

## 公開講演会要旨

# 安全性評価の統計モデル

統計数理研究所 柳 本 武 美

(昭和62年11月6日, 統計数理研究所 講堂)

### 1. 序

安全性の評価は究極的にはヒトの安全が目的となる。ところが人間は人様々で、生理学的に多様であると共に社会生活においても多様である。そしてこの多様性は、測定誤差とは違って、原則的にはありのまま認める必要がある。この事情は“みかけ上変動しているデータから、そのデータにひそんでいる法則性を探る”方法としての統計学の役割の重要性を示している。本講演では安全性の評価に用いられる統計的な考え方とモデルについて述べる。特に過去においては、統計手法の進歩が安全性評価の手法の進歩に資してきており、逆に安全性評価の手法が統計手法の進歩を促してきた事実を指摘する。そしてこの過去における事実が現在においても等しく真であることを強調する。

講演で述べられなかった詳細については巻末に掲げた演者による解説、論文等を参照されたい。

### 2. 安全性評価の考え方

#### 2.1 安全性

安全性という言葉をここでは食品、医薬品、環境汚染物質といった化学物質に対するヒトの安全性に限定する。一般に安全であるとは、ある物質が全く害作用を及ぼさないということである。統計モデルを用いて表現すれば害作用がでる確率が0である。ところが害作用の出現の有無のデータから確率0であることを証明することは不可能なことである。

#### 2.2 用量-反応関係

毒性学、あるいはもう少し広く、薬理学において普遍的に受け入れられている概念に用量-反応関係がある。既に16世紀にParacelsusは“Was ist das nit giftt ist? alle ding sind giftt und nichts ohn giftt/Allein die dosis macht das ein ding kein giftt ist.”と毒であるか否かは量で決まることを述べている。このことから動物実験では高い濃度で実験を行い、結果を外挿して安全基準を決める考え方が裏づけられる。

科学的側面からみれば、化学物質のヒトに対する影響を評価するクライテリアを求めることになる。しかし現実に行政的側面から考えると、実際には規制値、スタンダードを設定することになる。特に規制の容易さから規制値を0として使用禁止としたり、 $\infty$ として安全とみなすことが多い。リスクアセスメントについては、吉村(名古屋大)の別の講演を参考にされたい。

### 3. 統計的方法

#### 3.1 得られるデータ

用量-反応関係に基礎をおくデータとしては、動物実験あるいは疫学調査が基本的である。動物実験では種差あるいは高用量から低用量への外挿が必要となるが、条件を制御することができる。疫学調査では現実の条件の下でヒトの影響を、しばしば大きい標本サイズで見ることができるが、条件を制御できない。他に構造活性解析、健康人のボランティアによる実験等があるが、現実には *in vitro* 試験によるスクリーニングが重要視されている。

#### 3.2 用量-反応モデル

Gosset, Fisher らによる標本分布理論に基づいたデータ処理に対する革新は直ちに用量-反応モデルを実用化させた。用量  $d$  に対して害作用が出る確率を  $P(d)$  として、動物  $n$  匹を用いたとき、害作用が認められる匹数  $x$  は2項分布  $B_i(n, P(d))$  とみなして統計的推測を行う。Bliss と Gaddum による推測統計学の薬理学への導入によって  $LD_{50}$  が算定されるようになった。毒物の強さを計る尺度になった。母集団と標本の区別という推測統計学の基本によって、安全性の尺度を統計モデルにおける母数（の関数値）として表現されるようになった。

一方、疫学データの解析には検定理論が援用されることになり、単なる作表に基づく抽象的な議論から脱皮した。統計モデルの導入は後述のようにより精密な推論を可能にした。

#### 3.3 絶対安全量

毒物の強さを測る量として  $LD_{50}$  が導入されたが、現実には、より重要なものは一見安全な物質の安全性評価である。ナイーブな方法として毒見的に摂取して害作用が認められなければ良いとみなす場合があり、より慎重には実験した用量を安全係数で割ったものを安全量とみなす。安全係数  $s$  で割ることは用量-反応モデルでは、 $P(d)$  から  $P(d/s)$  へ落とすことに対応する。

ところが奇妙にも現実には  $P(d/s)=0$  を期待することが多い。安全係数を使う論理には用量-反応モデルは必ずしも必要としない。結局、安全係数の意味が曖昧で、それが故に得られた結果が神秘的にみえるものらしい。生体の反応に対する閾値の存在がその背景にあるが、ヒトの集団に閾値を仮定する根拠が確立されることは殆どない。現代でも規制値が閾値以下であるとみなされていることが多い。閾値の存在を仮定した統計モデルとしてホッカー・スティックモデル、一部の薬力学モデルがある。

#### 3.4 実質的安全量

閾値の存在を仮定しないで、絶対安全量を求めようとする、 $d=0$  となってしまう。そこで発想を変えて、害作用の出現確率をデータから高めに推定して、その値が別に与えた値  $p > 0$  より小さければ安全であるとみなす、という考え方が生まれる。高めに推定することは出現確率の上側信頼限界を求めることになる。

この方法はモデルを明示しなかったり、根拠のない閾値の存在を仮定した方法に比べて、推論としてはより科学的である。演者らもホッカー・スティック法を批判して“危険水準を用いる方法”として提唱していた。この方法は、Mantel-Bryan (1961) にその端緒を見る。その後、多段階モデルが導入されて、米国・カナダで受け入れられるようになった。近年我が国でも導入が検討されている。

#### 4. 統計モデルの進歩

実質的安全量の導入はデータを統計モデルを通して解析することによって安全性を評価することを明確にしている点が大きな長所である。統計モデルの利用による推論は、データを蓄積することによって、そしてまた、統計モデルを精密にすることによって、更には統計解析法を改善することによって、より改善される。

##### 4.1 データの蓄積

動物実験の精度の向上によって、過去のデータを利用する期待がもてるようになってきた。プロトコルが一定で、それが実験の実施の際に守られていれば、原理的に過去のデータ、特にコントロール群が利用できる。実際には繰り返し誤差が入ることが多い。繰り返し誤差の要因を探り、繰り返し誤差をランダム変動とみなした時の利用法の研究がなされており、柳川(九州大)らの貢献がある。

データの蓄積は種差の評価、構造活性解析、用量-反応曲線などに大きく資する可能性がある。そして GLP の実施は民間企業でのデータの蓄積を促している。

##### 4.2 時間データの利用

長期あるいは亜急性試験の場合、結果が用量-反応モデルで解析されていたとすると、害作用の出現した時刻を利用しない。現実の動物実験では定期的にかつ頻繁に観察しているから、害作用の出現時刻が得られることがある。全く労力と費用を必要としない情報である。そこで用量-反応モデルから用量-反応時間モデルに発展させる試みがある。生存時間モデルは近年著しく進歩したが、これはいわゆるセミパラメトリックモデルである。実際的なパラメトリックモデルが必要となる。

##### 4.3 多数の母数を含むモデル

データを十分に記述できるモデルには無限の母数が含まれていると考えられる。良い統計モデルは少ない母数で本質的な構造を捉えたモデルである。しかし最近の理論統計学の進歩とそれを支える計算機器の普及によって、従来考えられていた以上に多くの母数をもつモデルが有効であることが認識されつつある。この進歩は直ちに安全性評価のためのモデルの構築に取り入れられようとしている。

周辺尤度の利用：同一プロトコルで実験したときに繰り返し誤差があり、多世代試験では同腹効果が認められる。これらを良く記述するためには混合モデル（変動効果モデルとも呼ばれる）が用いられる。解析的にみた場合、経験 Bayes モデルと同一であり、この分野はここ数年急速に進歩している。混合モデルに関しては越智（大分大）、山本（岡山理大）らの貢献がある。

条件付尤度の利用：生存時間データの解析における比例ハザードモデル及び疫学データの解析における多重ロジットモデルでは条件付尤度（部分尤度）が利用されている。この場合、多くの母数は Nuisance であると仮定されるが、この仮定は必要なさそうである。条件付尤度は元来、直感的あるいは哲学的な動機によって導入されていたが、その数理的裏づけを得つつある。現在、演者は当研究部門の安楽氏と共に負の 2 項分布の推測における条件付尤度の利用について研究を遂行中である。

## 文 献

(以下は安全性の評価とその近周辺に関する演者の文献である。  
解説, 研究論文, その他に分類し発表順に並べた)

## 1. 解 説

- (1) 柳本武美 (1973). 新しい蛋白源は安全か, オペレーションズ・リサーチ誌, 第18巻, 第10号, 16-21.
- (2) 柳本武美 (1975). 大気汚染をめぐるデータ解析の話題, 日本統計学会誌, 第5巻, 第2号, 105-117.
- (3) 柳本武美, 山本英二 (1978). 安全さを測る統計モデルと解析, 医学のあゆみ, 第106巻, 第8号, 550-554.
- (4) 柳本武美 (1979). データの評価と解析——安全基準の設定を例にして, 公害と対策, 第15巻, 第3号, 430-434.
- (5) 柳本武美 (1982). 動物実験から発癌性の強さを評価する統計手法, 変異原と毒性, 第5巻, 第2号, 162-168.
- (6) 柳本武美 (1984). 閾値——データ解析面から, 新しい発癌のメカニズムと評価(林裕造, 黒木登志雄 編), サイエンス・フォーラム社, 84-90.
- (7) Kamakura, T. and Yanagimoto, T. (1985). Point process models in asthma attacks for environmental factors, *Env. Health Persp.*, **63**, 203-210.
- (8) Krewski, D., Yanagimoto, T. et al. Statistical methods for developmental toxicity studies, *Reproductive Toxicity* (ed. M.G. Grice) (to appear).

## 2. 研究論文

- (1) 柳本武美, 山添史郎 (1972). サリドマイドによる催奇性の統計的論証, 統計数理研究所シンポジウム記事, **4**, 13-20.
- (2) 柳本武美 (1974). SCPの安全性に関する統計学的考察, 統計数理研究所彙報, 第22巻, 第1号, 45-57.
- (3) 柳本武美 (1978). 安全な量を推定する手法の数値比較, 応用統計学, 第7巻, 第1号, 11-22.
- (4) 柳本武美, 山本英二, 竹島克朗 (1978). 安全基準としての閾値とホッケー・スティック回帰法, 統計数理研究所彙報, 第25巻, 第1号, 29-40.
- (5) 山添史郎, 福富和夫, 柳本武美 (1978). キノホルムとスモンの量・効果関係について, 日本公衆衛生雑誌, 第25巻, 第10号, 565-570.
- (6) Yanagimoto, T. and Yamamoto, E. (1979). Estimation of safe doses: Critical review of hockey-stick regression method, *Env. Health Persp.*, **32**, 193-199.
- (7) Yamazoe, S., Fukutomi, K. and Yanagimoto, T. (1980). Dose-response problem in drug-induced diseases after consecutive intake of long duration, *Ann. Inst. Statist. Math.*, **28**, 329-342.
- (8) Yanagimoto, T. and Hoel, D.G. (1980). Comparisons of models for estimation of safe doses using measures of the heaviness of tail of a distribution, *Ann. Inst. Statist. Math.*, **32**, 465-480.
- (9) Yanagimoto, T. and Hoel, D.G. (1980). Measures of the heaviness of tail of a distribution and their applications to comparisons of models for estimation of safe doses, *Recent Developments in Statistical Inference and Data Analysis* (ed. K. Matusita), North Holland, 330-340.
- (10) 柳本武美 (1980). 用量, 反応モデルにおける許容分布とその母数の制約, 京大数理解析研究所講究録, **384**, 生命科学データの統計解析と数学モデル, 48-58.
- (11) 柳本武美, 坂本淑子 (1980). 用量・反応モデルの下での安全係数の性質, 統計数理研究所彙報, 第27巻, 第1号, 55-61.
- (12) 渋谷政昭, 柳本武美 (1980). 量・反応模型と分布の裾の比較, 応用統計学, 第9巻, 第1号, 13-21.
- (13) 柳本武美, 坂本淑子, Hoel, D.G. (1981). 安全な量を推定するモデルの比較, 日本統計学会誌, 第11巻, 第1号, 94-99.
- (14) 柳本武美, 鎌倉稔成 (1981). 発癌性を評価するモデルと適用例, 統計数理研究所彙報, 第29

- 巻, 第1号, 11-23.
- (15) 柳本武美, 鎌倉稔成 (1982). 点過程モデルによる継続観察データの解析——特に喘息の発作に及ぼす環境要因の評価——, 応用統計学, 第11巻, 第1号, 39-47.
- (16) 鎌倉稔成, 柳本武美 (1983). 確率的死亡モデルとその応用, 統計数理研究所彙報, 第31巻, 第2号, 107-116.
- (17) 渡辺直大, 吉田 亮, 柳本武美 (1984). 気管支喘息に及ぼす環境因子の急性影響の解析方法に関する検討, アレルギー, 第33巻, 第10号, 868-878.
- (18) Yanagimoto, T. and Kamakura, T. (1985). Experiences in analysing survival data with some remarks, *Statistical Methods in Cancer Research* (eds. W.J. Blot, T. Hirayama and D.G. Hoel), 153-157.
- (19) Yamamoto, E. and Yanagimoto, T. Litter effect to dose-response curve estimation, 日本統計学会誌 (印刷中).
3. その他
- (1) 柳本武美 (1970). 原因追求の為の調査法の理論的検討, 統計数理研究所シンポジウム記事, 2, 63-69.
- (2) Yanagimoto, T. (1971). A statistical study of the hypothesis of thalidomide etiology, *Argument of Lecture Biomet. Soc. Group of Japan*, XXI and XXII, 26-31.
- (3) 柳本武美 (1973). 石油たんばくの安全性——特に吉田氏の問題提起について——, 応用統計学, 第3巻, 第3号, 137.
- (4) 柳本武美, 野田一雄, 村上征勝 (1977). 安全性の評価と安全基準に関する統計学的研究, 統計数理研究所, 昭和51年度特別研究報告.
1. 山本英二, 竹島克朗, 柳本武美: 安全基準としての閾値とホッケー・スティック回帰法の検討, 1-13.
2. 柳本武美: SCPの安全性の評価としての鶏を用いた累代試験の再解析, 36-43.
3. 柳本武美, 柏原大造, 山本英二: 慢性呼吸器症状の要因の分析と要因の補正, 44-50.
- (5) 柳本武美 (1978). 生物学・医学のための統計, 78-Fコース, 統計数理研究所公開講座テキスト.
- (6) 柳本武美 (1979). 統計処理のすすめ, ファルマシア (日本薬学会誌), 第15巻, 第8号, 698-702.
- (7) 柳本武美 (1980). 統計モデルとデータ解析——坪田, 前田の誌上討論へのコメント——, 日本公衆衛生雑誌, 第27巻, 第12号, 659-660.
- (8) 柳本武美 (1981). 臨床試験における統計的検定, 薬効の評価基準(七川敏次編), 永井書店, 73-80.
- (9) 柳本武美 (1982). 複合大気汚染に係る健康影響調査, 閉塞性肺疾患患者調査中間結果報告(分担), 東京複合大気汚染調査報告書.
- (10) 柳本武美 (1982). 発癌物質の発癌性の定量的評価, 昭和56年度がん研究助成金による研究報告書, 936.
- (11) 鎌倉稔成, 柳本武美 (1984). 比例ハザードモデルのパラメータの推定量の比較, 統計数理研究所彙報, 第32巻, 第1号, 35-43.
- (12) 柳本武美 (1984). 安全性評価のための統計的手法, 84-Fコース, 統計数理研究所公開講座テキスト.
- (13) 柳本武美 (1986). 危険解析のための統計モデル, 統計数理, 第34巻, 第1号, 159-160.