# 歯科疫学統計
## －第3報　重回帰分析、多重ロディスティック回帰分析モデルの適合度判定指標の解釈－
## －SPSS，STATAの利用に際して－

瀧　口　　　徹

# A review of oral epidemiological statistics
## －Part III：Interpretation of various goodness of fit indicators for the Multiple Regression Model and Multiple Logistic Regression Model.－
## －When using the statistical software SPSS and STATA－

Toru Takiguchi

要旨：統計学的モデル解析においては実際値とモデルによる予測値との誤差が可能な限り小さいこと、すなわち適合度（goodness of fit）が高いことが必須条件であることを前報で解説した。しかし適合度を計測する指標は多数あることと統計ソフトウエアー間で往々にして採用指標や表現法が全く異なるためユーザが混乱する可能性がある。そこで第3報では数多の多変量解析法の中で利用頻度の点で双璧である重回帰分析：MRAと多重ロディスティック回帰分析：LRAの各種適合度判定指標を代表的な統計解析ソフトSPSS13.0JとSTATA8.2の出力に沿って解説した。MRAにおいては決定係数：$R^2$、一元配置分散分析、Durbin－Watson統計量およびVIF：分散拡大係数（許容度）を解説した。LRAにおいては擬$R^2$：Cox & Snellの$R^2$とNagelkerkeの$R^2$、HosmerとLemesshowの検定、Omunibus　検定、変数選択法の最大対数尤度による方法とWald法の比較、AIC：Akaike's Information Criterion、およびdeviance：デビアンスの各指標を用いた適合度判定の概念と使用法を解説した。関連でLRAについてはodds比の計算法を合わせて解説した。さらに理解を容易にするため「歯科医師10万比」と「住民1人当たり月額歯科保険医療費」および「歯科診療所当たり月額保険収入」との関連を事例として用いた。この結果、MRAでもLRAでも前者は歯科医師10万比が増加すると増加する傾向が、後者はその逆に減少する関係が交絡因子を調整しても高度に有意であった。また閾値を境にデータを2区分してLRAを用いて計算されたodds比によって要因の重みがより明確になることが示された。

キーワード：重回帰分析、多重ロディスティック回帰分析、モデル適合度、－2対数尤度、AIC

## はじめに

第1報[1)]で各種統計分布の特徴と利用に際しての注意点について概説し、第2報[2)]においてはより厳格な統計分布への適合を統一的に行う一般化線形モデル：GLIM[3)]の概念と実例を通して利用の要点について概説した。第3報では第1報で示したように利用頻度の点で双壁[1)]である重回帰モデル、多重ロディスティック回帰モデルに絞ってこの使い分けと代表的な統計ソフトであるSPSS13.0JおよびSTATA8.2に用いられている各種適合性指標の解釈と利用の実際について解説する。なお、利用の実際では解析結果の理解を深める意図から「都道府県単位でみて歯科医師10万比

【著者連絡先】
〒341-0003　埼玉県三郷市彦成3-86
深井保健科学研究所
主席研究員　瀧口　徹
TEL&FAX：048-957-3315
E-mail：taki8020@mth.biglobe.ne.jp

と住民の歯科医療費および歯科診療所の保険収入との関係」を事例とした。

### １．事例のプロフィール

図1は47都道府県の住民1人当たり月額歯科保険医療費の分布を示す。全国加重平均値は月額1,619円である。ただしこの金額は公費負担を含んだ総額であり窓口で支払う治療費（dental fee）ではない。直接医療機関の窓口で支払う治療費だけではなく保険料であっても税金からであっても医療費は全て元をたどれば国民が負担していることを基礎においた指標とした。図2は都道府県の歯科医師10万比と住民1人当たり月額歯科保険医療費との関係を示す散布図である。相関係数：rは0.759（P＝0.000）であり両者の関係は一見して相当強い直線的関係があることが推定される。しかし背景に社会経済学的な交絡因子（confounding factor）があって現象が実際以上に強調されているかも知れないし、またその逆にもっと強い関係があるが背景因子の影響で見かけ上弱められているかも知れない。これを確かめるためには多変量解析の手法を用いて交絡因子の調整を行う必要がある。図3は47都道府県の歯科診療所当たり月額保険診療収入（2004年6月分）の分布を示す。全国加重平均値は月額310万円である。ちなみに2年ごとの医療経済実態調査によると個人経営の歯科診療所の場合このうちの約30％強が収支差額になりその額は月額約100万円となる。図4は都道府県歯科医師10万比と歯科診療所当たり月額保険診療収入の関係を示す散布図である。相関係数：rは－0.800（P＝0.000）であり歯科医師10万比が増加すると歯科診療所収入が減少する関係は相当

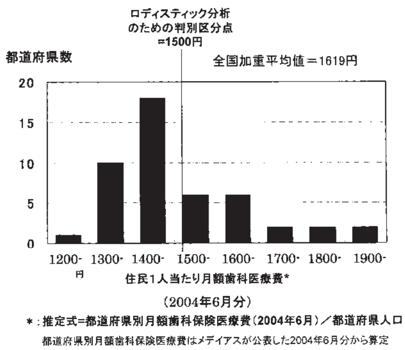



図1　47都道府県の住民1人当たり月額歯科保険医療費の分布

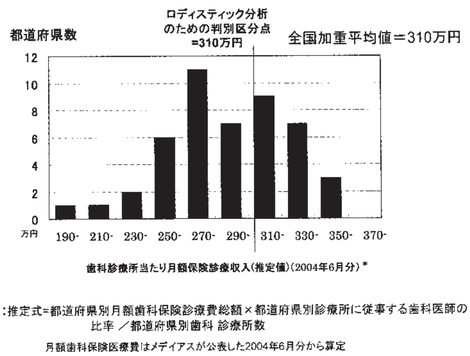



図3　47都道府県の歯科診療所当たり月額保険診療収入の分布

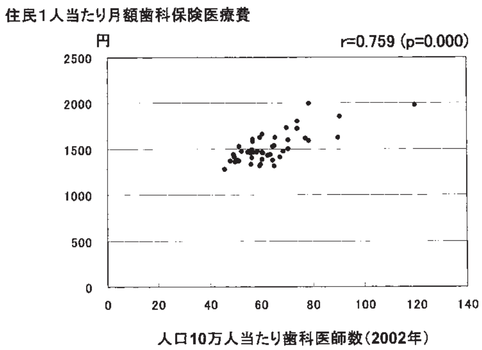



図2　都道府県歯科医師10万比と住民1人当たり月額歯科保険医療費

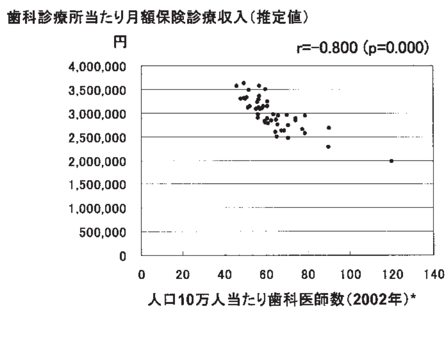



図4　都道府県歯科医師10万比と歯科診療所当たり月額保険診療収入

表1　都道府県特性因子分析（バリマックス法）結果

| 指標番号 | 調査年 | 都道府県特性指標 | 単位 | Factor | 因子負荷量 | 固有値 寄与率%(累積%) | 因子の特性 |
|---|---|---|---|---|---|---|---|
| 1 | '02 | 人口密度 | 1km$^2$ | | 0.89 | | |
| 2 | '02 | 人口 | 千人 | 1 | 0.89 | 8.08 | |
| 3 | '00 | 人口集中地区人口密度 | 1km$^2$ | | 0.82 | 20.7(20.7)% | 都市化 |
| 4 | '02 | 現金給与総額 | 千円 | | 0.81 | | |
| 5 | '00 | 1世帯人員数 | 人 | | −0.75 | | |
| 12 | '99 | 一人当医療費 | 円 | | −0.88 | | |
| 13 | 90-'00 | 世帯増加率 | % | 2 | 0.87 | 7.77 | |
| 14 | '02 | 人口性比 | 女性100 | | 0.79 | 19.9(40.6)% | 非高齢化 |
| 15 | '02 | 年齢別人口(65歳以上) | % | | −0.76 | | |
| 16 | '02 | 死亡率 | 千人 | | −0.75 | | |
| 22 | '01 | 一人当電気使用量 | kwh | | 0.85 | | |
| 23 | '02 | 有効求人倍率 | 倍 | 3 | 0.80 | 3.95 | 雇用活性 |
| 24 | '02 | 第2次産業就業者割合 | % | | 0.79 | 10.1(50.7)% | |
| 25 | '02 | 完全失業率 | % | | −0.66 | | |
| 26 | '02 | 1世帯1月当実収入 | 円 | | 0.96 | | |
| 27 | '02 | 1世帯1月当可処分所得 | 円 | 4 | 0.95 | 3.73 | 収入支出活性 |
| 28 | '02 | 1世帯1月当実支出 | 円 | | 0.89 | 9.6(60.3)% | |
| 29 | '02 | 年齢別人口(14歳以下) | % | | −0.53 | | |
| 30 | '02 | 可住地面積 | km$^2$ | | 0.98 | | |
| 31 | '02 | 総面積 | km$^2$ | 5 | 0.96 | 3.28 | 土地面積等 |
| 32 | '01 | 農家1戸当生産農業所得 | 千円 | | 0.91 | 8.4(68.7)% | |
| 33 | '00 | 平均寿命(男) | 歳 | 6 | 0.89 | 2.47 | 平均寿命 |
| 34 | '00 | 平均寿命(女) | 歳 | | 0.59 | 6.3(75.0)% | |
| 35 | '02 | 乳児死亡率 | 千人 | | −0.67 | | |
| 36 | '02 | 耕地率 | % | 7 | −0.61 | 2.00 | 乳児死亡率等 |
| 37 | '02 | 水道普及率 | % | | 0.55 | 5.1(80.2)% | |
| 38 | '01 | 都市ガス普及率 | % | 8 | 0.59 | 1.55 | 物価等 |
| 39 | '02 | (総合)物価指数 | '00年を100 | | −0.53 | 4.0(84.1)% | |

注1）データ出典:「データでみる県勢2004」（財）矢野恒太郎記念会
注2）因子負荷量(=相関係数)≧0.5および各因子の因子負荷量の絶対値上位5までを表示(表では11指標を省略)

強い直線関係があることが推定されるがこれも同様に交絡因子の調整を行ったうえで解釈する必要がある。そこで都道府県別の各種社会経済学的指標を交絡因子の候補とし、具体的には表1に示すように社会経済的な39指標[4)]を用いて因子分析（バリマックス直交回転）を行い固有値≧1の通法で8因子を得た。8因子（Factor1～8）それぞれの性格については表1を参照されたい。直交回転による因子分析によって標準化された39指標の全分散39のうちの84.1％を互いに無相関（r＝0）な8因子で説明できることを意味している。これら8因子の因子得点と歯科大学の有無を加えた9変数を交絡因子の候補として解析に使用した。

## 2．適合度（goodness of fit）に関する指標

代表的な統計ソフトであるSPSS13．0JおよびSTATA8．2の重回帰分析（以下、MRA）と多重ロディスティック回帰分析（以下、RLA）で用いられている各種の適合度判定指標（odds比以外）は下記のとおりである。これらを順次解説する。

Ⅰ．MRA：重回帰分析
- ⅰ）$R^2$：決定係数
- ⅱ）one-way ANOVA：一元配置分散分析
- ⅲ）Durbin-Watson統計量
- ⅳ）VIF：分散拡大係数、許容度

Ⅱ．LRA：ロディスティック回帰分析
- ⅰ）擬$R^2$：Cox & Snellの$R^2$とNagelkerkeの$R^2$
- ⅱ）HosmerとLemesshowの検定
- ⅲ）Omunibus検定
- ⅳ）最大対数尤度（変数選択）
  尤度、対数尤度、－2対数尤度、
  log likelihood
- ⅴ）Wald検定（変数選択））
- ⅵ）odds比、exp（B）
- ⅶ）AIC：Akaike's Information Criterion
- ⅷ）deviance：デビアンス

## 3．適合度に関する指標の解説

Ⅰ．MRA：重回帰分析

ⅰ）$R^2$：決定係数

MRAの適合度指標を理解するには単回帰分析における適合度を図形的に理解することが近道である。図5は各点からy軸に平行な回帰直線までの距離の二乗の総和、すなわち最小二乗法（The least squares method）で回帰直線が求められることを表している。ここで重要なことは直線回帰に限らずあらゆる曲線回帰においても（実際値と予測値の差）$^2$を最小とすることが適合度を最大にすることと同義であるということである。$SS_E$は残差平方和（誤差平方和）と呼ばれる。図6は$SS_T$：「全体の平方和」を示し各点とyの平均値間の差の二乗をΣしたものである。次の図7は$SS_R$：「回帰による平方和」を示し回帰直線（各点のx値に対応するy値）と平均値間差の二乗をΣしたものである。表2に$SS_T$、$SS_R$、$SS_S$と決定係数：$r^2$および回帰直線の分散分析のF値の関係を示した。数理的に$SS_T$＝$SS_R$＋$SS_E$の関係があるので両辺を$SS_T$で割ると$SS_R$／$SS_T$の項が出てくるがこれが決定係数：$r^2$であり、回帰直線と各点間の偏差の二乗値を各点と平均値からの偏差の二乗値の比である。この比の意味は回帰直線によって全体のバラツキのどのくらいが説明できるかを表しており、仮に全ての点が直線上に乗っていれば1となる。一方、$SS_T$＝$SS_R$＋$SS_E$の両辺を$SS_E$で割ると$SS_R$／$SS_E$の項が出てくるがこれが自由度（p：説明変数の数、n：例数－p－1）のF分布する。これが回帰直線の分散分析に用いられるF値である。従って回帰直線の決定係数（寄与率）：$r^2$　（または$R^2$）とF値は極めて近似した意味を持っている。ここで調整済み（修正済み）決定係数は表3の注1に示すようにデータ数：Nと説明変数の数：pによって決まりデータ数が少

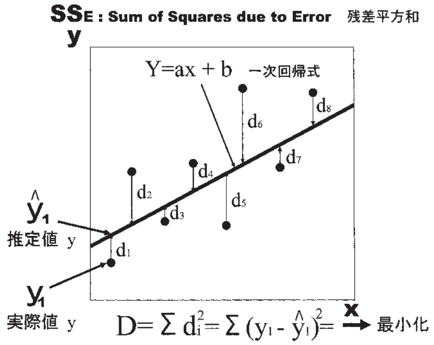



図5　直線回復適合度判定－残差平方和と最小2乗法－

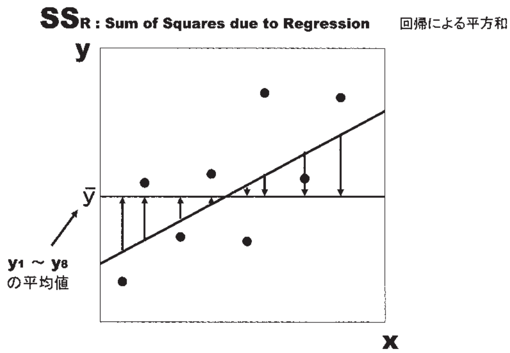



図7　直線回帰の適合度判定－回帰による平方和－

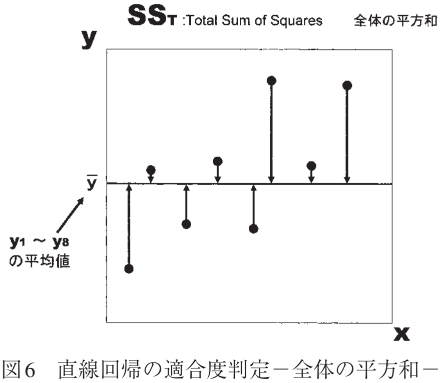
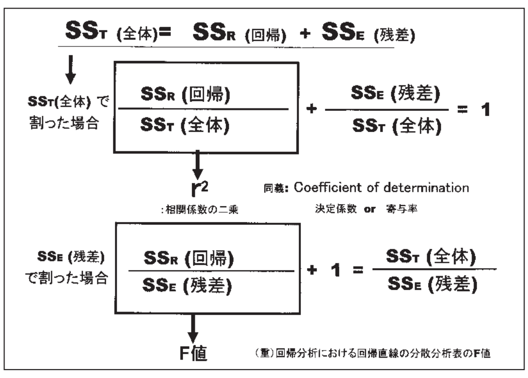



図6　直線回帰の適合度判定－全体の平方和－

表2　回帰直線における相関係数：$r^2$とF値（分散分布）の関係



なく説明変数が多い場合に精度を低めに見ることになる変換である。例えば決定係数：$R^2$＝0.64で説明変数の数：p＝10と同じであってもデータ数が100と20とでは前者の調整済み$R^2$は0.60なのに対し後者は0.24と非常に小さくなり大きな差が出る。なお、重回帰式を予測モデルとして用いる場合は一般に重相関係数：Rは0．7以上あることが目安となる。この理由は決定係数×100がほぼ50％となることに起因している。一方、要因分析モデルの場合は、Rの絶対値が大きいことは要因分析モデルの説明力が高いことを意味して好ましいことであるが測モデルのように0.7以上が必要という目安はない。

事例：

表3に、住民1人当たり月額歯科保険医療費（以下、「歯科医療費」）を目的変数とする段階式重回帰分析（以下、step-MRA）をFin＝2.00、Fout＝1.99の通法で行った結果を示す。

この例では選択された説明変数の数が5個と比較的多いためR＝0.8257、決定係数：$R^2$＝0.6818、調整済み$R^2$＝0.6429となっている。この解釈は調整済み決定係数により全体の平方和のうち64％が回帰直線で説明できることを意味している。これに対し表4の歯科診療所当たり月額保険収入（以下、「歯科保険収入」）の場合3つの変数が選択されR＝0.8484、決定係数：$R^2$＝0.7197、調整済み$R^2$＝0.7002となっている。いずれの分析においても「歯科医師10万比」は高度に有意であり見掛けの現象ではないことが確認された。

表3　住民1人当たり月額歯科保険医療費の増加要因分析

－ expenditure：step-MRA －

| ﾓﾃﾞﾙ | 重相関係数：R | 決定係数：$R^2$ | 調整済み $R^2$ | 標準誤差：SE | Durbin-Watson の検定：DW |
|---|---|---|---|---|---|
| 5変数 | 0.8257 | 0.6818 | 0.6429 | 97.8197 | 1.6525 |

| | 非標準化係数 | 標準誤差 | 標準化係数 | t値 | 危険率 | B 95%CL | | 共線性 | |
|---|---|---|---|---|---|---|---|---|---|
| | B | SE | β | t | p | 下限 | 上限 | 許容度 | VIF |
| (定数) | 1134.93 | 101.12 | | 11.22 | 0.00 | 930.72 | 1339.15 | | |
| 歯科医師10万比 | 5.96 | 1.58 | 0.48 | 3.77 | 0.00 | 2.77 | 9.15 | 0.48 | 2.09 |
| Factor 1 | 54.08 | 20.23 | 0.33 | 2.67 | 0.01 | 13.23 | 94.93 | 0.50 | 2.01 |
| Factor 4 | −30.37 | 14.28 | −0.19 | −2.13 | 0.04 | −59.21 | −1.52 | 1.00 | 1.00 |
| Factor 2 | −28.97 | 14.76 | −0.18 | −1.96 | 0.06 | −58.77 | 0.84 | 0.93 | 1.07 |
| Factor 3 | −21.22 | 14.32 | −0.13 | −1.48 | 0.15 | −50.13 | 7.70 | 0.99 | 1.01 |

注1)　調整済み決定係数;$R^2$ = $R^2$(N-1)/(N-p-1) − p/(N-p-1)　N:例数　p:説明変数の数

ii）one-way ANOVA：一元配置分散分析

分散分析（F分布）は概念が高尚にもかかわらず汎用性が高く様々なところで使われるため時にユーザの混乱を招くことがあるようである。例えばstep-MRAにおいて変数選択の基準として用いられるFin、Foutは分散分析によるF値であり、回帰直線の有意性をみるのも分散分析である。要は特定の種類の分析に分散分析が登場すると考えるより分散が分割されて分母と分子にある指標（＝偏差平方和が分母と分子にある指標）はF分布する、つまり分散分析であると考えるといい。i）で解説（表2）したように回帰直線の分散分析のF値は$SS_T$＝$SS_R$＋$SS_E$の両辺を$SS_E$で割った式の1つの項である$SS_R$／$SS_E$がF分布することを利用して回帰直線の適合度を判定している。

事例：

「歯科保険収入」を目的変数としたstep-MRAによって表4の3指標が選択されたがその回帰直線の適合度を示す分散分析結果を表5示す。F値

表4　歯科診療所当たり月額保険収入の減少要因分析

－ income：step-MRA －

| ﾓﾃﾞﾙ | 重相関係数：R | 決定係数：$R^2$ | 調整済み $R^2$ | 標準誤差：SE | Durbin-Watson の検定：DW |
|---|---|---|---|---|---|
| 3変数 | 0.8484 | 0.7197 | 0.7002 | 191837.11 | 2.0086 |

| | 非標準化係数 | 標準誤差 | 標準化係数 | t値 | 危険率 | B 95%CL | | 共線性 | |
|---|---|---|---|---|---|---|---|---|---|
| | B | SE | β | t | p | 下限 | 上限 | 許容度 | VIF |
| (定数) | 4158143 | 179384. | | 23.18 | 0.00 | 3796379 | 4519906 | | |
| 歯科医師10万比 | −17575.72 | 3156.39 | −0.66 | −5.57 | 0.00 | −23941.17 | −11210.26 | 0.46 | 2.16 |
| 歯科大学の有無 | −174171.11 | 82779.24 | −0.25 | −2.10 | 0.04 | −341111.35 | −7230.86 | 0.47 | 2.11 |
| Factor 2 | −60999 | 29847 | −0.18 | −2.04 | 0.05 | −121190.74 | −807.26 | 0.88 | 1.14 |

表5　重回帰分析の分散分析表（一元配置ANOVA）

－ income：step-MRA －

目的変数：1歯科診療所当たりの月額保険収入
説明変数：予測値：歯科医師10万比，歯科大学の有無，Factor 1

分散分析[d]

| ﾓﾃﾞﾙ | | 平方和 | 自由度 | 平均平方 | F 値 | 有意確率 |
|---|---|---|---|---|---|---|
| 1 | 回帰 | 3.610E+12 | 1 | 3.610E+12 | 79.822 | .0000000001569[a] |
| | 残差 | 2.035E+12 | 45 | 4.523E+10 | | |
| | 全体 | 5.646E+12 | 46 | | | |
| 2 | 回帰 | 3.910E+12 | 2 | 1.955E+12 | 49.542 | .0000000000541[b] |
| | 残差 | 1.736E+12 | 44 | 3.946E+10 | | |
| | 全体 | 5.646E+12 | 46 | | | |
| 3 | 回帰 | 4.063E+12 | 3 | 1.354E+12 | 36.805 | .0000000000605[c] |
| | 残差 | 1.582E+12 | 43 | 3.680E+10 | | |
| | 全体 | 5.646E+12 | 46 | | | |

a. 予測値：(定数)、歯科医師10万比。
b. 予測値：(定数)、歯科医師10万比，歯科大学の有無。
c. 予測値：(定数)、歯科医師10万比，歯科大学の有無，Factor 1。
d. 従属変数：歯科保険収入

は回帰と残差それぞれの平方和を自由度で割った平均平方の比であり自由度（3,43）のF表から有意性を判定している。結果はF＝36.805（p＝0.000）で高度に有意であり適合度は満足すべきものと言える。ここで有意確率を小数点以下14桁でみてみると表5示すように僅差であるが最も有意確率が良好なのは変数が2つの場合であることがわかる。このようにstep-MRAにおける分散分析は最終モデルが最も有意性が高いというわけではない。すなわちF検定は最適モデルを選択するためではなくあくまでも変数増減法によって得られたモデルの精度を評価している。また後述するようにF値では「実際値と予測値の差」が正規分布してMRAの前提を満たしているかということ、また変数間に多重共線性（collinearity, colinearity）が存在しないかということはわからない。

iii）Durbin-Watson統計量

表3、表4の上部右端にDurbin-Watsonの検定（以下、DW）結果が示されており「歯科医療費」の分析ではDW＝1.6525、「歯科保険収入」ではDW＝2.0086であるがそれぞれ何を意味しているのであろうか。これを理解するには重回帰モデルの構造をみる必要がある。

重回帰モデル：$Y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_p X_{pj} + \varepsilon_j \ (j = 1, 2, 3, \ldots n)$

において最後部の$\varepsilon_j$が残差である。重回帰モデルの最も重要な前提はデータが連続量であれば残差が平均値0、分散が$\sigma^2$の正規分布に従うことである。このことは残差$\varepsilon_1, \varepsilon_2, \varepsilon_3 \ldots \varepsilon_n$が互いに独立（＝ランダム）であるということと同義である。この残差が互いに独立であるかを確かめるのがDWテストである。このことを最も注意すべき種類のデータとしては時系列データがある。例えば、ある銘柄のビールの売り上げが経年的に線形の有意の上昇傾向を示したとしてもビールの売り上げに季節変動や趨勢変動（時代効果）があれば残差間には相関関係（内部相関）が生じる。季節変動であれば夏に近づけば前の月より売り上げは上昇し、冬に近づけばその逆の関係が出て誤差は近接するデータとの関係が「互いに独立」ではなくなるであろう。こうした場合、決定係数：$R^2$やF値が過大に評価される恐れがある。DW値は正の内部相関があると0に近づき負の場合は4に近づく。誤差が「互いに独立」であると、すなわち正規分布の前提が保障されると2に近接する性質を持つ。DWの有意性まで言及する報告は比較的少ないが厳密な使い方としてはDWの検定表[5]（note参照）からdLとdUを求め図9に示す基準で判定すればよい。

なお、誤差（実際値と予測値の差）がガウスが発見した誤差分布すなわち正規分布すると仮定するか否かによって分析法が大きく異なる。MRAの場合は正規分布を、LRAの場合はロディスティック分布を仮定するというように仮定する分布ごとに尤度計算法[3]（後述）が異なる。この点を厳密処理しようというのがGLIM法である。

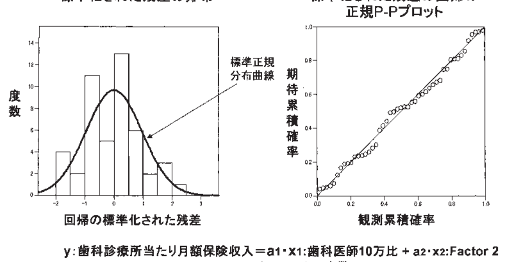



図8　歯科診療所当たり月額保険収入モデルの標準化誤差（実際値－推定値）の分布

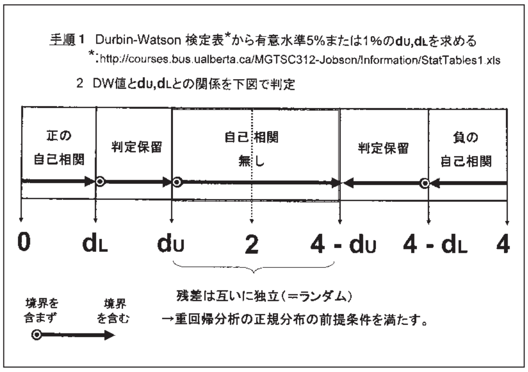



図9　Durbin-Watson統計量：DWによる自己相関の検定法

事例：

「歯科保険収入」のstep-MRAを例にとると表4に示すごとくDW＝2.0086と理想的な値を示している。これを視覚的に確認したのが図8である。ここに示した標準化された誤差の分布（左図）はデータ数が47と比較的少ないので頻度の乱れはあるが正規性が高い。また観測値の累積確率をx軸に期待値の累積確率をy軸にとったXYプロットであるP-Pプロット（右図）も各データが45度の直線上に概ね配置されている。それではDWが2からかけ離れた値をとった場合にはどうなるであろうか。表6は目的変数を「住民1人当たり月額医科保険医療費」としたstep-MRAの結果表であるがDW＝1.2697で表6下部に示す判定法により正の自己相関が有意の事例である。本事例を図示した図10と図8を比較するとDWの意義が理解していただけるであろう。

注）

DW＝Σ（$u_i$－$u_i$－1）$^2$／Σ $u_i^2$（分子のΣはi＝2からn、分母のΣはi＝1からn）で表され0≦DW≦4である。

表6　誤差の正規性の確認（Durbin-Watson統計量の見方）

分析法：step-MRA　a

| モデル | R | R2乗 | 調整済みR2乗 | 推定値の標準誤差 | Durbin-Watsonの検定 |
|---|---|---|---|---|---|
| 最終 5 | 0.8049 | 0.6478 | 0.6232 | 1474.7978 | 1.2697 |

| | 非標準化係数 | 標準誤差 | 標準化係数 | t値 | 危険率 | B 95%CL | | 共線性の統計量 | |
|---|---|---|---|---|---|---|---|---|---|
| | B | SE | β | t | p | lower | upper | 許容度 | VIF |
| (定数) | −29663.86 | 7398.48 | | −4.01 | 0.00 | −44584.33 | −14743.40 | | |
| 第3次産業比率 | 276.99 | 53.01 | 0.54 | 5.23 | 0.00 | 170.09 | 383.90 | 0.77 | 1.30 |
| 出生率 | 1029.64 | 393.27 | 0.33 | 2.62 | 0.01 | 236.55 | 1822.74 | 0.52 | 1.91 |
| 死亡率 | 2301.94 | 289.68 | 1.09 | 7.95 | 0.00 | 1717.75 | 2886.14 | 0.43 | 2.31 |

Durbin-Watson統計量：DWによる自己相関の検定

DW表の5%点　$d_L$=1.38　$d_u$=1.67

$\boxed{1.2697}$ ≦dL：1.38　→　正の自己相関がある

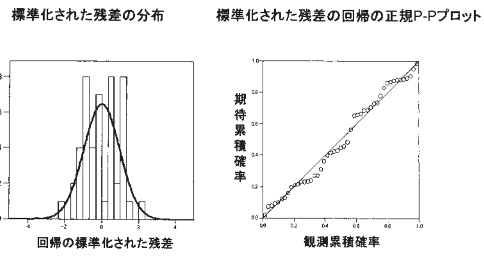



図10　モデルの誤差が正規分布しない場合の例（住民1人当たり月額医科保険医療費）

なお、DW検定表を掲載している統計学参考書は少ないのでもし手元に無ければ巻末の参考Web-site WS01からダウンロードするとよい。

iv）VIF：分散拡大係数

VIF：Variance Inflation Factors[6] は分散拡大係数とよばれ多重共線性（Multi-collinearity）を検出する指標である。また許容度（torerance）はVIFの逆数であり例えばVIF＝2のとき　許容度＝0.5となる。以下に多重共線性の現象、原因、問題点および対策を示す。

現　象：

①多重共線性がある変数の一方が極めて大きな値（inflationの言葉を用いる所以）になり、他方が逆に極めて小さな値になり、データの僅かな増減に対して係数が大きく揺れ動き不安定になる。

②標準化係数絶対値の異常な大きさに比してt値が小さいという異常現象が起きる

③目安はVIF値が10以上となる場合とされる。

原　因：説明変数間の高い（内部）相関関係によって回帰係数の計算アルゴリズムの行列式が0となり計算不能になるため生じる。

問題点：

見かけ上高い重相関係数や他回帰係数値が常識的に得られるので気づかないで潜在化し易い。

対　策：

多重共線性を示す2変数のうちどちらか一方を

表7　VIFによる多重共線性の検出と対策例

例）目的変数を歯科診療所当たり月額保険診療収入とするロディスティック分析

| | 非標準化係数 | 標準化係数 | | 有意確率 | 共線性の統計量 | |
|---|---|---|---|---|---|---|
| | B | ベータ | t | p | 許容度 | VIF |
| (定数) | 3913931 | | 17.330 | 0.000 | | |
| 2001年dmft | 70864 | 0.169 | 1.789 | 0.081 | 0.736 | 1.359 |
| 歯科大学の有無 | −245900 | −0.348 | −2.961 | 0.005 | 0.476 | 2.101 |
| 歯科医師10万比 | −18658 | −0.702 | −5.206 | 0.000 | 0.362 | 2.764 |
| 人口密度 | 83 | 0.266 | 1.026 | 0.311 | 0.098 | 10.217 |
| 1人当総医療費 | −6891 | −8.319 | −1.080 | 0.286 | 0.000 | 9011.377 |
| 1人当医科医療費 | 7777 | 8.351 | 1.111 | 0.273 | 0.000 | 8582.734 |

↓ 対策：総医療費を除外　　多重共線性の原因→総医療費と医科医療費間の高い相関：r=0.999850

| | 非標準化係数 | 標準化係数 | | 有意確率 | 共線性の統計量 | |
|---|---|---|---|---|---|---|
| | B | ベータ | t | p | 許容度 | VIF |
| (定数) | 3948075 | | 17.619 | 0.000 | | |
| 2001年dmft | 73802 | 0.176 | 1.863 | 0.070 | 0.739 | 1.353 |
| 歯科大学の有無 | −255110 | −0.361 | −3.082 | 0.004 | 0.481 | 2.079 |
| 歯科医師10万比 | −18800 | −0.707 | −5.238 | 0.000 | 0.362 | 2.760 |
| 人口密度 | 14 | 0.046 | 0.287 | 0.776 | 0.255 | 3.921 |
| 1人当医科医療費 | 219 | 0.235 | 1.282 | 0.207 | 0.197 | 5.078 |

説明変数から外す。

事例：

表7に多重共線性のためVIFが8500～9000にもなる極端な例を示した。総医療費と医科医療費の係数がそれぞれ-6891と7777と正負に分かれてかつ極端な絶対値が近似した値を示しており、しかも有意でない。この原因は両者の相関係数が$r=0.999850$とほぼ1であることに起因している。多重共線性を示す2変数のうちどちらか一方を説明変数から外すのが対策であり一般には目的変数との相関が高い方を残すのが合理的であるが、この場合は「歯科保険収入」は総医療費の一部を構成しており分析のルール違反の様相があるので直接関係がない医科医療費を残している。その結果VIFは5.078と多重共線性の目安である10を下回り許容範囲になった。ただし、VIF≦4を目安としているマニュアルや研究が散見されるので実際には4以下を目安とすることがより安全であろう。

## Ⅱ．LRA：ロディスティック回帰分析

### ｉ）擬$R^2$：Cox & Snellの$R^2$とNagelkerkeの$R^2$

これらの$R^2$は線形式において通常使われるPearsonの積率相関係数の二乗値である$R^2$に似せてLRA非線形式に使えるように開発されたものである。その特徴は以下のとおりである。

①重回帰分析：MRAにおける決定係数：$R^2$に似せて（imitate）作られた総計指標で尤度（likelihood）に基づいて計算される。従って使い方もMRAと同じ扱いをして差し支えない。擬（pseudo）$R^2$ともいわれる。

②Nagelkerkeの$R^2$はCox & Snellの$R^2$を変形してとりうる値の範囲を0-1にすることにより線形式の$R^2$にさらに似せたもので扱いやすい。このためNagelkerkeの$R^2$はCox & Snellの$R^2$より通常大きくなる。

事例：

表8に「歯科保険収入」を目的変数とするLRAの結果のうちの擬$R^2$を示す。Nagelkerkeの$R^2$に基づくと本モデルの回帰による分散は全分散の81.09％を説明できることを意味している。

表8　Cox&Snellの$R^2$とNagelkerkeの$R^2$

ー　income：backward-LRA (log-likelihood)ー

例）目的変数を歯科診療所当たり月額保険診療収入とするロディスティック分析

| ステップ | -2 対数尤度 | Cox & Snell R 2 乗 | Nagelkerke R 2 乗 |
|---|---|---|---|
| 1 | 13.93902184 | 0.6510 | 0.8791 |
| 2 | 13.99801838 | 0.6506 | 0.8785 |
| \| | | | |
| 9 | 20.28899266 | 0.6006 | 0.8109 |

### ii）HosmerとLemesshowの検定

Hosmer and Lemeshow's Goodness of Fit Test（以下、H-L検定）と呼ばれ1989年に開発されたモデルの適合度判定法である。特徴は以下のとおりである。

①データ値が少なく点在（sparse）する場合、deviance差やPeasonの$\chi^2$による近似が悪くなるがH-L検定はこのような場合の適合度判定として用いられる。

②データ値の大小から10前後の群に分割し、判別区分ごとの各群における観測値と期待値の差に対して自由度＝分割数-2の$\chi^2$検定を行う。

③目的変数の実際値と予測値に差が無いという棄無仮説検定であり$p<0.05$の場合は差あり、すなわちモデルの適合度は悪いという結論になる。有意差がないと適合が許容（acceptable）される。しかし例えば$p=0.06$で有意でないからといって即モデルの適合度が満足すべき状況だと言っているわけではなく、あくまでも許容なので注意する必要がある。むろんpが大きいほど適合度が良いことを意味する。また例数が少ないと差の検出力が弱い。

事例：

表9に「歯科保険収入」を目的変数としたLRA

表9　HosmerとLemeshowの検定

ー　income：backward-LRA (log-likelihood)ー

例）目的変数を歯科診療所当たり月額保険診療収入とするロディスティック分析

| ステップ | カイ 2 乗 | 自由度 | 有意確率 |
|---|---|---|---|
| 1 | 12.04 | 7 | 0.099 |
| 2 | 10.07 | 7 | 0.184 |
| 3 | 9.93 | 7 | 0.192 |
| 4 | 3.05 | 7 | 0.880 |
| \| | | | |
| 9 | 1.92 | 7 | 0.964 |

のH-L検定結果を示す。変数減少法でステップごとに$\chi^2$値による有意確率が増加していき3変数が残った最終の第9ステップにおいてp＝0.964になり本モデルの適合度が極めて高いことを示している。

iii）Omunibus検定

本検定の特徴は以下のとおりである。

①線形モデルの有意性判定の分散分析法：ANOVAに相当する検定法で非線形モデルにおいてANOVAと同様な扱いの検定ができる。

②定数項のみのモデルと変数を加えたモデルの尤度（likelihood）を比較し$\chi^2$検定で適合度（goodness of fit）を判定する。この検定の場合はANOVAと同様に$p < 0.05$の場合にモデルの適合度が高いことを意味する。

事例：

表10に「歯科保険収入」を目的変数としたLRAのOmunibus検定結果を示す。3変数が残った最終の第9ステップにおいてp＝0.000になり本モデルが極めて高い適合度になることを示したが、

表10　モデル計数のオムニバス（Omnibus）検定

－ income : backward-LRA (log-likelihood)－
例）目的変数を歯科診療所当たり月額保険診療収入とするロディスティック分析

| | | $\chi^2$ | 自由度 | 有意確率 |
|---|---|---|---|---|
| ｽﾃｯﾌﾟ 1 | ｽﾃｯﾌﾟ | 49.483 | 11 | .000 |
| | ﾌﾞﾛｯｸ | 49.483 | 11 | .000 |
| | ﾓﾃﾞﾙ | 49.483 | 11 | .000 |
| ｽﾃｯﾌﾟ 9(a) | ｽﾃｯﾌﾟ | −1.265 | 1 | .261 |
| | ﾌﾞﾛｯｸ | 43.133 | 3 | .000 |
| | ﾓﾃﾞﾙ | 43.133 | 3 | .000 |

a 負のｶｲ2乗値は、ｶｲ2乗値が前のｽﾃｯﾌﾟから減少していることを示す。

MRAの項の表5に示したMRAの分散分析結果と同様に最終ステップが最も適合が良いわけではない。

iv）最大対数尤度による変数選択

およびv）Wald検定による変数選択

ここではLRAモデルの適合度を決める核心部分にある変数選択法をSPSSの利用を主体に解説する。なお関連語である尤度、最大（対数）尤度、log likelihood、devianceのキーワードとしての説明は第2報2）を参照していただきたい。ここでは-2対数尤度およびWald検定の追加説明をする。LRAにおける変数選択法はSPSSの場合、強制投入法、変数減少法または変数増加法（条件付、Wald、尤度比）の7種類がある。これらのうち強制投入法は変数選択法ではない。また「条件付」の方法は標本数に比してパラメータが多い、場合や例数が少ない場合に用いる特殊な方法[7]であり一般には他法と同様な結果を得るのでここでは省略する。本項では最も一般的な方法である変数減少法：尤度比と変数減少法：Waldの2つを解説する。

方法1：「変数減少法：尤度比」による方法
（＝deviance差による方法）

①SPSSのlogistic分析の7つの変数選択法のうちから「変数減少法：尤度比」法を選択する。本法は結果出力表にある「-2対数尤度の変化」を用いた選択法である。

②変数入出力基準を設定する。(例えば通法では変数投入の有意確率p＝0.15、除去p＝0.16、これはMRAにおけるFin＝2.00、Fout＝1.99の有意水準に相当する)

注1）SPSSの「-2対数尤度の変化」による方法はSTATAのdeviance差による方法と呼び方が違うだけで全く同じである。余談であるが同じ解析法でも統計ソフト間で使用する指標や呼び方が相当異なっているのが現実である。統計学にもまた統計ソフト開発に関しても様々な流派があることは仕方ないにしても最終的にユーザの理解を妨げたり混乱させていることは改善されるべきであろう。ここで、SPSSの尤度比による変数選択法の優れた点は結果表のうち「項を削除した場合のモデル」で示されるとおり各ステップで相互比較、すなわち全ての変数を一つづつ削除したモデルを比較した上で最もd「-2対数尤度の変化（＝deviance差)」が小さくなる変数を削除しているので単純な変数減少法ではない。

注2）本法はモデルを構成する個々の変数ではなくモデル全体の尤もらしさ（尤度：likelihood）を最大にする方法である。

③各説明変数の「-2対数尤度」値の自由度1の$\chi^2$検定による「変化量の有意確率」が最も低いものから逐次変数が除去されていく。

④「変化量の有意確率」が②変数入出力基準に収まると計算が終了する。

方法2：「変数減少法：Wald」による方法

①SPSSのlogistic分析の「変数減少法：Wald」法のWaldの有意確率による方法を選択。

②変数入出力基準を設定する。

②各説明変数のWald値が最小になるものから逐次変数が除去されていく。

③Wald値の有意確率が②変数入出力基準に収まると計算が終了する。

事例：

表11と表12に「歯科医療費」と「歯科保険収入」をそれぞれ目的変数として-2log likelihoodを用いた変数減少法：尤度比によるLRA分析結果を示す。係数選択の有意水準は$p = 0.15$、除去$p = 0.16$である。表11の「歯科医療費」については歯科医師10万比のみが$p = 0.002$で選択され表3に示すstep-MRAの結果が5変数を選択したのに比較して非常にシンプルな結果となっている。表12の「歯科保険収入」では主因の歯科医師10万比とFactor2は変わらないが他は入れ替わる。しかし歯科医師10万比以外は$p < 0.05$の水準を満たしておらず図2および図4は見掛けの現象ではないという仮説の検証に関してメリハリがつく。ここでMRAとの比較でLRAをどういう場合に選択するかの目安を表13に示す。本事例はケース4の場合であるが、臨床領域ではケース3に該当する事例が多いであろう。例えばある要因と最高血圧との

表11　住民1人当たりの月額歯科保険医療費の大小に係わる要因分析

ー　expenditure：backward-LRA (-2log likelihoodによる変数選択)ー

| 変数 | 回帰係数 B | 標準誤差 SE | Wald | d.f | 有意確率 | odds比 Exp (B) | 95.0% 信頼区間 下限 | 上限 |
|---|---|---|---|---|---|---|---|---|
| 歯科医師10万比 | 0.170 | 0.054 | 9.974 | 1 | 0.002 | 1.185 | 1.067 | 1.317 |
| 定数 | −11.206 | 3.402 | 10.854 | 1 | 0.001 | 1.35877E−05 | | |

※ロディスティック回帰係数：Bと単位：Kからオッズ比が求まる

Odds 比＝自然対数の底：$e^{(k \times (B))} = +\exp(k \times (B))$

95%信頼区間 (下限-上限)＝ $+\exp(k \times (B \pm 1.96 \times SE))$

本表の例では単位を10とすると上式から

odds ratio=5.474　95%CL = 1.900 - 15.775

解釈）歯科医師10万比が10上昇するごとに住民1人当たりの月額歯科保険医療費が大の群に属するodds比（優比）は5.474倍となる。

表12　歯科診療所当たりの月額保険診療収入の大小に係わる要因分析

ー　income：backward-LRA (-2log likelihoodによる変数選択)ー

| 変数 | 回帰係数 B | 標準誤差 SE | Wald | d.f | 有意確率 | odds比 Exp (B) | 95.0% 信頼区間 下限 | 上限 |
|---|---|---|---|---|---|---|---|---|
| 歯科医師10万比 | -.547 | .186 | 8.704 | 1 | .003 | .578 | .402 | .832 |
| Factor 2 | -1.415 | .729 | 3.761 | 1 | .052 | .243 | .058 | 1.015 |
| Factor 6 | .890 | .631 | 1.988 | 1 | .159 | 2.434 | .707 | 8.382 |
| 定数 | 32.041 | 10.860 | 8.704 | 1 | .003 | 8.230E+13 | | |

表13　logistic回帰分析はどんな場合に使うか

重回帰分析：MRAではどんな時適合度が悪いか，利用しづらいか

ケース1：データ値が2値（あるいは不連続）の場合

ケース2：データ値は連続量だが分布が2層性、多層性でMRAの適合度が悪い場合

ケース3：介入（医療行為等）すべきか否か意志決定しづらい場合
（客観的判別指標を得たい場合）

ケース4：急増、急減の要因分析等MRAの結果にメリハリを付けたい場合

ケース5：多変量解析データにおいてオッズ比を必要とする場合

※モデル誤差（実際値と予測値の差）に関して重回帰モデルかlogistic回帰モデルが果たして最適かどうかをGLIM法によって確認する方が望ましい。

注1）SPSSのWald値による方法はSTATAのZ値による方法と同義であり変数選択順序が同じになるがZ値はB/SE（係数／標準誤差）が標準正規分布にすることに基づく確率計算を行うが狭義のWald値は（B/SE）2が自由度1の$\chi^2$分布することを用いて確率計算する違い[7]がある。従って有意性の判定に微妙な違いがでることがあるが両者は基本的には同一のものであるからいずれもWald検定と呼んで差し支えない。

注2）本法はモデル全体ではなくモデルを構成する個々の変数の有意性を確保することを主眼とした方法である。また計算が複雑な尤度比による方法の代用[7]としての意味が大きい。

注3）方法1と方法2は変数の減少の基準指標と方法が異なるので結果も微妙に異なる場合がある。ステップごとに対象となる変数を1つづ外して比較している方法1の方が優れているのは言うまでもない。

関係がMRAで相関が高く高度に有意であってもその要因への対応がどのように必要かが臨床的に要求される。この場合最高血圧を高血圧とそれ以下に2区分してLRAを適用することにより要因の臨床的重みが判定できる。このようにLRAは意志決定理論と極めて関連性が強い統計手法である。

vi）odds比、exp（B）

odds比はモデルの適合度を判定する指標ではないがLRAの結果において中核的役割を果たすので解説することとした。ただし、基礎的なodds比（優比）の意味と解釈上の注意点については著者の総説[8]において相対危険度との関係で解説しているので参照いただきたい。ここではLRAにおけるodds比の計算と活用に留めて解説する。LRAはロディスティック分析の係数：Bから簡単にodds比が計算できる利点がある。その計算式は

Odds比＝自然対数の底：$e^{(k\times(B))}$＝exp（k×（B））

95％信頼区間（下限－上限）

＝exp（k×（B±1.96×SE））

である。ここでexpはexponential（指数関数）の略語であるがexp（x）はExcelなどで一般的に自然対数の底：e（＝2.71828）のx乗を表す関数である。SPSSのLRAにおいてはexp（B）はodds比であるが単位：k＝1の場合のodds比であることに注意する必要がある。すなわちkが1でないときには入力データのカテゴリー幅をkに変えるかExcel等で計算し直さなければならない。STATAの場合はgeneralized linear models分析等においてLogit分析を選択し「Reporting」オプションの「Report exponanentiated coefficients」を選択すると出力がodds比に変わる。単位：kをどのように決めるかは重要でありこの単位：kを適切にしないと見かけ上要因の影響力がかなり小さいように見えたりその逆であったりするので注意を要する。「ある説明変数の値がk増えると△倍増える」という表現の妥当性に注意しそれにふさわしい正数値にすることが誤解を与えない解釈をする上で望ましい。例えば喫煙と疾患のodds比の議論を1日当たり喫煙本数の単位を1、2、7、10、18、20、25、50、100本単位のいずれにするかという場合である。無論一般的には1本では比較の単位として妥当でなく10本か20本が妥当であろう。次に95％信頼区間（信頼限界）の計算に用いる1.96はt分布において自由度∞のときの両側5％になるt値（1.95996）であり、厳密には自由度の大きさによりt値が変わるのに合わせて変化させなければならないが自由度に関係なく1.96を用いて計算されることが通例である。なお、odds比の有意性の簡便な見方としてodds比が有意であれば95％信頼区間は1をまたがない。

事例：

表11の例では単位を10とすると上式からodds ratio＝5.474　95％CL＝1.900－15.775となる。解釈は「歯科医師10万比が10上昇するごとに住民1人当たりの月額歯科治療費が大の群に属するodds比は5.474倍となり、95％信頼区間は1.900から15.775倍」となる。

なお、歯科医師10万比が40から50になっても、90が100になっても変化量の単位が同じであればodds比は同じである。一方、表12の例ではodds比および95％信頼区間が0.578（0.402－0.832）で高度に有意であり解釈は「歯科医師10万比が1増加すると収入大群（保険診療の月収310万円以上）に属するodd比が0.578倍となる」となるがこの場合は倍率が1以下のため一般的でなく理解しづらい表現になってしまう。むろんこのままでも間違いではないがodds比が1より小さい場合はexp（B）の代わりにexp（-B）とすると1.728（1.200-2.488）となり、解釈は「歯科医師10万比が1減少すると収入大群に属するodd比が1.728倍となる」となる。収入小群1を2、収入大群を2を1と逆にしてから解析しても同じことになる。

vii）AIC：Akaike's Information Criterion

赤池の情報量基準を略してAIC基準という。現在の文部科学省統計数理研究所の元所長赤池弘次氏が開発した国際的に知られた使用頻度の高い統計指標である。ちなみに、Google検索（英語の条件でAIC＋Akaike）で利用頻度を確認すると50,100件（2005/4/2）抽出された。うち「dental」の用語があるのは168件であった。このように国

際的に普及しているのは線形、非線形統計モデル開発において変数選択の適否を判断する汎用性の高い基準であることに起因すると思われる。回帰分析の変数選択のように変数の個数が増えるほど当てはまりの指標である残差平方和は小さくなる．このため、変数の個数が異なるモデル間でのモデルの良し悪しを比較するためには残差平方和は不適である。モデル選択の基準として赤池の開発したAIC規準では

AIC＝-2 log（最大尤度）＋2（説明変数の個数＋1）が最小になるモデルを最適とするものであり、新たな変数を加えたとき右項：2（説明変数の個数＋1）はAICを大きくする方向に、すなわち変数増加に対するペナルティとして働く。AIC値が小さくなれば「従属変数の推定にノイズ情報と比べてプラスの情報が多い」と判断する。情報量の観点からAICが最小になるモデルが最適モデルとなる。重回帰分析の場合は変数選択基準 $F_{in}$＝$F_{out}$＝2.0でAICが最小[5]となる。ここでAICを変形したものにAIC＊nがある。AICを単純に例数：nで割ったものである。STATAで使用されているAICはこれである。通常のAICは例えばnが10倍違うとAICも10倍違ってくるがAIC＊nはそのまま比較できる利点がある。例えば研究Aではモデルaが、研究Bではモデルbが、そして研究Cではモデルcが最適モデルで例数が異なる場合の比較が容易である。、ただし、一般的にはAICはnで割らないので他の統計ソフトで求めたAICとの比較では注意を要する。どうしてもnで割らなければならない場合を除いてはnを掛けて通常のAICとして使用する方が分かりやすいと思われる。

MRAもLRAも前述のようにそれぞれ独自の適合度判定指標が開発されており、これらを用いた解析にAICは必需品ではない。しかし例えば同一のデータでMRAとLRAとどちらが適合度がいいかを比較するときはAICが力を発揮する。ある散布図の傾向を最も精確に捉えた多項式を求める場合にかという問題の場合、1次回帰よりも2次回帰が、n-1次回帰よりもn次回帰の残差（実際値－予測値）が小さくなり見かけ上適合がよくなる。このような複数の多項式の適合度を比較するときAICは必需品である。

表14に尤度（Likelihood）関数と対数尤度（Log Likelihood）関数の関係を示す。ここで自然対数を用いるのはあくまでもコンピュータ計算をし易くするためである。ある推定式の実際値と予測値の個々の差：残差がランダムである、つまり独立であると仮定すると残差の尤度関数は正規分布の式のΣを掛け算を表すΠ（パイ）にした式となる。残差1、残差2、…残差nという条件を全て満たす確率は掛け算となる。この式が最小になるときの標本分散：$\sigma^2$が母集団の分散の「最も尤もらしい」最尤推定量となる。計算を容易にするために両辺に自然対数を掛けると次の式となる。Lnは自然対数をnは例数、$\pi$は円周率を意味する。なぜ$\pi$が出てくるかといえば正規分布の確率密度関数に$\pi$が出てくるからである。

対数尤度：LnL＝（-n/2）（Ln2$\pi$＋Ln$\sigma^2$＋1）

この式の（-n/2）に着目して扱いやすくするため-2を掛けたものが－2対数尤度である。例えばある予測式を作るため、1次式：y＝b＋axから5次式：$y = b + ax + ax^2 + ax^3 + ax^4 + ax^5$までの5つの予測式の適合度を判定する場合を考えると対数尤度関数に残差の分散$\sigma^2$をいれるとそれが-2×MLE：最大対数尤度推定値となる。5つの式ごとに計算すると1次式の変数は1個、5次式の変数は5個であるから5次式の-2対数尤度が最も大きくなる。しかしそこにペナルティを課すのが2（説明変数の個数＋1）の項でありこの項の効果によ

表14　対数尤度とは？

尤度関数：the likelihood function

$$L=\prod_{i=1}^{n} f(X_i \mid \theta)=f(X_i \mid \theta)f(X_i \mid \theta)\cdots f(X_n \mid \theta)$$

対数尤度関数：log-likelihood function

$$LogL=\sum_{i=1}^{n} \log f(X_i \mid \theta)$$

例）正規分布の場合

$$L=\prod_{i=1}^{n} 1/\sqrt{2\pi\sigma^2}\exp(-(x_i-\mu)^2/2\sigma^2)$$

対数尤度：$LnL=(-n/2)(Ln2\pi + Ln\sigma^2 + 1)$

Ln：$Log_eX$：自然対数

り必ずしも5次式が最適モデルとはならない。それでは-2対数尤度は単身では使い道が無いかというとそうではない。前述のようにLRAの変数選択法で使われているとおり重要な役割がある。また-2対数尤度＝devianceであり次項で解説する。

事例：

次のデビアンスの項でまとめて解説する。

viii）deviance：デビアンス

Deviance（差）：デビアンスは本来逸脱とか乖離（かいり）、または偏差値を表す言葉であるが、P.McCullaghとJ.A. Nelder[3]によって提案されたGLIM法において最適変数モデルを選択するための特別な統計的意味を持った指標として提案された。仮定したモデルと完全なモデルとの最大対数尤度の差の-2倍で定義される尤度比検定統計量である。デビアンスは、実際値とモデルとの誤差が必ずしも正規分布ではない場合に用いる指標でモデルの誤差の分布を画一的に考えないGLIM法において変数選択の主役の役割を果たしている。疫学調査等で「ネスト化した影響評価モデル」[2]の中から、実際値に、より適合するモデルを選ぶためのモデル選択指標として利用されている。ネステッドモデルの変数の出し入れはデビアンス差が$\chi^2$分布することを利用して決める。SPSSのbackword-LRAにおいてはデビアンスの呼び方はされておらず-2対数尤度と称されているが取り扱いは同じである。変数の数の差が1であるモデル間のdeviance差は自由度1の$\chi^2$分布することを利用してモデル間の差が予め決めてある有意水準に達しない場合は変数を減少させる方法で変数選択している。AICとの使い分けは変数が入り子（nested）の関係にある場合は有意差検定ができるデビアンスを、モデルのFamily（正規分布、二項分布、ポアソン等）あるいはLink（log変換、logit変換、指数変換等）が異なるモデルの比較はAICが適切である。

事例：

表15、表16、表17にSPSSで用いられている「-2対数尤度」とSTATAで用いられている「log likelihood」devianceおよびAICの相互関係（互換性）を示す。表15はSTATAにおける関係を示し、deviance＝-2×Log Likelihoodで計算され、モデファイされたAIC（AIC＊n）が計算される。しかしSPSSにおいては表16に示す指標のうち「-2対数尤度」しか登場しない。ここで尤度関数に最も

表15　Logistic回帰分析における最大対数尤度とAIC、devianceとの関係

－ 統計ソフト STATA8.2 利用に際して －

－ income：backward-LRA（-2log likelihoodによる変数選択）－

| n | Maximum Log Likelihood（最大対数尤度） | 説明変数の数：k ＋ 1 K+1 | AIC*n | deviance =-2最大対数尤度 |
|---|---|---|---|---|
| 47 | -9.526627487 | 6 | 0.6607076 | 19.05325497 |
| 47 | -9.875266655 | 5 | 0.6329901 | 19.75053331 |
| 47 | -10.14449719 | 4 | 0.6018935 | 20.28899438 |
| 47 | -11.43206368 | 3 | 0.6141304 | 22.86412736 |
| 47 | -14.47475702 | 2 | 0.7010535 | 28.94951404 |
| 47 | -31.71087039 | 1 | 1.3919519 | 63.42174078 |

注1）Logistic回帰分析におけるMLEs:最大対数尤度推定値の計算

$L_1$:実際値が0の場合→Ln(1-推定値)
$L_2$:実際値が1の場合→Ln(推定値)　} $L_1$と$L_2$の総和が最大対数尤度

注2）AIC*n=(-2最大対数尤度＋2×(k+1)) / n

注3）deviance=-2最大対数尤度

表16　-2対数尤度とdevianceおよびAICとの関係

－ SPSS12.0使用の場合 －

－ income：backward-LRA（-2log likelihoodによる変数選択）－

| n | -2対数尤度 =deviance | 説明変数の数+1 K+1 | AIC |
|---|---|---|---|
| 47 | 20.28899266 | 4 | 28.28899266 |

AIC=(-2最大対数尤度＋2×(k+1))
=(20.28899266+2×4) = 28.28899266

AIC＊n=(-2最大対数尤度＋2×(k+1)) / n
=(20.28899266+2×4) / 47 = 0.6018935

注1）SPSSの「-2対数尤度」の標記は「-2最大対数尤度」と同じ

注2）STATAのAICは厳密にはAIC･nでありAICを例数nで除したものである。

表17　最適モデルはAICが最小になるときに一致

－ SPSS12.0J使用の場合 －

－ income：backward-LRA（-2log likelihoodによる変数選択）－

| ｽﾃｯﾌﾟ | -2 対数尤度 MLE | Cox & Snell $R^2$ | Nagelkerke $R^2$ | 変数の数 k | 2(k+1) | AIC |
|---|---|---|---|---|---|---|
| 1 | 13.9390 | 0.65 | 0.88 | 11 | 24 | 37.94 |
| 2 | 13.9980 | 0.65 | 0.88 | 10 | 22 | 36.00 |
| 3 | 14.3930 | 0.65 | 0.87 | 9 | 20 | 34.39 |
| 4 | 14.6969 | 0.65 | 0.87 | 8 | 18 | 32.70 |
| 5 | 14.7923 | 0.64 | 0.87 | 7 | 16 | 30.79 |
| 6 | 16.4626 | 0.63 | 0.85 | 6 | 14 | 30.46 |
| 7 | 17.5027 | 0.62 | 0.84 | 5 | 12 | 29.50 |
| 8 | 19.0244 | 0.61 | 0.83 | 4 | 10 | 29.02 |
| 9 | 20.2890 | 0.60 | 0.81 | 3 | 8 | 28.29 |
| 10 | 22.8641 | 0.58 | 0.78 | 2 | 6 | 28.86 |
| 11 | 28.9495 | 0.52 | 0.70 | 1 | 4 | 32.95 |

注）SPSS12.0Jの場合AICは「-2対数尤度」と「変数の数」から計算しないと求められない

尤もらしい母数の推定値：MLEを代入して計算しているから「-2対数尤度」は前述のように取りも直さず「-2最大対数尤度」である。よってSPSSでもAICおよびAIC＊nは表16の計算式で簡単に計算できる。表17は「歯科保険収入」を目的変数とするSPSSのLRAで最適モデルで示された第9ステップと前述の擬R2およびAICの関係を示している。SPSSは「-2対数尤度」の差が自由度1の$\chi^2$分布をすることを利用して最適モデルを変数の数が3つの第9ステップとした。このときのAICが第1～11ステップ中最小値になっていることがわかる。このようにモデルが入り子関係にある場合は-2対数尤度（deviance）であれAICであれ最適モデルは同じものが得られる。しかし擬R2では最適モデルの選定はできない。

## 4. おわりに

MRAはもはや時代遅れの分析法で時代はLRA分析だという風潮があるやに聴く。しかし事実とすれば統計分布を無視した暴論である。確かに前報でも確認したようにLRAは優れた分析法であることは間違いないが扱う数値が連続量で残差（誤差）がきれいな正規分布をするものをLRAを適用するためだけの理由でわざわざカテゴリー化して機械的にLRAを適用することは重大な情報損失を招くので戒められなければならない。本シリーズで解説してきたように第1にその用途により、その上で生物学的妥当性と分布の適合度でMRAとLRAの利用を使い分けるべきである。またできればモデル誤差（実際値と予測値の差）に関して重回帰モデルかlogistic回帰モデルが果たして最適かどうかをGLIM法によって確認した方がより望ましい。

最適な統計分布、最適モデルの選択の必要性とその方法を3回のシリーズを通して解説してきた。多分に舌足らずで十分に語り尽くせたか自信がないが医学・歯学の分野においても疾病構造の変化と医療の概念の進化に後押しされてこの問題を曖昧に出来ない時代が到来しているとことを改めて強調したい。

謝辞：本総説を執筆するにあたりご指導いただいた歯科疫学研究会の顧問である重松逸造、蓑輪眞澄、高江洲義矩、境　脩、伊藤学而、佐々木英忠の各先生に深謝申し上げます。さらに貴重なご助言をいただいた深井穫博同研究会会長をはじめとする幹事の方々に感謝申し上げます。

### 参照Web-site（2005/4/2現在）

WS01）http://courses.bus.ualberta.ca/MGTSC312-Jobson/Information/StatTables1.xls

WS02）http://gcrc.uchsc.edu/Documents/StatsClass/PresentSlides/GCRC％20Data％20Analysis％20notes％205．ppt＃298，16，スライド　16

WS03）http://www.bucknell.edu/img/assets/4778/SPSS_Manual.pdf

WS04）http://www2.chass.ncsu.edu/garson/pa765/logistic.htm

WS05）http://endeavor.eng.toyo.ac.jp/~yoshino/lecture/keikaku2/node5.html＃SECTION00054100000000000000

### 文　献

1）瀧口　徹：歯科疫学統計－第1報　各種統計分布の相互関係と利用の潮流－Health Science and Health Care, No1 vol 4：19-27，2004.

2）瀧口　徹：歯科疫学統計－第2報　一般化線形モデルの意義と潮流－Health Science and Health Care, No1 vol 4：29-36，2004.

3）P McCullagh, J. A. Nelder：Generalized Linear Models. Chapman & Hall/CRC.., USA., 2nd ed., pp.7-8，30-31，33-36，1999.

4）矢野恒太記念会編：データでみる県勢　2004，矢野恒太記念会，東京，第3版，1-512頁，2004.

5）杉山高一：多変量データ解析入門，朝倉書店，東京，第1刷，134-136，150-151頁，1983.

6）STATA CORPORATION（College Station Texas）STATA REFERENCE MANUAL G-M P35-36

7）小野寺孝義，山本嘉一郎編：SPSS辞典－BASE編－，ナカニシヤ出版，東京，142頁，2004.

8）丹後俊郎，山岡和枝，高木晴良：ロジスティック回帰分析　SASを利用した統計解析の実際－，朝倉書店，東京，第1刷，195-196，200-202頁，1996.

9）瀧口　徹：EBMのための（臨床）疫学・統計学的基礎（2）第2章　疫学の基礎；流行病の法則性を見つけ予防する，障害者歯科学雑誌，23巻：89-98，2002.

ヘルスサイエンス・ヘルスケア　Volume 5，No.1（2005）

# A review of oral epidemiological statistics

## － Part III ： Interpretation of various goodness of fit indicators for the Multiple Regression Model and Multiple Logistic Regression Model. －
## － When using the statistical software SPSS and STATA －

Toru Takiguchi
（Fukai Institute of Health Science）

Abstract: As I mentioned in the previous two reviews, a very small variance of residuals between the observed values and the expected values （that is to say, a high goodness of fit） is a required condition for obtaining the optimum statistical model. However, due to wide variation in usage and technical terminology, not to mention the various indicators used to measure goodness of fit, it is no simple task for the user to gain an accurate understanding of the concepts and usage of statistical software.

In the third review, goodness-of-fit indicators for Multiple Regression Analysis（ MRA） and Multiple Logistic Regression Analysis（LRA）, which have typical multivariate analysis usage frequencies, were interpreted according to the outputs of two commonly used statistical software programs, SPSS 13. 0J and STATA 8. 2.

For the MRA, R-squared（ Coefficient of Determination）, ANOVA, Durbin-Watson Statistic and VIF（ Variance Inflation Factor）（Tolerance） were interpreted. For the LRA, Pseudo-R-squared（Cox & Snell R-squared and Nagelkerke R-squared）, Hosmer-Lemeshow Test and Omnibus test were interpreted. For variable selection, a comparison of he maximum log likelihood method and the Wald method was conducted. In addition, methods of assessing each model's goodness of fit using Akaike's Information Criterion（AIC） and Deviance were discussed, and for LRA the odds ratio calculation method was also explained.

For a deeper understanding of concrete application, a case study was conducted to assess the interrelationship of the following three factors：“The number of dentists per 100,000 people（Dentists-Ratio）”, “Monthly dental insurance expenditure per person（Expenditure/Person）”, and “Monthly insurance income per dental clinic（Income／Clinic）”. The results of both MRA and LRA indicated that Expenditure/Person very significantly increased with Dentists-Ratio. On the contrary, Income/Clinic very significantly decreased with Dentists-Ratio after adjustments for confounding factors. The odds ratio with LRA indicated that the weight of Dentists-Ratio as a contributing factor to both Expenditure/Person and Income/Clinic became clearer.

Key words: Multiple Regression Analysis, Multiple Logistic Regression Analysis, Goodness of Fit, -2 Log Likelihood, AIC

Reprint requests to T. TAKIGUCHI, Fukai Institute of Health Science, 3-86, Hikonari, Misato-shi, Saitama 341-0003, Japan
TEL:048-957-3315/FAX:048-957-3315/E-mail:taki8020@mth.biglobe.ne.jp