川崎医療福祉学会誌　Vol. 11　No. 2　2001　307-312

原　著

最小二乗法と最尤推定法による集団食中毒の平均潜伏期間と曝露時点の比較検討

小池大介[*1]　格和勝利[*2]　近藤芳朗[*2]

要　約

1996年，腸管出血性大腸菌 O157：H7による食中毒患者の集団発生が岡山県邑久町をはじめとし全国的に発生した．特に1996年５月に発生した岡山県邑久町（患者数 428名）の集団食中毒，また同年６月に発生した新見市（患者数 355名）の集団食中毒に対しては，我々の提唱した新しい方法[2,3]によって食中毒の曝露時点，平均潜伏期間，分散因子などをいち早く推定した．

今回は1993年から1996年まで全国で発生したO157：H7によるもの８例，O118：H2によるもの１例とサルモネラによるもの１例の計10例の集団食中毒の発生分布に対して，潜伏期間の分布に対数正規分布を仮定した最小二乗法による推定法，潜伏期間の分布にワイブル分布を仮定した最小二乗法による推定法と潜伏期間の分布に対数正規分布を仮定した最尤推定法による推定法の３つの方法により曝露時点・平均潜伏期間の推定をおこなった．

その結果，最小二乗法と最尤推定法の結果はどちらの方法が良いと優劣をつけることはできなかったが，潜伏期間の分布として対数正規分布よりワイブル分布を仮定したほうが良い結果が得られることがわかった．

はじめに

1996年，腸管出血性大腸菌 O157：H7による食中毒患者の集団発生が岡山県邑久町，新見市をはじめとし岐阜市，石川県，堺市など全国的に発生した．食中毒の集団発生時において曝露時点・平均潜伏期間を推定することは感染源の特定，伝播様式，感染経路の追求に有効である．我々はいち早く邑久町，新見市の事例に対して患者発生分布から曝露時点・平均潜伏期間を推定する新しい方法を提案し推定を試みた[2,3]．また，1997年７月に大阪府堺市にて発生した患者数 4,026名という大発生に対しては，最小二乗法と最尤推定法を用いた２つの手法により推定を行い両推定法の比較をおこなった[4]．そして，医学における数理科学的諸問題という見地から提案した推定法の一般式化と細菌性食中毒の潜伏期間が対数正規分布を示す利用について理論的に考察をおこなった[6]．

今回は1993年から1996年まで全国で発生したO157：H7によるもの８例，O118：H2によるもの１例とサルモネラによるもの１例の計10例の集団食中毒の発生分布に対して，潜伏期間の分布に対数正規分布を仮定した最小二乗法による推定法，潜伏期間の分布にワイブル分布を仮定した最小二乗法による推定法と，潜伏期間の分布に対数正規分布を仮定した最尤推定法による推定法の３つの方法による推定結果の比較検討をおこなったので報告する．

推定方法

今回は，潜伏期間の分布として対数正規分布，ワイブル分布を仮定した最小二乗法による推定法と，潜伏期間の分布として対数正規分布を仮定した最尤推定法による推定法の３つの方法により曝露時点・平均潜伏期間の推定をおこなった．Sartwell により“各種感染症についての各宿主の潜伏期間に対数をとるとその頻度分布は正規分布をなす”．つまり，各宿主の潜伏期間の分布は対数正規分布であるという報告[1]があることから潜伏期間の分布として対数正規分布を仮定した．また，対数正規分布と同様な形状をなす分布としてワイブル分布がある．入院患者の在院日数の分布がワイブル分布をなすとされており，人の病気の回復等に関係が深い．そこで今回ワ

*1 川崎医療福祉大学　医療技術学研究科　医療情報学専攻　*2 川崎医療福祉大学　医療技術学部　医療情報学科
（連絡先）小池大介　〒701-0193　倉敷市松島288　川崎医療福祉大学

イブル分布も潜伏期間の分布として試行錯誤的に仮定した．

1．最小二乗法による推定法

我々の方法[2-4,6]は，従来の確率紙を用いる図式法を改良し確率紙上での最小二乗法を用いる解析的に発展させたものである．この方法は種々の非心の確率分布たとえば正規分布，対数正規分布，ワイブル分布，指数分布などに適用できる応用範囲の広い一般的な方法である．我々が初めてこの方法を適用した1996年の岡山県邑久町の事例が，この種の研究としてコンピュータを用いる近代化されたものとしては最初である[2)]．

この方法は累積相対度数を確率紙（確率紙には，正規分布用，対数正規分布用，ワイブル分布用，指数分布用などがある）に順次プロットし，プロットしたデータから最小二乗法を用いて１つの直線を得ることによりパラメータを推定するものである．

確率変数を x として，パラメータを a,b,c，分布関数を $F(x)$ とすると，ある写像 Ψ に対して，

$$\Psi(F) = aX(x,c) + b$$

となり得るとき，このような分布関数に対して我々の提案する解析的確率紙の方法[6)]が適用できる．具体的には，与えられた点 $x_1, x_2, \cdots, x_n$ における累積相対度数を $F_1, F_2, \cdots, F_n$ とし，

$$X_i = X(x_i, c) \quad (i = 1, 2, \cdots, n)$$

とするとき，残差平方和 Ω^2 を，

$$\Omega^2 = \sum_{i=1}^{n} \{\Psi(F_i) - (aX_i + b)\}^2$$

とし，この残差平方和 Ω^2 が最小となるようにパラメータ a,b,c を決定する．つまり，次の連立方程式

$$\frac{\partial\Omega^2}{\partial a} = 0, \quad \frac{\partial\Omega^2}{\partial b} = 0, \quad \frac{\partial\Omega^2}{\partial c} = 0$$

を数値的に解きパラメータ a,b,c を得るのが提案する方法である．この方法は従来から用いられていた確率紙の方法を精密化し，解析的方法へ発展させたもので最尤推定法と肩を並べる方法である．

今回はこの方法に対数正規分布とワイブル分布を適用した．それぞれの分布の確率密度関数 $f(x)$，写像関数 $\Psi(F)$，および関数 $X(x,c)$ は次の通りである．またいずれも非心の分布である．

・対数正規分布（パラメータ： $c,\ \mu,\ \sigma$）

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}\sigma(x-c)} \exp\left[-\frac{\{\ln(x-c)-\mu\}^2}{2\sigma^2}\right] \\ F(x) &= \Phi(X) \\ X &= \frac{\ln(x-c)-\mu}{\sigma} \\ \Psi(F) &= \Phi^{-1}(F) \end{aligned}$$

・ワイブル分布（パラメータ： $c,\ \lambda,\ m$）

$$\begin{aligned} f(x) &= \lambda m(x-c)^{m-1} \exp\{-\lambda(x-c)^m\} \\ F(x) &= 1 - \exp\{-\lambda(x-c)^m\} \\ X &= m\ln(x-c) + \ln\lambda \\ \Psi(F) &= \ln\ln\frac{1}{1-F} \end{aligned}$$

上式を利用し，対数正規分布，ワイブル分布を仮定して残差平方和 Ω^2 を最小にするパラメータを算出し曝露時点を推定する．

2．最尤推定法による推定法

最尤推定法による曝露時点・平均潜伏期間については丹後氏[5)]よって提案されている．我々は丹後氏の提唱する最尤推定法に基づき，対数正規分布を仮定した最尤推定法による曝露時点・平均潜伏期間の推定をおこなった．非心対数正規分布の確率密度関数は，

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma(x-c)} \exp\left[-\frac{\{\ln(x-c)-\mu\}^2}{2\sigma^2}\right]$$

であり，患者発生日時を $x_1, x_2, \cdots, x_N$ とすると，尤度関数 L は，

$$L = \prod_{i=1}^{N} f(x_i)$$

となる．最尤推定法ではこの尤度関数 L が最大となるようにパラメータを算出し曝露時点を推定する．

結　　果

集団食中毒の曝露時点・平均潜伏期間の推定へ応用する場合，両推定によって得られた理論分布の原点が曝露時点であり，曝露時点から累積相対度数が50％となる点までの距離が平均潜伏期間となる．つまり，曝露時点は初発日$+c$により得られ，平均潜伏期間 L_{50} は，対数正規分布では $L_{50} = e^{\mu}$，ワイブル分布では $L_{50} = \left(\frac{\ln 2}{\lambda}\right)^{\frac{1}{m}}$ により得られる．

今回，1993年から1996年までに全国で発生したO157：H7によるもの８例，O118：H2によるもの1例とサルモネラによるもの１例の計10例の集団食中毒の発生事例に対して，上に概説した潜伏期間の分布について対数正規分布，ワイブル分布をそれぞれ

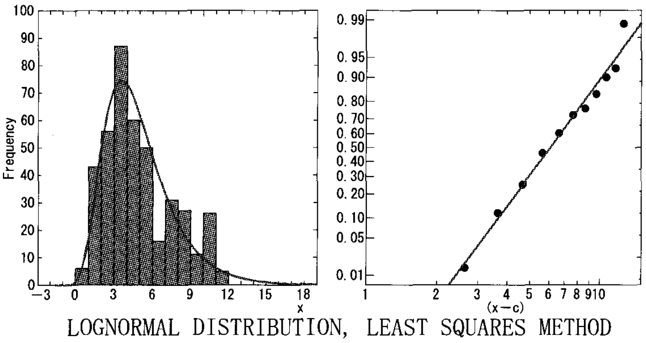

図1 邑久 [O157：H7]

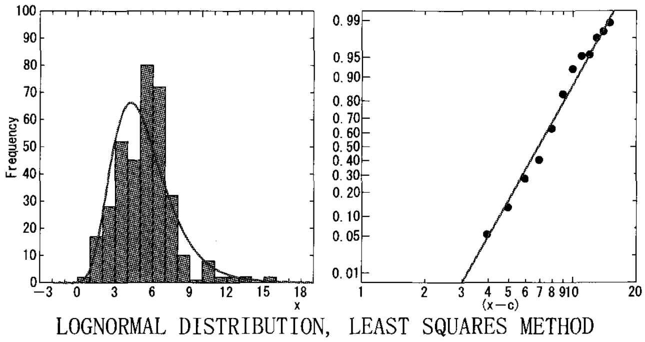

図2 新見 [O157：H7]

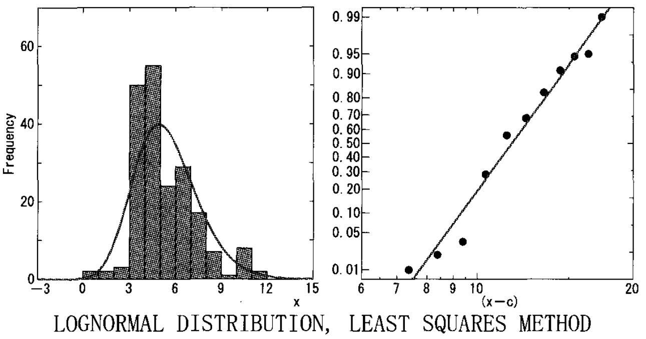

図3 岐阜 [O157：H7]

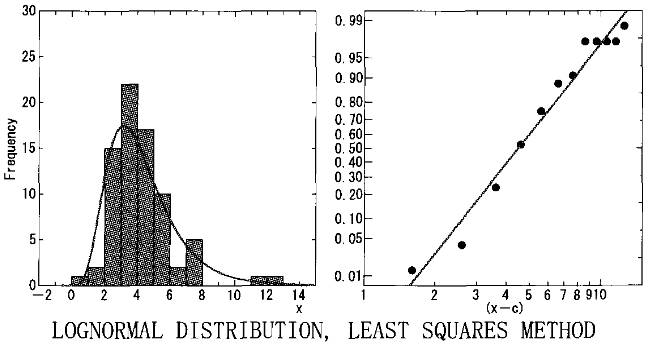

図4 石川 [O118：H2]

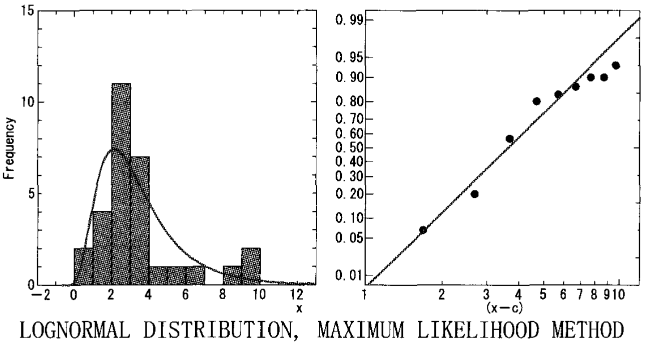

図5 東京 [O157：H7]

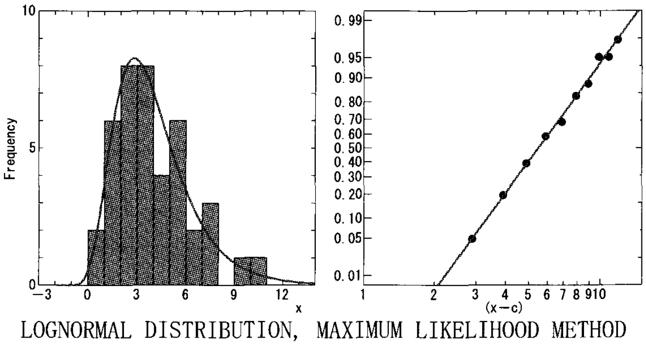

図6 盛岡 [O157：H7]

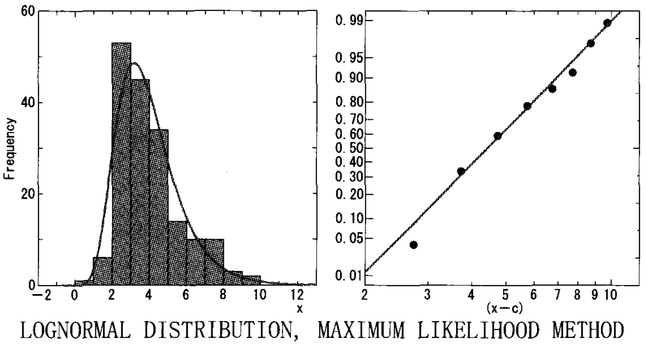

図7 岐阜 [salmonellae]

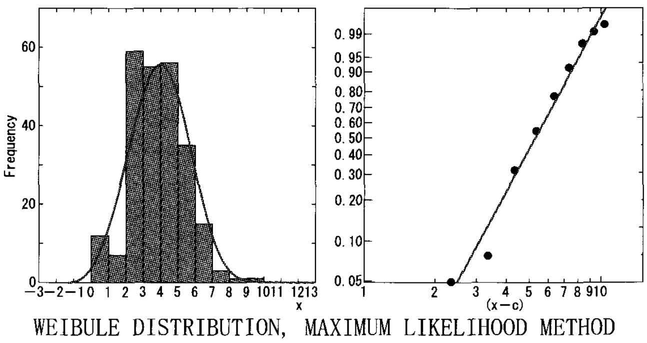

図8 奈良 [O157：H7]

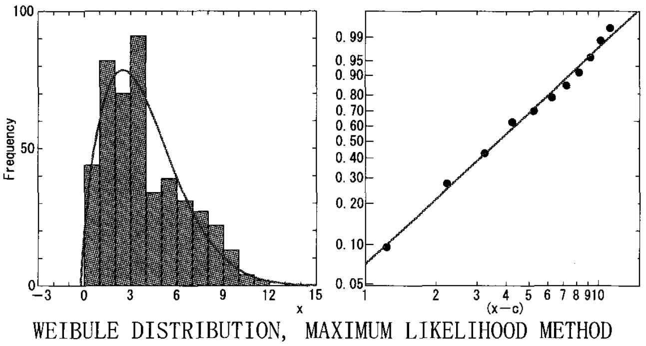

図9 岐阜 [O157：H7]

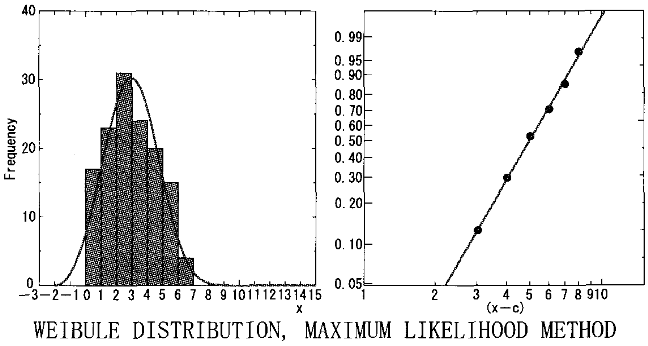

図10 北海道 [O157：H7]

表1　各推定法によるパラメータの推定結果

発生場所	原因菌	患者数	初発日	対数正規分布 最小二乗法			対数正規分布 最尤推定法			ワイブル分布 最小二乗法		
				C	μ	σ	C	μ	σ	C	m	λ
1 邑久	O157：H7	418	1996/05/24	22.3598	1.8085	0.3977	22.4080	1.7916	0.4132	24.6368	1.6396	0.0759
2 新見	O157：H7	355	1996/06/11	9.0738	1.9256	0.3295	5.2549	2.3860	0.2000	11.5366	2.0840	0.0269
3 岐阜	O157：H7	200	1996/06/07	0.6193	2.4544	0.1748	3.6708	2.1248	0.2263	5.9891	3.7883	0.0005
4 石川	O118：H2	76	1996/07/10	9.3998	1.5026	0.4219	8.5127	1.7021	0.3098	10.4557	2.0878	0.0392
5 東京	O157：H7	30	1993/08/28	28.2910	0.9011	0.7883	27.3215	1.2890	0.5054	28.6807	1.1978	0.2572
6 盛岡	O157：H7	41	1996/09/20	18.1964	1.6862	0.3985	18.1088	1.7025	0.3888	20.3904	1.5829	0.1082
7 岐阜	salmonellae	178	1996/09/13	11.8158	1.6135	0.3165	12.2683	1.4878	0.3514	13.5266	2.4156	0.0312
8 奈良	O157：H7	244	1994/09/30	17.2740	2.8010	0.0981	7.7459	3.2620	0.0602	28.6545	3.4797	0.0021
9 岐阜	O157：H7	459	1996/06/07	−14.4714	3.2269	0.1028	5.6692	1.5240	0.5052	6.7702	1.7106	0.0740
10 北海道	O157：H7	134	1996/10/24	−28699.96	10.2656	0.0001	16.4901	2.3399	0.1565	21.9730	3.2493	0.0038

表2　各推定法による平均潜伏期間 L_{50} の推定結果

発生場所	原因菌	患者数	対数正規分布 最小二乗法		対数正規分布 最尤推定法		ワイブル分布 最小二乗法
			L_{50}	γ	L_{50}	γ	L_{50}
1 邑久	O157：H7	418	6.1012	1.4885	5.9991	1.5116	3.8534
2 新見	O157：H7	355	6.8591	1.3902	10.8696	1.2214	4.7562
3 岐阜	O157：H7	200	11.6391	1.1910	8.3709	1.2540	6.7617
4 石川	O118：H2	76	4.4935	1.5248	5.4855	1.3632	3.9590
5 東京	O157：H7	30	2.4624	2.1996	3.6292	1.6576	2.2882
6 盛岡	O157：H7	41	5.3990	1.4895	5.4875	1.4752	3.2330
7 岐阜	salmonellae	178	5.0205	1.3723	4.4272	1.4210	3.6104
8 奈良	O157：H7	244	16.4606	1.1030	26.1011	1.0620	5.2942
9 岐阜	O157：H7	459	25.2008	1.1083	4.5907	1.6573	3.6981
10 北海道	O157：H7	134	28726.9180	1.0001	10.3801	1.1695	4.9643
平均潜伏期間 $\bar{L}_{50} \pm$ SD			5.9964 ± 2.8476 (1–7)		6.3241 ± 2.4907 (1–7)		4.2419 ± 1.2473 (1–10)

表3　各推定法による曝露時点 C の推定結果と疫学調査による曝露時点

発生場所	原因菌	患者数	初発日	対数正規分布 最小二乗法	対数正規分布 最尤推定法	ワイブル分布 最小二乗法	疫学調査による曝露時点
				$C \pm \sigma_C$	C	C	
1 邑久	O157：H7	418	1996/05/24	22.3598 ± 1.2116	22.4080	24.6368	22, 23 [学校給食]
2 新見	O157：H7	355	1996/06/11	9.0738 ± 0.7734	5.2549	11.5366	10, 11 [学校給食]
3 岐阜	O157：H7	200	1996/06/07	0.6193 ± 5.2725	3.6708	5.9891	5 [学校給食]*
4 石川	O118：H2	76	1996/07/10	9.3998 ± 0.7745	27.3215	10.4557	10頃 [学校給食]
5 東京	O157：H7	30	1993/08/28	28.2910 ± 0.4545	27.3215	28.6807	23〜25 [プールの水]
6 盛岡	O157：H7	41	1996/09/20	18.1964 ± 0.6302	18.1088	20.3904	19 [学校給食]
7 岐阜	salmonellae	178	1996/09/13	11.8158 ± 0.7860	12.2683	13.5266	13 [学校給食]
8 奈良	O157：H7	244	1994/09/30	17.2740 ± 10.7839	7.7459	28.6545	28 [学校給食]
9 岐阜	O157：H7	459	1996/06/07	−14.4714 ± 22.0477	5.6692	6.7702	5 [学校給食]
10 北海道	O157：H7	134	1996/10/24	−28699.96 ± 39.0055	16.4901	21.9730	22 [学校給食]*

* 原因菌の検出等により曝露時点が確定された事例

仮定した最小二乗法による推定法，潜伏期間について対数正規分布を仮定した最尤推定法による推定法の3つの方法により推定をおこないその結果の比較検討をおこなった．その結果は表1～3，図1～10に示す．

表中の C は曝露時点（日），L_{50} は平均潜伏期間であり，μ, σ, γ は対数正規分布における平均，標準偏差，分散因子であり，m, λ はワイブル分布におけるパラメータである．

表1は各事例の原因菌，患者数，初発日と3つの推定法によりそれぞれ推定されたパラメータである．各事例について1つの推定法を取り上げ表1に示したパラメータを利用して描いたグラフが図1～10である．図中の左図は集団食中毒の日別患者発生分布に推定したパラメータを用いた理論曲線を重ねたものであり，右図はそれを確率紙に描いたものである．図1～4は対数正規分布を仮定した最小二乗法による推定結果によるものであり，図5～7は対数正規分布を仮定した最尤推定法による推定結果によるものである．また図8～10はワイブル分布を仮定した最小二乗法による推定結果によるものである．これらの図で各事例の推定結果の適合度を視覚的に確認

できる．確率紙上で実測値がほぼ直線となり，理論曲線に重なっているものほど推定したパラメータが実測値に適合していることになる．

表2は各事例の3つの推定法により得られた平均潜伏期間である．対数正規分布を仮定した2つの手法については分散因子γを求めることができるのでそれも示す．O157：H7の平均潜伏期間は3～7日と言われており，3つの推定法はともに平均潜伏期間から考えればまずまずの推定結果が得られたことが判断できる．ただし，対数正規分布を仮定した最小二乗法による推定法では，3岐阜，8奈良，9岐阜，10北海道の事例について推定した平均潜伏期間がO157：H7の平均潜伏期間からはずれた．また対数正規分布を仮定した最尤推定法についても，2新見，8奈良，10北海道の事例について推定した平均潜伏期間がO157：H7の平均潜伏期間からはずれた．これに対してワイブル分布を仮定した最小二乗法による推定法ではO157：H7の平均潜伏期間から考えれば，他の2つの対数正規分布を仮定した推定法よりも安定した結果を得られていることが確認できる．

表3は各推定法による曝露時点Cの推定結果と疫学調査による曝露時点である．対数正規分布を仮定した最小二乗法による推定にのみパラメータの標準偏差σ_Cを算出したので曝露時点とともに示す．疫学調査による曝露時点が真の曝露時点にほぼ近いとすれば，表2同様，対数正規分布を仮定した2つの推定結果については推定結果にむらがあり，ワイブル分布を仮定した推定法のほうが安定した結果を得られていることが確認できる．

この結果，最小二乗法と最尤推定法についてはデータによって一長一短があり両推定法のどちらが良いとはいえないことがわかった．最尤推定法はデータ（度数）を密度関数に合わせる方法であるのに対して，最小二乗法はデータ（累積度数）を分布関数に合わせる方法である．したがって，与えられたデータが度数あるいは累積度数のどちらに誤差が入り易いかで互いに長所にも短所にもなり得るため，今回はどちらが良い推定法であるという判断はできなかった．

また，曝露時点の推定結果については対数正規分布よりもワイブル分布の場合が概して良い結果が得られた．これは両分布の密度関数について，原点からの立ち上がりの緩急の違いのためと推察される．対数正規分布は分布の両端の影響を大きく受けて推定値が大きく変動するが，ワイブル分布は対数正規分布のように分布の両端の影響を受けて大きく推定値が変動しないためである．

おわりに

今回の結果により，最小二乗法と最尤推定法はどちらが良い方法であるとはいえないこと，対数正規分布よりワイブル分布のほうが安定した結果が得られることがわかった．しかしこれは今回適用した事例に対して主観的な判断により得た結果である．今後，最小二乗法と最尤推定法のシミュレーションによる検討，ワイブル分布を仮定するための感染症発症の現象の理論面からの仮説を立てるとともに，推定結果の客観的な判定のためのシミュレーションによる適合度検定へ研究を進める．

文　献

1）Sartwell PE（1950）The Distribution of Incubation Periods of Infectious Disease. *American Journal of Hygiene*, **51**, 310–318.

2）格和勝利，緒方正名，近藤芳朗，発坂耕治（1996）感染症の平均潜伏期の計算法について－腸管出血性大腸菌O–157：H7による食中毒患者発生を例にして－．川崎医療福祉学会誌，**6**(2)，381–387.

3）格和勝利，緒方正名，近藤芳朗，小池大介，万代素子（1997）感染症の平均潜伏期の計算法についてII－新見市における腸管出血性大腸菌O–157による食中毒患者発生を例にして－．川崎医療福祉学会誌，**7**(1)，199–203.

4）格和勝利，小池大介，緒方正名，近藤芳朗（1998）感染症の平均潜伏期の計算法についてIII－堺市における腸管出血性大腸菌O–157による食中毒患者発生を例にして：最小二乗法と最尤推定法の比較－．川崎医療福祉学会誌，**8**(1)，181–185.

5）丹後俊郎（1998）潜伏期間に対数正規分布を仮定した集団食中毒の曝露時点の最尤推定法，日本公衆衛生雑誌，**45**(2)，129–141.

6）近藤芳朗，格和勝利，小池大介（2000）医学における数理科学的諸問題1　O–157食中毒・感染症の曝露日・平均潜伏期の推定，川崎医学会誌 一般教養篇，**26**，59–66.

（平成13年11月16日受理）

Estimation of the Mean Incubation Period and Exposed Point of an Outbreak of Food Poisoning Using the Least Squares and Maximum Likelihood Methods

Daisuke KOIKE, Katsutoshi KAKUWA and Yoshiro KONDO

(Accepted Nov. 16, 2001)

Key words : OUTBREAK OF FOOD POISONING, LEAST SQUARES METHOD, MAXIMUM LIKELIHOOD METHOD, MEAN INCUBATION PERIOD, EXPOSED POINT

Abstract

The exposed point, mean incubation period and dispersion factor of an outbreak of food poisoning by Enterohemorrhagic *Escherichia coli* O157 : H7 in Oku (Okayama, Japan, May 1996, 418 patients) and Niimi (Okayama, Japan, June 1996, 355 patients) were investigated. A new method of estimating the exposed point and the mean incubation period is proposed. A comparative study of the results was done by three methods: the Least Squares Method assuming both a Lognormal Distribution and a Weibull Distribution, and the Maximum Likelihood Method assuming a Lognormal Distribution.

The result showed that the estimation by the Least Squares Method is similar to that by the Maximum Likelihood Method.

Also, it was found that estimation by Weibull Distribution is better than by Lognormal Distribution.

Correspondence to : Daisuke KOIKE Doctoral Program in Medical Informatics, Graduate School of Medical Professions, Kawasaki University of Medical Welfare
Kurashiki, 701-0193, Japan
(Kawasaki Medical Welfare Journal Vol.11, No.2, 2001 307–312)