

BASIC 3

1991

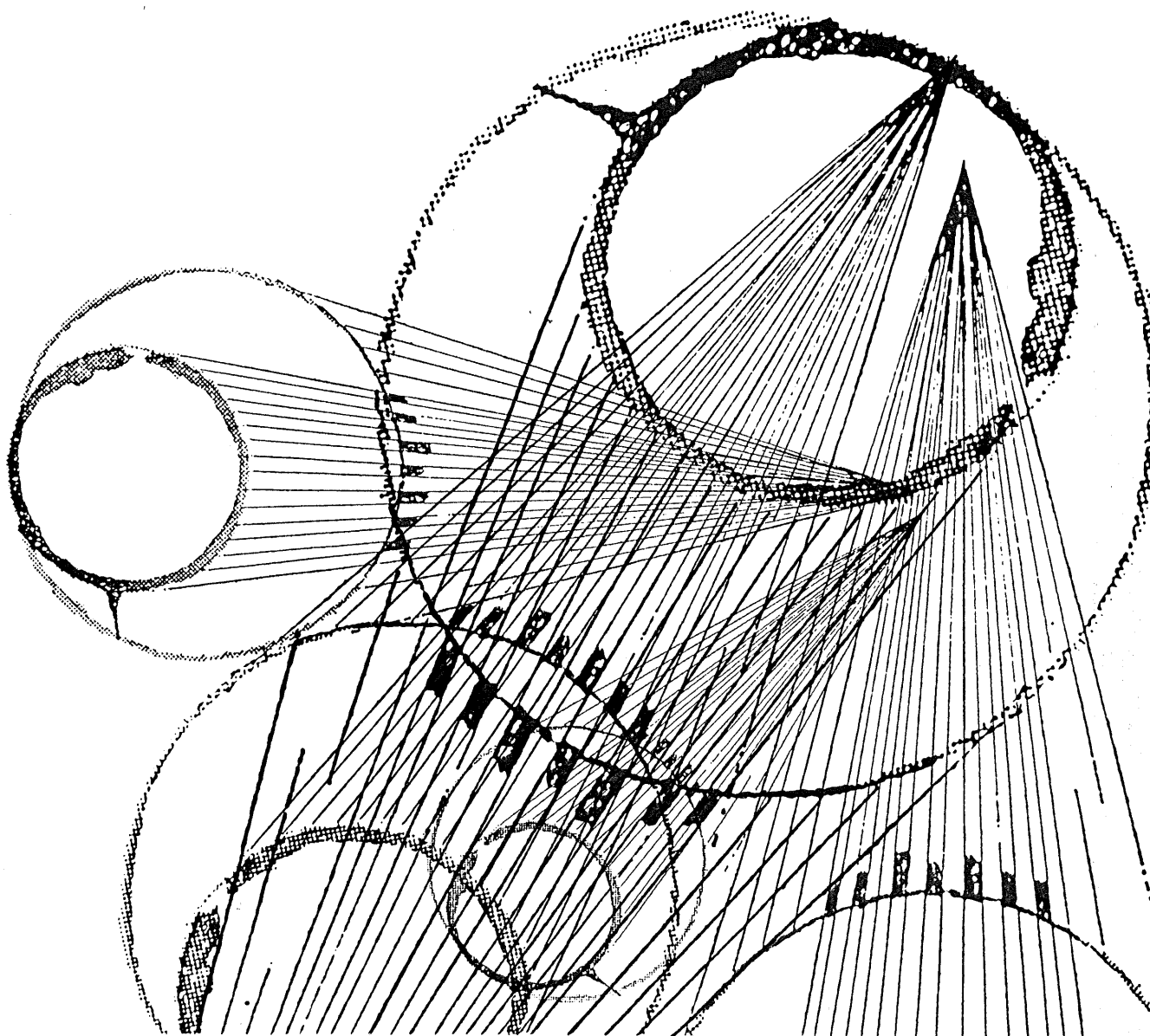
数学

社会科学部
学芸学部
数学科
の
専攻科

■特集／統計学は有効か？

- | | |
|---------------|-------|
| ■ 統計学の新しい考え方 | 鈴木義一郎 |
| ■ 正規分布とは | 安藤 洋美 |
| ■ 地理学と多変量解析 | 村山 祐司 |
| ■ 医学分野における統計学 | 高木 廣文 |

楽しく学べる OR 教室／意思決定理論 田畑吉雄



統計学は有効か

医学分野における統計学

—とくに疫学研究で用いられる統計学について—



高木 廣文 (統計数理研究所)

1. はじめに

ここでは、医学分野で用いられている統計学の方法について考えてみたい。しかし、医学分野といってもその範囲は極めて広い。今回は、公衆衛生学的な研究でよく用いられている「疫学」の方法を中心に解説したい。もっとも、「疫学」といっても馴染みのない読者が大部分であろうと思われるので、まず簡単に疫学とは何かを述べることにしよう。

疫学 epidemiology は、一昔前までは「流行病に関する学問」とされており、人口集団で流行するコレラなどの感染症の感染経路やその原因などを研究するものとされていた。ところが、現在では死亡の原因は癌や心疾患などに移ってきたため、感染症は先進諸国では、AIDS 以外はあまり問題ではなくなってきた。このため、疫学も癌などの非感染性疾患の原因を探索するための学問として考えられているといえる。

疫学研究は、基本的に非実験的な場合が多く、特定の人口集団についての調査に基づくデータを、統計的に解析することになる。したがって、用いられる解析方法の多くは、他の研究分野と共通するデータ解析的なものが多い。例えば、平均値や分散、標準偏差を求める、ヒストグラムを描く、幹葉表示や箱ヒゲ図を描くなどである。

しかし、特定の疾患などの発生やそれによる死亡が、どのような要因によるのかを探索、検討するために、「リスク risk」を評価するための各種の有用な方法がある。ここで、リスクとは特定の人口集団でのある疾患等の発生確率を示す用語である。

今回は、他分野と共通して用いられるような解析方法は、簡単に紹介するとどめ、要因のリス

ク評価に関する統計学的方法を中心に、独特な調査方法とともに説明し、医学分野での統計学の果たす役割の一端を覗いてみたい。

2. リスクに関する基本的な解析方法

喫煙は肺癌の原因であるなどと言われるが、実際にはどのような方法でそのようなことが確認されるのだろうか。ある疾患の発生の有無と特定の要因の有無を調べるには、有用ないくつかの調査方法と分析方法がある。

2.1 リスク比について

調査の形態として、要因をもつグループ(曝露群)ともたないグループ(非曝露群)の2グループを長期間追跡し、どちらが興味の対象となっている疾患の発生の度合を比較する方法がある。このような調査は、「曝露—非曝露調査」と呼ばれ、時間的な流れからは、「前向き調査」と呼ばれる。

実際に得られるデータは、表1のようにまとめられる。曝露群は n 人中 x が疾患の発生が見られたので、そのリスクは $p_1 = x/n$ のようになり、同様に非曝露群でのリスクは $p_2 = y/m$ となる。2つのリスクの比が「リスク比 risk ratio」と呼ばれるものである。実際には、 p_1 と p_2 とは、各群の母集団での発生割合 π_1 と π_2 の推定値となっており、母リスク比 $\phi = \pi_1/\pi_2$ の推定値は、 $\hat{\phi} = p_1/p_2$ により推定される。

表1. 曝露—非曝露調査

要 因	疾 患		計
	有	無	
曝 露	x	$n-x$	n
非曝露	y	$m-y$	m
計	l	$N-l$	N

$$l = x + y, \quad N = n + m$$

これらの計算は極めて簡単なものであるが、ある要因を有することで、特定の疾患の発生や死亡のリスクが何倍になるかを推定できるので、有用なものといえる。

実際の手順として、帰無仮説「母リスク比 $\phi=1$ 」の検定を行う必要がある。このためには、通常の四分表のカイ 2 乗検定を行うか、フィッシャーの直接確率（正確な確率）を求めて行えばよい。帰無仮説が棄却された場合、リスク比の信頼区間を求める必要がある。

リスク比 ϕ の分布は歪んでいるので、対数変換を行なった方が、正規分布への近似がよいことがわかっている。 $\log \hat{\phi}$ の期待値は $\log \phi$ となり、その分散は、

$$V(\log \hat{\phi}) = \frac{1}{x} + \frac{1}{n} + \frac{1}{y} + \frac{1}{m}$$

により近似できるので、リスク比の $100(1-\alpha)\%$ 信頼区間は、

$$\hat{\phi} \exp[\pm z_{\alpha/2} \sqrt{V(\log \hat{\phi})}]$$

で与えられる。ここで、 $z_{\alpha/2}$ は標準正規分布の上側 $\alpha/2$ 点である。

表 2. 糖尿病の治療薬と生死

グループ	死亡	生存	計
フェンフォルミン	34	170	204
プラセボ	7	57	64

The University Group Diabetes Program (1975). Diabetes 24, Suppl. 1 の表 5 より一部引用

表 2 は、糖尿病患者にフェンフォルミンと呼ばれる経口治療薬を投与した群とプラセボ（偽薬）を投与した群を長期間追跡し、その生死を調査したものである。この表からリスク比を求めると、 $\hat{\phi}=1.52$ となり、 $V(\log \hat{\phi})=0.152$ となる。したがって、リスク比の 95% 信頼区間は、 $0.71 < \phi < 3.27$ となる。この結果は、フェンフォルミン群の方がプラセボ群に比べ、死亡のリスクがやや高いものの、信頼区間に 1 を含むため、治療薬が危険であるとは断定できないことを示している。ただし、年齢などの対象者の背景要因を一致させた分析ではないので、結論は保留する方が無難であろう。

2.2 オッズ比について

曝露一非曝露調査は時間がかかり、かつ疾患の発生や死亡が比較的高頻度に起きなければ、その

費用に見合った成果は期待できない。稀にしか発生しない疾患などの原因や、ある要因の影響を調べるために、すでに病気にかかっている患者と健康者を比較することが、よく行われる。これは「患者－対照調査 case-control study」と呼ばれる調査研究方法である。このような調査では、表 3 のように患者数 n と対照者数 m があらかじめ決められている。表 1 の場合と比べると縦と横が入れ替わっていることがわかるだろう。したがって、要因の有無ごとに、患者の割合を求めることには意味がなく、そのままではリスク比を求めることはできない。しかし、「オッズ比 odds ratio」と呼ばれる統計量が、リスク比のよい近似値となることがわかっている。

表 3. 患者－対照調査

群	要因		計
	有	無	
患者	x	$n-x$	n
対照	y	$m-y$	m
計	t	$N-t$	N

$$t=x+y, N=n+m$$

患者群と対照群の母集団での要因を有する割合をそれぞれ π_1 , π_2 とすると、母オッズ比 ϕ は、

$$\phi = \frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}$$

と定義される。母オッズ比の推定値には、

$$\hat{\phi} = \frac{x(m-y)}{y(n-x)}$$

を求めればよい。また、信頼区間の推定は、リスク比の場合と同様に行える。すなわち、オッズ比の対数、 $\log \hat{\phi}$ を考えればよい。なお、 $\log \hat{\phi}$ は「対数オッズ比」と呼ばれる。対数オッズ比の期待値は母対数オッズ比に一致し、分散は、

$$V(\log \hat{\phi}) = \frac{1}{x} + \frac{1}{n-x} + \frac{1}{y} + \frac{1}{m-y}$$

となる。したがって、真のオッズ比の $100(1-\alpha)\%$ 信頼区間は、

$$\hat{\phi} \exp[\pm z_{\alpha/2} \sqrt{V(\log \hat{\phi})}]$$

により求められる。

なお、帰無仮説「 $\phi=1$ 」の検定は、リスク比の場合と同様に行えばよい。

表 2 のデータについて、オッズ比を求めてみよう。簡単な計算から、 $\hat{\phi}=1.63$ となり、 $V(\log \hat{\phi})$

$=0.196$ と求められる。先に見たようにリスク比は約1.5であったので、オッズ比はそれと大差ない数値を与えている。ただし、分散はやや大きくなっているため、95%信頼区間は $0.68 < \psi < 3.88$ とやや範囲も広がっている。

このようにオッズ比を用いることにより、発生の稀な疾患に関する各種要因のリスクも量的に把握することが可能となる。しかし、実際には特定の要因についての推論を行う場合、結果に影響を与える可能性のある変数をコントロールする必要がある。例えば、性や年齢の相違は各種疾患の死亡率の違いをもたらすことは、ほとんど自明といってもよい。したがって、この問題については適当な解決策が必要となる。

2.3 層別四分表の共通オッズ比

疫学研究において、特定の要因のリスクを評価する場合、疾病の発生に関係する第3の要因が存在すれば、研究下にある要因の評価は歪められる恐れがある。そのような要因は「交絡要因 confounding factor」と呼ばれる。交絡要因として、性、年齢、職業などが考えられるが、その影響を避けるためには、クロス表をそれらの要因ごとに別々に作成すればよいだろう。例えば、性別・年齢階級別に4分表を作成するなどがそうである。このような手順は「層化 stratification」と呼ばれている。最も極端な場合は、患者1人に対し、調べようとする要因以外が全て同一の対照者を1人選定し、分析を行う場合であろう。このような対応付けは「マッチング matching」と呼ばれるもので、すでに説明した患者一対照研究では、よく用いられる手法の一つである。ただし、人間に付随する属性は多数あるので、実際には性、年齢、居住地などをマッチするだけである。当然ではあるが、必要な場合に可能ならば他の要因もマッチすることもある。

層化を行った場合、表2のような四分表が複数個できることになる。個々の層ごとにオッズ比を求めるのではなく、各層のオッズ比が全て等しいものと仮定して、「共通オッズ比 common odds ratio」を推定するのが一般的な方法である。

推定のために通常よく用いられるのが「Mantel-Haenszel 推定量 (Mantel and Haenszel, 1959)」と呼ばれるものである。また、帰無仮説「共通オッズ比 = 1」を検定する方法も、

Mantel と Haenszel により与えられている。共通オッズ比の信頼区間の構成方法に関しては、Mehta ら (1985)、Robins ら (1986)、Sato (1990)、Takagi (1990) により与えられている。これらの詳細については、紙面の都合上省略させて頂くが、特定の疾患と要因間の関連を考える上で、優れて有用な情報を与えてくれる。

3. ロジスティックモデル logistic model

ここまで、疾病の発生に関係する要因が質的な場合について述べてきたが、要因が量的な場合について考えてみよう。まず、要因が1つの場合を考えてみよう。

ある疾患の発生、死亡と変数 X の値 x に関係がある場合、どのようなモデルを考えればよいだろうか。各種のモデルが考えられているが、ここでは、疾患の発生率 p の対数オッズと変数 X に線形な関係があるものと仮定する。すなわち、

$$\log \left[\frac{p}{1-p} \right] = \beta x + \alpha$$

と仮定する。対数オッズは、通常「ロジット logit」と呼ばれる。上式を書き直すと、

$$p = [1 + \exp \{-(\beta x + \alpha)\}]^{-1}$$

と表わすことができる。これは、シグモイド状の量-反応関係を解析するのによく用いられるロジスティック曲線である。 β は x にかかる重みであり、回帰係数と呼ばれ、 α は定数であり、 $x=0$ 、すなわち曝露がない場合の疾患の発生率に関係している。

変数 X が質的な場合、とくに要因の有無を示すような場合を考えてみよう。要因があれば $x=1$ 、なければ $x=0$ とおくと

$$\text{曝露群: } \text{logit}(p_1) = \beta + \alpha$$

$$\text{非曝露群: } \text{logit}(p_2) = \alpha$$

となる。2群のロジットの差をとると、

$$\text{logit}(p_1) - \text{logit}(p_2) = \log \left[\frac{p_1(1-p_2)}{p_2(1-p_1)} \right] = \beta$$

となる。これは対数オッズ比 $\log \psi$ が β であることを示している。すなわち、 $\psi = \exp(\beta)$ である。

したがって、ロジスティックモデルは、曝露因子が質的な場合のリスク評価の方法を、量的な要因に対する拡張したものと考えてよいだろう。

一般に、変数が $\mathbf{x} = (1, x_1, x_2, \dots, x_k)'$ と k 個あ

表4. 糖尿病患者の眼底網膜症発症に関する多重ロジスティックモデル

変数	係数	オッズ比**
空腹時血糖値 (mg/dl)	8.258×10^{-3}	1.008
罹病期間 (年)	7.320×10^{-2}	1.076
血清尿素窒素 (mg/dl)	1.072×10^{-1}	1.113
収縮期血圧 (mmHg)	1.332×10^{-2}	1.013
最大体重時年齢 (歳)	-2.286×10^{-2}	0.977
高血圧の家族歴*	-4.040×10^{-1}	0.668
定数項	-5.433	—

* なし = 0, あり = 1.

** 1測定単位当りのオッズ比.

高木廣文 (1979), 日本公衆衛生雑誌, 24(4), 185-194の表6より一部引用.

る場合,

$$\text{logit}(p) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k = \mathbf{x}'\boldsymbol{\beta}$$

とロジスティックモデルの拡張を考えることができる. ここで, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)'$ である.

上式は「多重ロジスティックモデル multiple logistic model」と呼ばれ, 特定の疾患について多数の要因が関与している場合, それらの要因の影響を調べるために用いられている.

表4は糖尿病患者のうち眼底所見が初診時と5年次ともに正常な367例と, 初診時は正常であるが5年次は異常となった96例の初診時データを用いて, 眼底網膜症のリスク因子を検討したものである. 各変数の値と係数を前述の式に代入すれば, その患者の発症リスクを求めることができる. また, 表中のオッズ比は各変数の測定で1単位分の増加に対応している. したがって, 空腹時血糖値が20 mg/dl増加し, 収縮期血圧が10 mmHg増加した場合は, オッズ比も, $1.008^{20} \times 1.013^{10} = 1.33$ となり, 眼底網膜症のリスクが以前よりも約1.3倍増加するようになる.

今回は, 実際の計算などに関する詳細は省略した. これまで見たように, 疫学研究においてはさまざまな変数と疾患の発症, 死亡などとの関係が調べられる. そのような場合, リスクに関する統計学的な解析方法が, 威力を発揮することが理解されたと思う.

4. おわりに

他の分野で用いられている多変量解析の各種の方法は, 医学分野でもよく用いられており, 実際面で多大な貢献をもたらしているといえよう. 今回解説した内容は, 疫学的な研究で用いられている統計学の方法のほんの一部についてのもの

である. また, 医学分野で多用される「生命表解析」などの重要な手法について, 紙面の都合上触れることができなかったが, 統計学が極めて重要な役割を果たしていることは, 理解されたものと思う.

人間集団を扱う学

問領域では, 研究に客観性を持たせようとすれば, 統計学を用いることは極めて自然なことに思われる.

[参考文献]

- Mantel, N. and Haenszel, W. (1959). *Journal of the National Cancer Institute* 22, 719-748.
 Mehta, C.R., Patel, N.R., and Gray, R. (1985). *Journal of the American Statistical Association* 80, 969-973.
 Robins, J.M., Breslow, N.E., and Greenland, S. (1986). *Biometrics* 42, 311-323.
 Sato, T. (1990). *Biometrics* 46, 71-80.
 Takagi, H. (1990). *Research Memorandum No. 385, the Institute of Statistical Mathematics*.
 高木廣文 (1979). 日本公衆衛生雑誌, 24(4), 185-194.
 The University Group Diabetes Program (1975). *Diabetes* 24, Suppl. 1, 65-184.

(たかぎ ひろふみ)

(p. 13 から)

う. 得点分布は正規分布とは限らないが, この場合分布は正規分布であるべきだと考え, 変換した得点 Z_i は $N(50, 10^2)$ に従うとする. それは得点の相対累積度数に対応する正規分布上の z 点を求め, Z に変換したものである. 粗点 X_i の分布が正規分布に近い時は, そのまま Z_i 点を求めればよい. 共通1次テストの得点分布はピアソンI型分布(ベータ分布)が一番よく適合するが, その形状は釣鐘型で, 正規分布とよく似ているので, Z_i 点が偏差値と考えることが慣習になっている.

たまたま受けたテストの個人得点が, 全体のどの辺にあるかは偏差値から分るが, それ以上の解釈は統計学の範囲を逸脱している. 逸脱した範囲で統計学の功罪, 有効性を論じて意味がない.

(あんどう ひろみ)