

## アザラシ状奇形の原因 I

——サリドマイド仮説の成立に関する統計学上の争点について——

吉 村 功\*

### サリドマイド仮説

いくつかの国の医者が、それ以前に、ほとんど経験したことのないある種の奇形の頻発に異常を感じたのは、1960年頃であった。現在、手があざらしのように短かいという特徴で紹介されている奇形である。この奇形には、他にもいくつかの特色があるが、それは医学書に譲ることにして、以下ではこれを単に奇形ということにする<sup>(1)</sup>。

医学の常道にしたがって、経験された奇形の症例は、散発的に医学会に報告された<sup>(2)(3)</sup>。症例報告の増加とともに、原因を究明するための系統的な調査も進められたようである。1961年11月18日、西ドイツのLENZ博士は、自らの研究経験と調査をもとにして、母親が妊娠初期に或る睡眠薬(サリドマイド)を服用した結果として、このような奇形が生じたという説を公けにした<sup>(4)</sup>。そこで提示されたいくつかの論拠のうち、もっとも重要なものは、

- (1) 奇形児と非奇形児の間で統計的にもっとも差のある因子は、妊娠初期におけるサリドマイドの服用である。すなわち、前者の母親には、妊娠初期にサリドマイドをのんだ確証のあるものが多いのに対し、後者の母親にはそれが少ない。

というものである。

LENZの発言のあったこの日の後、サリドマイドと奇形の関係についての調査報告や実験報告が次々と発表されるようになった<sup>(5)</sup>。その中には、LENZの説(以下ではサリドマイド仮説とよぶことにする)に焦点を合わせて調査や実験をしたものもあれば、WEICKERらのように、サリドマイド仮説の発表以前に、LENZと独立に系統的調査を進めてサリドマイド仮説に達していたものもあった。

報告の中には、サリドマイド仮説を否定するものもあったが、その論拠はほとんど、サリドマイド仮説で説明できない事実があるというものであって、対置しうる仮説を提示したものは稀であった。せいぜい流産すべき奇形胎児がサリドマイドの流産防止作用によって出生したという抗流産説や、昔から奇形児が存在したという天然

\* Isao YOSHIMURA 名古屋大学工学部工業数学教室

因子説ぐらいであったといわれている。

報告の増加に対応して、サリドマイドを含む薬の製造販売は、各国で逐次停止されていった。そして販売停止に対応して奇形児の出生は報告されなくなった。あるいは、奇形児の出生がほとんどなくなったことが報告されるようになった<sup>(7)</sup>。そればかりでなく、奇形児出生数とサリドマイド販売量に関してデータが得られたあらゆるところで次の事実が認められた<sup>(8)(9)</sup>。

- (2) サリドマイド販売量の時間推移に対して、約9カ月のおくれをもって、ほぼ同じ形の奇形児出生数の増減があった。

奇形児の出生が目につかなくなることによって、事件は一区切りした。しかし、生存する奇形児に対する責任の追及と、医学上および生物学上の関心とから、サリドマイド仮説の成否に関する研究や論争はつづき、

- (3) 妊娠動物にサリドマイドをのませるという実験を行なうと、ある種族の動物では、人間と同じ傾向の奇形症状が発生する<sup>(10)</sup>。
- (4) サリドマイド仮説におけるサリドマイド服用日と奇形形態の関係は、生物学上の知識・経験と矛盾しない。あるいは、サリドマイド仮説から予想されるような結果が実験によって得られる<sup>(11)</sup>。

などのことが確かめられていった。また(1)、(2)の事実を裏づける資料が増加していった。抗流産説などはこれらの事実の前に消えてゆき、サリドマイド仮説のみが事実を積極的に説明する唯一の仮説となった。こうして現在、たとえば日本では、製薬企業およびそれとつながりのある人たちがサリドマイドの催奇形作用に異論をとなえているのを除けば、生物学上、医学上の論争は、奇形を生じる細かいメカニズムの問題に集中するようになっていく。

さて、筆者は、以前から、サリドマイド仮説が統計学的にみても批判に耐えうるかどうか、仮説確立の過程において統計的方法が正しく用いられていたかどうかに関心を持っていた。そして、自ら吟味する機会をうかがっていた。そこへ、1969年2月11日、朝日新聞が、“西独のレンツ教授説は統計学的にみても間違い。大日本製薬が準備書面”という見出しの記事をのせた<sup>(12)</sup>。これは統計学にかかわっている者として見過せないものであった。筆

者はこれをきっかけに、実際に調べはじめ、上に概観したようなことを確かめた。ところがこの間に筆者と同じ関心のもとで、筆者の見解と反する主張が杉山によって公表された<sup>(13)</sup>。その主張は要約すると、“LENZの調査には方法論上の欠陥がある、調査内容にも信頼性がない、ゆえにLENZの主張は批判的に受け取られねばならない”というものであった。

この主張に対して、しばらくたってから反論が出された。建田<sup>(14)</sup>および高橋<sup>(15)</sup>からであって、それは“LENZなどの調査は、医学者にとって最善の努力で行なわれた、方法も結果も信頼に値する、したがってサリドマイド仮説は十分な根拠を持っている”という趣旨のものであった。

こうして日本でも、統計学的側面からのサリドマイド仮説の評価が必要になり、筆者も発言せざるをえないと感じるようになってきた。この小論は、以上のようなききつのもとで、サリドマイド仮説の成否を、主として統計学的側面から考察しようとしたものである。

#### 必要な論証の程度

サリドマイド仮説に疑問を提出する側の論理は、要約すると、先にのべた各論拠などが信頼できる資料にもとづいていないし、仮に資料が信頼できるとしても、先にあげたような論拠だけでは仮説を承認するのに不十分だということである。

そこで、まず、仮説が承認されるための絶対的基準があるかどうか考えてみよう。たとえば、先にのべた論拠に何かを追加すれば、サリドマイド仮説は絶対の真理であるということになるだろうか。

少し考えればわかるように、一つの仮説を絶対の真理と認めさせるための客観的な論証基準は存在しえない。たとえ人体実験がなされて奇形が生じたとしても、あげ足をとることはいくらでもできる。たとえば、サリドマイド発売以前にもこの奇形が存在していたから、サリドマイド以外にこの奇形を発生させる因子を見出して、その因子が現在問題になっている奇形児になかったことを証明しなければならぬというように。したがって、仮説承認の絶対的基準があるからそれに対して十分な論証を試みよと要求することは、科学の名のもとに、不可能なことを要求することになる。これは現在直面している問題においては、害を被っている者を苦しめるだけで、何の益もないであろう。

では、仮説を承認するのに目安は全くないのであるか。そうではない。我々が経験的に行なっているように、あるいは科学方法論の分野で論じられているように<sup>(16)</sup>、仮説承認の目安は人間がとろうとする行動、決定、ある

いは処置の現実的意味に応じて存在する。

サリドマイド仮説に即してのべよう。薬の販売を停止するという処置のためには、論拠の(1)で十分である。あるいは(1)が認められるかぎり、まず妊婦にサリドマイドを服用させることを止めなければならない。他の薬や他の手段を用いるよりもサリドマイド服用の方が眠りやすいという利益に比べて、奇形発生の可能性の方がずっと大きな損害だからである。

販売停止という処置の当否を判断するには、論拠の(2)があれば十分である。たとえサリドマイド以外の何かが奇形発生に介在していたとしても、処置をとったことだけによって奇形発生がほとんどなくなったのだから、処置は正しかったといえる。たとえてみれば、ニュートンの法則が実用性を持つ場合において、ニュートンの法則が正しいと判断するようなものである。

二度とこの世に奇形を発生させまいというためには、どうであろうか。おそらく、先にのべた論証以上のものをさらに追及すべきであろう。分子生物学的連鎖反応が明らかにされないかぎり、サリドマイドと分子構造の似た物質がほとんど同じ奇形を生じさせる可能性を否定できないからである。

では、生じてしまった奇形に関連した損害を誰が償うかという場合はどうであろうか。もし仮説成立の論拠が、作業仮説なしに集められた資料を整理しただけで、サリドマイド仮説がそれに矛盾しないというだけののであれば、この仮説にもとづいて責任を確定するのは危険である。論拠を支える資料において、何か特殊な因子が見落されている可能性があり、導かれた仮説もまた、“犯人に仕立てあげられた”可能性を持つからである。これに対して、サリドマイド仮説が提起されたのち入手されたあらゆる資料が、サリドマイド仮説に整合する事実を示し、かつ、それにとって代る仮説が提示されないならば、サリドマイド仮説は責任を指定する根拠となる。すなわち、サリドマイド仮説が真であるという前提で責任を追及してよい。ここで、とって代る仮説とは、存在する各事実との整合性がよりよい仮説のことであって、原則として、すでに提出されている仮説が奇形発生の詳細なメカニズムを記述していないとか、基礎資料の中に情報量の少ないものがあるという批判のことではない。また、薬の作用について十分な吟味をする能力と責任を持つものが、資料を準備しないで提出する異論のことでもない。これらのものは仮説の真実性を弱める何の力もないからである。

上の原則の例外は、仮説のもとで責任を追及されるものが、とって代る仮説や矛盾する事実を用意する能力を持たないときだけである。

要するに、サリドマイド仮説を責任追及の前提にしようかどうかは、仮説確立の過程において、各論拠がとって代りうるあらゆる仮説に対して一つの仮説を選択させる機能を果たしたかどうかによる。これを実際に確かめなければならない。

いま直面している事件では、実際にどうだったのだろうか。筆者は、その機能を果たすと断定する。前節で素描した歴史が、後者の状態になっているからである。

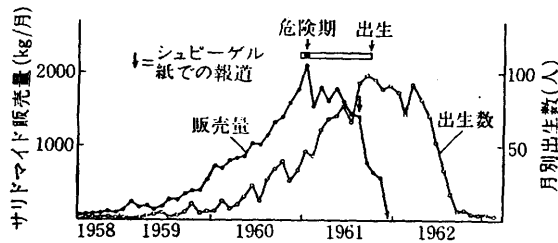
もちろん杉山などの主張は、この断定に矛盾する。だがそれは、筆者の断定が誤まっているからではなく、杉山などの主張が、錯覚にもとづく誤解または故意のごまかしにもとづいているからである。次節以下にそれを具体的に指摘しよう。

### カーブの同形性

統計学上の議論は主として、論拠の(1),(2)に集中する。そこで、ここでの議論もその二つにわけ、まず、論拠(2)について論じよう。日本において、サリドマイド仮説に異論をとこなしている例が、積極的な主張を行なっているのが(2)の方だからである。論拠(2)の典型的な例は、第1図に示される。これは THALHAMMER の著書<sup>(17)</sup>から転載したもので、販売量のカーブと奇形児出生数のカーブの同形性がよく見てとれる。

このデータがサリドマイド仮説を支える一つの科学的根拠であるというのは、次の論理による。すなわち、仮説Aのもとで生じる現象が現実におこっており、他にこの現象を生じさせる仮説が存在しないならば、仮説Aの真象性は大きいという論理である。もちろん、現在、サリドマイド仮説以外にこのカーブの同形性を説明する仮説は存在していない。

これに対して杉山は<sup>(13)</sup>、“(第1図のような図が)かりに販売量、奇形発生数ともに地域は一定地域に指定して行なわれた観察結果と仮定しても、二つの時間的年次的



第1図 西ドイツにおけるサリドマイド販売量と奇形児出生数。(THALHAMMER<sup>(17)</sup>) 約9カ月ずらすと、二つのカーブがほとんど重なる。西ドイツにおいて、サリドマイドは、まず多発神経炎を生ぜしめることが問題にされた。1961年4月以後は処方が必要とする薬品となった。図中の矢印のシブビーゲル紙での報道内容はまだよくわからない。

経過の曲線

$$x=f_1(t), \quad y=f_2(t) \quad (1)$$

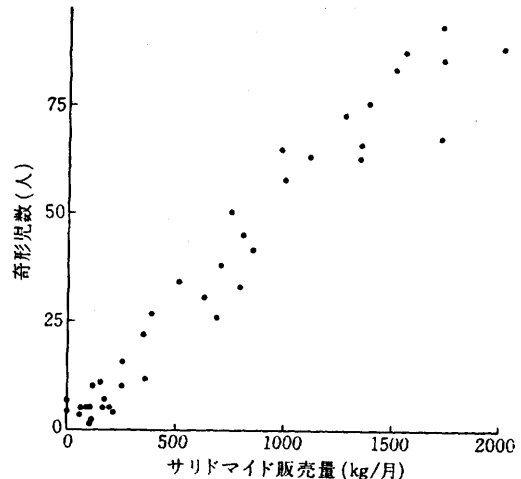
から  $t$  を消去して、

$$y=f_2[f_1^{-1}(x)]=f(x) \quad (2)$$

のような関係を求めることが必要であって、ただ二本の曲線を眺めているだけでは、このような  $x$  と  $y$  との関係は明らかにならないであろう。但し、ここに述べた函数関係は、統計的変動を捨象した上でのお話であり、現実には二つの曲線といっても、統計的変動を伴うのであるから、 $x$  と  $y$  の対応関係といえは当然のことながら、相関図により観察されなければならないのである”とのべている。

また大日本製薬(以下では大日本と書く)は<sup>(14)</sup>“カーブの同形が得られたからといって、そのことから、サリドマイドと奇形との間に相関関係があると速断することはできない。二つの変量の相関の有無を調べるためには、相関図による方法が直接的であり、かつ根本的である。相関関係の存在が相関図によって肯定されている場合には(ただし、その相関関係が意味のあるものであるかどうかは別として)カーブの同形が得られるであろう。そこで始めてカーブの同形性で相関関係を説明することが許される。しかし、これはあくまで相関関係の存在がすでに肯定されていることが前提であって、相関関係が肯定されていない以上カーブの同形性は、何事も語ることはできないのである”とのべている。

後者では、相関関係という概念と因果関係という概念とが混同されていて論旨がわかりにくいのが、要するに両者とも、相関図を書かなければいけないと主張している



第2図 西ドイツにおけるサリドマイド販売量と9カ月後の奇形児出生数とを第1図から読みとり、相関図にプロットしたもの。点が直線上に並んでいる。すなわち比例関係が明らかである。

のである。もしこの主張が正しければ、それが行なわれていない文献はデータ解析不十分ということになる。はたしてこの主張は正しいだろうか。

第1図から、各月の販売量とその9カ月後の奇形児出生数とを読みとり、対応させてグラフを画くと、第2図が得られる。これが杉山などの相関図である。第1図で見られたカーブの同形性は、そのまま第2図で点が直線状にならぶこととして表現されている。逆に第2図から第1図を復元させたらどうだろう。それには時間順序が必要である。仮にそれが定まったとすると、第1図のような2本のカーブが得られるが、点が直線状にならんでいるかぎり、このカーブの同形性は常に保存される。以上の考察から、時間推移カーブの同形性と相関図における点の直線状分布とは、ともに相関関係(ここでは比例関係)という一つの現象の異なった表現であることがわかる。

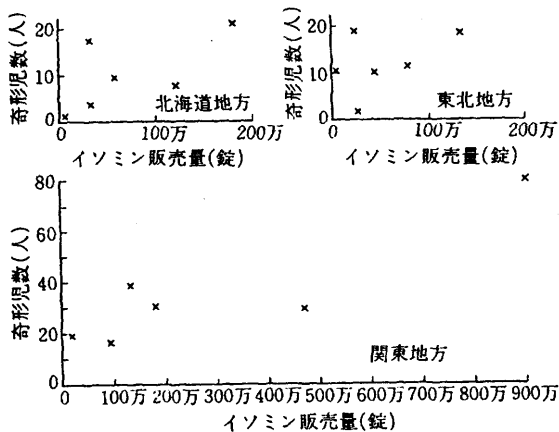
では時間推移カーブを画くことと、相関図を画くこととは全く同じことなのだろうか。そうではない。そもそも図による表現は、人間の視覚に直観的に訴えるものである。そして、同じ内容のものであっても、図によって違った印象を与えることは考えられる。それゆえ、直観的把握を目的とする場合に、第2図の方が第1図より相関関係を正しく認識させやすいという相手には、相関図を画くべきである。

しかしながら、ここで注意すべきことが二つある。一つは、時間推移カーブの方が相関図より多くの情報を含むことである。第1図から第2図を画くことはできるが、逆はできないということである。相関図においては、時間順序が無視されるからである。これは杉山などの主張が誤りであることを意味する。相関図に現われていることは必ず時間推移カーブの方にも現われているはずだし、相関図だけから判断すると、時間経過という重要な情報を失うことがあるからである。相関図を画いたときには、原則として、第1図のようなカーブあるいはそれと同等の内容のデータ表、またはグラフをつけておかなければならない。

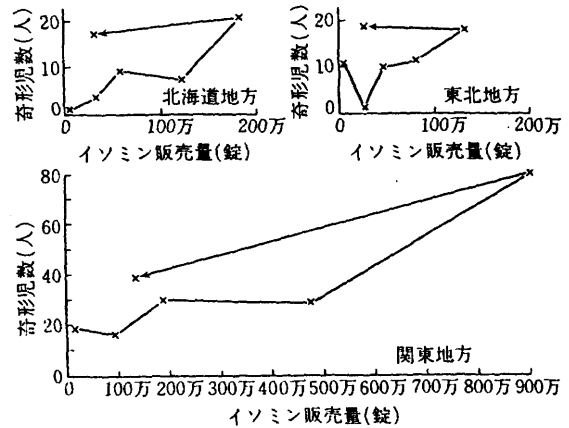
参考のために、時間が無視されるかどうかで、同じデータから異なった結論が得られる例を示そう。第1表は、大日本が最近公表したデータである。この表の最初の3地域(北海道、東北、関東)について、奇形児数Bとその出生に影響をおよぼしたとみられるイソミン(サリドマイド系睡眠薬の商品名)の販売量Dの相関図を画くと、第3図が得られる。これらの図の全体から、何か明瞭な傾向、法則性が見出せるだろうか。おそらく何も見出せないだろう。ましてこれがサリドマイド仮説に有利なデータであるとは考えられない。関東地方では、たしかにイソミン販売量の大小と奇形児数の大小とが正の相関を

第1表 大日本製薬が公表したイソミン販売量と奇形児出生数。数字はすべて大日本製薬が示したもので、数字の素性で説明がついていたのは次のものだけである。A=イソミン販売量(錠)、B=奇形児数、C=出産児数、D=年度別奇形児発生に關与したと思われる販売量(錠)、X=妊婦のサリドマイド服用率の推定値=D/C、Y=奇形発生率=B/C×100,000。

年	33	34	35	36	37	38	
北海道	A	253,250	469,424	983,238	1,823,956	761,410	
	B	1	3	9	8	21	17
	C	23,607	95,324	93,852	92,401	91,949	15,554
	D	63,312	307,293	597,882	1,193,421	1,824,810	304,564
	X	2.7	3.2	6.4	12.9	19.8	19.6
	Y	4.2	3.1	9.6	8.7	22.8	109.3
東 北	A	235,652	281,580	594,956	1,275,720	601,374	
	B	11	3	10	13	19	19
	C	48,099	183,019	174,151	166,333	160,704	26,104
	D	58,914	247,137	359,928	765,159	1,317,615	240,550
	X	1.2	1.4	2.1	4.6	8.2	9.2
	Y	22.9	1.6	5.7	7.8	11.8	72.8
関 東	A	786,268	1,343,764	3,280,660	9,213,978	3,476,994	
	B	19	17	30	29	80	39
	C	94,923	384,326	395,754	404,363	424,349	75,190
	D	196,566	925,638	1,827,984	4,763,988	8,996,685	1,390,798
	X	2.1	2.4	4.6	11.8	21.2	18.5
	Y	20.0	4.4	7.6	7.2	18.9	51.9
中 部	A	456,100	507,702	1,210,152	2,798,666	1,716,754	
	B	11	11	13	16	46	20
	C	49,374	193,959	194,823	195,757	201,152	34,706
	D	114,024	468,999	683,319	1,607,280	3,129,051	686,702
	X	2.3	2.4	3.5	8.2	15.6	19.8
	Y	22.3	5.7	6.7	8.2	22.9	57.6
北 陸	A	195,826	274,644	611,042	1,076,442	465,176	
	B	3	4	4	9	17	16
	C	23,274	88,251	86,135	81,964	81,219	13,706
	D	48,957	215,532	358,743	727,392	1,086,441	186,070
	X	2.1	2.4	4.2	8.9	13.4	13.6
	Y	12.9	4.5	4.6	11.0	20.9	116.7
近 畿	A	1,121,444	1,342,252	3,652,892	8,722,654	5,321,336	
	B	7	6	8	22	57	31
	C	62,043	249,495	254,337	260,852	282,519	48,962
	D	280,362	1,176,648	1,919,910	4,920,336	9,734,793	2,128,534
	X	4.5	4.7	7.5	18.9	34.5	43.5
	Y	11.3	2.4	3.1	8.4	20.2	63.3
中 国	A	362,996	477,202	926,020	1,663,482	918,662	
	B	7	8	3	8	31	18
	C	29,478	114,208	108,437	104,582	104,599	17,202
	D	90,750	391,551	589,407	1,110,384	1,798,812	367,464
	X	3.1	3.4	5.4	10.6	17.2	21.4
	Y	23.7	7.0	2.8	7.6	29.6	104.6
四 国	A	149,014	246,992	431,376	973,526	402,138	
	B	5	4	6	8	23	7
	C	18,330	70,227	64,559	61,651	60,472	10,052
	D	37,254	173,511	293,091	566,913	971,436	160,856
	X	2.0	2.5	4.5	9.2	16.1	16.0
	Y	27.3	5.7	9.3	13.0	38.0	69.6
九 州	A	509,274	645,572	1,155,606	2,455,184	1,216,788	
	B	9	7	15	36	40	52
	C	64,245	247,279	233,993	221,419	211,653	35,110
	D	127,320	543,354	773,085	1,480,506	2,571,465	486,716
	X	2.0	2.2	3.3	6.7	12.1	13.9
	Y	14.0	2.8	6.4	16.3	18.9	148.1



第3図 大日本製薬<sup>(18)</sup>が公表したデータのうち、各年度の奇形児数Bと、それに寄与したと思われるイソミン販売量Dとの対応。

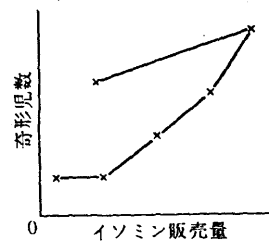


第4図 イソミン販売量Dと奇形児数Bとの対応関係の時間推移。第3図に時間順序を入れたもの。矢印の終点が昭和38年。

持つが、北海道や東北ではそうとは見えないからである。

そこで、第3図に時間順序をいれてみよう。このとき第4図が得られる。線のつながりが時間の推移を表わし、矢印の先が昭和38年である。この図からは、第3図と全く異なり、どの地域でも同じ傾向の存在することが見てとれる。概念的には、第5図で表わされるような傾向である。もちろん、この傾向が、物理的にどんな実体や本質を反映しているかは、次節でのべるような検討をまたねばならないが、少なくとも第3図で発見できなかった法則性が、第4図では発見できたということは確かである。簡単のために、3地域のみをとりあげたが、他の6地域でも全く同じ傾向が存在することを注意しておこう(読者が実際にグラフを書いて確かめることを期待する)。

相関図についてのもう一つの注意は、合理的な層別を行なうことで、たとえば性質の異なる地域の資料をまぜないことである。たしかに、サリドマイド仮説はどの地域でも同じである。しかし、それから生じる現象としての、サリドマイド販売量と奇形児出生数との関係には、地域差が生じる。たとえば都会であれば、1人が飲む薬の量が多いから、販売量の割に薬を飲んだ人の数が少く、したがって奇形児出生数も少ないであろう。このとき、相関図において、点の分布の勾配はゆるくなる。農山村の多い所では逆に、相関図において点の分布の勾配は急になる。また日本では、ある先生に教育を受けた医者がある地域に偏在する。したがって、奇形の判断基準に地域差が生じる可能性が大きい。判断基準がルーズで、いま問題にしている奇形の特徴をよく認識しようとしないう人が多いと、他の異常児が統計に混在する。このような地域では、点が全体として少し上に分布することになる。このようなわけで、自然的因果関係に地域差がなく



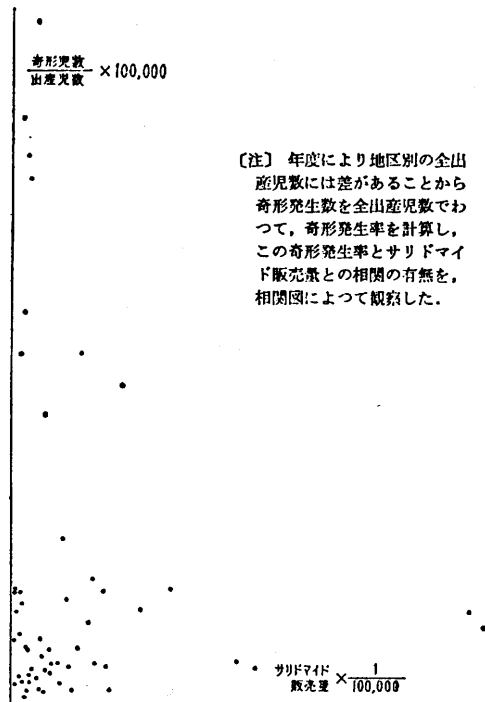
第5図 各地域におけるイソミン販売量Dと奇形児数Bとの間に見られる傾向の概念図。

ても社会的要因によって、地域差が生じるのである。地域差があるとき、異なった地域のもを無条件にまぜると、データに現われている法則性はわからなくなることがしばしばある。たとえば、第3図の三つの図をまぜるとそうなる(読者自ら確かめていただきたい)。層別を行なうのは、このような場合に、データに現われている法則性を見失わないためである。これはきわめて初等的な常識であるから、もし統計家がゆえなく層別を怠っているならば、そのような統計家の仕事は、あまり信用できないと考えてよい。

新たに公表されたデータが意味するもの

前節でのべたように、奇形発生とサリドマイド販売量との関係を示すには、時間推移カーブで十分である。ところが、杉山は前節で引用した文章につづけて、第6図を示し次のような論理展開を行なった。

“(第6図は)厚生省の委嘱により我国において行なわれた全国調査、いわゆる森山データとサリドマイド販売量にもとづいて、地区別かつ年次別に出生した奇形児数と9カ月さかのぼってそれに対応するサリドマイド販売量を対応させて作った相関図である(図には9(地区)×6(年次)=54点が全く区別なく記入されている)。地区、年次を考慮しなかったのは、同じ日本人であり、妊娠期



第6図 杉山が示した相関図。注なども原典のまま。

間中のある時期におけるサリドマイド服用と奇形児出生との間に相関関係があるとしても、それはこの間の年次や地区には無関係な医学的現象に違いないからである。えられた相関図をみると、プロットした点は右上りの傾向を示すどころか、むしろ右下りの傾向さえ示している。これからしてもサリドマイド販売量と奇形発生率との間に、問題になる相関関係は認められないということになる。以上の状況からしても、ただ単なる見かけ上のカーブの一致が必しも意味のある相関関係を示唆するものではないことは明らかである。

要するに、奇形とサリドマイドに正の関連がない、これは相関図によってのみ判断される、というのである。

これはおどろくべきことである。ありえない結論だからである。もしこんなことがあるとすれば、それは、どこかに新しいデータが加えられているか、さもなければ、一見正しそうなものどこかにごまかしがあるかである。実際はどちらであろうか。

実際は後者である。ごまかしの第一は、たて軸を奇形発生数から奇形発生率におきかえたことである。変数をこのようにおきかえると、得られる相関図は、サリドマイド仮説に関して何の情報ももたらさなくなる。サリドマイド仮説が成り立っていても、奇形発生率とサリドマイド販売量の間には必ずしも正の相関関係が生じないから

である。

数値で説明しよう。仮に、サリドマイドを使用する人は、1年に100錠あてのむとしよう。また全人口における妊婦の割合は年1%であるとしよう。このとき、人口4,000万人のA地区に1年間で100万錠のサリドマイドを売れば、サリドマイド仮説のもとでおおよそのところ、次のようになる。

$$\text{出産数} = 4,000 \text{ 万} \times 0.01 = 40 \text{ 万}$$

$$\text{奇形児出生数} = 100 \text{ 万錠} / 100 \text{ 錠} \times 0.01 = 100 \text{ 人}$$

$$\text{奇形発生率} = 100 / 40 \text{ 万} = 0.00025$$

これに対して、人口1,000万人のB地区に1年間で50万錠を売れば、おおよそのところ、次のようになる。

$$\text{出産数} = 1,000 \text{ 万} \times 0.01 = 10 \text{ 万}$$

$$\text{奇形児出生数} = 50 \text{ 万錠} / 100 \text{ 錠} \times 0.01 = 50 \text{ 人}$$

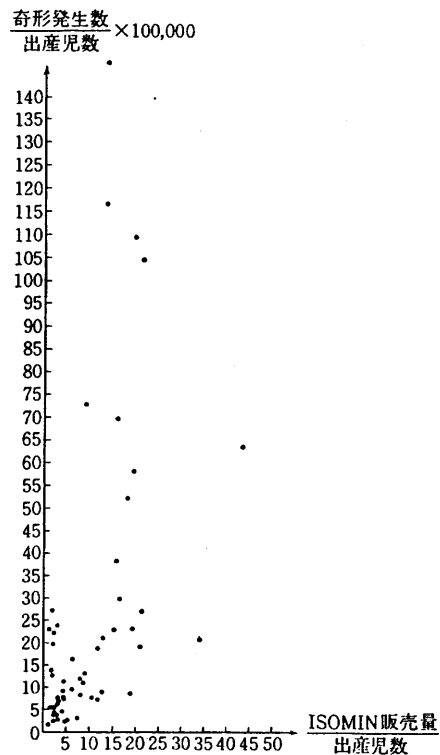
$$\text{奇形発生率} = 50 / 10 \text{ 万} = 0.0005$$

サリドマイド仮説が成立するとしたこの例で、たしかにサリドマイド販売量と奇形児出生数が比例しているのに対し、販売量と奇形発生率とは反比例している。絶対量と比率という無茶な対応をさせるかぎり比例関係という現象はおこらないのである。

ごまかしの第二は層別を行わないことである。層別を行わないときの危険性は前節でのべたとおりである。したがって杉山は、層別しないでよいという理由を、引用文にあるようなまやかしの理由でなく、データに即した判断で示さなければならない。すなわちデータの生の素性、性質を公表しなければならない。ところが、杉山はこれを全く行なわなかった。データの出所を示さず、第6図には目盛さえいれなかった。統計学の教授という立場でありながら、このような基礎的なことさえしていないというのは、上に指摘した意味のない相関図が、錯覚や誤解ではなく、特定の意図にもとづいて作られたものであることを明示している。そうでないかぎり、論文を発表してから一年半の間、このような初歩的な誤まりを訂正しないことは考えられないからである。

杉山のごまかしが誰の目にも明らかになった時点で、大日本は、新たな相関図を第1表とともに公表した。第7図がそれで、次の説明がつけられている。“プロットした点は右上りの傾向を示していない。したがって、この資料と相関図に関する限り、サリドマイド服用率と奇形発生率の間には、問題になる相関関係(正の相関関係)は認められない。…相関図により視覚的に相関が認められない以上、相関係数を求め出す必要はない”。

この表および図は、大日本が統計学的側面において公表した唯一の積極的、具体的な自らのための証拠である。サリドマイド仮説が誤まっていると主張するならば、大日本は具体的に自らの主張に整合する多くの証拠を、少



第7図 大日本製薬の示した相関図。第1表の(X, Y)を6年次×9地域=54点プロットしたもの。

くとも LENZ 以上に出さねばならないし、その能力も持っている。それにもかかわらず、これしか出さないとすることは、このデータが非常に重要な価値を持つことを意味する。これがサリドマイド仮説を否定するものでなければ、大日本はもはやサリドマイド仮説を否定する統計学的根拠を完全に失なうからである。

さて、実際に吟味してみよう。大日本が提出した数値には、その素性についてほとんど何の説明もつけられていなかった。しかし、数値相互の関係や、人口動態統計、森山の報告<sup>(1)</sup>などから、次のことがいえる。

(1) 第1表中33年の出産児数Cは9地域とも、33年の出産児数の3/12である。また38年の出産児数Cはその年の出産児数の2/12である。これに対して33年および38年の奇形発生数Bはその年の全数である。したがって、表中の33年と38年の奇形発生率Yは意味のない数値である。大日本は原データが未公表であることを利用して、このように歪んだ数値を提出したのである。

(2) 表中33年の奇形に関与したと思われる販売量Dは、33年の販売量の3/12である。同じく38年のDは、同年の販売量の2/12である。同じく37年のDは前年販売量 $\times \frac{9}{12} + 37$ 年販売量 $\times \frac{3}{5}$ である。他の年のDは、

$$\text{前年販売量} \times \frac{9}{12} + \text{同年販売量} \times \frac{3}{12}$$

である。このやり方で求めた数値は、月別販売量を9カ月ずらして得られる値と違はずである。その違いは、初年度と最終年度で特に大きいはずである。初年度に関するその明瞭な理由は、初年度の発売日が1月20日であること、最初の1~2カ月の販売量は、販売が軌道にのったときと全く違うこと、などである。最終年度に関するその明瞭な理由は、出荷停止が昭和37年5月17日であること、それから9月13日の回収決定までの間、イソミンが店頭にあったにもかかわらず、それが考慮されていないこと、回収終了はさらに1年以上あとであることなどである。これらの歪みは、月別販売量、出荷停止後の月別回収量、発売以前の知見薬品、試供品の配布量などが提出されれば、かなり修正されるはずである。

この他、実情がわからないため次の疑問点が残る。

(3) サリドマイド含有胃腸薬プロバンMの販売量が未公表であり、これは36年から37年にかけて増加している可能性がある。

(4) サリドマイドを危険期に服用してから出産までの正味期間は約8カ月である。それにもかかわらず、9カ月ずらして対応させるのは、買ってからひままでの時間おくれを統計的に考慮したものといえる。この考慮は、発売初期および出荷停止以後において少し変えねばならない。発売初期はどちらかというときすぐ使用するし、出荷停止後は、在庫品、回収残りを徐々に消費すると思われるからである。

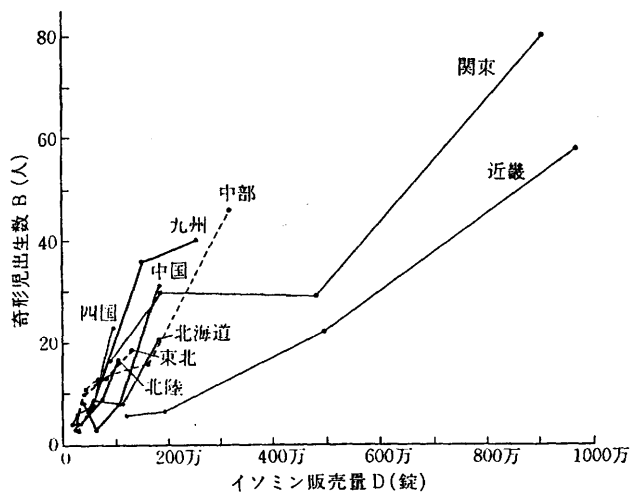
(5) 販売量Aが、どこで測定された数値かわからない。大日本からの出荷量である可能性が大きい。もしそうであれば、発売前後と出荷停止前後には、特別な政策的配慮をしている可能性が大きい。

以上、いずれも、データが精密にかつ素性を明らかにして公表されれば、かなり解決される問題点ばかりであるが、少なくとも現在公表されているかぎりではそれははっきりしない。明らかなことは、初年度と最終年度の値をそのまま使うことができないことぐらいであろう。このように、データに不備があるときの取扱いは、二通りある。一つは、明らかに不備なデータを除いた残りのデータに含まれている情報を最大限追求することであり、他の一つは、考えられる場合を推定してデータを修正し、そこで現われる現象を観察し評価することである。ここではその両者を行なおう。

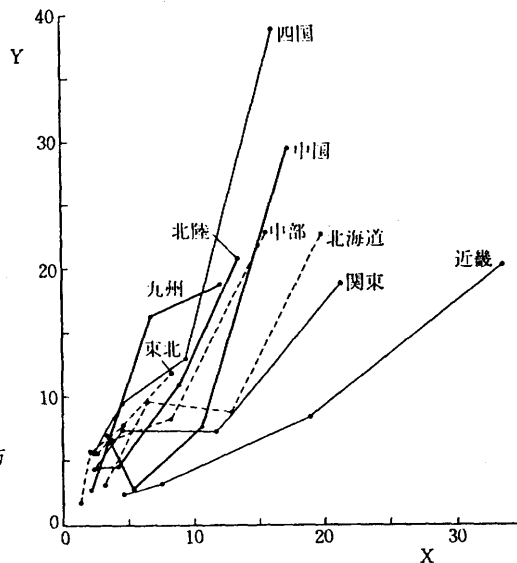
ところで、これらのデータから、サリドマイド仮説の論拠(2)を評価しようとするとき、相関図には二通りの書き方がありうる。一つは、奇形児数Bとそれに対応する販売量Dを、絶対量のまま相関図にプロットするもの

誤植: 同年→37年  
2/12→2/5

33年1月服用→33年1  
0月出産、2月服用⇒1  
1月出産、3月服用→1  
2月出産、4月服用→3  
4年1月出産



第8図 第1表のイソミン販売量Dと奇形児出生数Bとの相関図。昭和33年と昭和38年は除いてある。地域ごとの比例関係はかなり明瞭である。



第9図 第7図から昭和33年と昭和38年の点を除き、地域ごとに線で時間推移を示したものの、地域ごとでは第8図とほとんど同じである。地域ごとの比例関係はかなり明瞭である。

であり、他の一つは、大日本がやったように、それらを出産児数で割った  $X=D/C$ ,  $Y=B/C$  を相関図にプロットするものである。この二つは、大日本が33年および38年のデータについて行なったような細工をせず、しかも地域ごとに層別していれば、ほとんど違いを生じない。出産児数Cが、各地域ごとにほぼ一定だからである。ただ、地域間の比較をする際には、X, Yを用いた方が、よいかもしれない。出産児数で割ることが、一種の標準化になるからである。

ここではまず、33年と38年のデータを除いて両方の図を画こう。第8図が前者で、第9図が後者である。どちらの図でも各地域ごとにほぼ比例関係が見られる。地域ごとの勾配の違いは、前節の注意が必要であり妥当であることを裏書きしている。これらは大日本の公表したデータが、その主張と正反対に、サリドマイド仮説を支持する有力な証拠になっていることを示している。

第7図と第9図を比較すると、第9図にない点、すなわち33年のデータと38年のデータが特定の傾向を持って集中していることがわかる。読者が実際に33年と38年の点を確認してみれば、その傾向はもっとはっきり認識できるであろう。大日本の“相関関係が認められない”という見せかけの判断に、先にのべたデータの歪みが大きく役立っていたことがわかる。

ここで筆者があえて大日本の判断を見せかけだといったのは、引用した大日本の主張が事実と反し、第7図でさえも、正の相関が存在するからである。すなわち、第7図のように点の分布状態が正規分布型でないときの常道に従って順位相関係数を計算すると、その値は0.558

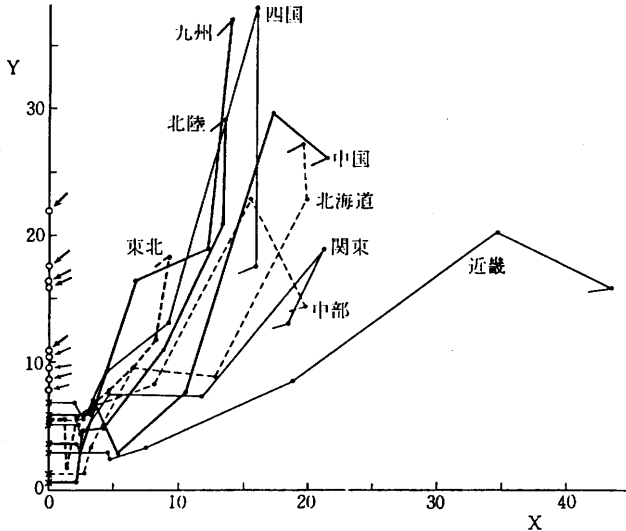
になる。この値にもとづいて、母相関係数が0という仮説を検定すると、有意差ありという結論に達する<sup>(19)</sup>。このように、層別が行なわれず歪んだデータが混入している第7図でも、統計的には相関が存在するのである。要するに大日本の主張は“驚を鳥といいくるめる”論理である。

杉山などが時間推移カーブではなく相関図を強調した理由がこれで明らかになった。相関図ならば、地域や時間順序をまぜて、データに存在している比例の関係をおからなくすることができるというわけである。

次に33年と38年のデータに考えられる補正をしてみよう。一つの方法は、大日本の主張になるべく有利なように、大日本が行なったやり方で、データの歪みだけを訂正するものであろう。実際には、33年の最後の3カ月に生まれた奇形児数、38年の最初の2カ月に生まれた奇形児数がわからないから、森山の報告を参考にして推定しよう。すなわち、森山の著書<sup>(20)</sup>の72頁図53にある四半期ごとの出生数のグラフから見当をつけ、33年の最後の3カ月の奇形児数も、38年の最初の2カ月の奇形児数も、それぞれの年の全数の1/4とする。こうすると、各地域のX, Yの値は、第2表のように修正される。これにもとづいてX, Yの相関図を画くと、第10図が得られる。

図の中で、×印は33年9月前、○印は38年3月後の値である。33年～34年のデータがほぼ水平にならぶ理



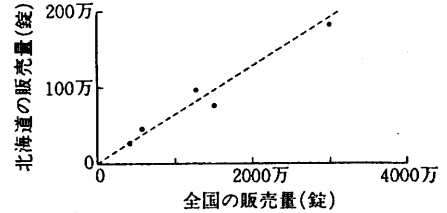


第10図 X, Y の相関図. 昭和33年と昭和38年の奇形児数をそれぞれ2分して、販売量と対応づけたもの. 2分するには推定値を用いた. 第7図と比べれば、第7図のごまかしの内容が理解できる.

由には (a) 森山調査がアンケート調査であるために、いま問題にしている奇形以外の異常児がはいってしまうこと、(b) 発売日以前に市場に出された試供品が存在すること、などが考えられる。データがかくされているため、これらを証明することが現在の段階ではできない。38年

第2表 大日本製薬の公表したデータの歪みを推定値で修正した場合の X, Y の値. 33年9月前, 38年3月後の出産児数はそれぞれその年の全出産児数の 9/12, 10/12 として、人口動態統計から求め、下表の X, Y の値を求めるのに用いた.

	33年9月前	33年10月後	38年2月前	38年3月後
北海道 X	0	2.7	19.6	0
北海道 Y	1.1	1.1	27.3	16.4
東北 X	0	1.2	9.2	0
東北 Y	5.7	5.7	18.2	10.9
関東 X	0	2.1	18.5	0
関東 Y	5.0	5.0	13.0	7.8
中部 X	0	2.3	19.8	0
中部 Y	5.6	5.6	14.1	8.6
北陸 X	0	2.1	13.6	0
北陸 Y	0.5	0.5	29.2	17.5
近畿 X	0	4.5	43.5	0
近畿 Y	2.8	2.8	15.8	9.5
中国 X	0	3.1	21.4	0
中国 Y	5.9	5.9	26.1	15.7
四国 X	0	2.0	16.0	0
四国 Y	6.8	6.8	17.4	10.4
九州 X	0	2.0	13.9	0
九州 Y	3.5	3.5	37.0	22.0



第11図 イソミン販売量の全国と北海道の比較. きれいな比例関係がある.

3月後の場合は、市場および消費者の在庫品によるものであろう。

いずれにせよ、大日本が提出した唯一のデータが、実際にはサリドマイド仮説を証明するのに都合なデータであることの意味は大きい。

なおついでに次のことを注意しよう。梶井は、北海道の小売量をもって、全国のサリドマイド販売量の時間推移を代用した。これは、北海道という地域が、薬品販売において、全国の中でそれほど特殊でないという梶井の体験的知識にもとづいていたようである。大日本が出した第1表でこれを吟味すると、第11図のように見事な比例関係が見られ、梶井のやり方は、“方法上の誤りがあるから無意味な分析だ”という大日本の評価に反して、梶井の分析の正しさが大日本のデータからも示されている。(未完)

文 献

(1) 西村秀雄他2名編: 先天異常, 金芳堂, 第5章(1966) (2) R. A. PFEIFFER & W. KOSENOW: Münch. Med. Wschr., 104, 68 (1962) (3) W. G. McBRIDE: Lancet, December, 16, 1358 (1961) (4) W. LENZ: Dtsch. Med. Wschr., 86, 2555 (1961) (5) G. SIEVERS: Arzneimittel-Forschung, 14, 6, 605 (1964) (6) H. WEICKER, K. D. BACHMANN, R. A. PFEIFFER & J. GLUEISS: Dtsch. Med. Wschr., 87, 1597 (1962) (7) W. LENZ: Ann. New York Acad. Science, 123, 228 (1965) (8) R. W. SMITHELLS & I. LECK: Lancet, May, 18, 1095 (1963) (9) T. KAJI: Ann. Paediat., 9月号, 205 (1965) (10) C. S. DELAHUNT & L. J. LASSEN: Science, 146, 1301 (1964) (11) 荒井良: 日本動物学雑誌, 79, 351 (1970) (12) 名古屋地裁, 昭和38年(ワ)第1152号, 同40年(ワ)第938号損害賠償請求事件被告準備書面 (13) 杉山博: 日本医事新報, No. 2351, 29 (1969. 5. 17) (14) 藤田恭一: 同上, No. 2363, 14 (1969. 8. 9) (15) 高橋純正: 同上, No. 2366, 14 (1969. 8. 30) (16) ポアンカレ著 河野伊三郎訳: 科学と仮説, 岩波文庫 (17) O. THALHAMMER: Pränatale Erkrankungen des Menschen, Georg Thieme, p. 296 (1967) (18) 昭和40年(ワ)第9994号, 同43年(ワ)第1393号損害賠償請求事件, 被告会社準備書面 (第18回) (19) 増山元三郎監修: 統計計算ポケットブック, 日本規格協会 (1964)

岩波書店『科学』

## アザラシ状奇形の原因 II

—サリドマイド仮説の成立に関する統計学上の争点について—

吉 村 功

## 5. 予見調査と回顧調査

次に論拠(1)に移ろう。前にのべた論拠(1)の確立は、まず奇形を発生させそうなあらゆる因子を列挙し、奇形児群と非奇形児群とで、各因子の出現率に差があるかどうか統計的に検討することからはじめられた。すなわち、標本として選ばれた調査対象を第3表のように分類し、奇形児群と非奇形児群とで、因子B(または $\bar{B}$ )の出現度数あるいは出現率を比べるのである。第3表では因子Bとして、サリドマイド服用をとりあげているが、これを、たとえば放射能をあげたかどうか、親や祖父母に奇形があるかどうかなどに換え、出現率の差の検定を行なえば、出現率に差のある因子がひろいだされる。あとでものべるように、出現率に差があるということは、その因子と奇形発生が関連あることを意味する。それゆえ、このようにして拾いだされた因子について、調査対象をさらに吟味し、より多元的な分類も行なって検討すれば、もっとも関連性の大きい因子を見出すことができる。これを比較的初期に、しかも詳しく実行したのは、LENZらとWEICKERらであった。LENZは、もともと先天奇形の研究者であったし<sup>(20)</sup>、WEICKERらは、石けん、腕時計、菌の治療経験など、ありとあらゆる因子を吟味する徹底性を持っていた。そして、彼等を中心にして、実に多数の報告がサリドマイドを指定したのである。

これに対して出された異論の第一は、ここにのべたような調査法についてであった。たとえば杉山は、LENZらの調査法に欠陥があったとして次のようにのべている。“…そこで大切な手続きは、胎生学上の該当時期にサリドマイドを服用した群と服用しなかった群を設定し、これら両群から奇形児がどのように発生するか、両群での発生率に有意差があるかを統計学的に検定することである。…このように、ある時点においてcomparableな両群を設定した上でこの両群から問題の症例がどのように

第3表 四分表。a, b, c, dなどは、調査対象となったn人中の各区分に属する人数である。

	A奇形	$\bar{A}$ 非奇形	行 和
Bサリドマイド服用	a	b	a+b
$\bar{B}$ サリドマイド非服用	c	d	c+d
列 和	a+c	b+d	n

発生していくかを観察し、両群における発生率の差を検討しようという方法を prospective study(以下では予見調査と書く)という。これがもし可能ならば、オーソドックスな方法であることは自然的であろう。しかしサリドマイド問題の場合、このような方法は現実には不可能である。そこで、これに代る方法として用いられるのは奇形児群と健常児群を設定し、この両群におけるサリドマイド服用状況の調査を行ない(第3表のような)、いわゆる四分表にまとめるのである。これを retrospective study(以下では回顧調査と書く)という。…回顧調査においては、通常調査時点における生残奇形児だけを対象としている点が、すでに予見調査の場合と根本的に異なるのである。しかもまた、回顧調査の場合には把握した生残奇形児および対照群での健常児の母親を対象として、…母親の記憶の糸をたぐり出すのであるから、出産後すでに数年という期間を経過した時点においては一般に極めて不正確なデータしか得られない…。さらに母親の記憶に関連して、母親のサリドマイドを服用したという証言をそのまま信用することは疑問がある。…奇形児の母親においては、サリドマイドの服用を誤って肯定する方が多く、サリドマイドの服用を誤って否定することは、はるかに稀であると考えられる。…第三者に責任があるのだというふうに、責任を転嫁できるからである。これに加うるに、調査者が(サリドマイド)仮説を確認しているならば、自分が薬剤を服用したかどうかは思い出しにくい母親達は…調査者にとって‘正しい答’に導かれることになる。…以上のべたように回顧調査においては種々の根本的欠陥を含んでいる…”。

要するに、予見調査の方がオーソドックスであり、やむをえず採用する回顧調査では、種々の根本的欠陥によって、みかけ上、サリドマイド仮説が作られてしまうというのである。この批判は、はたして正しいのであろうか。この批判の特徴は、とにかく回顧調査は欠陥があるという印象を与えようとしていることである。これは、調査経験の全くない人には通用するかもしれない。しかし多少とも経験のある人には、誤りであることがすぐわかる。すぐれた医学者 MAINLAND がその著書<sup>(21)</sup>でのべているように、杉山があげた回顧調査の欠陥なるものは、予見調査でも同様に存在するからである。

たとえば薬の効果の臨床実験を考えよう。ふつう望ましいとされているのは、一群の人を、ランダムにあるいは系統的に2分して、同等と思われる2組の被験者群を作り、一方に薬を投与し結果を見る方法である。これは典型的な予見調査である。それでも杉山があげているような欠陥が存在する。薬を飲んだあとで余病を併発したり行方不明になったりすることをコントロールできないのは、生残奇形児しか調べられないのと同じである。被験者が頭痛や吐き気など生じた現象を告げ忘れたり、述べる必要を感じなかったりして、重要な情報が見逃されることも多い。患者が指示どおり薬を飲んでいないにもかかわらず飲んだとつげることもしばしばある。実験者が自分に都合のよい結果を期待して判断を誤ることは、問診や写真の読影でよくあることである。

だが、回顧調査では与えられた条件のもとでの観測しか行なえないのに対し、予見調査では特定の条件を設定して観察が行なえるのではないかと考える人もいるだろう。たしかに予見調査において、そういう条件の設定が行なえる場合がある。その場合は予見調査を行なうべきだろう。たとえば新薬の効果の実験などの場合である。しかし、これは、予見調査がオーソドックスで回顧調査が代用品ということの意味するのではない。予見調査でも現実上の制約からそういう条件の設定ができない場合があるし、そもそも回顧調査しか意味がない場合もある。この場合には回顧調査がオーソドックスで正しい調査方法となる。たとえば、ある寮で集団赤痢が発生し、注意深い回顧調査によって、“患者がすべてあるサラダを食べており、そのサラダの調理人が保菌者であり、他に保菌調理人がいない”ことがわかったという場合、赤痢の伝染源は、保菌調理人と断定できる。この場合、予見調査がオーソドックスだという人はいないだろう。

要するに、予見調査であろうと回顧調査であろうと、あるいはまた計画された実験であろうと、一般には同様な困難が生じるのであって、回顧調査のみが特別な欠陥を持つというわけではない。だからこそ真実に対して忠実な研究者は、現実の制約の中で調査資料の信頼性を確保するために、苦労するのである。我々は関係者が実際に行なったその努力を一つ一つ吟味することを通して、結論を正しく評価するのでなければならぬ。杉山のように、一般的な空想や質の低い研究者の例を一般化して、欠陥があるはずだときめつけてはならない。たとえば LENZ や WEICKER らの調査では、死産児や生後すぐ死亡した奇形児も可能なかぎり調べられていること、医師の処方箋やサリドマイドの空箱のような客観的証拠で服用を確かめる努力がなされていること、そのような確認が十分でないものについては、サリドマイド仮説に不利

になるようにデータを整理しなおしていることなどは重要な点である。

LENZ らのこのような努力が存在しているからこそ、論拠(1)は成り立ち、杉山の批判は誤りなのである。もちろん、多くの場合、我々は調査報告書だけからそのような努力を読みとらねばならない。ごまかしや錯覚の存在も発見せねばならない。それは困難なことである。しかし不可能なことではない。それを発見するコツは、報告者が詳しいデータを公開するかどうか、調査者がどれだけ密接に調査対象と接触しているかという点に注目することである。真実の発見に努力するものは、必らず可能なかぎりのデータ公開を行なうし、対象の追及を行なう。それが錯覚や見落としを発見する一つの有力な方法だからである。これに対して、特に理由なくデータの公開を拒否する報告者はまず信用できない。必らずデータに弱点があるからである。その典型が杉山であることは、すでに指摘したとおりである。

#### 四分表における標本のとり方

異論の第二は、論拠(1)の裏づけとなる標本に偏りがあるというものである。この異論の展開は、人によって異なるが、杉山の場合は次のようにのべている。

“…非奇形群はむしろ場あたりで適当につかまえてきた集団である場合が多いのではなからうか。たとえば、LENZ がとりあげている四分表を第4表に掲げよう。この表の非服用の行を横にみると、計208例中奇形児が22例出生したことになるが、この10%以上の高率はどうみても到底納得し難いデータであろう。…このような四分表にカイ二乗検定を行ない、カイ二乗値が非常に大きいことを論拠に無相関仮説…は著しく小さな危険率で棄却され、サリドマイドの服用、非服用と問題の奇形の出現、非出現の間には相関ありといっても、もとのデータが根本的に怪しいのであるから、この検定自体およそ意味のないものになってしまう。…さらにサリドマイド服用率についてはどうであろうか。…(第4表に)おける比

第4表 四分表の例(LENZ のデータ)。

	A奇形	A非奇形	行和
B服用	90	2	92
B非服用	22	186	208
列和	112	188	300

第5表 説明のための例。ある地域、ある時期の出生児の構成。

	奇形	非奇形	行和
服用	90	2,000	2,090
非服用	22	186,000	186,022
列和	112	188,000	188,112

率は…92/300 となっている。…およその見当として西ドイツで妊娠初期にサリドマイドを服用した母親の割合は約1~3% というように SCHEUCH は推測している。このような調査と推測からすれば、先に掲げた四分表における調査対象集団での母親の服用率は、統計的変動を見込んだとしてもなおかつ異常に高いものであることがわかる。即ち、先の四分表では母親の服用状況も正しく反映されていないということになるのである。…” 要するに、調査対象となった 300 人中の奇形発生率やサリドマイド服用率は世間一般に比べて高すぎる。これは調査対象に偏りがある証拠であるというのである。

この主張の誤まりは、数値を用いて説明するとわかりやすい。いま、ある地域である時期に生まれた人数が 188, 112 人で、その区分構成が第 5 表のとおりであったとしよう。このとき、この地域、この時期の非服用者の奇形発生率は約 0.01% であり、サリドマイド服用率は約 1% である。いま、ここで、奇形とサリドマイド服用との関連を調べるのに全数を調べるのは大変なので、奇形児は全数調べ、非奇形児はランダムに 1/1,000 抽出したとすると、その調査結果は、ほぼ第 4 表になるであろう。この調査された 300 人においては、標本抽出に何の偏りもないのに、非服用者中の奇形児の割合は 10% という高率であり、サリドマイド服用者の割合も 30% という高率になる。これは標本に偏りがあるからではなく、こういう標本のとり方に必然的に伴う現象なのである。

ではこの場合、服用者中、あるいは非服用者中の奇形児発生率はどうやって推定するのであるか。ここで、統計数字についての次のことを思いださなければならない。すなわち、ある目的のためにとられたデータを他の目的に流用するのは一般に困難であり、時には不可能であるということである。いま直面している例は、その不可能な場合に相当する。すなわち、サリドマイド服用と奇形発生との関係を吟味するためにとられたデータであるがゆえに、上にのべたような現象が伴ない、したがって、奇形児発生率などが推定できないという場合なのである。これはごく初等的な常識であって、たとえば土屋<sup>(22)</sup>はこのようなデータにおいて、まちがって推定をしないために行和を記入しないよう配慮をしているくらいである。

高橋<sup>(15)</sup>は以上の内容のことを簡明に説明した。これによって杉山の誤まりは明示された。ところがこのあと、同じ異論があらたな展開でなされた。大日本製薬<sup>(23)</sup>の主張がそれで、要約すると次のようになる(記号は第 3 表参照)。

“四分表で二つのカテゴリーの関連性を追及するには、時間と場所を同一にして調査対象となる母集団を限定し、

I 型: その母集団からランダムに  $n$  個の個体をえらび

各区分内に属する個体数  $a, b, c, d$  を調べるのが一般的である。しかし、技術的な制約から I 型が利用できない場合には、

II 型: 母集団中で属性 A を持つ全個体(以下では集団 A とよぶ。集団 B などと同様)からランダムに  $a+c$  個をえらび、集団  $\bar{A}$  からランダムに  $b+d$  個をえらび、 $a, b, c, d$  の値を観測する、

III 型: 集団 B からランダムに  $a+b$  個、集団  $\bar{B}$  からランダムに  $c+d$  個をえらび、 $a, b, c, d$  を観測する、が用いられる。どの型の場合でも、統計的検定に用いられる計算公式は同一であるが、検定される仮説は異なる。I 型では無相関仮説であるのに対し、II 型、III 型では、母出現率に差がないという仮説である。それゆえ、II で無相関仮説を検定するためには、母出現率に差がないという仮説が無相関仮説に転化する条件、すなわち  $a+c$  と  $b+d$  の比が母集団における集団 A と集団  $\bar{A}$  の比にほぼ一致するという部分母集団の条件がみたされなければならない。III 型においても同様である。”

この主張には、術語がたくさん含まれている。このため、ふつうの人にはわかりにくく、それがもっともらしい雰囲気をもたらしている。しかし、この主張が誤まりであることは、四分表のカイ二乗検定における仮説の意味を示せば明らかになる。以下にそれを説明しよう。

いま、ある限定条件(たとえばある病院で出産すること)のもとで、ランダムに個体を一つえらぶとき、その個体が四分表のそれぞれの区分に入る確率は第 6 表の記号で表わされるとする。このとき、

$$H_1: \frac{p_{11}}{p_{.1}} = \frac{p_{1.}}{1} \text{ すなわち } p_{11} = p_{.1} p_{1.}$$

で表わされる仮説  $H_1$  を四分表における独立性の仮説という(杉山はこれを無相関仮説とよんでいるようである)。仮説  $H_1$  がなりたたないとき、二つの分類因子は独立でない、あるいは関連があるという。

独立性の仮説を検定するのに、標本のとり方が I 型であれば、周知のカイ二乗検定は、次の内容を持つ。すなわち、観測した結果において、

$$\frac{a}{n} - \frac{a+b}{n} \times \frac{a+c}{n} \text{ すなわち } \frac{a}{n} - \frac{(a+b)(b+c)}{n^2} \neq 0$$

がなりたっていたら、独立性の仮説  $H_1$  を否定せず、この近似的等号が大きいくずれるときには仮説  $H_1$  を否定する。事実、カイ二乗検定は

第 6 表 確率の記号の関係。

	A	$\bar{A}$	行 和
B	$p_{11}$	$p_{12}$	$p_{1.}$
$\bar{B}$	$p_{21}$	$p_{22}$	$p_{.2}$
列 和	$p_{.1}$	$p_{.2}$	1

$$\left\{ \frac{a}{n} - \frac{(a+b)(a+c)}{n^2} \right\}^2 = \frac{(ad-bc)^2 n}{(a+c)(b+d)(a+b)(c+d)} \cdot \frac{(a+c)(b+d)(a+b)(c+d)}{n^5}$$

がある値より大きければ、二つの因子の間に関連ありと断定する方式になっている。この検定方式の合理性は、

$$\frac{a}{n}, \frac{a+b}{n}, \frac{a+c}{n}$$

がそれぞれ、 $p_{11}, p_{1\cdot}, p_{\cdot 1}$  の値を反映することで直観的に納得できる。

ところで、 $p_{12} = p_{1\cdot} - p_{11}$ 、 $p_{\cdot 2} = 1 - p_{\cdot 1}$  であるから、仮説  $H_1$  がなりたつとき、

$$\frac{p_{12}}{p_{\cdot 2}} = \frac{p_{1\cdot} - p_{11}}{1 - p_{\cdot 1}} = \frac{p_{1\cdot}(1 - p_{11})}{1 - p_{\cdot 1}} = \frac{p_{1\cdot}}{1 - p_{\cdot 1}} = \frac{p_{11}}{p_{\cdot 1}}$$

がなりたつ。したがって、仮説  $H_1$  は次の仮説  $H_2$  と全く同じである。

仮説  $H_2$ :  $p_{11}/p_{\cdot 1} = p_{12}/p_{\cdot 2}$

つまり、“集団Aに属するという条件でランダムに個体をえらんだときそれが属性Bを持つ確率  $p_{11}/p_{\cdot 1}$  は、集団 $\bar{A}$ に属するという条件でランダムに個体をえらんだときそれが属性Bを持つ確率  $p_{12}/p_{\cdot 2}$  に等しい”ことが独立性の仮説である。これは、標本のとり方がII型のときの“集団AにおけるBの出現率と集団 $\bar{A}$ におけるBの出現率とが等しい”という仮説が、無条件に独立性の仮説と同等であることを意味する。母集団の部分集団などという条件は全く不用である。無条件であるからこそ検定方式が同じになるのだということは、LEHMANN<sup>(24)</sup>によって数学的に示されている。

同じことはIII型でもいえる。仮説  $H_1$  は次の仮説  $H_3$  に等しい。

仮説  $H_3$ :  $p_{11}/p_{\cdot 1} = p_{21}/p_{\cdot 2}$

これは標本のとり方がIII型であるときの“集団Bと集団 $\bar{B}$ とでAの出現率に差がない”という仮説が、独立性の仮説と同等であることを意味する。

以上のべたことからわかるように、標本のとり方がI型であろうとII型やIII型であろうと、カイ二乗検定で有意と判定されれば、それは二つの分類因子AとBが統計的に関連あることの証拠である。したがって、実際にどの型をえらぶかは、現実の制約条件と、そのもとの効率、つまり少ない労力で多くの情報をうることを基準にしてきめればよい。現実の制約条件というのは、たとえば製薬会社がこれから新薬の効果を調べるときにはII型が採用できないとか、奇形の発生に注目した医師が原因を追及しはじめるときにはIII型が採用できないという条件である。

効率というのは、見方によって異なるが、数学的には

“検出力が大きいこと”などと定義される。仮に標本の大きさが与えられているとすると、 $(p_{\cdot 1}, p_{\cdot 2})$ と $(p_{1\cdot}, p_{2\cdot})$ とのよりアンバランスな方の因子で二つの集団を作り、それぞれから  $n/2$  ずつの個体をえらぶのが最良である。たとえば、サリドマイド服用率  $p_{\cdot 1}$  が 0.01 ぐらい、奇形発生率  $p_{1\cdot}$  が 0.0001 ぐらいとすれば、母集団を奇形か非奇形かでわけて集団Aと集団 $\bar{A}$ を作り、それぞれからランダムに  $n/2$  ずつえらぶII型が最良である。もちろん正確にできなければ、それに近い数でよい。とにかく集団Aからの個体数と集団 $\bar{A}$ からの個体数を  $p_{\cdot 1}$  と  $p_{\cdot 2}$  の比にするというのがきわめて効率の悪い方法であることは、LEHMANN<sup>(24)</sup>が明快に記しているとおりでである。

#### 対照のえらび方と結論の普遍性

異論の第三は対照のえらび方に向けられた。たとえば、LENZ などの調査では対照が場あたりのにとられているから、そのデータをいかに統計的に処理しても、結論は無効であるという批判である。ここで対照とは、奇形児に対する非奇形児のように、注目している集団に対比される集団のことである。

この批判を考えるために、まず、直接例にあげられた LENZ のデータを説明しよう。LENZ は、Hamburg 大学病院小児科と、Hamburg, Braunschweig, Stade の各病院で見出された全奇形児 129 例を調べ、

サリドマイド服用が確かなもの 90 例

服用が確かとはいえないもの 22 例

調査がそれまでに終わっていないもの 17 例

という結果を得た。もちろん死亡例も含めている。対照には同時期に Hamburg の病院で入院、出産した非奇形児から層別してランダムに 188 例を選び、その母親が妊娠初期にサリドマイドを服用しなかったもの 186 例、妊娠初期に服用したが日付け不明のもの 2 例という結果を得ている<sup>(25)(26)</sup>。第4表はこの結果を表にしたもので、非服用の 22 例も、ほとんどが、服用したとみなせる証拠を伴っているが、慎重を期して非服用に分離しカイ二乗検定をしたと LENZ はのべている。

これくらい結果がはっきりしていると、仮に杉山がいうように、非奇形児に対する追及がゆるくて服用者の一部が非服用者に分類されていることがあっても、有意差ありという結果が変ることはありえないであろう。しかし、この LENZ の対照のえらび方には、吟味すべき点がある。奇形児は、いくつかの病院の全数を対象としているのに対し、非奇形児は Hamburg の病院で入院出産した者に限られていることである。もし LENZ が、奇形児を調べたのと同じ病院で見出された非奇形児全体からランダムに 188 例えらんだのであれば、そのえらび方は前

節のII型に相当し、前節でのべたように良いえらび方をしたことになる。実際は、一部の病院からランダムにえらんだのであるから、全体からランダムではない。このような場合には、全体からランダムでないことの影響を吟味しなければならない。

一般に、統計家は、標本抽出のランダム性を強調する。それは、標本を観測して得られた結論を母集団に普遍化してよいという保証を得るためである。ところが実際には、技術的・物理的制約があって、ランダムな標本抽出は困難な場合が多い。たとえば一医師が、ある病気の原因を追究して患者および健康人の標本を得ようとするとき、自分が関係している病院とか、その病気で苦しんでいる人とかからは、多少の便宜を得られるが、一般の協力を得ることは困難である。特にそれが、大きな力を持つものの利益に反する場合には、はなはだしい。このため、標本はある限られた範囲になり、ランダム化も部分的にならざるをえない。このような場合には、結論を無条件に母集団に普遍化することは許されない。しかし、だからといって、全く普遍化が許されないというわけではない。抽出の手続きがランダムでなくても、標本が母集団を公平に代表していると見なすことが既存の知識と矛盾しないならば、ランダムだと想定して矛盾に出会うまで次の仕事を進めてよい。それが科学の常道である。たとえば、作物のできと土壌成分の関係を調べるのに、ある畑の数カ所から、適当な見当をつけて試料をとり、土壌成分を分析したとしよう。この分析データは、それだけでは畑全体を代表すると見なせない。しかし、その畑の作物のでき具合が、毎年畑全体でほぼ均一であるという事実があれば、試料採取がランダムでなくても、分析データは畑全体に普遍化できるのである。

同じように、LENZの場合も、Hamburgの病院の入院出産者が、Hamburg, Braunschweig, Stadeなどの他の病院のものと異質でないかどうか、あらゆる既存の知識および追加データで確かめなければならない。LENZ自身もそのことは意識していたようで、1962年の報告には<sup>(2)</sup>、“この報告は暫定的報告(vorläufige mitteilung)である”と書いている。そして、大学産科で出産した非奇形児の集団、大学小児科に来院した非奇形児の集団、自分の知人での集団などを比較してサリドマイド服用率がほとんど変わらなかったことを後で公表している<sup>(27)</sup>。また他人の調査報告とも比較して、大学病院で出産した人が特に高いサリドマイド服用率を持っているわけではないことを確かめている。この吟味はWEICKERらもやって、同じ結論に達している。このような確認がなされているがゆえに、杉山などの非難は、全く根拠のない非難といわざるをえない。

以上はLENZの結論をある限られた地域に普遍化するという議論であるが、同じ論理は、この結論を人類全体に普遍化する場合にも用いられる。もしLENZあるいはWEICKERらのデータがこれだけのものならば、それをもって人類一般を論じるのは危険を伴なり。しかし、サリドマイド仮説を根拠づける同様なデータが、他の地域、他の階層にも共通にみられるならば、それらのデータは、相まって単独ではなしえない普遍性を獲得することができる。それが科学一般の認識法である。

関連しているので、対照のとり方と結論の普遍性に関する他の例についてもふれておこう。

竹内ら<sup>(28)</sup>は、都立築地病院で、1959年8月から1961年9月までの2年間に、111人の妊婦にサリドマイドを処方し、3例の奇形発生を見た報告している。この事実は、これだけではサリドマイド仮説の有力な根拠といえない。サリドマイド仮説とは矛盾しないが、対照について何もふれていず、したがって奇形発生例に併用されていた黄体ホルモンの影響などが分離できないからである。しかし、この報告に、WEICKERら<sup>(6)</sup>、森山<sup>(1)</sup>、高橋<sup>(29)</sup>らの調査結果が追加されると、サリドマイド仮説を支持する非常に有力な証拠が得られたことになる。もし黄体ホルモンの併用が必要ならば、WEICKERや森山の報告中に、サリドマイドと同程度の重みでそれが現われているはずなのに、実際はそれが無い。また、同時期に同産院で入院分娩した数千例中で、危険期に処方された4名(LENZの基準では8名)中3名に奇形が発生しているのに、他には発生を見ていない。これらの事実が、不完全ながらも対照の設定になっており、かつその結果を普遍化するための吟味になっているからである。

サルによる動物実験でも同様なことがいえる。DELAHUNTら<sup>(10)</sup>はその実験において対照群を設定しなかった。したがってこの実験のみでサリドマイド仮説を支持する有力な証拠ということとはできない。しかし実際は、これが有力な証拠となっている。それは、このような奇形のサルにおける発生が、サリドマイド仮説を攻撃する側からも典型例を出せないくらい稀であるのに対し、この実験では、ほぼ全数に奇形が生じているからである。対照となるべき状態がほぼ自明という場合に、対照を省略するというのは、物理や化学の場合によく用いられている科学の方法なのである。

蛇足であるが、薬効試験の場合に、対照群の設定が特にきびしく要求されるのは、人間の自然回復力および擬薬効果をはなはだしいからであり、薬効に他の効果がまじることをその実験だけからも排除しようとするからであるのを注意しよう。

サリドマイド仮説に対して、最近出されている統計学上の異論を吟味した。その結果、サリドマイド仮説はますます確かなものになった。反証と思われて出されたデータが、実はサリドマイド仮説の有力な証拠になっていたからである。異論の検討の過程で、故なくデータを隠すのはごまかしがある証拠だという認識はこの場合にも正しかった。本文では推論でしかなかったいくつかの点においても、より精密なデータが出れば、もっと真実がはっきりするであろう。森山調査報告の医学上の再検討も必要な段階ではなからうか。大日本製薬には調査結果の生データを知らせながら、他にはそれを知らせていない森山調査班の全メンバーの責任は決して軽くないと考えなければならない。

最後に、わが国の奇形発生状況についての情報を持っている人がおいでになったら、その情報を教えていただきたいということを、とくにお願いして筆をおく。

追記 この小論の前半に対して、多くの人から資料、意見、質問がよせられた。その一つはサリドマイド仮説における統計的認識の意義に関するものである。例えば、

“サリドマイド仮説の成否を問題にするのなら、奇形児の母親が100% サリドマイドを服用していることを示さなければならないのではないか。また、危険期にサリドマイドを服用したもののみを対象にして調査を行ない、すべてに奇形児がうまれたことを示さねばならないのではないか”

といった疑問である。そこで以下に追記として、統計的認識の意義について簡単にふれておこう。

もし我々が自然の状態を、自由自在に、完全に観察できるのなら、おそらく統計的推論とか統計的認識といったものは不要であろう。例えば上で問われているように、サリドマイド服用率が100%でないということで、サリドマイド仮説をすてるということになるだろう。

ところが、現実の我々は、そういうことができない。ある制約のもとで、実体を反映する現象の一部分をやっと観測し数値化するにすぎない。たとえば、サリドマイド問題でも、我々は次のような条件に支配されることを拒否できない。

(1) 妊娠期における薬物の服用歴を完全に知ることはめったにできない。“○○を△△の頃のんだように思う”というような情報で満足せざるをえないのが、むしろふつうである。

(2) 受胎日、胎児の成長速度などに個体差があるにもかかわらず、これを測定する確実な方法は、存在しない。

(3) いま問題になっている奇形をまちがいに判別できるのは、症例を多数吟味した経験を持つ、ごく少数

の人だけである。我々が人手できる多くのデータは、この点で誤まりを含むものであることが普通である。

(4) 限られた労力と時間しかないので、必要な情報を十分集めることは不可能である。

このようなわけであるから、我々が行なう観察で得られるデータは、必ず観察のていねいさと、観察対象の条件に対応したある程度の誤差あるいは変動を伴う。そして、100%とか、すべてにとこいいえない情報のみを我々にもたらす。たとえば、危険期にサリドマイドを服用した母親が100% 奇形児を出産したという実体があったとしよう。もし、その母親たちが、服用日を忘れ、危険期外にサリドマイドを服用して服用日を忘れた人たちと一緒にされれば、奇形児出生率が5割にもならないサリドマイド服用者の集団ができあがる。このとき我々は、この集団からサリドマイド仮説の成否を評価しなければならないという状態におかれる。

このようなとき、我々のなしうることは、データのとり方や対象に関する既存の知識を十分に利用して、誤差、変動の原因、程度、傾向をつかみ、そのうえで誤差をこえて現われている法則性を可能なかぎり追及すること、すなわち統計的推論を行なうことである。このような推論は、個々には、弱点の多い、一面的なものである。しかし、そのようにして認識される法則性がすべて一つの方向を示し、ある整合的説明を可能にするとき、それらのデータは、単独では持ちえない論拠の質を得る。そして確かに本質的認識をもたらす。サリドマイド問題で大きな役割を果たしているのはこのような認識の方法である。

もちろん、誤差、変動は大きいより小さい方がより確かな認識を我々にもたらす。そのゆえに、我々は誤差、変動の小さいデータをとる努力をせねばならない。しかし、いかに誤差などが小さくても、それがあらかぎり、上にのべた認識の方法は必須である。読者は、誤差、変動のない観測はめったにないという経験的知識から、上にのべた認識法を肯定するであろう。事実、この認識法は現在、あらゆる科学の分野で科学的認識法として用いられているのである。

## 文 献

- (20) W. LENZ: Acta Genetica et Statistica Medica, 9, 249 (1959) (21) D. MAINLAND 著 堀山元三郎監修 柏木力訳: 医学統計の基礎—臨床医学のための—, 岩波書店 (22) 土屋健三郎: 疫学入門, 医学書院, pp. 88~89 (1968) (23) 東京地裁昭和40年(ワ)第9994号, 同43年(ワ)1393号事件被告会社準備書面(第15回) (24) E. L. LEHMANN: Testing Statistical Hypothesis, Wiley (1959) [渡谷政昭・竹内啓訳: 仮説検定論, 岩波書店 (1969)] (25) W. LENZ からの返信 (1970. 5. 14) (26) W. LENZ & K. KNAPP: Dtsch. Med. Wschr., 87, 1232 (1962) (27) W. LENZ: Expertise presented to the Court at Alsdorf (アルストルプ裁判所に出された鑑定書) (28) 竹内啓啓他4名: 日産婦誌, 15, 278 (1963) (29) 高橋純正: 日本医事新報, 2400, 46 (1970)