t_{h-1})}$	$\hat{P}(t_h)$	$q(t_h)$	$w_h, v_h$	$t_h$	$f_h$	$n_h$	$\frac{\hat{P}(t_h)}{\hat{P}(t_{h-1})}$	$\hat{P}(t_h)$	$q(t_h)$	$w_h, v_h$
29	1	0	0.983	0.983	0.017	0.017	49	4	3	0.871	0.660	0.098	0.025
38	1	0	0.982	0.965	0.018	0.018	50	2	1	0.917	0.605	0.055	0.028
39	2	0	0.964	0.931	0.034	0.017	51	0	1	—	—	—	—
41	0	1	—	—	—	—	52	3	1	0.850	0.514	0.091	0.030
42	1	1	0.981	0.913	0.018	0.018	53	0	1	—	—	—	—
43	4	2	0.922	0.842	0.071	0.018	55	1	1	0.933	0.480	0.034	0.034
44	1	1	0.978	0.823	0.019	0.019	56	1	0	0.923	0.443	0.037	0.037
45	1	4	0.977	0.804	0.019	0.019	57	2	1	0.833	0.369	0.074	0.037
46	0	4	—	—	—	—	58	1	2	0.889	0.328	0.041	0.041
47	1	0	0.971	0.781	0.023	0.023	59	1	3	0.833	0.273	0.055	0.055
48	1	1	0.970	0.758	0.023	0.023	60	1	1	0.500	0.137	0.136	0.136
							計	29	29	—	—	0.863	—

表3 カテゴリ・スコア  $x_{jk}^{(1)}$

アイテム	カテゴリー	スコア
1	1	~39
	2	40~49
	3	50~
	レンジ	2.1246
2	2	110~139
	3	140~159
	4	160~179
	5	180~
	レンジ	2.1588
3	1	~79
	2	80~89
	3	90~99
	4	100~
	レンジ	0.9860
4	1	~159
	2	160~229
	3	230~
	レンジ	3.2889
5	1	正常
	2	やや正常
	3	異常
	4	非常に異常
	レンジ	3.1552
6	1	以前から吸わない
	2	1日に19本以下
	3	1日に20本以上
	4	現在はやめている
	レンジ	3.8803
7	1	以前から飲まない
	2	現在はやめている
	3	ときどき飲む
	4	毎日飲む
	レンジ	2.2234
8	1	やせている
	2	普通
	3	太っている
	4	肥満
	レンジ	1.00369

表4  $b^{(3)}(k_0, k)$  の値

$k_0 \backslash k$	1	2	3	4
1	1/2	1/3	1/6	0
2	2/7	3/7	2/7	0
3	2/9	2/9	1/3	2/9
4	0	1/6	1/3	1/2

点  $T$  までの発病データに対しては一番誤差が小さい。しかし、実際の推定対象である時点  $T$  以降の発病データに対しては [方法3], [方法2], [方法1] の順に誤差が小さい。

次に分類のあいまいさを考慮してみよう。式(24)で定義された  $a_H(jk)$  はシステム  $H$  のほぼこれだと思われるカテゴリ  $k_0$  がわかればシステム  $H$  にかわらず同じ値  $b^{(3)}(k_0, k)$  を与えられるものとする。

すなわち [方法1] で  $d_H(jk)=1$  となる  $k_0$  を中心に適当に値を与える。[方法1] を用いた場合のカテゴリ・スコア  $x_{jk}^{(1)}$  を表3に示す。アイテム3はカテゴリ・スコアのレンジからみると説明変数としての効果が小さい量であるが、それに対して表4に示すような  $b^{(3)}(k_0, k)$  の値を与えると、

絶対推定誤差平均=7.457,

推定誤差分散=82.07

と [方法3] よりも若干、劣る。他のアイテムに

表5(1) 交換機属性データ

No.	A	B	C	D	E	F	No.	A	B	C	D	E	F	No.	A	B	C	D	E	F
1	1	1	3	0	3	31	15	1	1	14	0	89	135	29	1	4	12	0	46	126
2	1	1	4	1	31	51	16	1	1	20	0	78	144	30	3	4	9	0	47	71
3	1	1	5	0	30	46	17	1	1	3	0	120	136	31	3	4	9	1	71	94
4	2	1	7	0	42	56	18	1	1	2	0	115	136	32	1	4	7	0	33	38
5	1	1	10	0	45	73	19	1	1	11	0	24	44	33	1	5	11	0	24	121
6	3	1	7	0	54	65	20	1	1	8	0	71	142	34	1	5	11	1	121	133
7	3	1	7	1	65	81	21	1	2	11	0	30	34	35	1	5	7	0	96	119
8	1	1	7	0	63	85	22	1	2	11	1	34	35	36	1	5	14	0	112	118
9	1	1	3	0	64	91	23	1	2	15	1	35	48	37	1	5	5	0	96	114
10	1	1	11	0	66	89	24	1	2	20	0	66	70	38	1	6	20	0	14	103
11	3	1	9	0	75	113	25	1	2	16	0	132	140	39	1	6	11	0	72	108
12	1	1	12	0	68	114	26	1	2	10	0	62	112	40	2	6	10	0	14	132
13	1	1	15	0	91	124	27	1	2	3	0	95	110	41	1	6	24	0	46	137
14	1	1	13	0	67	125	28	1	2	7	0	79	138	42	1	6	12	0	47	143

表 5(2) 交換機属性データ

No.	A	B	C	D	E	F	No.	A	B	C	D	E	F	No.	A	B	C	D	E	F
43	3	6	4	0	108	112	55	1	7	17	0	71	141	67	1	10	2	0	100	138
44	1	6	27	0	110	116	56	1	7	11	0	23	48	68	1	10	5	0	117	120
45	1	6	27	1	116	118	57	1	7	7	0	108	125	69	1	10	5	1	120	124
46	1	6	10	0	119	128	58	1	8	8	0	117	143	70	2	10	1	0	95	113
47	1	6	32	0	117	133	59	3	9	7	0	101	133	71	1	10	2	0	31	36
48	1	6	10	0	136	138	60	1	9	15	0	108	110	72	1	10	2	1	36	130
49	1	6	3	0	104	138	61	1	9	15	0	80	107	73	2	10	6	0	44	52
50	1	6	4	0	96	140	62	3	9	7	0	60	106	74	1	10	16	0	17	60
51	1	6	13	0	58	63	63	2	9	6	0	61	79	75	1	11	7	0	29	41
52	1	6	13	1	63	106	64	1	9	4	0	39	50	76	1	11	28	0	42	72
53	1	6	19	0	108	109	65	1	9	4	1	50	69	77	1	11	9	0	118	135
54	3	7	3	0	130	133	66	2	9	5	0	56	59							

A : 交換機種別 (1 : 市内, 2 : 市外, 3 : 併合)  
 B : 地域番号, C : 加入者および中継回線数  
 D : 改善歴 (0 : 無, 1 : 有)  
 E : サービス開始または再開時点  
 F : 故障時点

表 6 故障時間の推定誤差

分類	方法 1	方法 3
1	M	30.31
	V	1731
2	M	29.52
	V	1631
3	M	28.71
	V	1454

についても [方法 3] が改善されたとはいえない。  
 以上の検討にもとづき、次節では [方法 1] と最も精度の高かった [方法 3] だけを比較する。

#### 4.2 電子交換機の故障時点推定問題

電子交換機の属性としては、

(A)交換機種別, (B)所属通信局, (C)加入者および中継回線数, (D)故障歴の有無, をもちいる。表 5 (1)~(2)に示す属性データおよびサービス開始・故障時点データが得られている。時点96でそれ以降の故障時点を予測するものとする。時点96以前に複数回の故障があった場合には2回目以降の故障は故障歴有りという属性をもたせ、前回の故障からその故障までの時間を故障時間とした。最後の故障から時点96までの期間は正常システムのデータに加えた。ここで属性(B)については故障システムのなかった通信局もあり、統合を行なう必要があり、(C)についてはいろいろな分類が考えられる。そこで(B), (C)については次の分類をとりあげ検討する。

[分類 1]

(B) : 1, 2, (3, 4), 5, 6, (7, 8, 9), (10, 11),

|M| : 絶対推定誤差の平均  
 V : 推定誤差の分散

ただし、( )内は同じ属性と見なす。

(C) : 5以下, 6~15, 16~25, 26以上,

[分類 2] (B)は [分類 1] と同じ,

(C) : 10以下, 11~15, 16~20, 21以上,

[分類 3] (B)で (3, 4, 5) とする以外は [分類 1] と同じ。

これらのデータおよび分類から得られた時点96以降の故障システムに対する推定誤差を表 6 に示す。これから [方法 3] のほうが数量化理論第 I 類のみをもちいる [方法 1] よりも若干、精度がいいが、推定精度そのものはあまりよくなく、説明変数のより詳細な検討が必要である。

## 5. むすび

Kaplan-Meier の方法と数量化理論第 I 類を組

み合わせるにより単純に数量化理論第I類をもちいるよりも推定精度が向上することを示した。しかし分類のあいまいさを扱う方法は必ずしも精度の改善をもたらさず、今後さらに検討を進める予定である。

電子交換機の故障時間推定ではシステム停止を経験したシステムのみを扱ったが、現在までまったくシステム停止を経験していないシステムがほとんどであり、全システムを通じての良好な寿命推定法になっているとはいえない。本稿で提案した手法は、むしろどのシステムも遭遇しそうな事象にいたるまでの時間の長さを推定した場合により有効となるであろう。たとえば結婚年齢の推定などには有効である。

最後に、[方法4]として提案した方法ははなはだ不完全ではあるが、いろいろな改善法があると考えられるのであえて公表し、さらに有効な手法が提案されることを期待するものである。

#### 参 考 文 献

- [1] 林 知己夫：数量化理論とその応用例(IV)，統計研彙報 第5巻，第2号(1958)，163-169
- [2] Kaplan, E. and Meier, P. : Nonparametric Estimation from Incomplete Samples. J. Am. Statist. Assoc., 53 (1958), 457-481
- [3] 徳山 五郎：電子交換機の信頼度設計，信学会論文誌 51-A, 12 (1968), 471-478
- [4] 丹野 武宣 他2名：D10形自動交換機の商用実績，施設 25-11 (1973), 87-98
- [5] 林，駒澤：数量化理論とデータ処理，朝倉書店(1982)，278-279