

最尤法による
計数値・比率データの統計解析
メタ・アナリシスを事例として

第 22 回 高橋セミナー

2005 年 10 月 29 日

高橋 行雄

10/31/2005 8:04 PM

表紙裏

2005 年 6 月 16 日，新規作成，

ファイル名：H:¥R1_進行中_from_2005_6_16¥R1_ANZ_S21_JMP による効力比¥S21_2 値_順序_尤度関数.doc

尤度関数を用いた 2 値・順序データの解析

目 次

1. はじめに	1
2. 統計モデルを用いた 2 値データの解析	3
2.1. インフルエンザの予防	3
2.2. アマンタジンを用いたインフルエンザの予防	4
SASによる解析	5
GENMODプロシジャによる解析	5
正規分布のあてはめの妥当性	8
JMPを用いた解析	9
2.3. オッズ比による有害事象の比較	11
オッズ比とは	11
LOGISTICプロシジャによる解析	12
ロジスティック分布のあてはめ	14
正規分布とロジスティック分布のあてはめ <<以下, zをwに変換する>>	16
JMPによるオッズ比	18
2.4. インフルエンザ・ワクチンの地域別の予防効果	19
GENMODプロシジャによる解析	19
対数オッズとオッズ比の推定	21
推定値のグラフ表示	23
モデルの妥当性	25
JMPによる解析	27
3. 層をなす分割表	29
3.1. オッズ比	29
3.2. 研究を層とする分割表	31
JMPによるオッズ比, 相対リスク	35
3.3. CMH法とロジット法による共通オッズ比	37
3.4. Petoの方法による共通オッズ比	39

3.5. さらなる共通オッズ比の推定方法と選択基準.....	41
JMPによる最尤法でのオッズ比.....	43
3.6. 対象集団間のオッズ比の比較.....	43
4. 反応が順序データ	45
4.1. 田口の累積法の考え方.....	46
4.2. 多カテゴリーの反応変数（カテゴリーの順序に意味あり）	50
JMPによる順序データの解析	55
3 水準への拡張, 水準間の比較.....	55
4.3. 2 因子実験.....	59
JMPでの解析	64
5. 繰り返し実験がある 2 値反応, モデルの適合	65
5.1. 繰り返し実験.....	65
5.2. 因子が質的な場合のロジスティック回帰の基礎.....	68
切片のみのモデル.....	68
対象集団ごとのあてはめ	69
7 試験をすべて個別の集団と仮定	70
分散分析表スタイル.....	71
5.3. 理解を深めるためのGENMODプロシジャ	72
切片のみのモデル.....	72
水準平均モデル.....	73
完全フィット.....	74
5.4. 最小 2 乗法による分散分析との対比.....	75

図 表 目 次

表示 2.1	アマンタジンのプラセボに対するインフルエンザの予防効果	4
表示 2.2	アマンタジン投与後のインフルエンザ発病率	5
表示 2.3	正規分布のあてはめ	6
表示 2.4	GENMODプロシジャの回帰係数推定値	7
表示 2.5	正規分布のあてはめ	7
表示 2.6	正規分布のパーセント点 y (プロビット) に対する用量反応関係	9
表示 2.7	アマンタジンの有害事象	11
表示 2.8	オッズ比の計算	12
表示 2.9	中枢神経系の有害事象についてのオッズ比の推定	14
表示 2.10	Calssステートメントの参照型ダミー変数	14
表示 2.11	プラセボ群と 200mg群にロジスティック分布のあてはめ	16
表示 2.12	ロジスティック分布と正規分布のあてはめ	17
表示 2.13	プラセボ群と 100mg群にロジスティック分布のあてはめ	17
表示 2.14	乳幼児を対象にしたインフルエンザ・ワクチンの予防効果	19
表示 2.15	ブロックを固定効果とした解析プログラム	20
表示 2.16	ワクチン投与後のインフルエンザ発病に対するロジスティック回帰	22
表示 2.17	非接種群に対する接種群の対数オッズ比とオッズ比の推定	22
表示 2.18	地域が岩手の場合のロジスティック分布	24
表示 2.19	地域別に同一形状のロジスティック分布のあてはめ	24
表示 3.1	オセタミビルインフルエンザの予防効果	29
表示 3.2	オッズ比の信頼区間	30
表示 3.3	オッズ比および相対リスクの計算結果	31
表示 3.4	複数の臨床研究	32
表示 3.5	層がある場合の共通オッズ比の推定	33
表示 3.6	研究ごとのオッズ比の 95%信頼区間	33
表示 3.7	FREQプロシジャによるCMH検定統計量, 共通オッズ比の推定	34
表示 3.8	第 k 層の 2×2 表	37
表示 3.9	共通オッズ比を求めるための計算シート	38
表示 3.10	Petoのオッズ比の計算シート	40
表示 3.11	SASのデータデータステップを用いたピトーの共通オッズ比の計算	40
表示 3.12	最尤法および層を考慮した最尤法 (条件付き)	42
表示 3.13	一元配置枝分かれ型分割・反復測定実験	44
表示 4.1	順序がある破壊データ	46
表示 4.2	2 元配置とみなしたデータ構造	47
表示 4.3	累積比率 $P1$, $P2$ と x の関係	47
表示 4.4		48
表示 4.5		49
表示 4.6	推定値のグラフ	50
表示 4.7	臨床試験データ (人工データ)	50
表示 4.8		51
表示 4.9		52
表示 4.10		53
表示 4.11		56
表示 4.12		57
表示 4.13	水準間の対数オッズの差およびオッズ比の推定	57
表示 4.14	2 因子実験のデータ	59
表示 4.15	累積パーセント	59
表示 4.16	2 因子実験	60
表示 4.17	交互作用の検討	60
表示 4.18	主効果のみでの解析結果	61
表示 4.19	オッズ比で 4.934 倍の反応率	62
表示 4.20	効果の推定値	63
表示 4.21	JMPの解との比較	63

表示 5.1	一元配実験・繰り返しあり	65
表示 5.2	過大な変動の組み込み	66
表示 5.3	一般成人を基準にしたオッズ比の推定	66
表示 5.4	集団内の変動を無視したオッズ比の95%信頼区間	67
表示 5.5	デビエンスの不均一性	67
表示 5.6	対数尤度より χ^2 統計量の算出図	71
表示 5.7	分散分析表スタイル	72
表示 5.8	切片のみのモデル	73
表示 5.9	水準平均のモデル	74
表示 5.10	完全フィットのモデル	75
表示 5.11	一元配実験・枝分かれ・反復測定	75
表示 5.12	一元配実験・枝分かれ・反復測定の分散分析表	76
表示 5.13	GLMプロシジャを用いた解析	76
表示 5.14	GLMプロシジャを用いた分散分析表	77
表示 5.15	ロジスティック回帰の分散分析表スタイル	77

1. はじめに

皆様へ、10月29日(土)の104回安全研の前に同じ会場で午前10時から第22回のセミナーを行います。テーマは、「最尤法による計数値・比率データの統計解析」です。

事例として取り上げるのは、アマンタジンの実験的インフルエンザ予防効果と一般成人を対象としたDBTの論文の結果を元に、オッズ比とその95%信頼区間による要約の適切性について平易かつ理論的なアプローチをします。有害事象についてもオッズ比による要約の適切性についても論じます。2番目の事例は、日本で行われた乳幼児を対象にしたインフルエンザ・ワクチンの効果について地域を層にしたデータを用いて、ダミー変数を含むロジスティック回帰からオッズ比を算出する考え方、オッズ比の活用法を論じます。これらは、第6回で取り上げた「計数値・比率のデータ解析」の改定版です。

さらに、第15回で取り上げたメタ・アナリシスについては、タミフルのインフルエンザの予防効果をさらに発展した内容とし、ロジスティック回帰における過大発散とその対処の考え方を論じます。これは、生殖試験における同腹効果と共通の切り口です。この例を用いて、分散分析のアプローチによる最尤法入門とします。

プラセボを対象にした複数のダブルブラインド試験の結果を併合する方法としたメタ・アナリシスが広く行われている。第15回のセミナー「比の統計学入門ー生物検定法は比の統計学であったー」で、反応が2値の場合のメタ・アナリシスで、オッズ比による統合が一般的に行われていることを示した。

中外製薬での臨床統計の教育にもメタ・アナリシスを取り入れ、現代のEBM（エビデンスに基づく治療）を支える統計的方法として、オッズ比を取り入れてきたのであるが、さらにブラッシュアップした教材の開発に取り組んできた。今回は、様々な比較におけるオッズ比、その解釈、使い方などについて述べたい。

第15回のセミナーでの内容は次のようなものであった。

- 2. 出現率の差および比の信頼区間 6
- 2.1. 信頼区間の計算、おさらい 6
- 2.2. 出現率の差および比の信頼区間 8
- 2.3. ロジスティック回帰分析によるオッズ比の併合 11

2.4.	オッズ比の併合	14
2.5.	EBM メタ・アナリシスでのオッズ比	17
2.6.	コクラン・ライブラリ	17
	コクラン ライブラリの中身	17
2.7.	丹後のメタ・アナリシスの紹介	23
2.8.	Peto の方法	24
2.9.	漸近分散法ーオッズ比	26
2.10.	Mantel-Haenszel のオッズ比	27
2.11.	DerSimonian-Lairdーの方法ーオッズ比	28
2.12.	ロジスティック回帰によるオッズ比の併合	29
	PROC GENMOD 母数モデル	29
	PORC LOGISTIC 母数モデル	29
	JMP によるロジスティック回帰	30

このセミナーでは、メタ・アナリシスで用いられている統計的方法のオーバービューあった。今回は、2 値データの統計解析の入門書となるように数理的な側面と理解を深めるための事例を充実させた。

2. 統計モデルを用いた 2 値データの解析

2.1. インフルエンザの予防

2 値反応データの典型例として、身近な急性感染症であるインフルエンザの発病の予防効果を取り上げる。インフルエンザが毎年のように世界的に流行している。マラリヤ、赤痢などの多くの感染症は、先進国で大流行することは極めてまれになったが、インフルエンザは、冬になると世界で同時多発的に流行を繰り返している。多くの人がインフルエンザ・ワクチンの予防接種を受けているが、その流行を抑えることが困難な状況が続いている。

このような背景により、インフルエンザを予防する抗ウイルス薬も新たに開発されてきた。これらの薬が、本当にインフルエンザを予防することができるのか、多くの国でプラセボ¹を対照とした臨床試験²がおこなわれている。一つ一つの臨床試験は、薬効を検証するため対象集団を限定することが一般的である。小児・子供・成人・高齢者それぞれに限定して臨床試験が行われている。また、家族のだれかが感染したときの両親・祖父母に対する予防効果はどうか、子供同士はどうかなど、実験条件を限定して、プラセボと実薬間での予防効果が測定されている。実際に抗ウイルス薬を予防投与する患者さまは多様であり、どのくらいの予防効果が見込まれるのであろうか。

一つ一つの臨床試験では、プラセボに対して実薬の効果が統計的な有意差が得られなくても、複数の臨床試験の結果を統計的に統合して、プラセボに対して実薬の効果があるのか。あるいは、複数の臨床試験でプラセボに対する実薬の平均的な効果を推定し、一般的な診療ガイドラインを作成したいとの要求もある。そのような社会的な必要性により、別々に行われた複数の臨床試験の結果を統計的に併合した解析が頻繁に行われている。この解析は、メタ・アナリシスといわれており、丹後 (2002)、Whitehead (2002)、Petitti (2000) による成書もあり応用統計の一分野となっている。

丹後 (2002), メタ・アナリシス入門, 朝倉書店

Whitehead, A.(2002), Meta-Analysis of Controlled Clinical Trials, Wiley.

Petitti,D.B. (2000), Meta- Analysis, Decision Analysis, and Cost-effectiveness Analysis 2nd ed.
Oxford university press.

¹ 有効成分が入っていないが外観・味などが実薬と同じ薬。

² 医師・患者がプラセボか実薬かが実験中に分からないように二重にブラインドしている試験をする。

2.2. アマンタジンを用いたインフルエンザの予防

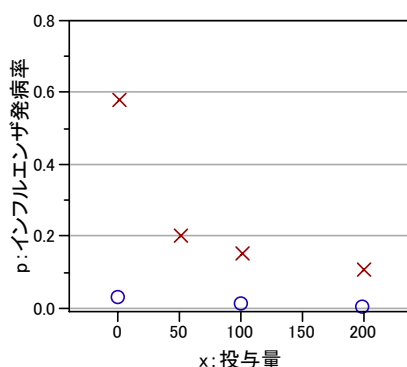
Reumanら（1989）は、抗パーキンソン氏病の治療薬として使われているアマンタジンがA型インフルエンザを予防する効果があるか、ダブルブラインドによる臨床試験を行った。アマンタジンのプラセボに対するインフルエンザの予防効果を 表示 2.1 に示す。この論文には、2つの対象集団が含まれている。第1は、18歳から55歳までの成人を対象にしたインフルエンザ発病予防試験であり、第2の試験は、18歳から40歳の主に大学の学生を対象とした実験的インフルエンザ発病予防試験である。

成人を対象にした予防試験のプラセボ群のインフルエンザ発病率が3.1%と低いが、この試験の計画時に、予備試験から発病率を30%と見積もって例数設計をしており、適切な検出力が確保された症例数であると述べられていた。実験的予防試験では、プラセボのインフルエンザ発病率は90%以上、50mg群では30%以下、100と200mg群では20%以下と見込んで、検出力が99%と少数例でも効果を検証できると述べられていた。

表示 2.1 アマンタジンのプラセボに対するインフルエンザの予防効果

Ex : 試験	x : 用量(mg)	発病あり	発病なし	n	発病率	相対リスク
Ex1 : 成人予防	プラセボ	5	154	159	3.1%	100.0%
	100	2	157	159	1.3%	40.0%
	200	1	157	158	0.6%	20.1%
Ex2 : 実験的予防	プラセボ	11	8	19	57.9%	100.0%
	50	4	16	20	20.0%	34.5%
	100	3	17	20	15.0%	25.9%
	200	2	17	19	10.5%	18.2%

実験条件によりプラセボの投与による発病率が明らかに異なるので、プラセボの発病率を100%とする相対リスクで比較してみる。200mg投与群では Ex1 では20%、Ex2 では18%と同程度の効果であることが確認できる。表示 2.2 に示すようにEx2 の用量反応関係は、シグモイド曲線があてはめられそうである。シグモイド曲線として正規分布をEx1とEx2にあてはめてみる。



表示 2.2 アマンタジン投与後のインフルエンザ発病率

SASによる解析

SASの多くのプロシジャ（プログラム）は、解析用のSASデータセットが1症例ごとの反応 y が1行となるような形式を前提にしている．あるいは、同じ反応 y の症例数が n 例あるというような集計されたデータセットを前提としている．表示 2.1 のデータは、インフルエンザが発病した場合を $y=1$ 、発病しなかった場合を $y=0$ として試験Exごと、投与量 x ごとの症例数をSASデータセット化する．

2 値反応データの場合には、異なるデータセットの形式が許されている．それは、反応ありの症例数を r とし、その分母となる症例数を n とし、 r / n を“反応 y ” とみなす形式である．

反応率に正規分布をあてはめる機能を持つ SAS のプロシジャには、PROBIT プロシジャ、LOGISTIC プロシジャ、および、GENMOD プロシジャである．PROBIT プロシジャは、正規分布をあてはめた後に、ある反応 y_0 となるような用量 x とその信頼区間推定する、いわゆる逆推定の機能を持っている．LOGISTIC プロシジャは、2 値データに対する“回帰分析”的アプローチに特徴があり、GENMOD プロシジャは、“分散分析”的アプローチに特徴がある．どちらのプロシジャも機能拡張により、使い方にかなり共通点があり、出力結果にも共通点が多い．本書では、GENMOD プロシジャと LOGISTIC プロシジャを計算の目的に応じてより簡潔な出力が得られプロシジャを例示する．この3つのプロシジャでも解析できない事例についてはNLMIXED プロシジャを用いる．

GENMODプロシジャによる解析

SAS データセットの作成と GENMOD プロシジャを用いた SAS プログラムを表示 2.3 に示す．SAS データセットの作成は、プログラムの中の datalines ステートメントとセミコロン ; の行間に埋め込んでデータから読み込む形式を用いている．変数は、input

ステートメントの変数リスト形式でデータを 1 行ごとに読み込み、SAS データセット `amantadine` に書き込んでいる。セミコロンを読み込むとデータが尽きたと判断され、SAS データセットが完成する。作られた SAS データセットは、7 行 5 列の矩形状である。

SAS データセットは、SAS により一時的な WORK フォルダに `amantadine.sas7bdat` (version9, Windows 版) として書き込まれており、GENMOD プロシジャでは、`data=amantadine` として使う。SAS データセットを表示するためには、PRINT プロシジャを用いるが、プログラムリストでは省略している。

説明変数として GENMOD プロシジャは、連続変数と分類変数が扱え、分類変数の場合は、class ステートメントで、`class Ex ;` のようにプログラムすると、分類数に応じたインディケータ型³ダミー変数が内部で生成される。Model ステートメントでは、`r/n = Ex x ;` のように応答変数として `r/n` を用い、説明変数として試験 `Ex` と投与量 `x` を用いている。試験 `Ex` はインディケータ型の 2 つのダミー変数として展開され、説明変数は、切片を含み 4 変数となり、7×4 のデザイン行列が内部で生成される。

Model ステートメントでは、さらに / の後に `link` と `dist` オプションで応答変数の変換方法と誤差の分布の指定をしている。応答変数の変換方法は正規分布なので `dist=probit` とし、誤差の分布は、2 項分布を仮定するので `link=binomial` を指定している。

表示 2.3 正規分布のあてはめ

```
data amantadine ;
  input Ex $ x r n_r n ;
datalines ;
Ex1_成人    0   5 154 159
Ex1_成人   100   2 157 159
Ex1_成人   200   1 157 158
Ex2_実験    0  11   8  19
Ex2_実験   50   4  16  20
Ex2_実験  100   3  17  20
Ex2_実験  200   2  17  19
;
proc genmod data=amantadine ;
  class Ex ;
  model r/n = Ex x / link=probit dist=binomial ;
run ;
```

表示 2.4 に GENMOD プロシジャの実行結果を示す。切片 `Intesept`, 試験 `Ex` については、

³ 変数 Ex の分類番号を j としたとき、 $Ex_i(j) = \{1 \text{ if } Ex = j, 0 \text{ if } Ex \neq j\}$ 。ここで、 i は症例番号。

Ex1_成人とEx2_実験の2水準(インディケーター型のダミー変数を w_1 および w_2 とする), 投与量 x の回帰係数の推定値が解析結果として示されている. これは, 次の線形回帰式

$$\eta_i = \beta_0 + \beta_1 w_1 + \beta_2 w_2 + \beta_3 x_i$$

の η_i が, インフルエンザの発現率 $p_i = r_i / n_i$ に正規分布,

$$p_i = \text{Normal Distribution}(\eta_i) + \varepsilon_i$$

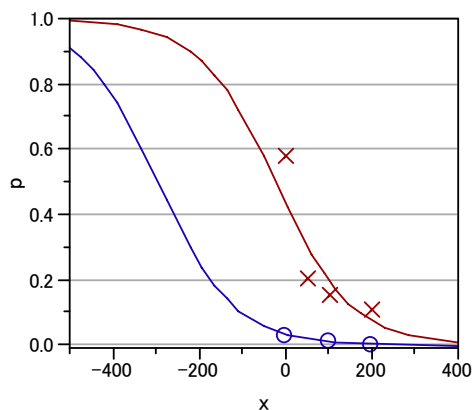
をあてはめたときの結果である. ここで, 誤差 ε_i は, 2 項分布に従うとしている.

回帰係数は, $\hat{\beta}_0 = -0.2725$, $\hat{\beta}_1 = -1.5068$, $\hat{\beta}_2 = 0.0000$, $\hat{\beta}_3 = -0.0050$ と推定値される.

表示 2.4 GENMODプロシジャの回帰係数推定値⁴

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept		1	-0.2725	0.1960	-0.6565 0.1116	1.93	0.1643
Ex	Ex1_成人	1	-1.5068	0.2199	-1.9378 -1.0759	46.97	<.0001
Ex	Ex2_実験	0	0.0000	0.0000	0.0000 0.0000	.	.
x		1	-0.0050	0.0016	-0.0081 -0.0018	9.70	0.0018
Scale		0	1.0000	0.0000	1.0000 1.0000		

Ex1_成人については, $\hat{p}_i = \text{Normal Distribution}(-0.2725 - 1.5068 - 0.0050 \cdot x_i)$, Ex2_実験については, $\hat{p}_i = \text{Normal Distribution}(-0.2725 - 0.0050 \cdot x_i)$ が推定値となる. この分布にしたがう推定値 \hat{p}_i と実験で得られた p_i を重ね書きした結果を 表示 2.5 に示す.



表示 2.5 正規分布のあてはめ

用量 x の回帰係数が $\hat{\beta}_3 = -0.0050$ と負であるので, 逆の正規分布となっている. 投与

⁴ Ex2_実験の自由度 DF が 0 の意味については第@@章で, Parameter 列の Scale については第@@章で詳しく述べる.

量 x が 0mg のときのインフルエンザ発現率の推定は、Ex1_成人については、

$$\hat{p}_{0\text{mg}} = \text{Normal Distribution}(-0.2725 - 1.5068) = \text{Normal Distribution}(-1.7793) = 0.0376,$$

Ex2_実験については、

$$\hat{p}_{0\text{mg}} = \text{Normal Distribution}(-0.2725) = 0.3926,$$

となることが 表示 2.5 から確認できる。

投与量 x の回帰係数 $\hat{\beta}_3 = -0.0050$ の SE は 0.0016 であり、その 95%信頼区間は $(-0.0081, -0.0018)$ と 0 を含まないので、統計的に有意な用量反応関係があるとみなされる。

正規分布のあてはめの妥当性

アマンタジンの臨床論文に含まれる 2 つの異なる条件設定の臨床試験を総合した結果から、アマンタジン投与量を増加するとインフルエンザの発病を統計的に有意に予防できるとの解析結果が得られたのであるが、さらに詳しく探索的に検討してみよう。探索的な解析は、標準治療を模索するための臨床上の疑問、さらなる臨床試験を計画する場合の科学的な根拠をあたえる。

一般成人を対象にした予防試験 Ex1 と、インフルエンザ・ウィルスを強制的に接種した予防試験 Ex2 場合に、プラセボ群（アマンタジンが含まれない製剤、アマンタジンの投与量は 0mg）のインフルエンザ発病率が、3.1%と 57.9%と明らかに異なる。試験 Ex1 だけに限定すると、どのような探索的な解析を行っても統計的に有意な差があるという結果を得られない。これは、実験計画でプラセボ群のインフルエンザ発病率を 30%と仮定した例数設計に起因する。明らかなインフルエンザ発病率の過大予測であり、そのために症例数設定が少なくなり統計的には有意な差が得られなかったのである。試験 Ex2 では、症例数の設定は各群 20 例ほどであるが、低用量の 50mg 群からプラセボ群に対して有意な差となる。期待以上にアマンタジンのインフルエンザ発病の抑制効果があったのであろう。これらの経験は、次の臨床試験の計画の基礎データを与えることになる。

用量反応関係に正規分布をあてはめたのであるが、表示 2.5 をみると奇異である。外挿とはいえ投与量が負になるときにインフルエンザ発病抑制が 100%となるような反応を仮定することは、適切な統計モデルといえるのであろうか。一般成人を対象にした試験でプラセボのインフルエンザ発病率が、100%になることは、実験前から想定もされていない。用量反応曲線としては、投与量が 0mg のインフルエンザ発現率に収束するような妥当のように思われる。

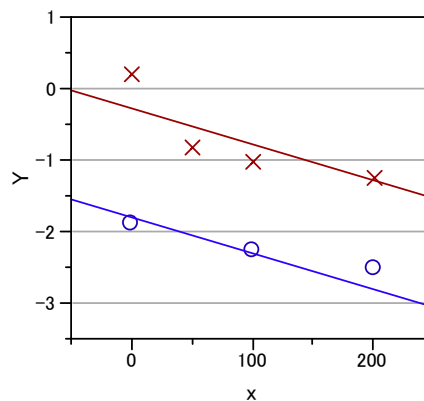
正規分布のあてはめの前の、線形回帰式、

$$\hat{\eta}_i = \hat{\beta}_0 + \hat{\beta}_1 w_1 + \hat{\beta}_2 w_2 + \hat{\beta}_3 x_i$$

の $\hat{\eta}_i$ についての結果を 表示 2.6 に示す. 実際の発現率 $p_i = r_i / n_i$ は, 逆正規分布

$$\eta_i = \text{Normal Quatile}(p_i)$$

によって変換した点をプロットしている. 2 つの試験の用量 x の直線の傾きは共通なので, 平行線のあてはめとなっている. この図だけ見ていると, 2 本の平行な回帰直線のあてはめは妥当のように思ってしまう.



表示 2.6 正規分布のパーセント点 y (プロビット) に対する用量反応関係

そもそも, 用量反応関係が, シグモイド曲線的になるとの生物学的な経験則は, 対数用量と反応の関係に対するもので, 投与量が 0mg は陰性対象としての役割であり, 対数用量反応関係では, 0mg の対数は負の無限大となるので, 他の用量とは別に扱うべきものである. これは, 0mg の反応の推定から, 極大用量での反応の推定値の間でシグモイド曲線をあてはめるのが妥当である. 詳細は@@章で論ずる.

Reuman,P.d., Bernstein,D.I., Keefer,M.C., *et al.* (1989) Efficacy and safety of low dosage amantadine hydrochloride as prophylaxis for influenza A, Antiviral Research, 11,27-40.

JMPを用いた解析

これまで, 正規分布にかえてロジスティック分布 (0 から 1 の範囲のロジスティック曲線) をシグモイド曲線として使ってきた. しかしながら, 正規分布を最初の説明としては, 統計の基礎知識を前提とするので, ロジスティック分布を先に出すよりも説明もしやすいし, 新規に学習する人たちに直感的な理解が得られると思われる. また, 出現率をプロビット変換して解く伝統的な記述は冗長であるので, 直接正規分布をあてはめるような手順に書きあらためた.

第 21 回のセミナーで、JMP で正規分布のあてはめを行うためには、非線形回帰分析を使うことを示してきた。復習をしてみよう。

データは、「アマンタジン_normal.JMP」にある。

アマンタジン_normal										
	条件	z1	z2	x	r	n	p	p [^]	like	
1	成人予防	1	0	0	5	159	0.031	0.03836357	22.3275348	
2	成人予防	1	0	100	2	159	0.013	0.01160379	10.7452944	
3	成人予防	1	0	200	1	158	0.006	0.00280281	6.31779087	
4	実験的	0	1	0	11	19	0.579	0.39358013	14.258638	
5	実験的	0	1	50	4	20	0.200	0.30153179	10.5373691	
6	実験的	0	1	100	3	20	0.150	0.22064995	8.77154787	
7	実験的	0	1	200	2	19	0.105	0.10204232	6.39448506	

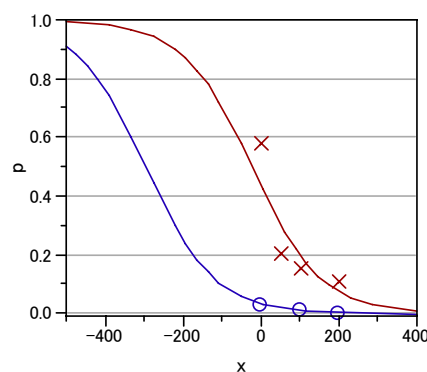
非線形回帰を実行して、推定値を求めなさい。

解

	損失	DFE	平均損失	平均損失平方根	
	79.351058677	4	19.837765	4.4539606	
パラメータ	推定値	近似標準誤差	下側信頼限界	上側信頼限界	
beta0	-0.272491037	0.19595031	-0.6560277	0.11317264	
beta0_1	-1.506828201	0.21986353	-1.9481885	-1.083067	
beta1	-0.004966075	0.00159421	-0.0082739	-0.0019927	
次の値で解く: 解析 NR					

SE は GENMOD プロシジャと比べると微妙に異なる。これは、最尤法の反復計算で JMP はフィsher のスコア法を、GENMOD は、ニュートン・ラフソン法を持ちいていることによる。LOGISTIC プロシジャはデフォルトではフィsher のスコア法であるが、ニュートン・ラフソン法をオプションで設定することもできる。どちらの方法でも推定値は一致する。

重ね合わせプロットを用いて、正規分布のあてはめをしなさい



2.3. オッズ比による有害事象の比較

薬には期待したい効果だけではなく、有害な作用を併せ持つ。アマンタジンがインフルエンザの発病を予防する効果が期待できそうであるが、神経系の有害事象も無視できないくらい多いとの観察研究もあった。

アマンタジンのインフルエンザの予防効果を保ちつつ、有害事象⁵を減らす用量があるのかを探索するのも、Reumanら（1989）のダブルブラインド法による臨床試験の目的であった。アマンタジンの中枢神経系（Central nervous system）の有害事象を軽減しつつ予防効果が期待できる投与量として 100mg 群が設定された。一般成人を対象にした場合には、プラセボのインフルエンザ発病率が期待した 30% より極端に低く 3.1% しか発現せず、アマンタジンの投与による予防効果について有意な差を得ることができなかったが、表示 2.7 に示すようにアマンタジン 100mg が中枢神経系の有害事象をプラセボと同程度の結果が得られたことは、この試験の研究の成果である。

表示 2.7 アマンタジンの有害事象

	Placebo		Amantadine 100mg			Amantadine 200mg		
	N	%	n	%	オッズ比	n	%	オッズ比
Total No. of subjects	159		159			158		
Total No. (%) of subjects with adverse experiences	49	30.8%	47	29.6%	0.94	71	44.9%	1.83
Central nervous	25	15.7%	23	14.5%	0.91	47	29.7%	2.27
Gastrointestinal	12	7.5%	12	7.5%	1.00	17	10.8%	1.48
Cardiovascular	1	0.6%	1	0.6%	1.00	1	0.6%	1.01
Whole body"	24	15.1%	16	10.1%	0.63	21	13.3%	0.86

オッズ比とは

複数の有害事象の出現率の比較には、オッズ比が適している。オッズ比（*odds ratio*）は、2つの群の反応率を p_1 と p_2 としたときに、それぞれのオッズ $p_1/(1-p_1)$ 、 $p_2/(1-p_2)$ の比、

$$OR = \frac{p_2/(1-p_2)}{p_1/(1-p_1)} \quad (\text{式 2.1})$$

として定義されている統計量である。2群の反応率の比は p_2/p_1 であるが、反応率は、より小さいほうに価値がある場合、より大きいことに価値がある場合の2面性がある。

⁵ 投与した薬との因果関係が否定できない場合を“副作用”という。有害事象は、すべての有害反応を含む。アマンタジン投与中におきた交通事故もその原因にかかわらず害作用があったとする。

反応率 $p_1=90\%$ と $p_2=95\%$ の差は 5% ということもできるが、 100% に近づけることが困難な場合に反応率を 90% から 95% にするのは、 $(1-0.90)/(1-0.95)=2.0$ 倍の価値があるともいえるが、 $0.95/0.90=1.06$ 倍では、その価値を過小評価しかねない。オッズ比の定義式 (2.1) から、 p_1 を分母とする比は 1.06 倍、逆に $(1-p_2)$ を分母とする比は 2.0 倍となり、その積は 2.11 になる。オッズ比は、これら 2 つの比の積、

$$OR = \frac{p_2/(1-p_2)}{p_1/(1-p_1)} = \frac{p_2}{p_1} \cdot \frac{1-p_1}{1-p_2} = \frac{0.95}{0.90} \times \frac{1-0.90}{1-0.95} = 1.06 \times 2.0 = 2.11$$

でもある。オッズ比は、反応率が小さい場合には、反応率の比と同様の何倍と同様であり、反応率が大きい場合は、反応を 100% に近づけるための労力の程度を表す統計量として用いることができる。反応率 $p_1=5\%$ と $p_2=10\%$ の場合は、

$$OR = \frac{0.05}{0.10} \times \frac{1-0.10}{1-0.05} = 0.50 \times 0.947 = 0.474$$

のように 0.474 倍となるが、 $p'_1=90\%$ と $p'_2=95\%$ であっても両者の関係は互いに逆数の関係となっていて、反応率の大小に関わらず統計量としてのオッズ比は、一意に定まる。

LOGISTIC プロシジャによる解析

プラセボ群の有害事象発現率を p_2 とし、アマンタジン 100mg と 200mg 群の有害事象発現率を p_{11} と p_{12} としたオッズ比は、すでに 表示 2.7 に示してある。 200mg 群の中枢神経系のオッズ比が 2.27 、消化器系は 1.48 とプラセボに比べて有害事象の発現率が 100mg 群の 0.91 、 1.00 に比べて多いことが一目でわかる。オッズ比とその信頼区間を LOGISTIC プロシジャ⁶で計算してみよう。

表示 2.8 オッズ比の計算

```
data amantadine_ad ;
  input ad $14. @ ;
  do x = 0, 100, 200 ;
    input r n @ ;
    output ;
  end ;
datalines ;
0:有害事象全体 49 159 47 159 71 158
1:中枢神経系 25 159 23 159 47 158
2:消化器系 12 159 12 159 17 158
3:呼吸器系 1 159 1 159 1 158
```

⁶ GENMOD プロシジャでも計算できるが出力結果が LOGISTIC プロシジャに比べていくぶん冗長である。

```

4:全身症状      24 159   16 159   21 158
;
proc logistic data=amantadine_ad ;
  class x (param=reference ref=first) ;
  model r/n = x / link=logit expb ; /* tech=newton */
  by    ad ;

```

SASデータセットの作成とLOGISTICプロシジャを用いたSASプログラムを 表示 2.8 に示す. SASデータセットの作成は, 有害事象の 1 項目ごとにプラセボ: 0mg, 100mg, 200mg群に対して 1 行ごとに r/n の形式となるように, do; end; ステートメントの間に inputステートメントとoutputステートメントを挟んでいる.

Input ステートメントの @は, 入力ポインタを“改行しないで留まれ”とする制御情報で, $x=0$ にして r と n を読み, そのまま入力ポインタをその場に留め, 次の outputステートメントで, その時点で入力されている ad , x , r , n を SAS データセットに出力する. 次に $x=100$ として次のデータリストを r と n として読み, これまでの r と n に上書きする. 再度 output ステートメントで ad , x , r , n を SAS データセットに出力する. このような手順で 5 行分の入力データリストから, 15 行 4 列の SAS データセット amantadine_ad が作成される.

説明変数として LOGISTIC プロシジャも, GENMOD プロシジャと同様に連続変数と分類変数が扱える. Class ステートメントで, ad の後の (param=reference ref=first) は, 分類変数から解析用のダミー変数を生成するときのオプションであり, 最初の水準を基準とした参照型の意味である. これは, プラセボを最初の水準として 100mg と 200mg 群との比較をするために適したダミー変数が生成される.

Model ステートメントでは, $r / n = x$ のように反応変数として r/n を使い, 説明変数としての投与量 x は, 参照型の 2 つのダミー変数として展開される. Model ステートメントでは, さらに / の後に link=logit でシグモイド曲線として, ロジスティック分布のあてはめを行い, expb はオッズ比の追加出力するオプションを設定している.

LOGISTIC プロシジャでは, 誤差の分布は常に 2 項分布を仮定しているので, GENMOD プロシジャのように誤差分布を設定する必要はない. さらに, By ステートメントにより, 有害事象の項目ごとに解析を繰り返し替えるように設定している

表示 2.9 に中枢神経系の有害事象についてLOGISTICプロシジャでの最尤法による推定値と $\exp(\beta_1)$, $\exp(\beta_2)$ によって計算されたプラセボを基準としたアマンタジン 100mgと 200mg群のオッズ比とその 95%信頼区間が, 0.906 (0.490~1.679), 2.270 (1.314~3.920) が出力されている. 200mg群の 95%信頼区間の下限が 1.314 と 1 を含まないので, 有意な有害事象の発現といえる. 他の有害事象についてもbyステートメントによ

て計算されているが，結果の表示は省略する．

表示 2.9 中枢神経系の有害事象についてのオッズ比の推定
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp (Est)
Intercept	1	-1.6790	0.2179	59.3923	<.0001	0.187
x 100	1	-0.0982	0.3135	0.0981	0.7541	0.906
x 200	1	0.8196	0.2788	8.6396	0.0033	2.270

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
x 100 vs 0	0.906	0.490 1.676
x 200 vs 0	2.270	1.314 3.920

ロジスティック分布のあてはめ

有害事象のclassステートメントで生成された2つのダミー変数を w_1 と w_2 としたときに表示 2.10 に示すように， w_1 は100mgのときにのみ1， w_2 は200mgのときにのみ1となるような変数である．

表示 2.10 Calss ステートメントの参照型ダミー変数

	r	n	率	w_1	w_2
プラセボ : 0mg	25	159	15.7%	0	0
100mg	23	159	14.5%	1	0
200mg	47	158	29.7%	0	1

LOGISTIC プロシジャでは，中枢神経系の有害事象について回帰係数が $\hat{\beta}_0 = -1.6790$ ， $\hat{\beta}_1 = -0.0982$ ， $\hat{\beta}_2 = 0.8196$ として推定されている．これらの推定された回帰係数から，次のロジスティック分布により有害反応率が推定される．

$$\hat{p}_i = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 z_1 + \hat{\beta}_2 z_2)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 z_1 + \hat{\beta}_2 z_2)} \quad (\text{式 2.2})$$

実際に計算してみると，

プラセボ : 0mg の場合は， $\hat{p}_1 = \exp(-1.6790) / \{1 + \exp(-1.6790)\} = 0.157$ ，

100mg の場合は， $\hat{p}_2 = \exp(-1.6790 - 0.0982) / (1 + \exp\{-1.6790 - 0.0982\}) = 0.145$ ，

200mg の場合は， $\hat{p}_3 = \exp(-1.6790 + 0.8196) / (1 + \exp(-1.6790 + 0.8196)) = 0.297$

となり， r_i / n_i で計算された反応率に一致する．プラセボに対する

100mg 群のオッズ比は, $e^{-0.0982} = 0.906$,

200mg 群のオッズ比は, $e^{0.8196} = 2.270$

のように推定された回帰係数について指数関数を用いて推定できる. プラセボ群のオッズは,

$$\frac{\hat{p}_1}{1 - \hat{p}_1} = \frac{\exp(\hat{\beta}_0) / \{1 + \exp(\hat{\beta}_0)\}}{1 - \exp(\hat{\beta}_0) / \{1 + \exp(\hat{\beta}_0)\}} = \frac{\exp(\hat{\beta}_0) / \{1 + \exp(\hat{\beta}_0)\}}{1 / \{1 + \exp(\hat{\beta}_0)\}} = \exp(\hat{\beta}_0)$$

であり, 100mg 群のオッズは同様に,

$$\begin{aligned} \frac{\hat{p}_2}{1 - \hat{p}_2} &= \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1) / \{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1)\}}{1 - \exp(\hat{\beta}_0 + \hat{\beta}_1) / \{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1)\}} \\ &= \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1) / \{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1)\}}{1 / \{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1)\}} = \exp(\hat{\beta}_0 + \hat{\beta}_1) \end{aligned}$$

であるので, オッズ比は,

$$\frac{\hat{p}_2(1 - \hat{p}_2)}{\hat{p}_1(1 - \hat{p}_1)} = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1)}{\exp(\hat{\beta}_0)} = \exp(\hat{\beta}_1)$$

となることから推定したのである.

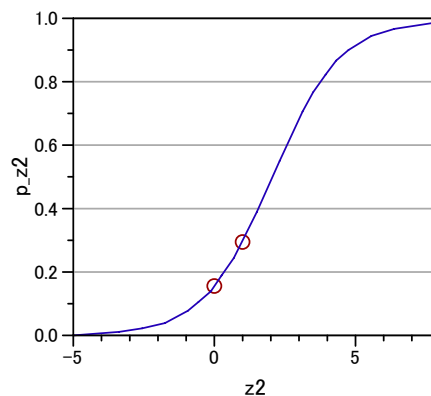
正規分布のあてはめと同様にロジスティック分布をあてはめてみよう. X 軸は, 100mg 群の場合は $w_2 = 0$ なので, w_1 の分布関数として,

$$p_{w1} = \frac{\exp(-1.6790 - 0.0982 \cdot w_1)}{1 + \exp(-1.6790 - 0.0982 \cdot w_1)}$$

200mg 群は $w_1 = 0$ なので, w_2 の分布関数として,

$$p_{w2} = \frac{\exp(-1.6790 + 0.8196 \cdot w_2)}{1 + \exp(-1.6790 + 0.8196 \cdot w_2)}$$

となる. w_2 が 0 の場合はプラセボ群の出現率, w_2 が 1 の場合は 200mg 群の出現率の推定値に一致する. w_2 を -5 から 8 の範囲でのロジスティック分布を表示 2.11 に示す.



表示 2.11 プラセボ群と 200mg 群にロジスティック分布のあてはめ

正規分布とロジスティック分布のあてはめ <<以下, z を w に変換する>>

正規分布のあてはめを考えてみよう. 正規分布のパーセント点を *probit* (プロビット) としたときに, z_2 が 0 の場合の下側 15.7%点は $probit_{z_2=0} = -1.0059$, z_2 が 1 の場合の下側 29.7%点は $probit_{z_2=1} = -0.0537$ である. *probit* を Y 軸, z_2 を X 軸となるような回帰直線 $\eta = -1.0059 + 0.4742 \cdot z_2$ を考えるととなる. プロビットは, $-\infty$ から ∞ の範囲であるが, $p = \text{Normal Distribution}(\eta)$ は, 0 から 1 の範囲となる.

ロジスティック分布

$$p = \frac{\exp(\beta_0 + \beta_2 z_2)}{1 + \exp(\beta_0 + \beta_2 z_2)} = \frac{\exp(\eta)}{1 + \exp(\eta)}$$

の指数内の回帰式を η として解くと,

$$\exp(\eta) = \frac{p}{1 - p}$$

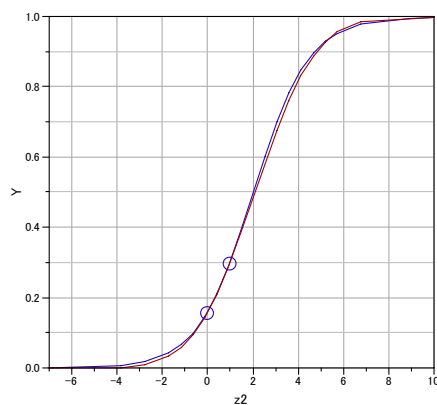
から, 対数を取ると,

$$\log\left(\frac{p}{1 - p}\right) = \eta \quad (\text{式 2.3})$$

が得られる. 反応率 p のオッズの対数はロジット (*logit*) といわれていて, プロビットと同様に, η は, $-\infty$ から ∞ の範囲である. η に関するロジスティック分布は,

$$p = \frac{\exp(\eta)}{1 + \exp(\eta)}$$

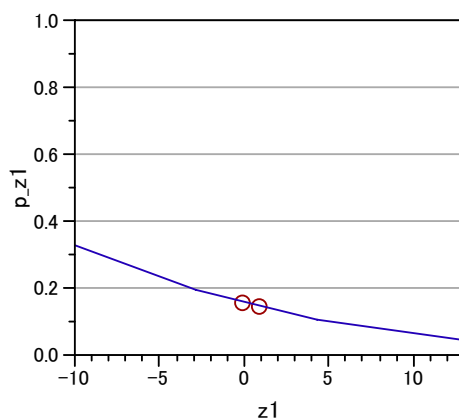
は, 0 から 1 の範囲となる. シグモイド曲線として, 表示 2.12 に示すようにロジスティック分布は正規分布に比べてやや裾が重いのであるが, 簡単な指数関数で現せることと, ロジットの回帰式の回帰係数の指数からオッズ比が推定できることから, 2 値データの解析に広く用いられている.



表示 2.12 ロジスティック分布と正規分布のあてはめ

ロジスティック分布は z_2 が0以下と1から5.374までが上側の曲線で、下側が正規分布であり、5.374以上は逆に下側がロジスティック分布である。

プラセボと100mg群に対するロジスティック分布は、 z_1 の係数が-0.0982と負なので表示 2.13 に示すように右下がりとなるが、表示されているX軸の範囲ではシグモイド状とは見えないが、範囲を拡大すればシグモイド曲線となる。



表示 2.13 プラセボ群と100mg群にロジスティック分布のあてはめ

歴史的には、プロビットが先に定式化されたこともあり、SASではPROBITプロシジャが古くからのサポートされ、その後LOGISTICプロシジャが加わり、それらを一般化し両方に対応するGENMODプロシジャが加わった。最新のSASバージョン9では、いずれのプロシジャでも、どちらのシグモイド曲線のあてはめが可能となっている。

JMPによるオッズ比

もとのテーブルからロジスティック回帰分析を行えるようにデータ変換を行え。

有害事象	r1	n1	r2	n2	r3	n3
0:有害事象	49	159	47	159	71	158
1:中枢神経	25	159	23	159	47	158
2:消化器系	12	159	12	159	17	158
3:呼吸器系	1	159	1	159	1	158
4:全身症状	24	159	16	159	21	158

JMP でロジスティック回帰を行う場合は、次のような形式のデータが必要である。

有害事象	dose	w2	w3	y	r_10	n
0:有害事象全	0	0	0	0	49	159
0:有害事象全	0	0	0	1	110	159
0:有害事象全	100	1	0	0	47	159
0:有害事象全	100	1	0	1	112	159
0:有害事象全	200	0	1	0	71	158
0:有害事象全	200	0	1	1	87	158
1:中枢神経系	0	0	0	0	25	159
1:中枢神経系	0	0	0	1	134	159
1:中枢神経系	100	1	0	0	23	159
1:中枢神経系	100	1	0	1	136	159
1:中枢神経系	200	0	1	0	47	158
1:中枢神経系	200	0	1	1	111	158
2:消化器系	0	0	0	0	12	159

中枢神経系の場合のオッズ比の推定

項	推定値	標準誤差	オッズ比	オッズ下側	オッズ上側
切片	-1.678964	0.2178593			
w2	-0.0981966	0.3135186	0.90647065	0.48793377	1.67755097
w3	0.81958135	0.2788338	2.2695495	1.32444487	3.96556488

2.4. インフルエンザ・ワクチンの地域別の予防効果

インフルエンザの予防法にワクチンの接種がある。神谷ら（2002）は、乳幼児を対象にしたインフルエンザ・ワクチンの予防効果を、全国7地域の小児科診療所の協力を得てワクチン接種を行い、郵送法（ハガキ方式）によって、発病調査、副反応調査を行った。結果を 表示 2.14 に示す。

表示 2.14 乳幼児を対象にしたインフルエンザ・ワクチンの予防効果

地域	非接種			接種			オッズ比
	<i>r</i>	<i>n</i>	<i>p</i>	<i>r</i>	<i>n</i>	<i>p</i>	
1:北海道	38	192	0.198	38	193	0.197	0.994
2:岩手	23	164	0.140	14	173	0.081	0.540
3:東京	57	192	0.297	36	203	0.177	0.511
4:三重	44	173	0.254	43	227	0.189	0.685
5:大阪	8	24	0.333	9	25	0.360	1.125
6:四国	64	185	0.346	36	189	0.190	0.445
7:福岡	35	198	0.177	43	199	0.216	1.284
全体	269	1128	0.238	219	1209	0.181	0.706

地域ごとにみると非接種群と接種群のインフルエンザ発現率には変動があり、接種群の非接種群に対するオッズ比を見ると、全体では、0.71 と接種群のインフルエンザ発現率が小さいことがわかるが、逆転している地域が2つある。統計的にはワクチンの接種は乳幼児に対する予防効果があるといえるのだろうか。

GENMODプロシジャによる解析

全国7地域をブロック因子⁷とする実験である。実験計画の観点でこの実験をみると、7つのブロックから全体で $1128 + 1209 = 2337$ 人をワクチン投与の非接種・接種群に分け、反応としてはインフルエンザの発病（1：有，2：無）とした。地域により例数はアンバランスであり、繰返し数が不揃いの2元配置型の実験とみなせる。

ブロック因子を変量効果とする解析は煩雑なので、第@@章で詳しく述べるので、ここでは、ブロック因子を固定効果とみなし、GENMODプロシジャを用いた解析事例を示す。

DATA ステップにより、データの入力、地域ごとに非接種・接種についてインフルエンザ発現数 *r*、対象症例数 *n* を入力し、14行4列のSASデータセットが作成される。

⁷ 取り上げた因子の性質を実験計画では幾つかに分類する。非接種・接種は実験者が設定できるので制御因子（固定効果）といい、地域は実験の局所管理と各地域による反応の変動を越えて、非接種・接種に差があるかを見るための因子であり、ブロック因子（変量効果）と分類される。

GENMOD プロシジャでは、説明変数として、地域 *block* とワクチン投与 *x* を共に `class` ステートメントで分類数として定義している．ダミー変数の生成オプションを指定していないので、デフォルトのインディケーター型の GLM タイプになる．地域 *block* は、7 つのダミー変数 $block_1, \dots, block_7$ に展開され、 $block_1$ は地域（北海道の場合に 1, それ以外は 0), \dots , $block_7$ は地域（福岡の場合に 1, それ以外は 0) であり、ワクチン投与 *x* の場合は、2 つのダミー変数 x_1 と x_2 に展開され、 x_1 はワクチンの投与（非接種の場合に 1, それ以外は 0), x_2 はワクチンの投与（接種の場合に 1, それ以外は 0) となる．

ロジスティック回帰の説明変数は、切片を含み 9 変数となり、 14×9 のデザイン行列が内部で生成される．Model ステートメントでは、主効果 **block** と **x** を指定しているが交互作用 **block*x** は、繰り返しなしの 2 元配置型なので解析モデルには含めずに誤差項になるように設定している．オプションとして / の後に `link` と `dist` オプションで応答変数の変換方法と誤差の分布の指定をしている．応答変数の変換方法はロジスティック分布のあてはめなので `logit` とし、誤差の分布は、2 項分布を仮定するので **binomial** を指定している．ワクチン投与の非接種群に対する接種群のオッズ比を求めるために `estimate` ステートメントで x_1 と x_2 についての対比を **x -1 1** とし、オプション **exp** を指定している．

表示 2.15 ブロックを固定効果とした解析プログラム

```
data vaccine ;
  input block $ @ ;
  do x = '1:非接種', '2:接種' ;
    input r n @ ;
    output ;
  end ;
datalines ;
1:北海道 38 192 38 193
2:岩手 23 164 14 173
3:東京 57 192 36 203
4:三重 44 173 43 227
5:大阪 8 24 9 25
6:四国 64 185 36 189
7:福岡 35 198 43 199
;
proc genmod data=vaccine ;
  class block x ;
  model r/n = block x / dist=binomial link=logit ;
  estimate 'x:接種(-1 1)' x -1 1 / exp ;
run ;
```

表示 2.16 に GENMOD プロシジャによるロジスティック回帰で推定された回帰係数を示す．地域 *block*, ワクチン投与 *x* の最後の水準, “福岡” および “接種” の自由度 DF

が 0 かつ推定値が 0 となっているのは、それらのダミー変数 $block_7$ と x_2 は表示されてはいるものの、ロジスティック回帰の計算からは除かれていることが示されている。ダミー変数を使った場合に、それらの回帰係数の推定値は、ダミー変数の設定に関連して異なる。次の式から、

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \eta_{ij} = \mu + \alpha_i x_i + \beta_j (block)_j + e_{ij}, \quad i=1,2, \quad j=1,2,\dots,7$$

表示 2.16 で示されている回帰係数が何を意味しているのか考えてみよう。非接種群 x_1 で北海道 $block_1$ の推定値、 $\hat{\eta}_{11}$ は、 $x_1=1$, $block_1=1$ なので、

$$\hat{\eta}_{11} = \hat{\mu} + \hat{\alpha}_1 x_1 + \hat{\beta}_1 (block)_1 = -1.5951 + 0.3550 \cdot x_1 + 0.0059 \cdot (block)_1 = -1.2342$$

であり、接種群 x_2 で北海道 $block_1$ の $\hat{\eta}_{21}$ は、

$$\hat{\eta}_{21} = \hat{\mu} + \hat{\alpha}_2 x_2 + \hat{\beta}_1 (block)_1 = -1.5951 + 0 \cdot x_2 + 0.0059 \cdot (block)_1 = -1.5892$$

である。推定値の差 $\hat{\eta}_{11} - \hat{\eta}_{21}$ は、

$$\hat{\eta}_{11} - \hat{\eta}_{21} = \hat{\alpha}_1 \cdot 1 = 0.3550$$

であり、非接種群 x_1 の回帰係数は、接種群との対数オッズの差となっている。これはすべての $block_i$ に関して同じである。

次に、 $block_i$ の回帰係数の推定値 $\hat{\beta}_j$ について考えてみよう。非接種群 x_1 で福岡 $block_7$ の $\hat{\eta}_{17}$ は、

$$\hat{\eta}_{21} = \hat{\mu} + \hat{\alpha}_1 x_1 + \hat{\beta}_7 (block)_7 = -1.5951 + 0.3550 \cdot x_1 + 0 \cdot (block)_7 = -1.2401$$

である。推定値の差 $\hat{\eta}_{11} - \hat{\eta}_{17}$ は、

$$\hat{\eta}_{11} - \hat{\eta}_{17} = \hat{\beta}_1 \cdot 1 = 0.0059$$

であり、北海道の回帰係数は、福岡との対数オッズの差となっている。これは接種群でも同じである。

切片は、単回帰分析では、X 軸の 0 のときの回帰直線が Y 軸を横切る点であるが、ダミー変数を用いた場合は意味が異なる。この例の場合は、

$$\hat{\eta}_{27} = \hat{\mu} + \hat{\alpha}_2 x_2 + \hat{\beta}_7 (block)_7 = -1.5951 + 0 \cdot x_2 + 0 \cdot (block)_7 = -1.5951$$

なので、接種群の福岡の対数オッズの推定値となっている。

対数オッズとオッズ比の推定

この臨床研究の仮説は、地域の差を越えてワクチンの接種群が非接種群に比べてインフルエンザの発現率が統計的に小さいかであり、オッズ比とその 95%信頼区間で結果を要約している。表示 2.14 で地域を込みした場合のオッズ比は、

$$OR = \frac{p_2 / (1 - p_2)}{p_1 / (1 - p_1)} = \frac{0.181 / (1 - 0.181)}{0.238 / (1 - 0.238)} = 0.706$$

表示 2.16 ワクチン投与後のインフルエンザ発病に対するロジスティック回帰
Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-1.5951	0.1389	-1.8673	-1.3229	131.93	<.0001
x 1:非接種	1	0.3550	0.1032	0.1528	0.5573	11.84	0.0006
x 2:接種	0	0.0000	0.0000	0.0000	0.0000	.	.
block 1:北海道	1	0.0059	0.1803	-0.3475	0.3593	0.00	0.9738
block 2:岩手	1	-0.6828	0.2156	-1.1054	-0.2602	10.03	0.0015
block 3:東京	1	0.2364	0.1737	-0.1041	0.5769	1.85	0.1736
block 4:三重	1	0.1525	0.1757	-0.1918	0.4968	0.75	0.3852
block 5:大阪	1	0.7839	0.3268	0.1435	1.4243	5.76	0.0164
block 6:四国	1	0.4042	0.1725	0.0660	0.7424	5.49	0.0191
block 7:福岡	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

として推定されている。

ロジスティック回帰の回帰係数の推定値は対数オッズとなっているので、地域を固定効果とした非接種群に対する接種群の共通のオッズは、

$$\exp(\hat{\eta}_{2j} - \hat{\eta}_{1j}) = \exp(-\hat{\alpha}_1) = \exp(-0.3550) = 0.7012$$

によって計算できる。GENMOD プロシジャでは、estimate ステートメントを用いて、

```
estimate 'x:接種(-1 1)' x -1 1 / exp ;
```

で対数オッズとその指数をとりオッズとその95%信頼区間を計算している。estimate ステートメントでは、class ステートメントで指定した分類に対応するダミー変数のパラメータの線形和が推定される。非接種群に対する接種群の対数オッズは、

$$-1 \cdot \hat{\alpha}_1 + 1 \cdot \hat{\alpha}_2$$

であるので、x -1 1 のように推定された回帰係数の線形和の計数（対比）を抜き出してして設定している。結果は、表示 2.17 に示すように対数オッズとその95%信頼区間が、-0.3550 (-0.5573~-0.1528) として表示され、それらの指数をとったオッズとその95%信頼区間は、0.7012 (0.5728~0.8583) となり、95%上限が1を越えていないので、有意な差があると結論される。

表示 2.17 非接種群に対する接種群の対数オッズ比とオッズ比の推定
Contrast Estimate Results

Label	Estimate	Standard Error	Alpha	Confidence Limits		Chi-Square	Pr > ChiSq
x:接種(-1 1)	-0.3550	0.1032	0.05	-0.5573	-0.1528	11.84	0.0006
Exp(x:接種(-1 1))	0.7012	0.0724	0.05	0.5728	0.8583		

推定値のグラフ表示

解析の結果を適切なグラフに描くことは、結果の解釈に役立つだけでなく解析方法の本質的な理解につながる。ダミー変数 x_1 は、ワクチンが非接種群のとき 1、それ以外は 0 としたので、それ以外の値では意味がないのであるが、解析上では、あたかも連続量として解析を行っている。地域が岩手の $block_2$ の推定値 $\hat{\eta}_{12}$ は、

$$\hat{\eta}_{12} = \hat{\mu} + \hat{\alpha}_1 x_1 + \hat{\beta}_2 (block)_2 = -1.5951 + 0.3550 \cdot x_1 - 0.6828$$

である。ダミー変数 $x_1 = 1$ 、非接種群の場合であるが、出現率の推定値は、

$$\hat{p}_{12} = \frac{\exp(\hat{\eta}_{12})}{1 + \exp(\hat{\eta}_{12})} = \frac{\exp(-1.9229)}{1 + \exp(-1.9229)} = 0.1275$$

ダミー変数 $x_1 = 0$ は、非接種群以外は接種群であるので、

$$\hat{p}_{12} = \frac{\exp(\hat{\eta}_{12})}{1 + \exp(\hat{\eta}_{12})} = \frac{\exp(-2.2779)}{1 + \exp(-2.2779)} = 0.0930$$

となる。X 軸を x_1 として、2 つの出現率の推定値を通るロジスティック分布をあてはめたことと同じである。2 つの出現率の推定値からオッズ比を計算してみよう。

$$OR = \frac{0.0930 / (1 - 0.0930)}{0.1275 / (1 - 0.1275)} = 0.7016$$

とうぜん結果は一致する。

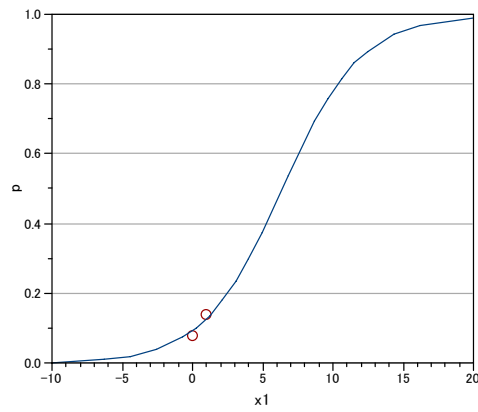
それぞれの地域についてのロジスティック分布

$$\hat{\eta} = \hat{\mu} + \hat{\alpha}_1 x_1 + \hat{\beta}_j (block)_j = -1.5951 + 0.3550 \cdot x_1 + \hat{\beta}_j (block)_j$$

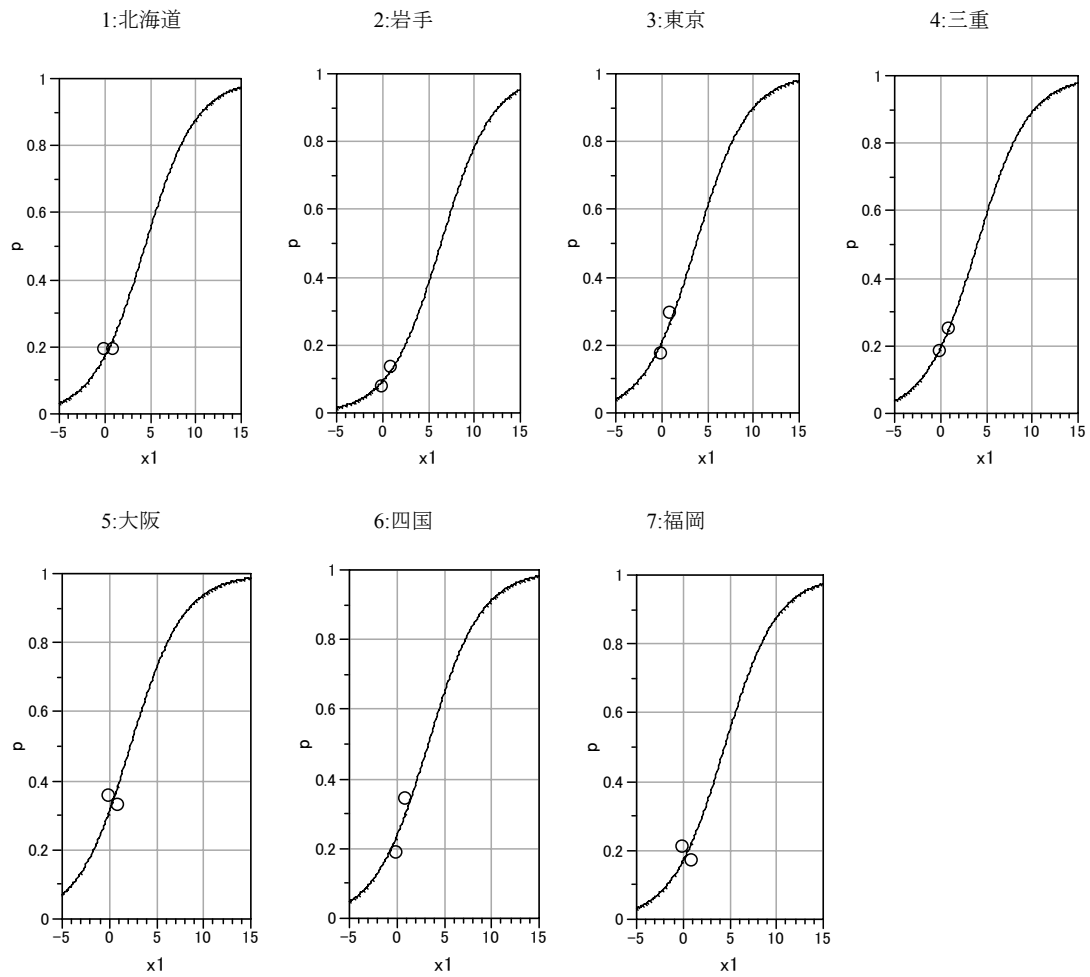
$$p = \frac{\exp(\hat{\eta})}{1 + \exp(\hat{\eta})} = \frac{\exp(-1.5951 + 0.3550 \cdot x_1 + \hat{\beta}_j (block)_j)}{1 + \exp(-1.5951 + 0.3550 \cdot x_1 + \hat{\beta}_j (block)_j)}$$

をあてはめた結果を 表示 2.19 に示す。この式から地域が変わって $\hat{\eta}$ の切片が変わるだけで、 x_1 の回帰係数は 0.3550 と一定なので、描かれたロジスティック分布は X 軸の位置は異なるが、曲線の形状は同じである。大阪と福岡は実データにまったくあっていないようにみえるが、それぞれのロジスティック分布の推定値からの尤度⁸の積が最大になるように回帰係数が推定されていることから、地域によってはあてはまりが悪いように見える場合もあるが、全体としては、最もあてはまりがよいロジスティック分布となっている。これは最尤法といわれている方法であり、第@@章で詳しく述べる。

⁸ 推定値を π としたときに、観測された点の二項分布の確率。大阪の非接種群は、 $\pi = 0.3879$ となるので、 $n=24$ 、 $r=8$ となる確率は 0.1464 である。



表示 2.18 地域が岩手の場合のロジスティック分布



表示 2.19 地域別に同一形状のロジスティック分布のあてはめ

モデルの妥当性

地域を固定効果として、ロジスティック回帰により、ワクチン投与によるオッズ比とその信頼区間を推定しているであるが、解析法は妥当なのであろうか。地域により、非接種群のインフルエンザ発病率が、14.0%から34.6%であり、地域を標示因子⁹そのような状況下でもワクチン接種のオッズ比が同じならば、ワクチンの効果を一般化できるのであるが、オッズ比自体が0.44から1.28と変動し、新たな地域で1群200例程度の実験を行ったときに、全データから得られたオッズ比の95%信頼区間(0.5728~0.8583)から外れることも多々あるだろう。地域間の変動を考慮した95%信頼区間も合わせて示す必要がある。このことについては、第@@章で、地域を変量効果としたロジスティック回帰分析で詳しく述べる。

ダミー変数 x_1 は、0と1にしかデータは存在しないのに、曲線のあてはめの妥当性はどのように考えるのであろうか。もっと単純に、非接種群のインフルエンザ発病率を100%にしたときの接種群の相対的な発病率(相対リスク, Relative Risk)に換算し各地域では99.5%, 57.7%, 59.7%, 74.5%, 108.0%, 55.1%, および、122.2%であり、全体では76.0%とすればよいのではないか。これは最もな考え方で、説明しやすい。問題は、相対リスクには2面性があることである。全体でのワクチンの接種群に対して非接種群でインフルエンザ非発病率の相対リスクは、 $(1 - 0.238)/(1 - 0.181) = 0.930$, 93.0%となる。76.0%とどちらが正しいのであろうか。どちらも正しいのである。オッズ比は、この2つの相対リスクを掛けた統計量で相対リスクの2面性はないが、直感的なわかりやすさが犠牲になっている。オッズ比の解釈は、相対リスクと同様に考えても差し支えないと割り切りてしまうことが、理解の助けになる。

ロジスティック回帰など行わなくても、 $2 \times 2 \times k$ 表でのオッズ比の統合方法で十分なのではないのかとの疑問もわく。地域によるデータ数の違いをロジスティック回帰では無視して、均一とみなした解析になっているようだが、いいのか。これらは、最もな観点なので、 $2 \times 2 \times k$ 表でのオッズ比の統合は、第@@章で、データ数を考慮したロジスティック回帰については、第@@章で詳しく論ずる。

神谷 齊 (2002), 乳平成12年度 厚生科学研究費補助金研究報告書—乳幼児に対するインフルエンザワクチンの効果に関する研究—

<http://www003.upp.so-net.ne.jp/manao/2000kenkyu.html>

⁹ 実際に製品が使われる様々な状況を実験条件とし組み込み、期待した性能を評価するための因

JMPによる解析

「ワクチン神谷_小児.JMP」に JMP データセットが用意されている。

	Area	r1_0	r1_1	n1	r2_0	r2_1	n2
1	0:全体	269	859	1128	219	990	1209
2	1:北海道	38	154	192	38	155	193
3	2:岩手	23	141	164	14	159	173
4	3:東京	57	135	192	36	167	203
5	4:三重	44	129	173	43	184	227
6	5:大阪	8	16	24	9	16	25
7	6:四国	64	121	185	36	153	189
8	7:福岡	35	163	198	43	156	199

課題 1) JMP で解析できるようにデータの構造を変換せよ。

	Area	ID	群	y	n
1	0:全体	r1_0	非接種	0	269
2	0:全体	r1_1	非接種	1	859
3	0:全体	r2_0	接種	0	219
4	0:全体	r2_1	接種	1	990
5	1:北海	r1_0	非接種	0	38
6	1:北海	r1_1	非接種	1	154
7	1:北海	r2_0	接種	0	38
8	1:北海	r2_1	接種	1	155
9	2:岩手	r1_0	非接種	0	23
10	2:岩手	r1_1	非接種	1	141
11	2:岩手	r2_0	接種	0	14
12	2:岩手	r2_1	接種	1	159
13	3:東京	r1_0	非接種	0	57
14	3:東京	r1_1	非接種	1	135

課題 2) ワクチンの接種と非接種のオッズ比を、地域を考慮して算出せよ。

項	推定値	標準誤差	オッズ比	オッズ下側	オッズ上側
切片	-1.2890093	0.0633877			
群[接種]	-0.1775143	0.0515986	0.70115335	0.57275819	0.85833073
Area[1:北海道]	-0.1226705	0.1256148	0.78243773	0.4743282	1.27152355
Area[2:岩手]	-0.8113919	0.1605139	0.19734856	0.10270739	0.36275687
Area[3:東京]	0.10779575	0.1187828	1.24059549	0.77451141	1.96738644
Area[4:三重]	0.02394426	0.1207602	1.0490537	0.64936788	1.67542141
Area[5:大阪]	0.65529287	0.2623126	3.70834518	1.27874202	10.1423592
Area[6:四国]	0.27561732	0.1175443	1.73539427	1.08976466	2.74120879

推定値は次の対数オッズに対するものです: 0/1

オッズ比は SAS と一致しているが、推定値は異なる。なぜか。

課題 3) SAS と同様の結果を出すためにどのようにしたらよいか方針を示せ。

Nominal Factors

Nominal factors are transformed into indicator variables for the design matrix. SAS GLM constructs an indicator column for each nominal level. JMP constructs the same indicator columns for each nominal level except the last level. When the last nominal level occurs, a one is subtracted from all the other columns of the factor. For example, consider a nominal factor A with three levels coded for GLM and for JMP as shown below.

Figure A.1

A	GLM			JMP	
	A1	A2	A3	A13	A23
A1	1	0	0	1	0
A2	0	1	0	0	1
A3	0	0	1	-1	-1

In GLM, the linear model design matrix has linear dependencies among the columns, and the least squares solution employs a generalized inverse. The solution chosen happens to be such that the A3 parameter is set to zero.

In JMP, the linear model design matrix is coded so that it achieves full rank unless there are missing cells or other incidental collinearity. The parameter for the A effect for the last level is the negative sum of the other levels, which makes the parameters sum to zero over all the effect levels.

Interpretation of Parameters

Note: The parameter for a nominal level is interpreted as the differences in the predicted response for that level from the average predicted response over all levels.

The design column for a factor level is constructed as the zero-one indicator of that factor level minus the indicator of the last level. This is the coding that leads to the parameter interpretation above.

Figure A.2

JMP Parameter Report	How to Interpret	Design Column Coding
Intercept	mean over all levels	1'
A[1]	$\alpha_1 - 1/3(\alpha_1 + \alpha_2 + \alpha_3)$	(A==1) - (A==3)
A[2]	$\alpha_2 - 1/3(\alpha_1 + \alpha_2 + \alpha_3)$	(A==3) - (A==3)

3. 層をなす分割表

分割表として結果が整理されている場合に、多くの統計手法が考案されてきたが、Koch ら (1988), Stokes ら(2000), はオッズ比を拡張することによりそれらを一般化して統一的に取り扱えることを示した. SAS の FREQ プロシジャでは、伝統的な分割表に対する多くの統計手法だけでなく、統一化した CMH 法が取り扱えるので、 $k \times 2 \times 2$ 表の解析事例を取り上げ、さらなる拡張についても解説したい.

分割表の解析の基本は $s \times r$ の 2 元表である. ここで、 r を反応の水準数とし、 s をある実験因子の水準数とする. 複数の 2 元表をセットにした 3 元表を $k \times s \times r$ 表という. CMH 法は、複数の 2 元表のセットを取り扱う方法である.

Peace, K.E. (1988), Biopharmaceutical Statistics for Drug Development (403-457), Koch, G.G., and Edwards, S., Clinical Efficacy Trials with Categorical Data, Dekker. 日本語訳：中上節夫, 森川俊彦 監訳 (1992), 医統計学 (373-424) .

Stokes, M.E., Davis, C.S., and Koch, G.G. (2000), Categorical Data Analysis using the SAS System, SAS Institute.

3.1. オッズ比

インフルエンザの予防薬としてオセタミビルを、2004 年に販売許可を日本の規制当局（厚生労働省）が出した. 表示 3.1 は、柏木ら (2000) の、プラセボに対するオセタミビルのインフルエンザの予防効果である. 反応は $r=2$, 実験因子は $s=2$ の 2×2 表である.

表示 3.1 オセタミビルのインフルエンザの予防効果

	インフルエンザ		計
	発病あり	発病なし	
オセタミビル	$n_{11}=2$	$n_{12}=153$	155
プラセボ	$n_{21}=13$	$n_{22}=140$	153

表示 3.2 に DATA ステップと FREQ プロシジャによるオッズ比とその 95% 信頼区間を計算するプログラムを示す. データの入力は、反応を y とし、その数を n とする 4 行 3 列の SAS データセットを作成する. FREQ プロシジャでは、TABLE ステートメントで要因 x 反応 y の 2 元表の作成を指示し、オプションで相対リスク RELRISK を付ける. NOCOL と NOPERCENT オプションは、分割表の出力から列の計に対するパーセント、合計に対

するパーセントの出力をするなどの意味である。

表示 3.2 オッズ比の信頼区間

```
data oseltamivir ;
    input x $14. y n ;
datalines ;
1:オセタミビル 0 2
1:オセタミビル 1 151
2:プラセボ 0 13
2:プラセボ 1 142
proc print data=oseltamivir ;
run ;

proc freq data=oseltamivir ;
    table x*y / nocol nopercnt relrisk ;
    weight n ;
run ;
```

FERQ プロシジャでは、 2×2 表のセル度数 n_{11} , n_{12} , n_{21} , n_{22} としたときに

$$OR = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{2/151}{13/142} = 0.1447$$

で計算をし、その対数オッズ比の分散¹⁰を

$$V(\ln OR) = \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}} = 0.5906$$

で計算し、オッズ比の 95%信頼区間を、

$$(\exp(\ln OR + 1.96\sqrt{V}), \exp(\ln OR - 1.96\sqrt{V})) = (0.0321, 0.6524)$$

で計算した結果が示されている。出力にケースコントロール研究（オッズ比）と標示されているのは、オッズ比が疫学研究デザインの一つであるケースコントロール研究¹¹で用いられている統計量であることに起因する。

¹⁰ 比の分散の導出の考え方については第@@章で取り扱う。

¹¹ ケースは、ある疾患をもっている患者集団で、コントロールは、患者集団の症例ごとに性、年齢などをマッチした対象集団で、過去にさかのぼって疾患の原因と思われる因子を調査し、ある疾患の発生の原因を特定する研究方法である。

表示 3.3 オッズ比および相対リスクの計算結果

x	y		
度数 行のパーセント	0	1	合計
1: オセタミビル	2 1.31	151 98.69	153
2: プラセボ	13 8.39	142 91.61	155
合計	15	293	308

x と y の統計量

相対リスクの推定値（行 1 / 行 2）

研究の種類	値	95% 信頼区間	
ケースコントロール研究（オッズ比）	0.1447	0.0321	0.6524
コホート研究（列 1 の相対リスク）	0.1559	0.0358	0.6791
コホート研究（列 2 の相対リスク）	1.0773	1.0237	1.1337

サンプルサイズ = 308

柏木征三郎，工藤翔二，渡辺彰，吉村功（2000），リン酸オセルタミビルのインフルエンザ発症抑制効果に関する検討—プラセボを対照とした第 III 相二重盲検並行群間比較試験成績—，感染症学雑誌，74（1062-1076）。

3.2. 研究を層とする分割表

2 群の比較で反応が 2 値の場合に，フィッシャーの正確検定，適合度のカイ 2 乗検定などを適用することも考えられるが，反応率の違い計量化するには，どうしたらよいのであろうか．表示 3.4 にオセタミビルのインフルエンザの予防効果について 7 つの臨床研究の結果を示したもので，柏木らの研究も 4 番目に含まれている．7 つの臨床研究は，独立に行われているが，研究デザインは先行する研究デザインを踏襲しており，異なる患者集団に対するオセタミビルのインフルエンザの予防効果があるかが，これらの研究の共通するテーマである．それぞれの研究のオッズ比は，誤差的な変動により一定ではない．新たな対象集団に対する期待される予防効果はどのくらいであろうか．また，その変動はどのくらいなのか．

表示 3.4 複数の臨床研究

No. 著者 年 対象	プラセボ群			オセタミビル群*			オッズ比
	<i>r</i>	<i>n</i>	<i>p</i>	<i>r</i>	<i>n</i>	<i>p</i>	
1:Hayden 1999 実験的 A 型	8	12	0.667	6	21	0.286	0.200
2:Hayden 2000 実験的 B 型	17	19	0.895	32	39	0.821	0.538
3:Hayden 1999 成人	25	519	0.048	13	1040	0.013	0.250
4:柏木 2000 成人	13	155	0.084	2	153	0.013	0.145
5:WV15825 高齢者	12	272	0.044	1	276	0.004	0.079
6:Welliver 2001 家族内	43	206	0.209	16	209	0.077	0.314
7:Hayden 2004 家族内	33	258	0.128	10	244	0.041	0.291
全体	151	1441	0.105	80	1982	0.040	0.359

* 研究 1, 2, および 3 は, オセタミビル用量を併合した結果を示している.

個々の研究は, 2×2 表であり, 研究全体として, 7 層の 2×2 表から構成されている. SAS の FREQ プロシジャの TABLE ステートメントで CMH オプションを付けることにより, k 層の 2×2 表の共通のオッズ比が行われる. CMH オプションは, k 層の $s \times r$ の 2 元表に対する解析法として広く適用できる. CMH オプションにより, 3 種類の CMH1, CMH2, CMH3 といわれる χ^2 統計量が得られる. それらは, 因子 S および反応 R の尺度によって選択する. 実験の因子 S が名義尺度で反応 R が名義尺度の場合は, CMH3 (一般型), 反応 R が順序尺度の場合は, CMH2 (ANOVA 型), 実験の因子 S が順序または連続尺度で反応 R も順序または連続尺度の場合は, CMH1 (相関型) というように選択する. 2×2 表の場合は, 3 種類の CMH 統計量は一致するので, 3 種類の CMH 統計量と使い分けと, $s \times r$ に対する各種の検定統計量との関係などについては, 第@@章で述べることにし, ここでは k 層の 2×2 表のオッズ比の統合に限定する.

k 層の 2×2 表のオッズ比とその 95%信頼区間の統合の方法は, 2.4 節で GENMOD プロシジャを用て最尤法による結果を示した. FREQ プロシジャでは, Mantel-Haenszel の共通オッズ比, 経験的ロジット推定による共通オッズ比を求めることができる. これらの 2 つの共通オッズ比は, 40 年以上前に提案された方法ではあるが, 計算が簡便であることから広く使われている.

表示 3.5 に層がある場合の共通オッズ比を推定する SAS プログラムを示す. データステップでは, 1 研究ごとに 1 行となっているプラセボ群, オセタミビル群の r/n 型のデータを, do, end のループで, FREQ プロシジャが要求する反応あり $y=0$, 反応なし $y=1$ のデータタイプにした $7 \times 4=28$ 行 5 列の SAS データセットを作成している. FREQ プロシジャでは, Table ステートメントで **ex*x*y** のように, 層*因子*反応の順とし, オプションで **cmh** を指定する.

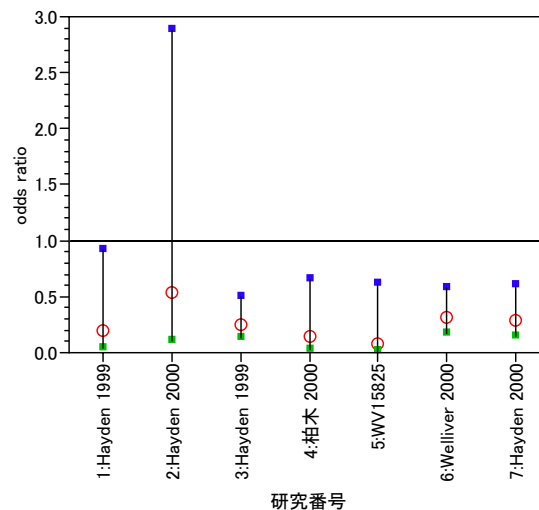
表示 3.5 層がある場合の共通オッズ比の推定

```

data oseltamivir7 ;
  length drug $12. ;
  input ex $16. @ ;
  do x = '2:placebo', '1:osetamivil' ;
    input r nn @ ;
    y=0 ; n=r ; output ;
    y=1 ; n=nn-r ; output ;
  end ;
  drop nn ;
datalines ;
1:Hayden_1999      8 12      6 21
2:Hayden_2000     17 19     32 39
3:Hayden_1999     25 519    13 1040
4:柏木_2000       13 155     2 153
5:WV15825        12 272     1 276
6:Welliver_2001  43 206     16 209
7:Hayden_2004     33 258    10 244
;
proc freq data=oseltamivir7 ;
  table ex*x*y / nocol nopercnt relrisk cmh ;
  weight n ;
run ;

```

FREQプロシジャでは, 表示 3.6 に示すように 7つの研究ごとに 2×2 表のオッズ比の推定が行なわれ, その後に 表示 3.7 に示すように共通オッズ比が出力される. Mantel-Haenszelの共通オッズ比 (95%信頼区間) は, 0.2590 (0.1829, 0.3669) で, 経験的ロジットの共通オッズ比 (95%信頼区間) は, 0.2677 (0.1888, 0.3797) となっている.



表示 3.6 研究ごとのオッズ比の 95%信頼区間

研究によりオッズ比が等質であるか Breslow-Day 検定結果も同時に出力され, 自由度

6 の $\chi^2=3.4080$, $p=0.7562$ であり, 研究によるオッズ比の差異は少ないことがわかる. しかしながら, この検定は, オッズ比の変動の指標ではないので, この目的で, 使うべきでないと警告が出力されている.

表示 3.7 FREQ プロシジャによる CMH 検定統計量, 共通オッズ比の推定

x と y の要約統計量 層別変数 : ex Cochran-Mantel-Haenszel 統計量 (テーブルスコアに基づく)				
統計量	対立仮説	自由度	値	p 値
1	相関統計量	1	65.2585	<.0001
2	ANOVA 統計量	1	65.2585	<.0001
3	一般連関統計量	1	65.2585	<.0001

相対リスクの推定値 (行 1 / 行 2)				
研究の種類	調整方法	値	95% 信頼区間	
ケースコントロール研究 (オッズ比)	Mantel-Haenszel ロジット	0.2590	0.1829	0.3669
		0.2677	0.1888	0.3797
コーホート研究 (列 1 のリスク)	Mantel-Haenszel ロジット	0.3786	0.2961	0.4843
		0.6540	0.5484	0.7800
コーホート研究 (列 2 のリスク)	Mantel-Haenszel ロジット	1.0713	1.0525	1.0906
		1.0511	1.0358	1.0665

オッズ比等質性に対する Breslow-Day 検定	
χ^2 乗値	3.4080
自由度	6
Pr > ChiSq	0.7562

1:Hayden 1999 実験的 A 型

Hayden,F.G., Atmar, R.L , Schilling,M., *et al.* (1999), Use of Theselective Oral Neuraminidase Inhibitor Oseltamivir to Prevent Influenza. N Engl J Mcd; **341**: 1329 – 35.

2:Hayden 2000 実験的 B 型

Hayden,F.G., Jennings, L., Robson,R. *et al.* Oral oseltamivir in human experimental influenza B. Antiviral Therapy **5**: 205-2.

3:Hayden 1999 成人

Hayden,F.G., Atmar, R.L , Schilling,M., *et al.* (1999), Use of the Selective Oral Neuraminidase Inhibitor Oseltamivir to Prevent Influenza. (N Engl J Mcd; 341 : 1336 - 43 .

4:柏木 2000 成人

5:WV15825 高齢者

公表されている申請概要書

6:Welliver 2001 家族内

Welliver, R., Monto,A.S., Carewicz,O., *et al.* Effectiveness of Oseltamivir in Preventing Influenza in Household Contacts A Randomized Controlled Trial. JAMA:**285**:748-754.

7:Hayden 2004 家族内

Hayden,F.G., Belshe,R., Villanueva,C., *et al.* Management of Influenza in Households: A Prospective, Randomized Comparison of Oseltamivir Treatment With or Without Postexposure Prophylaxis. The Journal of Infectious Diseases 2004;189:440-9.

JMPによるオッズ比，相対リスク

JMP の Ver.6 から「2 変量の関係」でオッズ比とその信頼区間を求めることができるようになった。次のファイルから実行してみよ。

メタ_群間オッズ比_T

メタ_群間オッズ比_T

Source

列(5/0)

研究番号

ラベル

薬剤

発病

n

行

すべての行 28

選択されている行 0

除外されている行 0

表示しない行 0

ラベルのついた行 0

	研究番号	ラベル	薬剤	発病	n
1	1:Hayden 1999	1_1:P+	1:P	1:+	8
2	1:Hayden 1999	1_2:P-	1:P	0:-	4
3	1:Hayden 1999	2_1:A+	2:A	1:+	6
4	1:Hayden 1999	2_2:A-	2:A	0:-	15
5	2:Hayden 2000	1_1:P+	1:P	1:+	17
6	2:Hayden 2000	1_2:P-	1:P	0:-	2
7	2:Hayden 2000	2_1:A+	2:A	1:+	32
8	2:Hayden 2000	2_2:A-	2:A	0:-	7
9	3:Hayden 1999	1_1:P+	1:P	1:+	25
10	3:Hayden 1999	1_2:P-	1:P	0:-	494
11	3:Hayden 1999	2_1:A+	2:A	1:+	13
12	3:Hayden 1999	2_2:A-	2:A	0:-	1027
13	4:柏木 2000	1_1:P+	1:P	1:+	13
14	4:柏木 2000	1_2:P-	1:P	0:-	142
15	4:柏木 2000	2_1:A+	2:A	1:+	2
16	4:柏木 2000	2_2:A-	2:A	0:-	151
17	5:Wong 2005	1_1:P+	1:P	1:+	18

薬剤と発病の分割表に対する分析 研究番号=4. 柏木 2000

度数: n

分割表

		発病		
		0:-	1:+	
行%	1:P	142 91.61	13 8.39	155
	2:A	151 98.69	2 1.31	153
		293	15	308

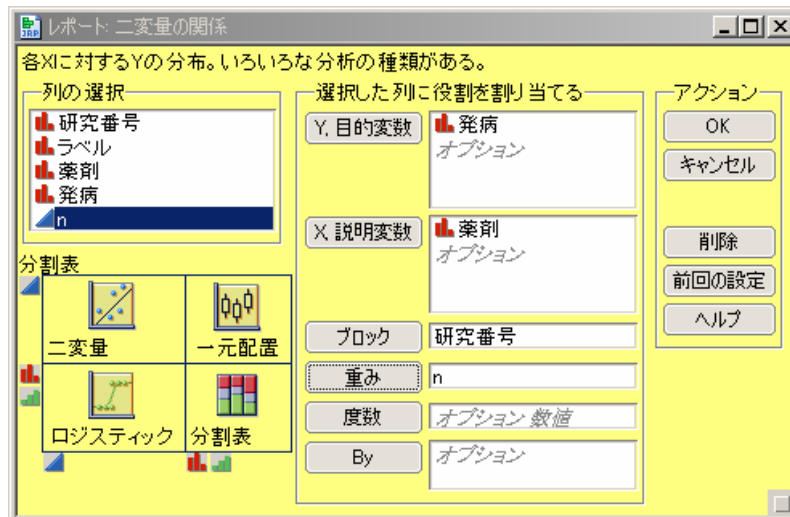
オッズ比

オッズ比	下側95%	上側95%
0.144677	0.032081	0.652445

相対リスク

説明	相対リスク	下側95%	上側95%
P(1:+ 2:A)/P(1:+ 1:P)	0.155857	0.035769	0.679113

層を考慮した CMH オッズ比



CMH 統計量は SAS と一致，オッズ比は不一致＜＜単純な合計で出されている＞＞

薬剤と発病の分割表に対する分析

重み: n

Cochran-Mantel-Haenszel検定

層別変数: 研究番号

Cochran-Mantel-Haenszel検定	カイ2乗	自由度	p値(Prob>Chisq)
スコアの相関	65.2585	1	<.0001 *
列カテゴリごとの行スコア	65.2585	1	<.0001 *
行カテゴリごとの列スコア	65.2585	1	<.0001 *
カテゴリの一般連関	65.2585	1	<.0001 *

オッズ比

オッズ比	下側95%	上側95%
0.359329	0.271548	0.475486

3.3. CMH 法とロジット法による共通オッズ比

表示 3.8 に示すように第 k 番目の層の 2×2 表が与えられていたときに、Mantel-Haenszel の共通オッズ比の推定量 $\hat{\phi}_{\text{MH}}$ は、

$$\hat{\phi}_{\text{MH}} = \frac{\sum_k n_{k11}n_{k22}/N_k}{\sum_k n_{k12}n_{k21}/N_k} \quad (\text{式 3.1})$$

である. 式 3.1 の $n_{k12}n_{k21}$ がゼロでなければ, $S_k = n_{k12}n_{k21}/N_k$ とおくと, 第 k 層のオッズ比の推定量 $\hat{\phi}_k$ を用いて,

$$\hat{\phi}_{\text{MH}} = \frac{\sum_k n_{k11}n_{k22}/N_k}{\sum_k n_{k12}n_{k21}/N_k} = \frac{\sum_k S_k \hat{\phi}_k}{\sum_k S_k} \quad (\text{式 3.1})$$

とあらわすことができ, S_k を重みとする重み付け平均となっている.

表示 3.9 の計算シートから, 共通オッズ比の推定量は,

$$\hat{\phi}_{\text{MH}} = \frac{\sum_h S_h \hat{\phi}_h}{\sum_h S_h} = \frac{18.113}{69.932} = 0.259$$

と計算できる.

表示 3.8 第 k 層の 2×2 表

	インフルエンザ 発病あり 発病なし		全体
オセタミビル	n_{k11}	n_{k12}	$n_{k1\cdot}$
プラセボ	n_{k21}	n_{k22}	-
全体	$n_{k\cdot 1}$	-	N_k

95%信頼区間については, Ver.9 の SAS では, Robins, Breslow, and Greenland (1986) の対数オッズ比の分散の推定量 $\hat{\sigma}^2$ を用いて,

$$(\hat{\phi}_{\text{MH}} \cdot \exp^{-1.96 \cdot \hat{\sigma}}, \hat{\phi}_{\text{MH}} \cdot \exp^{1.96 \cdot \hat{\sigma}}) = (0.1829, 0.3669)$$

で計算されている.

ロジット法による共通オッズ比の推定量 $\hat{\phi}_L$ は, それぞれの層の対数オッズ比 $\ln(\phi_h)$ について分散 $v_k = 1/n_{k11} + 1/n_{k12} + 1/n_{k21} + 1/n_{k22}$ の逆数 $w_k = 1/v_k$ の重み付け平均であり,

$$\ln \hat{\phi}_L = \frac{\sum_k w_k \ln(\hat{\phi}_k)}{\sum_k w_k} = \frac{-41.445}{31.452} = -1.318$$

$$\hat{\phi}_L = \exp(\ln \hat{\phi}_L) = 0.268$$

で計算されている．95%信頼区間は，

$$\left(\hat{\phi}_L \cdot \exp\left(\frac{-1.96}{\sqrt{\sum_k w_k}}\right), \hat{\phi}_L \cdot \exp\left(\frac{1.96}{\sqrt{\sum_k w_k}}\right) \right)$$

であるので，

$$\left(0.268 \cdot \exp\left(\frac{-1.96}{\sqrt{31.425}}\right), 0.268 \cdot \exp\left(\frac{1.96}{\sqrt{31.425}}\right) \right) = (0.189, 0.380)$$

となり，FREQ プロシジャの結果と等しくなることが確認できる．

CMH 法による共通ロジットは，ゼロのセルがあっても計算できるが，ロジット法の場合は，ゼロのセルがあった場合には，FREQ プロシジャでは，すべてのセルに 1/2 を加えて計算している．

表示 3.9 共通オッズ比を求めるための計算シート

試験		CMH 法			ロジット法		
h	N_k	$\hat{\phi}_k$	S_k	$S_k \hat{\phi}_k$	$\ln \phi_h$	w_k	$w_h \ln \phi_h$
1	33	0.200	3.636	0.727	-1.609	1.644	-2.646
2	58	0.538	2.052	1.103	-0.620	1.364	-0.846
3	1559	0.250	16.469	4.119	-1.386	8.339	-11.556
4	308	0.145	6.373	0.922	-1.933	1.693	-3.273
5	548	0.079	6.022	0.474	-2.541	0.917	-2.329
6	415	0.314	19.998	6.284	-1.158	10.302	-11.925
7	502	0.291	15.382	4.482	-1.233	7.193	-8.870
全体	3423	-	69.932	18.113	-	31.452	-41.445

佐藤ら (1998) は，層の数が限られていて 2 群の標本数が多い場合には，ロジット法による共通オッズ比が漸近有効性に優るが，層の数が増え各層の標本数が減るような場合にロジット法は，共通オッズ比の一致推定量とならないとしている．したがって，複数の実験あるいは研究結果を層とするような場合には，どちらの方法でもさしつかえないが，一つの実験あるいは研究のデータを幾つかの層に分けるような場合には共通オッズ比の推定方法としては，CMH 法が薦められる．

Robins, J.M., Breslow, N., and Greenland, S. (1986), “Estimators of the Mantel-Haenszel Variance Consistent in Both Sparse Data and Large-Strata Limiting Models,” *Biometrics*, 42, 311–323.

佐藤俊哉，高木廣文，柳川堯，柳本武美 (1998)，Mantel-Haenszel の方法による複数の 2×2 表の要約．統計数理，**46**，153-177.

3.4. Peto の方法による共通オッズ比

複数の臨床研究を層とする 2×2 表の共通オッズ比の推定方法として計算が簡便であることからピトー (Peto) の方法も広く使われている。ピトーの方法は、各層の推定オッズ比とその 95%信頼区間に基いて共通オッズ比を求めることに特徴がある。ゼロのセルがあるために計算不能となる場合でも各層のオッズ比の推定値とその 95%信頼区間を求めることができることもあり、広く使われている。

第 k 層の n_{k11} に度数を着目し、その期待度数との差、

$$n_{k11} - E(n_{k11}) = n_{k11} - \frac{n_{k1\cdot} n_{\cdot k1}}{N_k}$$

を求め、 n_{k11} の超幾何分布の分散

$$\text{var}(n_{k11}) = \frac{n_{k1\cdot} n_{k2\cdot} n_{\cdot k1} n_{\cdot k2}}{N_k^2 (N_k - 1)}$$

で割ったものが近似的に対数オッズ比となることが知られている。

$$\ln(\phi_k) = \ln\left(\frac{n_{k11} n_{k22}}{n_{k12} n_{k21}}\right) = \frac{n_{k11} - E(n_{k11})}{\text{var}(n_{k11})}$$

対数オッズ比の標準誤差 SE_k を

$$SE_k = \frac{1}{\sqrt{\text{var}(n_{k11})}}$$

で計算し、各層のオッズ比の 95%信頼区間を次式で求める。

$$\left(\hat{\phi}_k \cdot \exp\left(\frac{-1.96}{SE_k}\right), \hat{\phi}_k \cdot \exp\left(\frac{1.96}{SE_k}\right) \right)$$

共通オッズ比は、

$$\ln(\hat{\phi}) = \frac{\sum_k (n_{k11} - E(n_{k11}))}{\sum_k \text{var}(n_{k11})}$$

95%信頼区間は、

$$\left(\hat{\phi}_P \cdot \exp\left(\frac{-1.96}{\sqrt{\sum_k \text{var}(n_{k11})}}\right), \hat{\phi}_P \cdot \exp\left(\frac{1.96}{\sqrt{\sum_k \text{var}(n_{k11})}}\right) \right)$$

で求める。

表示 3.10 Peto のオッズ比の計算シート

試験	n11	N	OR	E(n11)	V(n11)	ln OR	OR [^]	L95	U95
1	6	33	0.200	8.9	1.924	-1.512	0.220	0.054	0.906
2	32	58	0.538	32.9	1.704	-0.556	0.573	0.128	2.573
3	13	1559	0.250	25.3	8.239	-1.499	0.223	0.113	0.442
4	2	308	0.145	7.5	3.579	-1.523	0.218	0.077	0.614
5	1	548	0.079	6.5	3.179	-1.745	0.175	0.058	0.524
6	16	415	0.314	29.7	12.683	-1.081	0.339	0.196	0.588
7	10	502	0.291	20.9	9.841	-1.108	0.330	0.177	0.617
全体	80			131.8	41.148				

表示 3.9 の計算シートの結果を用いて、次のようにピトーの共通オッズ比とその95%信頼区間が計算される.

$$\hat{\phi}_P = \exp\left(\frac{80 - 131.8}{41.148}\right) = 0.284$$

$$\left(0.284 \cdot \exp\left(\frac{-1.96}{\sqrt{41.148}}\right), 0.284 \cdot \exp\left(\frac{1.96}{\sqrt{41.148}}\right)\right) = (0.209, 0.385)$$

表示 3.11 SAS のデータデータステップを用いたピトーの共通オッズ比の計算

```

data oseltamivir_peto ;
  input ex $16. @ ;
  input n11 n1x n21 n2x ;
  n12 = n1x - n11 ;
  n22 = n2x - n21 ;
  nx1 = n11 + n21 ;
  nx2 = n12 + n22 ;
  N = n11 + n12 + n21 + n22 ;
  odds = n11*n22/(n12*n21) ;
  O_E = n11 - (nx1*n1x)/N ;
  V = n1x*n2x*nx1*nx2 / (N**2*(N-1)) ;
  ln_odds = O_E / V ;
  E_odds = exp(ln_odds) ;
  L95 = E_odds*exp(-1.96/sqrt(V)) ;
  U95 = E_odds*exp( 1.96/sqrt(V)) ;
  keep ex n11 N odds O_E V ln_odds E_odds L95 U95 ;
datalines ;
1:Hayden_1999      6   21    8  12
2:Hayden_2000     32   39   17  19
3:Hayden_1999     13 1040   25 519
4:柏木_2000        2  153   13 155
5:WV15825         1  276   12 272
6:Welliver_2001   16  209   43 206
7:Hayden_2004     10  244   33 258
;
proc summary data=oseltamivir_peto ;
  var n11 O_E V ;

```

```

output out=work2  sum( O_E V)=Sum_OE Sum_V ;
run ;

data work3 ;
  set work2 ;
  ln_OR_P = Sum_OE / Sum_V ;
  OR_P    = exp(ln_OR_P) ;
  L95 = OR_P*exp(-1.96/sqrt(sum_V)) ;
  U95 = OR_P*exp( 1.96/sqrt(sum_V)) ;
proc print data=work3 ;
run ;

```

3.5. さらなる共通オッズ比の推定方法と選択基準

層が 1 元的で共通オッズ比の算出だけが目的であるならば、FERQ プロシジャの適用による共通オッズ比の推定は簡便である。共通オッズ比の推定方法には、さらに、LOGISTIC プロシジャ、GENMOD プロシジャを用いた最尤法による方法もある。各層の例数が少なくなったときには、LOGISTIC プロシジャのSTRATA ステートメントで層を指定することにより、条件付ロジスティック回帰による共通オッズ比の推定もできる。SAS のプログラムを 表示 3.12 に示す。LOGISTIC プロシジャおよび GENMOD プロシジャから得られた最尤法による共通オッズ比 (95% 信頼区間) は、0.258 (0.182, 0.365)、LOGISTIC プロシジャで研究 ex を STARATA ステートメントで層とした場合には、0.259 (0.184, 0.367) となる。

層を固定効果とした場合に 5 種類の共通オッズ比の推定法を示したのであるが、推定値は微妙に異なる。どれを用いたらよいのだろうか。複数の研究を統合するメタ・アナリシスの標準としては、Excel などでも手軽に計算でき、各層のオッズ比の推定をベースにした Peto の方法に魅力を感じる。CMH 法およびロジット法による共通オッズ比の推定は、Peto の方法が Ver.9 の SAS でプロシジャとしてサポートされていないために、Peto の方法の代わりとしての使用にとどめたい。これは、層の数が増えデータ数が少なくなるような場合には、ロジット法でなく CMH 法を使用すべきであるが、メタ・アナリシスの場合には、このようなことはないので、FERQ プロシジャの使用を積極的に薦める理由が見当たらない。

LOGISTIC プロシジャおよび GENMOD プロシジャによる最尤法は、次節で述べる理由からさらなるメタ・アナリシスを探索的に行う場合に行う場合に層の数が増えた場合の交互作用を解析、あるいは、特定の水準間の差について対比を用いて細かな検討ができる。各層の例数が非常に小さい場合は、条件付き最尤法の適用すべきである。しかし

ながら、メタ・アナリシスの解析で各層の例数が極端に小さくなることはよくないの
で、条件付き最尤法を使うことはない。

表示 3.12 最尤法および層を考慮した最尤法（条件付き）

```
proc logistic data=oseltamivir7 ; /* LOGISTIC */
    class ex x ;
    model y = ex x / link=logit ;
    freq n ;
    run ;

proc genmod data=oseltamivir7 ; /* GENMOD */
    class ex x ;
    model y = ex x / dist=binomial link=logit ;
    estimate '1:0se. vs 2:Pl.' x 1 -1 / exp ;
    freq n ;
    run ;

proc logistic data=oseltamivir7 ; /* 条件付き */
    class ex x ;
    model y = x / link=logit ;
    strata ex ;
    freq n ;
    run ;
```

これらの 5 種類の方法は、層を固定効果とみなした解析であり共通オッズ比の 95% 信頼区間は、解析に用いたすべての研究をプールしたときの 95% 信頼区間が構成されている。これに対して、あらたな同程度の研究を計画したときに得られるオッズ比の予測には、層を変量効果にした解析が必要になる。これは、非線型混合効果モデルの解析のための NLMIXED プロシジャにより可能となる。SAS ではまだサポートされていないが、モーメント法による Delsaiminian-Laiyad 法もある。どれを選択するか、対象とするデータの性質、解析目的によって前もって選択する必要がある。これらについては、第@@章であらためて議論する。

JMPによる最尤法でのオッズ比

パラメータ推定値

項	推定値	標準誤差	オッズ比	オッズ下側	オッズ上側
切片	1.90964999	0.1025779			
薬剤[1:P]	-0.6771646	0.0884554	0.25812037	0.18084602	0.36234817
研究番号[1:Hayden 1999]	-1.7686144	0.3318749	0.02909384	0.00797075	0.10953661
研究番号[2:Hayden 2000]	-3.9665329	0.337425	0.00035869	8.62729e-5	0.00124706
研究番号[3:Hayden 1999]	1.76230507	0.1723691	33.9405393	17.5671162	68.0805551
研究番号[4:柏木 2000]	1.26269398	0.2490147	12.4957422	4.96309021	35.3911405
研究番号[5:WV15825]	2.00856744	0.2606127	55.5417443	21.3158797	166.976108
研究番号[6:Welliver 2000]	0.04357742	0.1595638	1.0910656	0.58873532	2.06067788

推定値は次の対数オッズに対するものです: 0:-/1:+

95%信頼区間は, SAS と不一致, これはプロファイル尤度を用いているからである. SAS の LOGISTIC プロシジャでもプロファイル尤度が計算できる. CLIDDS=PL | WALD | BOTH, 確認する予定.

3.6. 対象集団間のオッズ比の比較

インフルエンザの予防では, 異なる対象集団に対して臨床試験が実施されているので, それらの対象集団間でオッズ比が異なるのか探索的に検討したい.

実験計画法の観点からこの臨床研究の“計画”を記述してみよう. 標示因子として因子Aを 4 水準 (A1:実験的研究, A2:一般成人, A3 高齢者, A4:家族内), 制御因子として因子Bでは 2 水準 (B1:プラセボ群, B2:オセタミビル群) を取り上げた. 因子Bは, 因子Aに対して入れ子にして反復を 2 回行うことにした. 入れ子になった因子B 内で, その対象集団をランダムに薬剤, B1:プラセボとB2:オセタミビルに割り付けて, 反応 (有, 無) を観測した. 実験は 5 年がかりで行われ, 表示 3.13 に示すように, B1 と B2 について, 反応があった人数を r , 全体を n として整理した. なお, A3:高齢者に対しては, 反復は 1 回しか行えなかった. 適切な解析計画が前もって立ててあり, 実際の解析作業を監督し, 解析報告書を作成することが, このプロジェクト統計家に課せられた責務である.

表示 3.13 一元配置枝分かれ型分割・反復測定実験

対象集団	No. 著者 年	B1:プラセボ群		B2:オセタミビル群	
		<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>
A1:実験的研究	1:Hayden 1999 A 型	8	12	6	21
	2:Hayden 2000 B 型	17	19	32	39
A2:一般成人	3:Hayden 1999	25	519	13	1040
	4:柏木 2000	13	155	2	153
A3 高齢者	5:WV15825	12	272	1	276
A4:家族内	6:Welliver 2001	43	206	16	209
	7:Hayden 2004	33	258	10	244
全体		151	1441	80	1982

FREQ プロシジャによる解析は、層が一つの場合に限られており、因子 A の中に反復があるような場合には対応できない。LOGISTIC プロシジャまたは GENMOD プロシジャでシグモイド曲線をあてはめる解析となりそうだが、誤差が複数ある実験データの解析にはたして対応できるのだろうか。著者たちも 1990 年代には難解な問題であったのであるが、その後の SAS のバージョンアップにより 2000 年代になって、このような 2 値データの解析に手が付けられるようになってきた。

4. 反応が順序データ

反応が (++, +, -) のように 3 以上の反応に分類できる場合に、ウィルコクソンの 2 標本検定、比較したい群の数が 3 水準以上の場合にはクリスカル・ワリスの検定、水準に順序がある場合には、ヨンキーの傾向性検定など、検定をベースにした統計手法が思い浮かぶが、これらの検定ベースの統計手法は、基本的には実験が 1 元配置の場合を対象にしている、共変量を考慮した解析への発展性がない。

直交表を用いた実験で反応が 3 水準以上の場合に、田口 (1976) の累積法といわれている分散分析の方法が日本の品質管理の分野で用いられている。この方法は、3 水準の反応を、++ 以上、+ 以上の 2 値反応 (1, 0) に再分類し、それぞれを計量値とみなして 2 つの分散分析表を求め、2 項分散の逆数を重みとして 2 つの分散分析表を 1 つにまとめる方法である。

田口の累積法は、シグモイド曲線を 2 値反応データにあてはめる解析法が普及する以前の方法であり、現在では、累積反応率へ複数のシグモイド曲線を直接あてはめる解析に移行すべきである。しかしながら、要因の組み合わせ効果の推定には、出現率を対数オッズに変換し要因効果の線形和から組み合わせ効果を推定することが、 Ω 法 (ロジット法) として定式化されている。この点に関して、田口の累積法は、2 値あるいは順序データの先進的な解析法であった。

なお、 Ω 法では、オッズの常用対数の -10 倍をデシベル単位と定義されている。田口 (1976) には、出現率をデシベルに変化するための細かさ数値表も付録として示されている。現代のロジスティック回帰では、オッズの自然対数をロジットというが、本書では、対数オッズとしている。

田口玄一 (1976), 第 3 版 実験計画法 (上, 下), 丸善.

4.1. 田口の累積法の考え方

表示 4.1 のデータは、強化ガラスに与えた衝撃の強さ (kg) とガラスが破壊する確率の関係を求めるために行なった仮想的な実験結果である。第 i 水準で n_i 個を検査し、破壊した個数を r_{i1} とし、ひび入り数を r_{i2} とし、破壊しなかった個数を r_{i3} とする。破壊数 r_{i1} と ひび入り数 r_{i2} の合計を R_{i2} とする。また、破壊数 r_{i1} を改めて R_{i1} とする。田口 (1976) の累積法は、まず R_{i1} と R_{i2} について反応がある場合を 1, ない場合を 0 としてそれを計量値とみなして別々に分散分析表を求め、次いで、それらを併合する方法である。反応変数の順序に意味のある質的変数であるとき、累積度数に基づいて解析するので累積法と呼ばれる。

表示 4.1 順序がある破壊データ

衝撃 x kg	破壊 $r_{i1}=R_{i1}$	ひび入 r_{i2}	非破壊 r_{i3}	合計 n_i	累積 $r_{i1}+r_{i2}=R_{i2}$
35	0	1	8	9	1
40	1	2	7	10	3
45	0	2	6	8	2
50	3	7	9	19	10
55	3	6	2	11	9
60	6	6	2	14	12
65	14	5	0	19	19

累積法の考え方は、結果が、製品を [良品, 格下げ品, 不良品] と分類する場合、消費者の満足度調査で [不満, やや不満, どちらでもない, やや満足, 満足] のように 5 段階で評価する場合、薬効を [著明改善, 中等度改善, 軽度改善, 無効] のように評価する場合などのように、反応変数が順序に意味のある質的変数であれば、広い分野で適用することができる。

田口の累積法は、 R_{i1} と R_{i2} の分散分析表を併合する際に、2 つの平方和が独立ではないのに、併合後の分散分析表の自由度を 2 倍している。これが、有意差が出やすくなる原因である。

衝撃 x を横軸に、 $P_{i1} = R_{i1} / n_i$ と $P_{i2} = R_{i2} / n_i$ 縦軸にをとったグラフを描くと表示 4.3 が得られる。田口の累積法の考え方では、表示 4.2 に示すように $R1$ と $R2$ についてガラスが破壊したときを $y=1$, しなければ $y=0$ とし、別々に単回帰分析を行い、得られた 2 つの分散分析表の重みをそれぞれ、2 項分散の逆数

$$W_1 = \frac{1}{\frac{28}{90} \cdot \left(1 - \frac{28}{90}\right)} = 4.050$$

$$W_2 = \frac{1}{\frac{57}{90} \cdot \left(1 - \frac{57}{90}\right)} = 4.306$$

として自由度と平方和を合算し、あらためて平均平方を算出する．回帰の平均平方和が有意ならば，衝撃 x の増加に伴い破壊率が統計的に増加すると総合的な判断を下す．

別々に単回帰を行なわなくても，2つの累積不良率を A_1 水準と A_2 水準とし，衝撃 x を連続量として，切片は異なるが傾きが共通な2本の平行線を当てはめる問題としても考えられ，GLM プロシジャの MODEL ステートメントで `y=A x / solution noin` として，

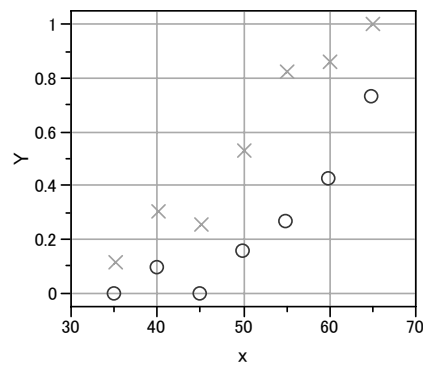
$$\hat{p}^{(1)} = -1.1436 + 0.0276 \cdot x$$

$$\hat{p}^{(2)} = -0.8214 + 0.0276 \cdot x$$

が得られる．この場合には傾きの自由度が1となるので，田口の累積法の問題が解決したかのように思われるが，衝撃が弱い場合に破壊率の推定値が負になることから，直線のあてはめを仮定する統計モデルは不適切である．

表示 4.2 2元配置とみなしたデータ構造

x	$R_1 : A_1$		$R_2 : A_2$		n
	R1	n-r1	r1+r2	n-(r1+r2)	
	y=1	y=0	y=1	y=0	
35	0	9	1	8	9
40	1	9	3	7	10
45	0	8	2	6	8
50	3	16	10	9	19
55	3	8	9	2	11
60	6	8	12	2	14
65	14	5	19	0	19



表示 4.3 累積比率 P_1 , P_2 と x の関係

○ : P_1 , × : P_2

R_1 と R_2 は同じテストピースから求めたものであって、独立ではない。累積出現率に複数のシグモイド曲線をあてはめる場合は観測された比率が 2 項分布に従うという前提で行う。表示 4.1 のデータは、1 枚のテストピースの実験結果が 3 つのカテゴリの何れかに該当するので、3 項分布（一般には多項分布）に従う。すなわち、 $p_1 = r_1 / n$, $p_2 = r_2 / n$, $p_3 = r_3 / n$ は独立ではない。したがって、累積比率、

$$\begin{aligned} P_1 &= p_1 \\ P_2 &= p_1 + p_2 \end{aligned} \tag{4.1}$$

も独立ではない。これは、 p_1 が両方に共通に入っていることから、 P_1 と P_2 の間に相関が生じ、非共通部分である p_2 が小さいとき相関が強くなる。反応変数のカテゴリ数が多く、中間のカテゴリの比率が小さいとき相関が強くなる。これらの相関を考慮して 2 つの累積破壊率について、GENMOD プロシジャを使って、正規分布をあてはめてみよう。

表示 4.4 に示すように、GENMOD プロシジャで解析するときには、順序データを反応とした場合には、2 値反応のように r/n タイプの MEDEL ステートメントは使えないので、反応 $y = 1, 2, 3$ に対して $r_i, i=1,2,3$ のデータがあるというように $3 \times 7 = 21$ 行 3 列の SAS データセットを作成する。GENMOD プロシジャでは、MODEL ステートメントで、`model y = x` のように“単回帰分析”とする。

仮定する分布として多項分布 `multinomial` をオプションで指定し、累積反応として、累積正規分布 `cprobit` をつかう。GENMOD プロシジャでは、他に累積ロジット `clogit`、累積補 2 重対数 `cumclog` が使えるので、目的に応じて使い分けることができる。累積ロジットは、累積正規分布と同じ上限と下限への収束スピードが同じであるが、累積補 2 重対数は、下限への収束が上限に比べて遅い。第 @@ 章で、使い分けについて詳しく述べる。

表示 4.4

```
data glass2 ;
  input x @@ ;
  do y = 1, 2, 3 ;
    input r @@ ;
    output ;
  end ;
datalines ;
35 0 1 8
40 1 2 7
45 0 2 6
50 3 7 9
55 3 6 2
60 6 6 2
65 14 5 0
```

```

;
proc print data=glass2 ;
run ;

proc genmod data=glass2 ;
  model y = x / dist=multinomial link=cprobit ;
  freq r ;
run ;

```

2つの正規分布のあてはめ結果を 表示 4.5 に示す. 切片が Intercept1 とIntercept2 のように2つ出力され, 衝撃 x の傾きの回帰係数として 0.1017 が推定されている.

$$\hat{p}^{(1)} = \text{Normal distribution}(\hat{\beta}_0^{(1)} + \hat{\beta}_1 x) = \text{Normal distribution}(-6.1097 + 0.1017 \cdot x)$$

$$\hat{p}^{(2)} = \text{Normal distribution}(\hat{\beta}_0^{(2)} + \hat{\beta}_1 x) = \text{Normal distribution}(-4.8735 + 0.1017 \cdot x)$$

衝撃を強くすればガラスが割れるのは当然である. この実験の結果を, 5%のガラスがひび割れをおこす衝撃は何 Kg なのか, 50%の場合はどうか, 90%のガラスが破壊される衝撃は何 Kg なのか, などを求めて他の条件で試作してガラスと比較したいとする. 5%がひび割れを起こす x を知りたいのであるから,

$$0.05 = \text{Normal distribution}(-4.8735 + 0.1017 \cdot x)$$

となる x は, 正規分布の下側確率のパーセント点は 1.64 であるので,

$$-1.64 = -6.1097 + 0.1017 \cdot x$$

$$x = \frac{-1.64 + 6.1097}{0.1017} = 31.8$$

31.8kg となる. 50%の場合も同様に 47.9kg, 90%が破壊される場合は,

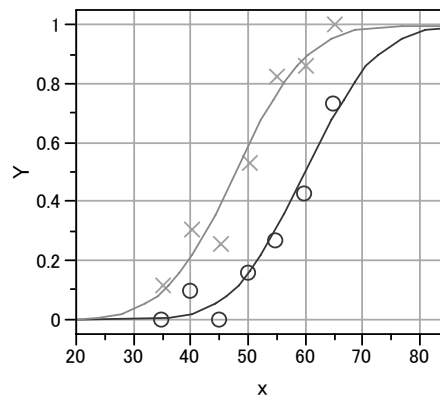
$$0.90 = \text{Normal distribution}(-6.1097 + 0.1017 \cdot x)$$

$$x = \frac{1.28 + 6.1097}{0.1017} = 72.7$$

72.7kg と推定される. 実験データの要約にさらに点推定値だけでなく 95%信頼区間を示したい. どのようにすればよいのだろうか. この問題は, 化合物の 50%致死量を推定する場合と同様に, 逆推定の問題として生物検定法の定式化されているので, 詳しくは第@@章を参照されたい.

表示 4.5

Parameter	DF	Estimate	Standard Error	Wald	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept1	1	-6.1097	0.8944	-7.8626	-4.3568	46.67	<.0001
Intercept2	1	-4.8735	0.8196	-6.4800	-3.2671	35.36	<.0001
x	1	0.1017	0.0158	0.0707	0.1326	41.47	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		



表示 4.6 推定値のグラフ

○ : P_1 , × : P_2

4.2. 多カテゴリーの反応変数（カテゴリーの順序に意味あり）

プラセボを基準としてA薬の2用量の薬効を二重盲検比較試験でヒトに投与して，表示 4.7 のデータを得た．反応変数は，+++：著明改善，++：中等度改善，+：軽度改善，-：無効，の順序に意味のある4カテゴリーである．この試験の第1の目的は，プラセボに対してA薬の低用量と高用量を併合した場合に効果が統計的に見出せるかである．

表示 4.7 臨床試験データ（人工データ）

薬剤	+++	++	+	-	合計
プラセボ	0	4	12	12	28
A 薬	8	19	18	18	63
(10mg)	3	10	8	10	31
(20mg)	5	9	10	8	32
合計	8	23	30	30	91

2値データの場合に定義されたオッズ比を順序データへの拡張を考えよう．そのために，

- +++：著明改善率に対するオッズ，
- ++：中等度改善以上の率に対するオッズ，
- ＋：軽度改善以上の率に対するオッズ

のように順序関係の方向に逐次的に累積度数の3つの“有効率”に対してオッズを考える．プラセボ群とA薬10mg群と20mg群を合わせてA薬群とし，プラセボ群に対する

A 薬群のオッズ比も 3 つ計算される．この 3 つのオッズ比の“平均”が A 薬のプラセボに対する効果の要約統計量となる．

表示 4.8 に 3 つの有効率に対して共通のオッズ比の推定を考える．+++の有効率は，プラセボ群で 1.8%，A 薬群で 12.7%であるが推定され推定有効率がそれぞれ 4.4%，10.9%となっているとしよう．これから計算される推定オッズ比が 2.688 となる．同様に++以上の有効率については，プラセボ群で 14.3%，A 薬群で 42.9%であるが推定され推定有効率がそれぞれ 20.6%，41.1%ならば，オッズ比は 2.688 と+++の有効率と同じになる．+以上の場合は，プラセボ群で 57.1%，A 薬群で 71.4%であるが推定され推定有効率がそれぞれ 52.2%，74.6%ならば，オッズ比は 2.688 と++以上，+++の場合と同じになる．

表示 4.8

		+++	++以上	+以上	全体
累積度数	プラセボ	0	4	16	28
	A 薬	8	27	45	63
累積有効率	プラセボ	0.018	0.143	0.571	1.000
	A 薬	0.127	0.429	0.714	1.000
オッズ オッズ比	プラセボ	0.018	0.167	1.333	-
	A 薬	0.145	0.750	2.500	-
		8.000	4.500	1.875	-
推定累積有効率	プラセボ	0.044	0.206	0.522	
	A 薬	0.109	0.411	0.746	
推定オッズ 推定オッズ比	プラセボ	0.045	0.260	1.091	-
	A 薬	0.122	0.698	2.933	-
		2.688	2.688	2.688	

累積度数が 0 の場合は，0.5 とする．

共通のオッズ比の推定は，次のような考え方で推定する．各薬剤群のセルの期待出現率 π_{i1} ， π_{i2} ， π_{i3} ， π_{i4} が多項分布に従うとしたときに，累積出現率のオッズ比

$$\frac{\hat{\pi}_{21}/(1-\hat{\pi}_{21})}{\hat{\pi}_{11}/(1-\hat{\pi}_{11})} = \frac{(\hat{\pi}_{21}+\hat{\pi}_{22})/(1-(\hat{\pi}_{21}+\hat{\pi}_{22}))}{(\hat{\pi}_{11}+\hat{\pi}_{12})/(1-(\hat{\pi}_{11}+\hat{\pi}_{12}))} = \frac{(\hat{\pi}_{21}+\hat{\pi}_{22}+\hat{\pi}_{23})/(1-(\hat{\pi}_{21}+\hat{\pi}_{22}+\hat{\pi}_{23}))}{(\hat{\pi}_{11}+\hat{\pi}_{12}+\hat{\pi}_{13})/(1-(\hat{\pi}_{11}+\hat{\pi}_{12}+\hat{\pi}_{13}))}$$

が等しくなるような π_{ij} を推定したい．累積比率は，次のロジスティック分布によって推定される

$$\pi_{i1} = \frac{\exp(\beta_0^{(1)} + \beta_1 x)}{1 + \exp(\beta_0^{(1)} + \beta_1 x)}$$

$$\pi_{i1} + \pi_{i2} = \frac{\exp(\beta_0^{(2)} + \beta_1 x)}{1 + \exp(\beta_0^{(2)} + \beta_1 x)}$$

$$\pi_{i1} + \pi_{i2} + \pi_{i3} = \frac{\exp(\beta_0^{(3)} + \beta_1 x)}{1 + \exp(\beta_0^{(3)} + \beta_1 x)}$$

出現度数を r_{i1} , r_{i2} , r_{i3} , r_{i4} としたときに, 対数尤度

$$\log L = \sum_i (r_{i1} \log \pi_{i1} + r_{i2} \log \pi_{i2} + r_{i3} \log \pi_{i3} + r_{i4} \log \pi_{i4})$$

を最大とするように, 切片 $\beta_0^{(1)}$, $\beta_0^{(2)}$, $\beta_0^{(3)}$, 傾き β_1 を推定する.

表示 4.9

		+++	++	+	-	合計
度数	プラセボ	r_{11}	r_{12}	r_{13}	r_{14}	n_1
	A 薬	r_{21}	r_{22}	r_{23}	r_{24}	n_2
期待値	プラセボ	π_{11}	π_{12}	π_{13}	π_{14}	1
	A 薬	π_{21}	π_{22}	π_{23}	π_{24}	1

LOGISTIC プロシジャでは, プラセボ群の場合 0, A 薬の場合は 1 となるダミー変数を CLASS ステートメントで `drug (param=ref ref=first)` として設定し, MODEL ステートメントで, `y = drug / link=clogit expb ;` のように リンク関数として `clogit` (累積対数オッズ), 分布関数は読み込んだデータの反応カテゴリー数が 3 以上ならば自動的に多項分布となる. 推定値は, 対数オッズ (ロジット) が出力されるので, 指数をとってオッズを合わせて出力するように `expb` を合わせて指定している.

```
data drug ;
  input drug $ @@ ;
  do y1 = 1, 2, 3, 4 ;
    input r @@ ;
    output ;
  end ;
datalines ;
1:Placebo 0 4 12 12
2:A      8 19 18 18
;
proc logistic data=drug ;
  class drug / param=ref ref=first ;
  model y1 = drug / link=clogit expb tech=newton ;
  freq r ;
run ;
```

LIGISTIC プロシジャの出力は多彩であるので, 最も基本となる回帰係数の推定値に限定して示す. 表示 4.4 の推定値は, 切片 $\hat{\beta}_0^{(1)} = -3.0887$, $\hat{\beta}_0^{(2)} = -1.3476$, $\hat{\beta}_0^{(3)} = 0.0877$, 傾き $\hat{\beta}_1 = 0.9883$ である. これらの推定値は, 対数オッズ (ロジット) であり, 変数 `drug`

は、A薬の場合に 1 となるダミー変数であるので、変数drugの回帰係数の推定値 $\hat{\beta}_1 = 0.9883$ は、A薬群とプラセボ群の対数オッズの差となっており、その指数をとるとオッズ比の推定値となり、2.687 がExp(Est)の欄に出力されている。オッズ比の 95%信頼区間は、別途出力されていて、(1.159, 6.227) と下限が 1 以上なのでA薬のプラセボに対するオッズ比 2.687 は統計的に有意であることがわかる。

オッズ比は 2 群間の有効率の比とみなすことができるので、A 薬の効果は、プラセボに対して、2.687 倍と推定する。プラセボの+++有効率が 10%であれば、A 薬は 26.87%であり、+以上の有効率が 70%の場合であれば、A 薬はオッズを単純に掛けると 188%となるので、プラセボが有効でない率（無効率）にし、オッズ比の逆数を掛け、A 薬の無効率を推定してそれから有効率に換算すると $1 - (1 - 0.70) / 2.687 = 0.888$ から 88.8%と推定される。

表示 4.10
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp (Est)
Intercept 1	1	-3.0897	0.4962	38.7708	<.0001	0.046
Intercept 2	1	-1.3476	0.3736	13.0120	0.0003	0.260
Intercept 3	1	0.0877	0.3456	0.0644	0.7997	1.092
drug 2:A	1	0.9883	<u>0.4182</u>	5.5852	0.0181	2.687

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
drug 2:A vs 1:Placeb	2.687	<u>1.159 6.227</u> /* tech=fisher */

理解を確かなものとするために、表示 4.8 ですでに示してあるプラセボおよびA薬の有効率の推定値を、推定された回帰係数を用いて確認してみよう。対数オッズ η_{ij} は、推定された回帰係数の線形和となり、有効率の推定値は、 $\hat{\pi}_{ij} = \exp(\hat{\eta}_{ij}) / (1 + \exp(\hat{\eta}_{ij}))$ で推定される。プラセボおよびA薬それぞれ 3 つの累積有効率は、次のように計算される。

薬剤	累積反応	対数オッズの推定値	累積有効率の推定値
			$\hat{\pi}_{ij} = \exp(\hat{\eta}_{ij}) / (1 + \exp(\hat{\eta}_{ij}))$

プラセボ	+++	$\hat{\eta}_{11} = \hat{\beta}_0^{(1)} + \hat{\beta}_1 \cdot 0 = -3.0897$	0.0435
	++	$\hat{\eta}_{12} = \hat{\beta}_0^{(2)} + \hat{\beta}_1 \cdot 0 = -1.3476$	0.2063
	+	$\hat{\eta}_{13} = \hat{\beta}_0^{(3)} + \hat{\beta}_1 \cdot 0 = -0.0877$	0.5219
A 薬	+++	$\hat{\eta}_{21} = \hat{\beta}_0^{(1)} + \hat{\beta}_1 \cdot 1 = -2.1014$	0.1090
	++	$\hat{\eta}_{22} = \hat{\beta}_0^{(2)} + \hat{\beta}_1 \cdot 1 = -0.3593$	0.4111
	+	$\hat{\eta}_{23} = \hat{\beta}_0^{(3)} + \hat{\beta}_1 \cdot 1 = 1.0760$	0.7457

推定されたそれぞれの有効率の A 薬に対するプラセボのオッズ比は,

$$\hat{OR}^{(1)} = \frac{0.1090 / (1 - 0.1090)}{0.0435 / (1 - 0.0435)} = 2.687$$

$$\hat{OR}^{(2)} = \frac{0.2063 / (1 - 0.2063)}{0.4111 / (1 - 0.4111)} = 2.687$$

$$\hat{OR}^{(3)} = \frac{0.5219 / (1 - 0.5219)}{0.7457 / (1 - 0.7457)} = 2.687$$

とすべて同じになることが確認される. 3つの累積有効率から直接算出したオッズ比は, 8.000, 4.500, 1.875 であり, オッズ比が大きく異なるように思われるが, これらの異なるオッズ比を一つの推定オッズ比で代表させても良いものなのか, どのくらいまでなら代表させて良いのか, その限界はどのくらいなのか, などについては, 第@@章で論ずる.

JMPによる順序データの解析

	A	z1	z2	y	n	n+	n_total
1	1	1	0	1:+++	0	0	28
2	1	1	0	2:++	4	4	28
3	1	1	0	3:+	12	16	28
4	1	1	0	4:-	12	28	28
5	2.3	0	1	1:+++	8	8	63
6	2.3	0	1	2:++	19	27	63
7	2.3	0	1	3:+	18	45	63
8	2.3	0	1	4:-	18	63	63

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片[1:+++]	-3.0900729	0.5039245	37.60	<.0001 *
切片[2:++]	-1.3481128	0.3900909	11.94	0.0005 *
切片[3:+]	0.08723843	0.3611119	0.06	0.8091
z2	0.9886992	0.4288864	5.31	0.0212 *

3水準への拡張，水準間の比較

比例オッズモデルの基礎を A 薬の 2 用量を込みにして述べてきたのであるが，3 水準のまま解析をして，興味のある水準間の比較を行なってみよう．実験前に立てた作業仮説が，結果から統計的に証明されたのか，あるいは予期しなかった結果なのかに忠実にならなければならない．

作業仮説は，様々である．取り上げた因子の水準間に何らかの統計的な差があることを検証したいのか，プラセボを 0mg として，A 薬 10mg，20mg の有効率が直線的あるいは比例的に上昇することを作業仮説とするのか，プラセボと A 薬の 20mg 群の間で有意な差を出すことを第 1 の命題とし，10mg 群は用量反応と副次的に探索する目的とするなど，などによって解析方法が異なる．ここでは，A 薬の 2 用量を合わせてプラセボ群と比較をすることを第 1 の命題とし，第 2 の命題として A 薬の用量間のオッズ比を推定しよう．

前節では，A 薬の 2 つの用量を合わせて 1 水準としてオッズ比を推定したのであるが，プラセボを含めて 3 水準として解析し，プラセボ対 A 薬の 2 用量のオッズ比，A 薬内の 2 用量間のオッズ比を推定してみよう．表示 4.11 に示すように，データステップで，A 薬の 2 用量のデータを r として読み込み全体で 3 水準とし，12 行 3 列の SAS データセットを作成する．LOGISTIC プロシジャの MODEL ステートメントは，2 水準の場合と同じであるが，実際に読み取ったデータの水準数になる．CLASS ステートメントでは，

3 水準の薬剤drugの最初的水準 1:Placeboを基準とするインディケータ型ダミー変数が内部で自動生成される. それらを, w_1, w_2, \dots としたとき, 薬剤drugの 2 番目の水準 2:A10 が w_1 となり, 3 番目の水準 3:A20 が w_2 としよう. CONTRASTステートメントの ESTIMATEオプションを用いて, オッズ比の推定を行う.

表示 4.11

```
data drug2 ;
    input drug $ @@ ;
    do y = 1, 2, 3, 4 ;
        input r @@ ;
        output ;
    end ;
datalines ;
1:Placebo 0 4 12 12
2:A10      3 10 8 10
3:A20      5 9 10 8
;
proc logistic data=drug2 ;
    class drug / param=ref ref=first ;
    model y = drug / link=clogit expb ; /* tech=Newton */
    contrast 'A(1+2) vs P' drug 0.5 0.5 / estimate=both ;
    contrast 'A1 vs A2' drug -1 1 / estimate=both ;
    freq r ;
run ;
```

表示 4.12 に示す結果から, 対数オッズは,

$$\hat{\eta}^{(j)} = \hat{\beta}_0^{(j)} + \hat{\beta}_1 w_1 + \hat{\beta}_2 w_2 = \hat{\beta}_0^{(j)} + 0.8518 \cdot w_1 + 1.1217 \cdot w_2$$

で推定される. CLASS ステートメントで, 最初的水準を基準にしたので, $z_1 = 0, z_2 = 0$ の場合がプラセボ群となり, $w_1 = 1, w_2 = 0$ の場合が A 薬の 10mg 群, $w_1 = 0, w_2 = 1$ の場合が A 薬の 20mg 群であるので, それぞれの水準の対数オッズは,

$$\text{プラセボ群: } \hat{\eta}^{(j)} = \hat{\beta}_0^{(j)}$$

$$\text{A 薬 10mg 群: } \hat{\eta}^{(j)} = \hat{\beta}_0^{(j)} + 0.8518 \times 1 + 1.1217 \times 0$$

$$\text{A 薬 20mg 群: } \hat{\eta}^{(j)} = \hat{\beta}_0^{(j)} + 0.8518 \times 0 + 1.1217 \times 1$$

で推定される.

表示 4.12
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-3.0940	0.5044	37.6297	<.0001	0.045
Intercept 2	1	-1.3492	0.3903	11.9524	0.0005	0.259
Intercept 3	1	0.0886	0.3611	0.0601	0.8063	1.093
drug 2:A10	1	0.8518	0.4857	3.0751	0.0795	2.344
drug 3:A20	1	1.1217	0.4861	5.3236	0.0210	3.070

A 薬の 10mg 群と 20mg 群の対数オッズの平均をプラセボ群との差であるので、

$$([A \text{ 薬 } 10\text{mg 群}] + [A \text{ 薬 } 20\text{mg 群}]) / 2 - [\text{プラセボ群}]$$

各水準の対数オッズから、

$$\begin{aligned} \log OR_{(A2+A1)/2-A1} &= \frac{(\hat{\beta}_0^{(j)} + \hat{\beta}_1) + (\hat{\beta}_0^{(j)} + \hat{\beta}_2)}{2} - \hat{\beta}_0^{(j)} \\ &= \frac{1}{2} \hat{\beta}_1 + \frac{1}{2} \hat{\beta}_2 \end{aligned}$$

が対数オッズの差となり、指数をとることにより、オッズ比となる。A 薬 10mg 群と A 薬 20mg 群の対数オッズの差は、

$$[A \text{ 薬 } 20\text{mg 群}] - [A \text{ 薬 } 10\text{mg 群}]$$

から、

$$\begin{aligned} \log OR_{A3-A2} &= (\hat{\beta}_0^{(j)} + \hat{\beta}_2) - (\hat{\beta}_0^{(j)} + \hat{\beta}_1) \\ &= -1 \cdot \hat{\beta}_1 + 1 \cdot \hat{\beta}_2 \end{aligned}$$

を対数オッズの差を推定し、その指数をとればオッズ比が推定できる。表示 4.11 の CONTRASTステートメントで **drug 0.5 0.5** と **drug -1 1** のように推定された回帰係数の線形和を求める際の定数を指定し、**estimate=both**で対数オッズの差とその指数を出力するように指定している。表示 4.13 にその結果を示す。差についてのWaldの χ^2 検定および対数オッズの推定値とその 95%信頼区間、オッズ比の推定値とその 95%信頼区間が出力されている。

表示 4.13 水準間の対数オッズの差およびオッズ比の推定
Contrast Test Results

Contrast	DF	Wald Chi-Square	Pr > ChiSq
A(1+2) - P	1	5.2916	0.0214
A2 - A1	1	0.3493	0.5545

Contrast Rows Estimation and Testing Results

Contrast	Type	Row	Estimate	Standard Error	Alpha	Confidence Limits	
A (1+2) - P	PARM	1	0.9867	0.4289	0.05	0.1460	1.8275
A (1+2) - P	EXP	1	2.6824	1.1506	0.05	1.1572	6.2180
A2 - A1	PARM	1	0.2699	0.4567	0.05	-0.6252	1.1650
A2 - A1	EXP	1	1.3098	0.5982	0.05	0.5352	3.2059

A 薬の 2 用量とプラセボの対数オッズの差は,

$$\log OR_{(A2+A1)/2-A1} = \frac{1}{2}\hat{\beta}_1 + \frac{1}{2}\hat{\beta}_2 = 0.5 \times 0.8518 + 0.5 \times 1.1217 = 0.9867$$

であり, $SE(\log OR_{(A2+A1)/2-A1}) = 0.4283$ からワルドの $\chi^2 = (0.9867/0.4283)^2 = 5.29$ が自由度 1 の χ^2 分布から $p = 0.0214$ が求められている. 対数オッズの差について指数をとってオッズ比 $OR_{(A2+A1)/2-A1} = \exp(0.9862) = 2.6824$ が求められている. オッズ比の 95%信頼区間は, 対数オッズの差について 95%信頼区間から指数をとって (1.1572, 6.2180) となっている. A 薬の 2 用量はプラセボに対して, オッズ比が 2.6824 でその 95%信頼区間の下限が 1.1572 と 1 を越えないので, 有意な差であることが示されている.

A 薬の 20mg の A 薬の 10mg に対するオッズ比は 1.3098 であり, 95%信頼区間は (0.5352, 3.2059) と 1 をまたいでいるので有意でない.

不良品が, 格下げ品と本当の不良品に分けられ, 「良品, 格下げ品, 不良品」の 3 カテゴリーである場合も同様である. このような場合に用いられる手法が複数の累積比率に対してロジット変換し, 平行な直線を当てはめるのがロジスティック回帰分析である.

まえがきに述べた, 「寸法が小さすぎる (-), ちょうど良い (+), 大きすぎる (++)」という場合は, 2つの考え方がある. 製造条件によって, 寸法が平均的に大きくなったり, 小さくなったりする場合は, カテゴリーに順序があると考え, 累積ロジットを用いる. それに対して, 製造条件によって, 寸法の安定性が悪くなり, 両側の規格外れが増えるという場合には, 通常のロジット変換を用いる. どちらを使うべきかは, 対象についての固有技術と, グラフを合わせて判断する. 統計の用語で表現すると, 「どのような対立仮説に対して検出力を高めるか」によって解析方法を選択することになる.

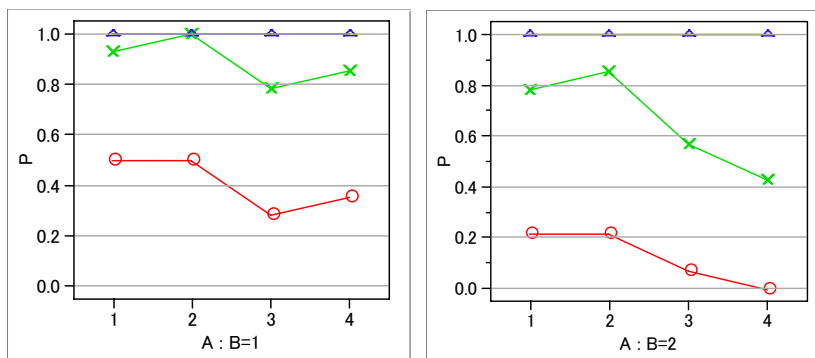
4.3. 2 因子実験

田口（1976）の 68 ページの例を使って，典型的な累積法データの解析例を示す．因子 A は，混紡糸の 4 種類で，因子 B は仕上げ方法の 2 種類で，共に質的因子である．実験は，2 元配置法で作った 8 種類の生地を 7 人の人間に 8 種類の生地の風合いを（手触り）の良さを 上中下 に評価してもらい，翌日にも同じように評価してもらった．「人」は変量因子として解析する必要があるが，ここでは簡単のために，それぞれの生地に 14 人ずつ 112 人がランダムに 上中下 に評価をしたとみなした．その結果を表示 4.14 に示す．

表示 4.14 2 因子実験のデータ

	度数			累積度数		
	上	中	下	上	中	下
$B_1 A_1$	7	6	1	7	13	14
$B_1 A_2$	7	7	0	7	14	14
$B_1 A_3$	4	7	3	4	11	14
$B_1 A_4$	5	7	2	5	12	14
$B_2 A_1$	3	8	3	3	11	14
$B_2 A_2$	3	9	2	3	12	14
$B_2 A_3$	1	7	6	1	8	14
$B_2 A_4$	0	6	8	0	6	14

2 因子以上の実験，直交表実験から得られた順序データに対しても，1 因子実験の考え方を拡張して，判定が上の率，中以上率にたいする累積反応率としてロジスティック回帰を適用できる．統計量として水準間のオッズ比を用いるので，基準となる水準の組み合わせを第 1 水準としておくとよい．ここでは， $A_1 B_2$ とする．表示 4.15 に累積反応率をグラフ化すると因子 B については第 1 水準が，因子 A については第 1 および第 2 水準での評価が高いように思われる．オッズ比による解析を試みよう．



表示 4.15 累積パーセント

○ : 上, × : (上+中)

2 因子以上の実験では、無視してよい交互作用の同定をすることかが解析の第 1 歩である。表示 4.16 に示すようにデータステップで、4 列 24 行の SAS データセットを作成し、LOGISTIC プロシジャの CLASS ステートメントで、因子 A は `ref=first` と第 1 水準、因子 B は `ref=last` のように最後の第 2 水準を基準としたインディケータ型のダミー変数を生成させる。MODEL ステートメントでは、 $y = A \ B \ A*B$ のように因子 A と因子 B の交互作用項を入れる。

表示 4.16 2 因子実験

```
data taguchi ;
  input A @@ ;
  do B=1, 2 ;
    do y = 1, 2, 3 ;
      input r @@ ;
      output ;
    end; end ;
datalines ;
1 7 6 1 3 8 3
2 7 7 0 3 9 2
3 4 7 3 1 7 6
4 5 7 2 0 6 8
;
proc logistic data=taguchi ;
  class A (param=ref ref=first)
        B (param=ref ref=last );
  model y = A B A*B / link=clogit expb ;
  freq r ;
run ;
```

ロジスティック回帰の結果は、表示 4.17 に示すように分散分析表のスタイルで出力される。 F 値ではなく、 χ^2 値が検定統計量となり、因子 A と B の交互作用は、 $\chi^2 = 1.6629$ が自由度 3 の χ^2 分布に従うことから $p = 0.6452$ となり、交互作用は検出されなかった。

表示 4.17 交互作用の検討
Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
A	3	8.5808	0.0354
B	1	3.2919	0.0696
A*B	3	1.6629	0.6452

MODEL ステートメントから交互作用 `A*B` を除き再実行し表示 4.18 を得る。因子 A は $p = 0.0149$ と有意、因子 B は $p < 0.0001$ と高度に有意である。ロジスティック回帰回

帰係数は、手触りのよさが [上] の場合の対数オッズの切片 1, [上+中] の場合の切片 2 が推定値として出力されている。これは参照型のダミー変数性質からの A_1B_2 の出現率の推定対数オッズとなっている。次いで因子Aの第 2, 3, 4 水準の第 1 水準との対数オッズの差, Exp(Est) に指数をとったオッズ比が出力されている。次のオッズ比の 95%信頼区間から、因子Aの混紡糸は、現行第 1 水準にくらべ第 2 水準はオッズ比が 1.118 と同程度であるが、第 2 水準と第 3 水準は、オッズ比がそれぞれ 0.332, 0.297 であり、95%の上限が 1 を含んでいないので、有意に生地風の合いが落ちている。因子Bは、仕上げ方法であり、現行の第 2 水準に対して、第 1 水準のオッズ比は 4.934 と有意に優れている。仕上げ方法の第 1 水準は、風合いに対する効果は大きく、混紡糸の種類によらず、現行の A_1B_2 の風合いより優れた効果を出すことが、表示 4.15 の累積比率の図から読み取れる。

表示 4.18 主効果のみでの解析結果
Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
A	3	10.4779	0.0149
B	1	15.7560	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept 1	1	-1.4984	0.4548	10.8539	0.0010	0.223
Intercept 2	1	1.2533	0.4413	8.0645	0.0045	3.502
A	2	0.1723	0.5274	0.1067	0.7439	1.188
A	3	-1.1032	0.5323	4.2954	0.0382	0.332
A	4	-1.2145	0.5349	5.1546	0.0232	0.297
B	1	1.5961	0.4021	15.7560	<.0001	4.934

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
A 2 vs 1	1.188	0.423 3.340
A 3 vs 1	0.332	0.117 0.942
A 4 vs 1	0.297	0.104 0.847
B 1 vs 2	4.934	2.243 10.850

因子Bの仕上げが風合いにおよぼす影響は、オッズ比で 4.934 倍と推定されるのであるが、実際の出現率に換算したときに現行の第 2 水準の反応率が 5%であった場合におよそ第 2 水準は 5 倍の 25%弱になると推定される。正確には、反応率が 0.05 であつ

た場合に、そのオッズは、 $odds_1 = 0.05 / (1 - 0.05) = 0.053$ であり、4.934 倍のオッズは、

$$odds_2 = 0.053 \times 4.934 = 0.260$$

となるので、改良水準の反応率は

$$odds_2 / (1 + odds_2) = 0.206$$

となる．表示 4.19 にいくつかの反応率に対して、オッズ比が 4.934 倍となる反応率を示す．50%の場合は 83.1%，80%の場合は 95.2%となることがわかる．

表示 4.19 オッズ比で 4.934 倍の反応率

現行水準		4.934 倍の改良水準	
反応率	オッズ	オッズ	反応率
0.01	0.010	0.050	0.047
0.05	0.053	0.260	0.206
0.10	0.111	0.548	0.354
0.20	0.250	1.233	0.552
0.50	1.000	4.934	0.831
0.80	4.000	19.735	0.952

因子 A と因子 B の水準を組み合わせた場合の効果の推定を行ってみよう．LOGISTIC プロシジャでは、OUTPUT ステートメントを用いて対数オッズと反応率の推定値を SAS データデータセットとして得ることができるが、推定された回帰係数から求めてみよう． A_1B_1 の風合いが [上] の場合は、

$$\log(odds) = [\text{Intecept } 1] + [A \ 1] + [B \ 1] = -1.4984 + 0 + 1.5961 = 0.09765$$

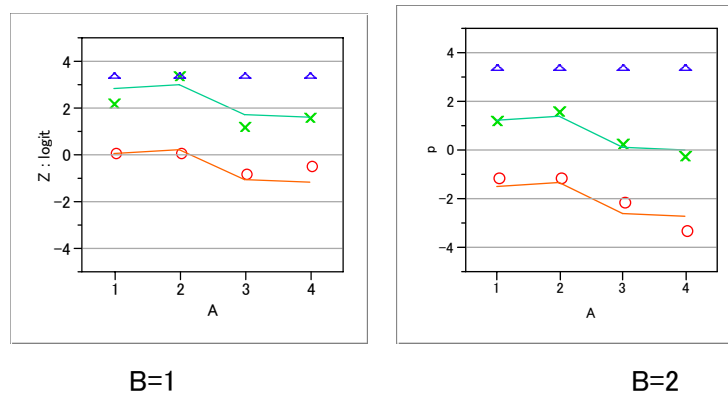
$$\hat{p}_{1,1}^{(1)} = \exp(0.09765) / (1 + \exp(0.09765)) = 0.52439$$

対数オッズは 0.097 となり、反応率は 0.524 と推定される．風合いが [上+中] の場合は、

$$\log(odds) = [\text{Intecept } 2] + [A \ 1] + [B \ 1] = 1.2533 + 0 + 1.5961 = 2.84936$$

$$\hat{p}_{1,1}^{(2)} = \exp(2.84936) / (1 + \exp(2.84936)) = 0.94529$$

対数オッズは 2.849 となり、反応率は 0.945 と推定される．対数オッズについて観測値と推定値のを重ね書きした図を表示 4.20 に示す．



表示 4.20 効果の推定値

累積ロジットを用いたロジスティック回帰分析の結果と田口の累積法の結果と比べたのが表示 4.21 である。累積法では、検定統計量として分散比 F が用いられる。他の方法の χ^2 と比較するため、 F 比に分子の自由度を掛けた値を示す。累積法の解は、 p 値が小さくなり、第 1 種の誤りの確率が増す。しかし、 p の推定値は 2 つの方法の間にはほとんど違いがない。これから、推定としては、全く問題ないが、検定としては注意を要するということが分かる。

従来の累積法の計算に比べて計算が極めて容易であるから、今後は、累積ロジスティック回帰を利用すると良いであろう。

表示 4.21 JMP の解との比較

	A		B		$\hat{P}_{1(2,1)}$	$\hat{P}_{2(2,1)}$
	χ^2	p 値	χ^2	p 値		
JMP	10.48	1.5%	15.76	0.0%	0.567	0.954
累積法	18.75	0.5%	27.67	0.0%	0.515	0.969

JMPでの解析

累積2元_混紡糸

累積2元_混紡糸
ロックされたファイル E:\R1_進行中

列(0/0)

B
A
y
r
R
n
P
P*
Z*

行

すべての行 24
選択されている行 0
除外されている行 0
表示しない行 0
ラベルのついた行 0

	B	A	y	r	R	n	P
1	1	1	1	7	7	14	0.50
2	1	1	2	6	13	14	0.93
3	1	1	3	1	14	14	1.00
4	1	2	1	7	7	14	0.50
5	1	2	2	7	14	14	1.00
6	1	2	3	0	14	14	1.00
7	1	3	1	4	4	14	0.29
8	1	3	2	7	11	14	0.79
9	1	3	3	3	14	14	1.00
10	1	4	1	5	5	14	0.36
11	1	4	2	7	12	14	0.86
12	1	4	3	2	14	14	1.00
13	2	1	1	3	3	14	0.21
14	2	1	2	8	11	14	0.79
15	2	1	3	3	14	14	1.00
16	2	2	1	3	3	14	0.21
17	2	2	2	9	12	14	0.86
18	2	2	3	2	14	14	1.00
19	2	3	1	1	1	14	0.07
20	2	3	2	7	8	14	0.57
21	2	3	3	6	14	14	1.00
22	2	4	1	0	0	14	0.00
23	2	4	2	6	6	14	0.43
24	2	4	3	8	14	14	1.00

Type2 の平方和 のようだ

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
A	3	3	10.5526337	0.0144 *
B	1	1	15.9963523	<.0001 *
B*A	3	3	1.66288354	0.6452

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
A	3	3	10.47787	0.0149 *
B	1	1	15.7560169	<.0001 *

5. 繰り返し実験がある 2 値反応, モデルの適合

第 3 章で取り上げたオセタミビルのインフルエンザの予防試験では, それぞれの臨床試験でオセタミビルのインフルエンザ予防効果を検証するためにプラセボ群との相対的な比較ができるようにデザインされていた. 比較対照のプラセボ群のインフルエンザ発病率は, 対象とする集団でかなり異なっているようである. この章では, プラセボ群のみに着目して対象集団間でインフルエンザ発病率の違いをオッズ比によって計量してみよう.

5.1. 繰り返し実験

プラセボ群のみのデータを表示 3.13 に示す. この実験のデザインは, 4 種類の対象集団に対して, A3: 高齢者を除いて 2 回の実験が繰り返されている. 2 回の繰り返しのインフルエンザ発病率が 2 倍程度あり, 同じ対象集団内でも無視できないそうもないインフルエンザ発病率の変動がある. 対象集団が家族内の場合には, 2 つの試験間でオッズ比は 1.799 となり 95%信頼区間は (1.095, 2.955) と下限が 1 を含まない. この変動は実験の繰り返しに伴う誤差変動そのものであり, オッズ比の 95%信頼区間を計算する上で考慮しなければならない.

表示 5.1 一元配実験・繰り返しあり

対象集団	No. 著者 年	プラセボ群		発現率	集団内 オッズ比
		r	n		
A1: 実験的研究	1: Hayden 1999 A 型	8	12	0.667	—
	2: Hayden 2000 B 型	17	19	0.895	4.250
A2: 一般成人	3: Hayden 1999	25	519	0.048	—
	4: 柏木 2000	13	155	0.084	1.809
A3 高齢者	5: WV15825	12	272	0.044	—
A4: 家族内	6: Welliver 2001	43	206	0.209	1.799
	7: Hayden 2004	33	258	0.128	—
全体		151	1441	0.105	

表示 5.2 示すように LOGISTIC プロシジャの MODEL ステートメントのオプションで SCALE=DEVIENCE オプションを指定することにより, 繰り返しに伴う発現率の過大な変動を対数オッズの差の SE に組み込むことができる. CLASS ステートメントでは, A2: 一般成人を基準になるような参照型のダミー変数を生成するようにしている. そうすることにより, A2: 一般成人を基準とするオッズ比が得られる.

表示 5.2 過大な変動の組み込み

```

data placebo4r ;
    input A $ ex $16. r n ;
datalines ;
A1 1:Hayden_1999      8 12
A1 2:Hayden_2000     17 19
A2 3:Hayden_1999     25 519
A2 4:柏木_2000       13 155
A3 5:WV15825         12 272
A4 6:Welliver_2001  43 206
A4 7:Hayden_2004     33 258
;
proc logistic data=placebo4r ;
    class A (param=ref ref='A2') ;
    model r / n = A / link=logit expb tech=newton scale=deviance;
run ;

```

A1:実験的研究に対しては、オッズ比は、69.734 であり、95%信頼区間の下限も 11.861 であり、統計的に判定するまでもなく、インフルエンザの発病率は高い。A3:高齢者に対しては、一般成人に対してインフルエンザの発病率は低いことから、オッズ比が 0.773 であり、1 より小さいが、95%信頼区間が 1 を含んでいるので、有意な差ではない。A4:家族内に対しては、3.278 となり、95%信頼区間の下限が 1.527 なので有意な差である。

SCALEオプションで集団内の変動を考慮にいと 95%信頼区間が広がるが、SCALEオプションを付けずに集団内の変動を無視した場合には、表示 5.4 に示すように、A4:家族内に対して、95%信頼区間の下限が 2.177 となり、95%信頼区間が狭くなり、有意な差が出やすくなる。

表示 5.3 一般成人を基準にしたオッズ比の推定
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept	1	-2.8176	0.3117	81.7378	<.0001	0.060
A A1	1	4.2447	0.9038	22.0565	<.0001	69.734
A A3	1	-0.2582	0.6330	0.1663	0.6834	0.772
A A4	1	1.1873	0.3898	9.2792	0.0023	3.278

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
A A1 vs A2	69.734	11.861 409.990
A A3 vs A2	0.772	0.223 2.671
A A4 vs A2	3.278	1.527 7.038

表示 5.4 集団内の変動を無視したオッズ比の 95%信頼区間
Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
A A1 vs A2	69.734	26.990	180.172
A A3 vs A2	0.772	0.397	1.502
A A4 vs A2	3.278	2.177	4.937

SCALEオプションを用いて集団内の変動を考慮にいたしたオッズ比の推定方法は、どのような手順で行われているのであろうか。LOGISTICプロシジャで、`scale=devience`とオプションを付けると表示 5.5 に示すように、因子Aの各水準の実験の繰り返しに起因するデビエンスが 10.4482 であり、実験の繰り返しがの自由度が 3 であることが示されている。デビエンスが自由度が 3 の χ^2 分布に従うとしたときの p 値は 0.0151 であり、各水準内のインフルエンザ発現率の変動は、2 項分布を前提とした誤差変動よりも大きく、統計的にも有意であることが示されている。

デビエンス 10.4482 を自由度 3 で割った 3.4827 を不均一性を計量する回帰係数として、共分散行列に 3.4827 を掛けるとのメッセージが出力されている。この結果、回帰係数の推定値の分散が 3.4827 倍となり、SE はその平方根 1.8662 倍が出力され、それに伴い 95%信頼区間も広がることになる。

表示 5.5 デビエンスの不均一性
Deviance and Pearson Goodness-of-Fit Statistics

Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	10.4482	3	3.4827	0.0151
Pearson	10.7745	3	3.5915	0.0130

Number of events/trials observations: 7

NOTE: The covariance matrix has been multiplied by the heterogeneity factor
(Deviance / DF) 3.48274.

これまでは、ロジスティック回帰の解法である最尤法の原理については、意図的に触れてこなかったのであるが、この事例について最尤法の考え方について詳しく示す。

5.2. 因子が質的な場合のロジスティック回帰の基礎

これまでの章では、ロジスティック回帰の数理については、概念的な記述にとどめ、応用事例を通じてロジスティック回帰の使い方とオッズ比をベースにした結果の解釈を優先してきたのであるが、さらなる現実の問題にロジスティック回帰を使って解析するためには、数理的な基礎知識が結果を適切に用いるために不可欠である。

ロジスティック回帰は、最尤法により回帰係数の推定を行っている。最適な推定値を得るために、求めたい回帰係数に初期値を与えて逐次的に最適な推定値を得るという方法をとっている。そのために、解析方法は煩雑かつ難解で、理解を深めるための計算手順を示すことも、かえって妨げになる。最尤法の原理は単純明快であるので、煩雑な繰り返し計算については、第@@章で述べることにして、ここでは、最尤法の考え方、得られた結果の見かた使い方に焦点をあてる。

切片のみのモデル

切片のみのモデルを考えよう。切片は、回帰直線 $y = \beta_0 + \beta_1 x + \varepsilon$ の場合には、 $x = 0$ の場合に直線が Y 軸を横切る点であるが、 $y = \beta_0 + \varepsilon$ のように切片だけのモデルでは、切片 $\hat{\beta}_0$ は y の総平均の推定値になる。最小 2 乗法は、総平均からの偏差平方和が最小になるように総平均を推定するのであるが、最尤法では、尤度が最大となるように総平均を推定する。総平均の推定値は、最小 2 乗法でも最尤法でも、算術平均になることが第@@章に示すように簡単に証明できる。2 値反応の場合に総平均は、全試験を込みにしたインフルエンザの発現率となるので、実際に切片のみのモデルでの尤度を計算してみよう。

表示 3.13 に示した 7 つの臨床試の症例数を n_i , $i=1, \dots, 7$, インフルエンザ発病数を r_i , $i=1, \dots, 7$ とする。7 つの集団全体の発病率の平均 p_0 は、単純に症例数を足し合わせて、

$$p_0 = \frac{\sum_i r_i}{\sum_i n_i} = \frac{8+17+\dots+33}{12+19+\dots+258} = 0.1048$$

である。 $p_0 = 0.1048$ を母集団の発現率 $\hat{\pi}_0$ としたときに、それぞれの集団の発病率が 2 項分布に従うと仮定したときに、その発病率が起きる確率 L_i を、

$$L_i = \text{Binomial Density}(\hat{\pi}_0, r_i, n_i)$$

とす。2 項分布の確率密度の計算は、SAS のデータステップで PDF 関数でもできるが Excel の関数¹²でも手軽に計算できるので実際に計算してみると、

¹² Excel の場合: BINOMDIST(成功数, 試行回数, 成功率, 関数形式), 関数形式が FALSE の場合は、

$$L_1 = \text{Binomial Density}(\hat{\pi}_0, r_1, n_1) = \text{Binomial Density}(0.1048, 8, 12) = 0.00000462$$

...

$$L_7 = \text{Binomial Density}(\hat{\pi}_0, 33, 258) = 0.03713035$$

であり、 $\pi_0 = 0.1048$ であった場合に試験 1 の発現率 0.667 となることは $L_1 < 0.00001$ なので極めてまれであることがわかり、試験 7 の発現率 0.128 は、 $L_7 = 0.037$ なので、起こりうる範囲と思われる。これらの 7 試験の発現率が得られる確率 L は、 $L = \Pi_i(L_i) = L_1 L_2 \cdots L_7$ である。 L が尤度 (likelihood) といわれているものであり、これからは確率の積ではなく、尤度ということにする。尤度のままでは取り扱いにくいので対数を取った対数尤度を最大となるように回帰係数の推定を行う。

$$\begin{aligned} \log L &= \log[\Pi_i(L_i)] = \log L_1 + \log L_2 + \cdots + \log L_7 \\ &= -12.2847 - 33.4285 - \cdots - 3.2933 = -85.9403 \end{aligned}$$

2 項分布の確率を正確に計算しなくても、尤度の本質を損ねないように簡略化した式が次のように定式化されている。2 項分布、

$$L_i = \text{Binomial Density}(\hat{\pi}_0, r_i, n_i) = \binom{n_i}{r_i} \hat{\pi}_0^{r_i} (1 - \hat{\pi}_0)^{n_i - r_i}$$

の組み合わせの項は、尤度 L を最大にするようなパラメータ π_0 を探索する際に定数項となるので、それを除いたもの

$$\log L'_i = \log[\hat{\pi}_0^{r_i} (1 - \hat{\pi}_0)^{n_i - r_i}] = r_i \log \hat{\pi}_0 + (n_i - r_i) \log(1 - \hat{\pi}_0)$$

が、広く使われている。実際に計算すると、

$$\begin{aligned} \log L'_1 &= r_1 \log \hat{\pi}_0 + (n_1 - r_1) \log(1 - \hat{\pi}_0) \\ &= 8 \cdot \log(0.1048) + (12 - 4) \cdot \log(1 - 0.1048) = -18.4893 \end{aligned}$$

...

$$\begin{aligned} \log L'_7 &= r_7 \log \hat{\pi}_0 + (n_7 - r_7) \log(1 - \hat{\pi}_0) \\ &= 33 \cdot \log(0.1048) + (258 - 33) \cdot \log(1 - 0.1048) = -99.3482 \end{aligned}$$

$$\begin{aligned} \log L' &= \log L'_1 + \log L'_2 + \cdots + \log L'_7 \\ &= -18.4893 - 38.5702 - \cdots - 99.3482 = -483.4244 \end{aligned}$$

$\log L' = -483.4244$ が得られる。

対象集団ごとのあてはめ

4 つの集団それぞれの母集団のインフルエンザ出現率の推定値として、2 つの試験の

確率密度関数となる。SAS の場合：PDF('BINOMIAL', m, p, n) で 2 項分布の確率が計算できる。

症例数を合わせて出現率とする．1 試験しかない A3：高齢者は，1 試験の出現率を母集団のインフルエンザ出現率の推定値とする．

$$\hat{\pi}_{A_1} = p_{A_1} = \frac{r_1 + r_2}{n_1 + n_2} = \frac{8 + 17}{12 + 19} = 0.8065$$

...

$$\hat{\pi}_{A_4} = p_{A_4} = \frac{r_6 + r_7}{n_6 + n_6} = \frac{43 + 33}{206 + 258} = 0.1638$$

対数尤度は，

$$\begin{aligned} \log L'_1 &= r_1 \log \hat{\pi}_{A_1} + (n_1 - r_1) \log(1 - \hat{\pi}_{A_1}) \\ &= 8 \cdot \log(0.8065) + (12 - 4) \cdot \log(1 - 0.8065) = -8.2898 \end{aligned}$$

...

$$\begin{aligned} \log L'_7 &= r_7 \log \hat{\pi}_{A_4} + (n_7 - r_7) \log(1 - \hat{\pi}_{A_4}) \\ &= 33 \cdot \log(0.1638) + (258 - 33) \cdot \log(1 - 0.1638) = -99.9498 \end{aligned}$$

$$\begin{aligned} \log L' &= \log L'_1 + \log L'_2 + \cdots + \log L'_7 \\ &= -8.2898 - 6.9413 \cdots - 99.948 = -417.4964 \end{aligned}$$

$\log L' = -417.4964$ が得られる．

7 試験をすべて個別の集団と仮定

7 試験をすべて個別の集団とみなし，母集団の出現率としたとき，

$$\hat{\pi}_1 = p_1 = \frac{r_1}{n_1} = \frac{8}{12} = 0.6667$$

...

$$\hat{\pi}_7 = p_7 = \frac{r_7}{n_6} = \frac{33}{258} = 0.1279$$

対数尤度は，

$$\begin{aligned} \log L'_1 &= r_1 \log \hat{\pi}_1 + (n_1 - r_1) \log(1 - \hat{\pi}_1) \\ &= 8 \cdot \log(0.6667) + (12 - 4) \cdot \log(1 - 0.6667) = -7.6382 \end{aligned}$$

...

$$\begin{aligned} \log L'_7 &= r_7 \log \hat{\pi}_7 + (n_7 - r_7) \log(1 - \hat{\pi}_7) \\ &= 33 \cdot \log(0.1279) + (258 - 33) \cdot \log(1 - 0.1279) = -98.6562 \end{aligned}$$

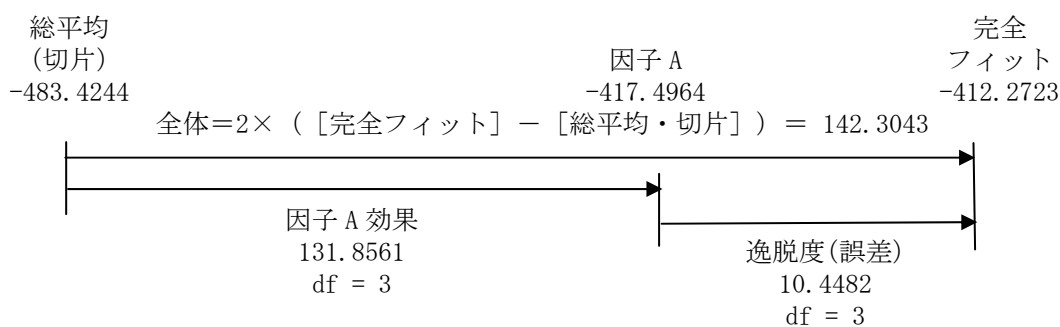
$$\begin{aligned}\log L' &= \log L'_1 + \log L'_2 + \cdots + \log L'_7 \\ &= -76382 - 6.3934 \cdots - 98.6562 = -412.2723\end{aligned}$$

$\log L' = -412.2723$ が得られる.

分散分析表スタイル

総平均（出現率）モデル，水準平均（出現率）モデル，および，個別平均（出現率）モデル，それぞれの対数尤度を，分散分析表のスタイルでまとめてみよう．表示 5.6 にこれらの尤度の関係を示す．パラメータを 1 つ用いた総平均モデルの対数尤度は $\log L'_{\text{Total}} = -483.4244$ であり，パラメータを 7 つ用いた個別平均モデルの対数尤度は $\log L'_e = -412.2723$ となり，尤度の増加は，71.1544 である．使用したパラメータ数の差を自由度 $df_{\text{Total}} = 7 - 1 = 6$ としたときに，対数尤度の差の 2 倍 $2 \times (-412.2723 - (-483.4244)) = 2 \times 71.1544 = 142.3043$ が，自由度 6 の χ^2 分布に従うことを利用して尤度比検定が行える．

パラメータを 3 つ加えた水準平均モデルの対数尤度は $\log L'_A = -417.4964$ と尤度が総平均モデルの尤度から 65.9281 増加する．したがって， $2 \times 65.9281 = 131.8561$ が自由度 3 の χ^2 分布に従うことを利用して，因子 A の効果について尤度比検定が行える．総平均モデルの効果（尤度の差）から因子 A の効果を差し引いた 10.4482 が逸脱度 Deviance（誤差）である．逸脱度の自由度は，総平均モデルの自由度 6 から因子 A の自由度 3 を引いた 3 が逸脱度（誤差）の自由度となる．自由度 3 の χ^2 分布の上側確率は 0.015 なので，この逸脱度（誤差）には，まだ誤差的な変動でない未知の要因効果が含まれていることが示唆される．



表示 5.6 対数尤度より χ^2 統計量の算出図

パラメータ数が多くなればなるほど対数尤度は大きくなるのであるが，追加されたパラメータが統計的に考慮すべき価値があるか否かは，パラメータ数の増加にみあう対数尤度の増加があるかである．対数尤度の差の 2 倍が，加えたパラメータ数を自由度とする

χ^2 分布に従うことを利用して、それが妥当かの判断を行うのである。

切片+因子Aのモデルで、因子Aの効果は尤度の増加分の2倍の $\chi^2 = 131.8561$ であり、自由度 3 の χ^2 分布の上側確率は $p < 0.0001$ で、対象集団間に明らかなインフルエンザの発現率に差があると結論される。ここで注意しなければならないのは、因子Aをあてはめた後にも、2項分布を逸脱する誤差の変動が残っていることである。これを考慮する方法の一つとして、因子Aの χ^2 値を誤差変動（逸脱度）の χ^2 値で割りF値とすると、分子の自由度が3、分母の自由度が3のF分布に従うことを利用してF検定を行う。表示 5.7 に示すように、 $F = 12.62$ 、 $p = 0.0330$ となる。したがって、因子Aの各水準内でのインフルエンザ発現率の変動を考慮しても、因子Aは統計的に有意な差があると結論される。

誤差的な変動とは認めがたい逸脱度である場合を、過大発散（Overdispersion）の問題といわれている。過大発散を考慮した解析の一つが、ここで示したF検定に持ち込む方法である。この方法は、 n_i がほぼ等しいことを前提としているので、インフルエンザの例では、例数の違いが20倍もあり、この方法の適用には問題＜＜どのような、示す＞＞があるので、第@@章では、 n_i が等しくない場合の解析方法について詳しく示す。

表示 5.7 分散分析表スタイル

要因	対数尤度	増加	df	χ^2	χ^2/df	F 値	p 値
切片モデル	-483.4244	0	6	—			
切片+因子 A	-417.4964	65.9281	3	131.8561	43.9520	12.62	0.0330
逸脱度・誤差	—	—	3	10.4482	3.4827		
完全フィット	-412.2723	71.1544	0	0			

5.3. 理解を深めるための GENMOD プロシジャ

切片のみのモデル

GENMOD プロシジャによる切片のみのモデルを次に示す。MODEL ステートメントでは、何も右辺に指定しないことにより切片のみのモデルとなる。オプションに **pred** を入れることにより、各試験の発現率の推定値を出力させ、すべて同じになることを確認する。

```
proc genmod data=placebo4r ;
  model r / n = / dist=bin link=logit pred ;
```

この結果、表示 5.8 に示すように、対数尤度として-483.4244 が出力され、さらに、完全にフィットしたモデルとの対数尤度との差の2倍、デビエンスが 142.3043 と出力さ

れている。各実験の出現率の推定値は、

$$p_0 = \frac{\sum_i r_i}{\sum_i n_i} = \frac{8+17+\cdots+33}{12+19+\cdots+258} = 0.1048$$

と推定されている。0.1048 の対数オッズは、Xbeta の欄で-2.1451 であることがわかる。

表示 5.8 切片のみのモデル
Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	142.3043	23.7174
Scaled Deviance	6	142.3043	23.7174
Pearson Chi-Square	6	221.1060	36.8510
Scaled Pearson X2	6	221.1060	36.8510
Log Likelihood		-483.4244	

Observation Statistics

Observation	r	n	Pred	Xbeta	Std	HessWgt
1	8	12	0.1047883	-2.145118	0.0860099	1.1256929
2	17	19	0.1047883	-2.145118	0.0860099	1.7823472
3	25	519	0.1047883	-2.145118	0.0860099	48.68622
4	13	155	0.1047883	-2.145118	0.0860099	14.5402
5	12	272	0.1047883	-2.145118	0.0860099	25.515707
6	43	206	0.1047883	-2.145118	0.0860099	19.324395
7	33	258	0.1047883	-2.145118	0.0860099	24.202398

水準平均モデル

水準平均モデルを次に示す。MODE ステートメントに因子 A を入れる。

```
proc genmod data=placebo4r ;
  class A ;
  model r / n = A / dist=bin link=logit pred ;
```

この結果、表示 5.9 に示すように、対数尤度として-417.4964 が出力され、さらに、完全にフィットしたモデルとの対数尤度との差の 2 倍、デビエンスが 10.4482 と出力されている。各実験の出現率の推定値は、

$$\hat{\pi}_{A_1} = p_{A_1} = \frac{r_1 + r_2}{n_1 + n_2} = \frac{8+17}{12+19} = 0.8065$$

...

$$\hat{\pi}_{A_4} = p_{A_4} = \frac{r_6 + r_7}{n_6 + n_7} = \frac{43+33}{206+258} = 0.1638$$

と推定されている。

表示 5.9 水準平均のモデル

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	3	10.4482	3.4827
Scaled Deviance	3	10.4482	3.4827
Pearson Chi-Square	3	10.7745	3.5915
Scaled Pearson X2	3	10.7745	3.5915
Log Likelihood		-417.4964	

Observation Statistics

Observation	r	n	Pred	Xbeta	Std	HessWgt
1	8	12	0.8064495	1.4271031	0.4546042	1.8730641
2	17	19	0.8064495	1.4271031	0.4546042	2.9656848
3	25	519	0.0563798	-2.817612	0.1669973	27.611393
4	13	155	0.0563798	-2.817612	0.1669973	8.246177
5	12	272	0.0441176	-3.075775	0.2952617	11.470588
6	43	206	0.1637931	-1.630272	0.1254401	28.214775
7	33	258	0.1637931	-1.630272	0.1254401	35.336951

完全フィット

完全フィットの対数尤度との差の2倍は、デビエンスとして出力されているのであるが、実際に対数尤度を求めて、確認をしてみる。MODEL ステートメントで、7つの実験をを7水準のまま次のように因子として指定する。

```
proc genmod data=placebo4r ;
  class ex ;
  model r / n = ex / dist=bin link=logit pred ;
run ;
```

この結果、表示 5.10 に示すように、対数尤度として-412.2723 が出力されている。完全にフィットしたモデルとの対数尤度との差の2倍のデビエンスは、当然のことであるが0である。各実験の出現率の推定値は、

$$\hat{\pi}_1 = p_1 = \frac{r_1}{n_1} = \frac{8}{12} = 0.6667$$

...

$$\hat{\pi}_7 = p_7 = \frac{r_7}{n_6} = \frac{33}{258} = 0.1279$$

と推定されていることが確認される。

表示 5.10 完全フィットのモデル
Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	0	0.0000	.
Scaled Deviance	0	0.0000	.
Pearson Chi-Square	0	0.0000	.
Scaled Pearson X2	0	0.0000	.
Log Likelihood		-412.2723	

Observation Statistics

Observation	r	n	Pred	Xbeta	Std	HessWgt
1	8	12	0.6666667	0.6931472	0.6123724	2.6666667
2	17	19	0.8947368	2.1400662	0.747545	1.7894737
3	25	519	0.0481696	-2.98366	0.2049983	23.795761
4	13	155	0.083871	-2.390878	0.2897677	11.909677
5	12	272	0.0441176	-3.075775	0.2952617	11.470588
6	43	206	0.2087379	-1.33255	0.1714374	34.024272
7	33	258	0.127907	-1.919593	0.1864067	28.77907

5.4. 最小 2 乗法による分散分析との対比

反応が計量値の場合には、誤差の分布が正規分布に従うことを仮定した最小 2 乗法による解散手順が標準的に適用されている。反応が 2 値の場合には、誤差の分布が 2 項分布に従うと仮定するので、最小 2 乗法を適用することはできないが、実験の繰り返しがあるような場合にロジスティック回帰の解析手順を理解するためには、最小 2 乗法による分散分析と対応することにより理解を深めることができる。

表示 3.13 は、繰り返しがある一元配置実験で、それぞれ反復測定が行われている。実験は、繰り返しを含む因子 A についてランダムに実験が行われ、それぞれ実験単位内で反復測定がされた。実験結果は、1 と 0 であり、それぞれの数をカウントして表示 5.11 に整理した。作成する分散分析のイメージは、要因と自由度をを使って、

表示 5.11 一元配実験・枝分かれ・反復測定

対象集団	No. 著者 年	1	0	n
A1:実験的研究	1:Hayden 1999 A 型	8	4	12
	2:Hayden 2000 B 型	17	2	19
A2:一般成人	3:Hayden 1999	25	494	519
	4:柏木 2000	13	142	155
A3 高齢者	5:WV15825	12	260	272
A4:家族内	6:Welliver 2001	43	163	206
	7:Hayden 2004	33	225	258
全体		151	1290	1441

この実験の分散分析表は 表示 5.12 としてまとめられる。1 次誤差の分散を 2 次誤差

の分散で割った F 値が有意でなければ、1次誤差の平方和を2次誤差の平方和にプーリングした誤差分散を再度求め、因子Aの分散を誤差分散を割って F 検定を行う。1次誤差の分散を2次誤差の分散で割った F 値が有意であれば、因子Aは、1次誤差で評価する。この F 値は、分子の自由度は3、分母の自由度は3である。

表示 5.12 一元配実験・枝分かれ・反復測定の分散分析表

要因	自由度	平方和	平均平方	F 値
A	3	S_A	$V_A = S_A / 3$	V_A / V_{e_1}
1次誤差	3	Se_1	$Ve_1 = Se_1 / 3$	Ve_1 / Ve_2
2次誤差	1434	Se_2	$Ve_2 = Se_2 / 1434$	
全体	1440	S_T		

実際の計算は、表示 5.13 に示すようにSASのGLMプロシジャで、逐次型の平方和（SASではType1 と称している）によって求める。DATAステップでは、インフルエンザ発病した場合を $y = 1$ 、しなかった場合を $y = 0$ としてそれぞれの症例数を n として入力し、14行4列のSASデータセットを作成している。

GLM プロシジャでは、MODEL ステートメントで因子 A と実験番号 ex を指定し、逐次型の平方和での解析を行うようにオプションで ss1 オプションで指定している。因子 A を先に MODEL ステートメントに入れることにより、切片、因子 A の順にモデルに取り込まれるので、因子 A と包含関係にある自由度が7ある実験順序 ex の自由度は、 $7-1-4 = 3$ となる。これは、因子 A1 内、A2 内、および A4 内の差に由来する平方和の自由度3に相当する。したがって、実験番号 ex が1次誤差となる。因子 A は1次誤差で検定するので、TEST ステートメントで仮説を $H=A$ として誤差を $E=ex$ とする F 検定を行うように指示している。

表示 5.13 GLM プロシジャを用いた解析

```
data placebo4_01 ;
  input A $ ex $16. @ ;
  do y = 1, 0 ;
    input n @@ ; output ;
  end ;
datalines ;
A1 1:Hayden_1999      8      4
A1 2:Hayden_2000     17      2
A2 3:Hayden_1999     25    494
A2 4:柏木_2000       13    142
A3 5:WV15825         12    260
A4 6:Welliver_2001   43    166
A4 7:Hayden_2004     33    225
```

```

;
proc glm data=placebo4_01 ;
  class A ex ;
  model y = A ex / ssl ;
  test H=A E=ex ;
  freq n ;
run ;

```

GLMプロシジャの出力を分散分析表として整理した結果を 表示 5.14 に示す。2 次誤差の平均平方 0.0798 は、実験番号ex内の分散であり、二項分布の誤差分散の近似値となっている。これに比べて、1 次誤差の平均平方 0.4277 は、F値が 5.36 と大きく、 $p = 0.0011$ と有意である。そのために因子Aは、1 次誤差でF検定を行うことになる。

表示 5.14 GLM プロシジャを用いた分散分析表

要因	自由度	平方和	平均平方	F 値	P 値
A	3	19.4584	6.4861	15.17	0.0256
1 次誤差	3	1.2831	0.4277	5.36	0.0011
2 次誤差	1434	114.4355	0.0798		
全体	1443	135.1770			

ロジスティック回帰の結果をGLMプロシジャの分散分析表のスタイルに合わせた結果を 表示 5.15 に示す。2 次誤差は、1 次誤差に相当するデビエンスの χ^2 値が自由度 3 の χ^2 分布に従うことから $p = 0.0151$ が得られるが、自由度で割って「平均平方」とした場合には、分子の自由度が 3、分母の自由度が ∞ の F 分布に従うことから $p = 0.0151$ が得られる。因子Aの χ^2 値も自由度 3 で割り、1 次誤差の「平均平方」から計算した、F値 11.62 が分子の自由度 3、分母の自由 3 の F 分布に従うことから、 $p = 0.0330$ と有意となる。

最小 2 乗法による場合は、平方和を自由度で割った平均平方の比である F 値を算出したのであるが、ロジスティック回帰の場合は F 分布の定義である 2 つの χ^2 値をそれぞれの自由度で割ったものの比が F 分布となることを用いて分散分析表スタイルにまとめることができる。

表示 5.15 ロジスティック回帰の分散分析表スタイル

要因	自由度	χ^2 値 (平方和)	χ^2 値/df (平均平方)	F 値	p 値
A	3	131.8561	43.9520	12.62	0.0330
デビエンス (1 次誤差) (2 次誤差)	3 ∞	10.4482	3.4827	3.48	0.0151
全体	6	142.3043			

