HOKKAIDO UNIVERSITY

Title	VDT障害と統計的仮説検定:VDT労働と妊娠帰結の関係について
Author(s)	木村, 和範
Citation	經濟學研究, 41(4), 34-44
Issue Date	1992-03
Doc URL	https://hdl.handle.net/2115/31896
Type	departmental bulletin paper
File Information	41(4)_P34-44.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

経済学研究 41-4
北海道大学 1992.3

ＶＤＴ障害と統計的仮説検定

――ＶＤＴ労働と妊娠帰結の関係について――

木 村 和 範

はじめに

16ビット・コンピュータの機能の充実，価格の低廉化，操作の容易性によって，オフィスのなかにＯＡ機器が入りこむようになって久しい。この結果，ディスプレィつき端末装置(VDT, video or visual display terminal)のオペレータの間で，ＶＤＴ障害という名前で総称される労働災害が多発するようになった。それは，コンピュータ・プログラマに現われることがあるテクノ・ストレス（神経過敏症，不眠症，さらには入院加療を要する精神的疲労）だけではない。ＶＤＴ障害の研究者ウテ・ボイカート(ドイツ・ハンブルク特別市労働・厚生・社会局）やドイツのＶＤＴ障害問題に詳しいジャーナリスト，カローラ・シェーベなどによれば，ＶＤＴ障害といわれる症状には，視力低下，眼精疲労，首，肩，背中，指，腕，腰などの筋肉系の障害，皮膚疾患（ＶＤＴ表面に発生する静電気によって室内の「ちり」や「ほこり」が顔面にあたって汗孔や毛孔をふさぐことによる)，口中に溶解した歯の充填剤に起因する中毒症（ＣＲＴ[cathode-ray tube：陰極線管]からの漏洩電波の作用による)，内臓疾患(アドレナリンや胃酸の過剰分泌，高血圧，低血圧，心悸昂進)，神経障害などがある[1]。さらに最近では，ＶＤＴオペレータのドライ・アイ(ズ)が注目されるようになった。これは，オペレータが原稿やディスプレィを凝視するために，まばたきの回数が減少し，眼球面が乾燥し，灼熱感・異物感を訴えるなどの症状を呈する疾病である。

このようなＶＤＴ障害はコンピュータ化の進展にともなって，わが国でも社会問題として注目され，その研究が積み重ねられている[2]。またアメリカやカナダ，北欧諸国においては女性ＶＤＴオペレータに流産や奇形出産が多発した1980年ころから，ＶＤＴ障害の研究が盛んになったようである[3]。この意味で生殖にかんするＶＤＴ障害は「古典的」な障害であるということができる。今日にいたるまで，この生殖の障害については，その原因がＣＲＴから漏洩する低周波電磁波であるとする数多くの動物実験がおこなわれている[4]。

ところが，このような実験結果にもかかわらず，オペレータの妊娠にたいするＶＤＴ労働の影響については，スウェーデンのエリックソンとケーレン[5]が統計的方法を用いて否定している。

1) Boikat, U.: Belastungen an Bildschirmarbeitsplätzen, in: Humane Produktion, 4-5/1988. Schewe, C.: Krank durch Computer, ... und wie man sich dagegen schützen kann, Reinbek bei Hamburg:Rohwolt Taschenbuch Verlag, 1989, Teil II.

2) たとえば梶原三郎監修（ＶＤＴ職場の労働と健康編集委員会編）『ＶＤＴ職場の労働と健康』労働安全衛生研修所,1984年。ＶＤＴ労働研究会編『ＶＤＴ労働と健康』労働基準調査会,1988年など。

3) DeMatteo, B., *Terminal Shock. The Health Hazards of Video Display Terminals*, Tronto:NC Press Limited, 1985.

4) Schewe,C., a.a.O., S.102-107.

5) Ericson,A. and Källén, B., "An Epidemiological

彼らによれば,「流産や児の奇形・異常の発生は,VDT作業経験が最多と考えられる群ではやや多い傾向があったが,統計的には有意でなかった」とされ,この研究結果はわが国にも肯定的に紹介されている[6]。多くの動物実験にもかかわらず,エリックソンとケーレンの結論は正当なものであろうか。本稿では,彼らの用いた統計的方法をとくに統計的仮説検定に着目して,その有効性を検討し,その結論の妥当性を吟味する。

1．VDT労働と妊娠帰結の関係にかんする統計的検証
—エリックソン＝ケーレン論文の概要—

エリックソン＝ケーレンの論文のタイトルは,「VDT労働と妊娠にかんする疫学的研究」であるが,これは2部からなり,1部のサブ・タイトルは「届出統計の研究」であり,2部は「症例対照研究」である。本稿ではこのうち第1部だけを検討する。

(1) 資料と資料加工

エリックソン＝ケーレンは,1976-77年と1980-81年の2つの時期について,VDT労働が妊娠や出産にたいして影響をあたえるかどうか,あたえるとすればいかなる影響であるかを調べる目的で,医療出生登録(Medical Birth Registry),疾病登録(Inpatient Registry),先天奇形登録(Swedish Registry of Congenital Malformation)に注目した。これらの記録には,職業についての記載がない。そのため彼らは「個人識別番号 (unique personal identification number)」を用いて,1975年と1980年のセンサスから,妊娠した女性の職業をつきとめた。さらにその女性の職業を「北欧職業分類コード(Nordic Classification Code of Occupations)」によって分類した。それは,上記の登録に記載された女性のなかからVDT労働の経験があるものを抜き出すためにであり,またVDT労働の強度の違いと妊娠・出産の関係を調べるためにである。

VDT労働の経験がある職業と認定されたのは,郵便局員,図書館職員など7つの職業であり,それらがVDT労働の強度によって,低位,中位,高位の3つのグループに分けられた。その結果,エリックソン＝ケーレンの「疫学的研究」の対象になったのは,1977-78年については5,130件の出産,1980-81年については4,895件であり,したがって全部で10,025件であった(**表1**)。

(2) 終了点(end points)と結論

エリックソン＝ケーレンは,全期間あわせて10,025件の出産について,死産,新生児死亡(生存が7日未満),出生時体重 (1,500g未満と2,500g未満),特定奇形の有無の4項目を終了点(最終確認点)として妊娠帰結を調べたが,1980-81年については,さらに自然流産が追加されている。

その分析結果は次のように要約されている。「女性3グループ (照射の程度が異なる低位,中位,高位の3グループ) のすべての妊娠帰結には,有意差はなかったが,自然流産と先天奇形については高位のグループに多いという傾向が弱いながら,うかがうことができる。しかし,この差はランダムなものである。これらのグループについて,1976-77年と1980-81年の時期別にその妊娠帰結を比較したが,この2つの時期の間に,職場ではかなりコンピュータ化が進行したにもかかわらず,定まったパターンというものは検出されなかった。」[7] すなわち,VDT労働の強度の違いが妊娠帰結に影響を与えることはないと言うのである。

Study of Work with Video Screens and Pregnancy Outcome : I. Registry Study,"[以下Ericson and Källén, I と略記] and "II. A Case-Control Study," *American Journal of Industrial Medicine*, Vol.9, 1987, pp.447-457, 459-475.

6) 渡辺真也「VDT作業の妊娠への影響」前掲『VDT労働と健康』p.164。

7) Ericson and Källén, I, p.447.

表1　エリックソン=ケーレンの「疫学的研究」の対象

		職業コード	職　業	出産数 1976-77年		出産数 1980-81年	
照射群	低位	651	郵便局員	869	1802	858	1626
		093	図書館職員など	460		535	
		203	銀行窓口業務担当	473		233	
	中位	209	保険会社事務員	1373		707	
	高位	291	コンピュータ専門職員	1389	1955	1120	2562
		293	旅行代理店事務員	211		216	
		297	社会保険事務員	355		1226	
				5130		4895	
				10025			

［出所］Ericson and Källén, I, p.448, p.451(Table I)より作成

(3) 分　析

上に述べた結論はどのようにして導かれたのであろうか。エリックソン＝ケーレンの作成した表にそくして，彼らの分析をたどることにする。彼らは4枚の表を掲載しているが，分析結果が表から明らかになる**表2**（原表は表III）と分析結果を比較的詳細に解説している**表3**（原表は表IV)について述べることにする。まず**表2**についてである。

表2では「自然流産」についてχ^2（カイ二乗）検定が行われている。照射グループ内の職業を区別した場合，χ^2は7.2(=1.1+1.0+1.7+0.0+0.9+2.2+0.3)である。このとき自由度は6であるから，χ^2分布表により，

$$\chi^2_{0.400}=6.211<7.2<\chi^2_{0.300}=7.231$$

となる。(ただし，$\chi^2_{0.400}$は，χ^2分布の上側確率が0.400となるχ^2値，すなわち $\Pr(\chi^2\geqq\chi^2_E)=0.400$が成立する$\chi^2_E$を示すものとする。)この

表2　1980-81年における職業コード別[a]の自然流産件数，妊娠件数（自然流産数＋分娩数），全妊娠件数における自然流産件数の割合[b]

	照射グループ						
	低位			中位	高位		
職業コード	651	093	203	290	291	293	297
妊娠件数	749	483	200	639	1023	212	1014
自然流産数	60	50	27	56	100	27	92
期待度数	68.9	56.9	21.0	56.1	91.0	20.3	97.8
自然流産率	8.0	10.4	13.5	8.8	9.8	12.7	9.1
χ^2	1.1	1.0	1.7	0.0	0.9	2.2	0.3
	0.7			0.0	0.5		

［原注］a 職業コードについては表Iを参照されたい。［本稿の表1参照――訳注］
b 人工流産を除く。
各グループの自然流産の期待度数は各グループの自然流産率に差がないという仮説のもとで母親の年齢について内的な標準化（internal standardization）をして計算した。
［出所］Ericson and Källén, I, p.452.

ときは，有意水準が5%でも1%でも上のχ^2値(7.2)は棄却域に落ちない。このため照射グループ内のそれぞれの職業にかんして，自然流産の発生率に有意な差がないという帰無仮説は棄却されない。もしも棄却されれば，有意な差がある，つまり照射グループ内の職業別に自然流産の発生率が異なり，したがってVDT労働の強度が自然流産に影響を与えると結論されるが，この場合はそうではないことになる。

さらにまた，同一照射グループ内における職業の違いを捨象した場合においても，その結論は変わらない。なぜならば，χ^2は1.2(=0.7+0.0+0.5)であり，自由度が2のときには，

$$\chi^2_{0.700}=0.713<1.2<\chi^2_{0.500}=1.386$$

となり，このχ^2値は有意水準が5%でも1%でも棄却域に落ちず，そのため上と同様に，照射グループごとの自然流産の発生率に有意な差がないという帰無仮説は棄却されないからである。

この結論は**表3**においても変わらない[8]。**表3**は，表の形式が**表2**と次の点で異なっている。

第1に，**表3**では，照射グループが低位，中位，高位の3つにまとめられており，職業コード別になってはいない。

第2に観測度数と期待度数の比率(O/E)——これをO/E比という——が，その95%信頼区間とともにあたえられている。このO/E比が1に近いほど，観測度数と期待度数との乖離が小さいことを意味する。このために，O/E比が1に近い値をとるときには，期待度数を計算するときに使った理論，すなわち，たとえば照射グループごとの自然流産の発生率にかんして有意な差がないという帰無仮説にもとづく計算の根拠の正当性が，観測度数によって与えられたと考えるのである。自然流産にかんして，エリックソン＝ケーレンは「表IV［本稿の**表3**のこと——引用者］が示すように，各グループの自然流産率の間には統計的に有意な差はない。」[9]と結論する論拠は以上のとおりである。

表3にある自然流産以外の妊娠帰結についてのエリックソン＝ケーレンの分析も，すべてその根拠を上に述べたような，観測度数と期待度数とのO/E比が1に近いかどうかを基準にして行われている。たとえば，周産期死亡（死産＋生後7日未満生存）について見れば，その観測度数は「すべてのグループ，すべての時期について，期待度数とほぼ同じか，それを下回っている。中位グループのO/E比は大きい(1.5)が，事例の数で言えば，わずかに2.5件だけ上回っているにすぎず，有意ではない。」[10]と結論されているのである。その他の妊娠帰結について明白な言及はないが，それは，上の分析手法をそのまま適用すれば，同様の結論が得られるものとして，叙述を省略したのかもしれない。

さらにエリックソン＝ケーレンは，1976-77年と1980-81年とでは，最近の方が事務労働のOA化は進展しているとしたうえで，**表3**を用いて，2つの期間をそれぞれの妊娠帰結ごとに，O/E比によって比較している。すなわちO/E比が(1)周産期死亡については，低位と高位のグループでは減少しているが，中位グループでは増大している，(2)奇形については3グループで増大している，(3)2,500g未満の軽量体重児については，3グループで減少し，(4)1,500g未満の軽量体重児については3グループで増大した。このようにして，エリックソン＝ケーレンはコンピュータ化の進展と妊娠帰結との関係に一般的な傾向がないことを指摘する。

なお，奇形の項目にかんして，**表3**は高位グループのO/E比が0.5(1976-77年)から1.2(1980-81年)に増大したことを示している。このことについて，これは「形式的な統計的有意性に達してはいたが，しかしながらそれは，1976-77年におけるO/E比が(おそらくランダムな)低率であったせいであろう。」[11]と述べている。しかし，この「形式的な統計的有意性」の判定には有意水準が用いられていない。そのためこの「統計

8) *Ibid*.
9) *Ibid*.
10) *Ibid*.
11) *Ibid*.

表 3　照射グループ別データの要約[a]

	照射グループ					
	低　位		中　位		高　位	
期　間	1976-77	1980-81	1976-77	1980-81	1976-77	1980-81
周産期死亡						
O(E[b])	17(18.2)	9(13.0)	15(14.0)	8(5.5)	19(19.9)	16(20.1)
O/E	0.9	0.7	1.1	1.5	1.0	0.8
信頼区間	0.5-1.5	0.3-1.3	0.6-1.8	0.6-2.9	0.6-1.5	0.5-1.3
意味ある奇形						
O(E[b])	22(28.9)	26(27.5)	14(21.9)	11(11.7)	14(29.2)	53(42.8)
O/E	0.8	0.9	0.6	0.9	0.5	1.2
信頼区間	0.5-1.2	0.6-1.4	0.3-1.1	0.5-1.7	0.3-0.8	0.9-1.6
軽量体重(<1500g)						
O(E[b])	9(11.5)	12(10.4)	14(9.1)	10(4.5)	10(12.9)	21(16.4)
O/E	0.8	1.5	1.5	2.2	0.8	1.3
信頼区間	0.4-1.5	0.6-2.0	0.8-2.6	1.1-4.4	0.4-1.5	0.8-2.0
軽量体重(<2500g)						
O(E[b])	69(74.9)	58(70.4)	86(59.0)	38(30.6)	99(84.1)	109(110.9)
O/E	0.9	0.8	1.5	1.2	1.2	1.0
信頼区間	0.7-1.2	0.6-1.1	1.2-1.8	0.9-1.7	1.0-1.4	0.8-1.2
自然流産						
O(E[c])	-	137(146.8)	-	56(56.1)	-	209(209.1)
O/E		0.9		1.0		1.1
信頼区間		0.8-1.1		0.8-1.3		0.9-1.2

[原注] a 高位グループの女性のなかには1976-77 年にもＶＤＴ労働をしていたと考えられる人もいる。
b 期待度数は，母親の年齢，パリティ，分娩施設について標準化した全国数字から計算した。
c 期待度数は，全グループのなかの女性の年齢について内的な標準化をして，計算した。
O は観測度数，Ｅは期待度数，信頼区間は厳密なポアソン分布から計算される95%信頼区間。
[出所] Ericson and Källén, I , p.453.

的有意性」は，通常の統計的仮説検定における統計的有意性という用語の用法とは異なっている。後述するように，検定における統計的有意性は，標本特性値の確率分布にもとづいて与えられる有意水準と密接不離の関係にある概念だからである。

それだけではない。「形式的な統計的有意性」の意味が明確ではないのである。エリックソン＝ケーレンはこの用語を，1980-81年のO/E比1.2が，そのものとして高率であり，「形式的な統計的有意性」を示すという意味で用いているのであろうか。あるいは1976-77年から1980-81年にかけての，O/E比の増加割合が2.4(=1.2/0.5)となり，この増加割合が他のどの項目よりも，特別に大きかったことをもって，「形式的な統計的有意性」を示すという意味で用いているのであろうか。このことは論文からは伺い知ることはできない。しかし，この「形式的な統計的有意性」がランダムネスのせいにされて，コンピュータ化の進展と妊娠帰結との関係を示す証拠と考えられていないことは確かである。こうして，エリックソン＝ケーレンは２つの期間について比較した結果，「いろいろな傾向があるにはあるが，1976-77年と1980-81年の２つの時期の間で事務室のコンピュータ化があったと思われるにもかかわらず，２つの期間の間にははっきりし

た変化(definite change)はないのである。」[12]と結論する。

2．統計的仮説検定の有効性
――エリックソン＝ケーレン論文の検討――

(1) 統計的有意性の概念

エリックソン＝ケーレンの用いる「統計的有意性」という言葉の意味をまず，問題にしたい。彼らは，ＶＤＴ労働の強度別の自然流産についてχ^2検定を適用している。そのときには，統計的仮説検定で用いられる場合と同じ意味で「統計的有意性」を使っている(**表2**参照)。そして，ＶＤＴ労働の強度や時期と妊娠帰結との関連を「統計的に」否定している。

しかし「統計的有意性」という言葉は，周産期死亡，奇形出産，軽量体重児の出産などの妊娠帰結にかんして，ＶＤＴ労働の強度別，観察期間別の分析を「観測度数／期待度数」，すなわちO/E比によって行うときにも用いられている。そこでは，O/E比が１であるときには，理論的に何らかの仮定をおいて割り出した期待度数と観測度数とが一致しており，したがってこのときには期待度数を与えた論拠としてのその仮説の正当性が経験的に確かめられたと考えるのである。そしてそうでないときには，有意な結果がえられたとして仮説を棄却する。

O/E比によるエリックソン＝ケーレンの分析は，高位グループの奇形出産にかんしてだけ，1976-77年と1980-81年においてO/E比が0.5から1.2に変化したことを捕らえる。そして，この点にかんしては「統計的有意性」に達しているとする。(それ以外の妊娠帰結については，そうではないと言う。)すでに述べたように，このとき「統計的有意性」を判定する根拠が，他に類を見ないO/E比の増加割合にあるのかどうか，あるいは他の論拠によるのかどうかは不明である。そのような問題はあるが，それをおくとしても，このようなときに「統計的有意性」の概念を使うことができるのであろうか。

一般に，「統計的有意性」の概念は，標本特性値の確率分布を前提とし，その確率分布の上方（または下方，あるいはその両方）の累積確率（通常は5%かあるいは1%）を有意水準として，その有意水準を規定する標本特性値の領域に，現実の標本特性値が落ちるときに用いられる概念である。したがって，仮説検定論の数理においては，統計的に有意であるかどうかは，5%とおくか1%とおくか，その根拠は何かという論議はあるにせよ，ひとたび有意水準をある値に決めてしまえば，あとは自動的に仮説が検定されるのである。

このように有意性の判定には，有意水準が不可欠である。しかし，エリックソン＝ケーレンのO/E比分析では，有意水準が使われておらず，その意味ではけっして，数理統計学的な意味で統計的仮説検定を行ったとは言いがたい。したがってそのような場面で「統計的有意性」の概念を用い，しかも分析の結論部分ではあたかも統計的仮説検定を適用して，有意な結果や有意でない結果がえられたかのような印象を与えるのは，人を惑わせやすい叙述の仕方であると言わざるをえない。

(2) 有意な結果の解釈

ここでエリックソン＝ケーレンにゆずって，「統計的有意性」の概念についての彼らの用法がゆるされるにしても，まだ問題は残る。それは，彼らが「統計的に有意」なO/E比であると判断した1976-77年と1980-81年における高位グループの奇形出産についてである。すでに引用したように，彼らはこれについては「形式的な統計的有意性に達してはいた」と言う。しかし，1976-77年のO/E比が偶然にも小さすぎたというランダムな結果であって，仮説を棄却させる実質的な根拠にはならないと解釈している。

統計的仮説検定の場合では標本特性値が棄却域に落ちたとき，帰無仮説は棄却されるのがふ

12) *Ibid*.

つうである。たとえば，コインを100回投げ上げた結果，100回とも表が出たとすれば，そのときにはこのコインは表と裏が半分ずつでるような物理的性質をもっているとか，あるいはそのような物理的性質の実現する条件があったという仮説が棄却される。そのような物理的性質をもつコインを投げて，連続して100回とも表が出るのは，$0.5^{100}=7.89\times10^{-31}$という非常に小さい確率であり，そのようなまれな出来事が生じたと考えるよりも，使ったコインの物理的性質が考えられたようなものではなかったとか，あるいは投げ方が公正でなかったとして，要するにこの場合は仮説が実現する条件はなかったと判断して，その仮説を棄却するのである。このような推論がふつうに行われるが，もちろんきわめてまれな出来事が奇跡的に生じたと考えて，帰無仮説を棄却しないこともできる。有意水準をどう設定するか(5%にするか1%にするか)によって同一の標本特性値が帰無仮説を棄却する証拠になることもあるし，棄却しない証拠にもなるという判断の不確実性は，統計的仮説検定につきまとう問題のひとつである。しかし，有意水準をある値に設定していても，上のコイン投げの例から明らかなように，判断結果としては帰無仮説の棄却でもよいし，そうでなくともよい。すなわちエリックソン＝ケーレンのように，「統計的に有意」な結果が得られたときにも，それをランダムネスのせいにして仮説を棄却しないことも可能である。しかし，なぜ偶然のせいで「統計的に有意」な結果がたんに「形式的」なものであり，仮説を棄却する証拠になりえないかの説明は必要であろう。この点についてエリックソン＝ケーレンは何も述べていない。高位グループの奇形出産について1976-77年のO/E比が偶然にも小さすぎたことがランダムな結果であると解釈する論拠は，明示されていない。

(3) 期待度数

χ^2検定では，χ^2値$\Sigma\{(O-E)^2/E\}$［ここにOは観測度数，Eは期待度数］が，分割表で規定される自由度のχ^2分布に従うという数学的性質を利用する。したがってχ^2値を計算するには期待度数を計算しなければならない。検定の対象になる特性（標識）にかんして特性が独立であるか，分布が同一であるという帰無仮説のもとで，期待度数を計算するには，さしあたり2とおりの仕方がある。

第1は，理論的に，あるいは別のデータにもとづいて帰無仮説を設定し，それが真であるとして，期待度数を計算する仕方である。たとえば10回の投銭を1組の試行として，その試行を10回行う。他方で理論的にか，あるいはそれ以前の試行経験にもとづいて，表の確率(p)が0.5であるとして，帰無仮説を$H_0: p=0.5$と設定する。期待度数はこの帰無仮説にもとづいて計算し，χ^2検定を行うのである。この例は分割表が一元的な場合であるが，二元的なときもこの方式で期待度数を計算することができる。エリックソン＝ケーレンはこの方式で期待度数をもとめた。

第2は，分割表にもとづいて期待度数を計算する仕方である。一般に，分割表が**表4**であたえられるとき，帰無仮説H_0のもとで期待度数(E_{ij})は

表4　$k\times l$ 分割表

グループ＼特性A	1	2	・・・	l	計
1	o_{11}	o_{12}		o_{1l}	n_1
2	o_{21}	o_{22}		o_{2l}	n_2
⋮					
k	o_{k1}	o_{k2}		o_{kl}	n_k
計	m_1	m_2		m_l	N

［出所］古川俊之監修・丹後俊郎著『医学への統計学』朝倉書店，1983年 p.113。

$$E_{ij}=\frac{n_i m_j}{N}$$

で与えられる[13]。

分割表が与えられたときは，このように期待度数は2とおりの仕方で計算できる。そこでエ

13) 古川俊之監修・丹後俊郎著『医学への統計学』朝倉書店，1983年　p.113。

リックソン＝ケーレンのデータから，われわれは独自に第2の方法で期待度数を計算することにしよう。**表3**から「周産期死亡」の観測度数を拾うと，次のようになる(**表5**)。

表5　周産期死亡の観測度数

	低位	中位	高位	合計
1976-77	17	15	19	51
1980-81	9	8	16	33
合計	26	23	35	84

1976-77年における高位グループについて上に掲げた期待度数の式にもとづき，「周産期死亡」の期待度数を計算すると，15.8(=26×51/84)になる。このような手続きで「奇形」，「軽量体重児」について期待度数を計算して，まとめたのが**表6**である。

ところでエリックソン＝ケーレンのデータの場合，期待度数は，彼らがもとめた方式と上に述べた方式以外に計算することはできないのであろうか。**表1　エリックソン＝ケーレンの「疫学的研究」の対象**から照射グループ別，期間別の出産数を拾うと，**表7**のようになる。

表7　照射グループ別，期間別の出産数

	低位	中位	高位	計
1976-77	1802	1373	1955	5130
1980-81	1626	707	2562	4895
計	3428	2080	4517	10025

この**表7**と**表5**を組み合わせてみると，全グループと全期間をつうじた周産期死亡率は［全周産期死亡数÷全出産数］により84÷10025であることが分かる。ここでもしも全照射グループと全期間をつうじた周産期死亡率が一定であるとすれば（数理統計学的に言えば，周産期死亡が全照射グループと全期間をつうじて同一の分布にしたがうとすれば，あるいは周産期死亡が照射グループと期間について独立でないとすれば)，全照射グループと全期間をつうじた周産期死亡率は0.0084(=84/10025)であると推定される。この周産期死亡率を1976-77年の低位照射グループの出産数(1802)に乗ずれば，周産期死亡率が一定であるという仮説のもとにおいて，1976-77年の低位照射グループの周産期死亡の期待度数が計算できる。このとき，この期待度数は15.1である。そして O/E比は［1976-77年の低位照射

表6　観測度数(O)からもとめた期待度数(E)［(　) 内数字］

	照射グループ					
	低　位		中　位		高　位	
期　間	1976-77	1980-81	1976-77	1980-81	1976-77	1980-81
周産期死亡						
O(E)	17(15.8)	9(10.2)	15(14.0)	8(9.0)	19(21.3)	16(13.8)
O/E	1.1	0.9	1.1	0.9	0.9	1.2
意味ある奇形						
O(E)	22(17.1)	26(30.9)	14(8.9)	11(16.1)	14(23.9)	53(43.1)
O/E	1.3*	0.8	1.6*	0.7	0.6	1.2
軽量体重(<1500g)						
O(E)	9(9.1)	12(11.9)	14(10.4)	10(13.6)	10(13.5)	21(17.5)
O/E	1.0	1.0	1.4	0.7*	0.8	1.2
軽量体重(<2500g)						
O(E)	69(70.3)	58(56.7)	86(68.6)	38(55.4)	99(115.1)	109(92.9)
O/E	1.0	1.0	1.3	0.7*	0.9	1.2

［注］　1) ＊はエリックソン＝ケーレンの計算した95%信頼区間に落ちないO/E比。
　　2) 自然流産についてはこの方法では期待度数を計算することができない。

表 8　発生比率にもとづく期待度数［(　) 内数字］

	照射グループ					
	低　位		中　位		高　位	
期　間	1976-77	1980-81	1976-77	1980-81	1976-77	1980-81
周産期死亡						
O(E)	17(15.1)	9(13.6)	15(11.5)	8(5.9)	19(16.4)	16(21.5)
O/E	1.1	0.7	1.3	1.4	1.2	0.7
意味ある奇形						
O(E)	22(25.2)	26(22.7)	14(19.2)	11(9.9)	14(27.3)	53(35.8)
O/E	0.9	1.1	0.7	1.1	0.5	1.5
軽体量重(<1500g)						
O(E)	9(13.7)	12(12.3)	14(10.4)	10(5.4)	10(14.8)	21(19.4)
O/E	0.7	1.0	1.3	1.9	0.7	1.1
軽体量重(<2500g)						
O(E)	69(82.5)	58(74.5)	86(62.9)	38(32.4)	99(89.5)	109(117.3)
O/E	0.8	0.8	1.4	1.2	1.1	0.9
自然流産						
O(E)		137(136.9)		56(59.5)		219(215.6)
O/E		1.0		0.9		1.0

グループの周産期死亡の観測度数÷期待度数］により1.1(=17/15.1)になる。このような計算手続きによって求めた期待度数と O/E比を一覧にしたものが**表 8** である。

χ^2検定の適用においても，O/E比による「統計的有意性」の判定においても，その考察のもっとも重要なかなめの位置にあるのは，期待度数の計算である。エリックソン＝ケーレンは，周産期死亡率，奇形発生率，妊婦の年齢，出産経歴などにかんする全国数字をもとにして，期待度数を計算したと述べている[14]。また分娩施設(delivery unit)が奇形診断率に影響を与えるから，これを「標準化」して期待度数を計算したとも言っている。しかし彼らはそう述べるだけで，その論拠を示していない。このため彼らの計算する期待度数は正しいとも，間違っているとも判断することができない。

エリックソン=ケーレンは，O/E比の95%信頼区間(**表 3** 参照)が自分たちの期待度数を正当化すると考えているのかもしれない。区間推定については，Pr {0.5 < O/E <1.5} =0.95の意味をめぐって，かつてネイマンとR．A．フィッシャーの間で論争があり，それはいまでも係争中の論点であるが，ここではそのことに触れない。そしてかりにゆずって，Pr {0.5 < O/E <1.5} =0.95という表記が許されるとしても，そのことは，O/E比が0.5から1.5までの間にある数値であれば，どのような値のO/E比であっても，数理統計学的にはこの式と矛盾しないということを意味するだけである。Pr {0.5 < O/E <1.5} = 0.95は，エリックソン＝ケーレンの期待度数から計算された，たとえば0.9というO/E比(1976-77年の低位グループの周産期死亡――**表 3**)がその他の仕方で計算されたO/E比(「発生比率」にもとづくと 1.1――**表 8**)よりも良好であるということの根拠にはなりえないのである。われわれが「発生比率」にもとづいて独自に計算した期待度数は，すべてエリックソン＝ケーレンの95%信頼区間をクリアしている。われわれの期待度数は，彼らの土俵の上においてさえ，数学

14) Ericson and Källén, *op.cit.*, p.450.

表 9　妊娠帰結とχ^2
——期間別・ＶＤＴ労働強度別——

	周産期死亡	奇形	軽量体重		自然流産
			1,500g未満	2,500g未満	
χ^2	5.4	17.2	8.6	16.9	0.3
α(%)					
0.5		10.597		10.597	
1.0			9.210		
2.5			7.378		
5.0	5.991				
10.0	4.605				
85.0					0.325
90.0					0.211

［注］αは有意水準(%)，その横の数値は棄却域の臨界値を示す。
いずれの妊娠帰結についても自由度は２。

的には何の遜色もないのである。Pr{0.5<O/E<1.5} =0.95という確率的言明は，O/E比0.9が1.1よりも好ましいと判断する根拠とはなりえず，そう主張するには別の，実質的な論拠が必要である。

以上，エリックソン＝ケーレンがχ^2値を計算したり，あるいはO/E比を計算したりするときに求めた期待度数のほかにも，期待度数をもとめるうることを述べ，彼らの期待度数は正当性を主張しうるものとしては，その根拠が脆弱であることを指摘した。では，別の期待度数をもちいると，どのような結論がえられるであろうか。次に節を改めて，このことについて述べる。

(4) χ^2検定

上でみたように期待度数はいくとおりも計算することができる。そのうちのどの期待度数を使ってもχ^2検定を適用することができる。そこで試みに，「発生比率」にもとづいて計算された期待度数(**表 8**)を用いて，χ^2検定をすることにする。「観測度数」からもとめた期待度数(**表 6**)を使わず，**表 8**によって検定をするのは，エリックソン＝ケーレンの基準(区間推定)からみて，**表 8**の方が**表 6**より遜色がないと考えられるからである。この場合，帰無仮説は妊娠帰結が期間とＶＤＴ労働の強度に関係なく，同一の分布をもつ（あるいは妊娠帰結が期間やＶＤＴ労働の強度と独立ではない）というものである。(期待度数はこれが真であることを前提して計算された。)

期間別・ＶＤＴ労働強度別の周産期死亡についての観測度数と期待度数（**表 8**）にもとづいて，χ^2を計算すると

$$\chi^2 = (17-15.1)^2/15.1+(15-11.5)^2/11.5 +(19-16.4)^2/16.4+(9-13.6)^2/13.6 +(8-5.9)^2/5.9+(16-21.5)^2/21.5 = 5.4$$

このような計算手続きでそれぞれの妊娠帰結についてχ^2を計算すると，**表 9**のようになる。

表 9の読み方は次のとおりである。たとえば周産期死亡についてχ^2値は5.4と計算されるが，表より有意水準５%で有意となるには，χ^2値が5.991以上でなければならず，したがってこのときは有意な結果ではない。しかし，有意水準を10%にすれば，この結果が有意となり，帰無仮説は棄却される。すなわち周産期死亡については，期間をつうじて，またＶＤＴ労働の強度のいかんを問わず，分布が同一であるという帰無仮説が棄却され，周産期死亡にたいしては，期間の相違が影響を与えるだけでなく，ＶＤＴ労働の強度の違いも影響を与えていることになる。

Ｒ．Ａ．フィッシャー以来，有意性検定で5%

が有意水準として慣習的に使用されていることにかんして，その根拠が薄弱であることが，これまでにも論議されている[15]。そのような問題はありつつも，ここであえて慣習的に「聖なる0.05」を使ったとすれば，奇形出産と軽量体重児出産（1,500g未満と2,500g未満）については，帰無仮説が棄却され，エリックソン＝ケーレンとは逆の結論になるのである。しかし棄却域に落ちるχ^2値を得たとしても，それをランダムネスのせいにして，棄却しないことも可能なのである。さらにまた有意水準を小さい値に設定して，帰無仮説を棄却しないことも可能なのである。こうして有意性検定の結果はきわめて不確実なものとなってしまう。

むすび

医学の領域に統計的推論を応用するようになったのは，グリーンウッド＝ユールが腸チフスとコレラの免疫製剤の有効性をχ^2検定で調べようとしたことに端を発する[16]。それ以降，統計的推論が医学の領域で利用されるようになった。ここで取り上げたエリックソン＝ケーレンの論文もそのひとつとして位置づけることができる。

エリックソン＝ケーレン論文はＶＤＴ労働が妊娠帰結に影響を与えないことを統計的に検証するものであった。しかし，ここでの検討の結果，数理統計学の慣習的な数理に依拠する限り，その結論とは逆の結論が得られた。この結論も，その判断の確率的性質により，必ずしも確言的なものではない。この不確実な推論の不確実性は，あらためて統計的推論の有効性という問題を提起するのである。

15) Skipper, Guenter and Nass, "The Sacredness of .05 : A Note on the Uses of Statistical Levels of Significance in Social Science," *The American Sociologist*, 2 (Feb. 1957), and also in Morrison and Henkel (eds.), *The Significance Test Controversy — A Reader*, Chicago : Aldine Pub. Co., 1970（内海庫一郎他訳『統計的検定は有効か』梓出版社，1980年　第8章）

16) Greenwood, M. and Yule, U., "The Statistics of Anti-typhoid and Anti-cholera Inoculations and the Interpretation of such Statistics in General," *Proc. of Royal Society of Medicine*, Vol. 8, 1915.