

計 数 値, 比 率 デ ー タ の 解 析
第 6 回・高橋セミナー

芳賀 敏郎 (元東京理科大学)
高橋 行雄 (日本ロシュ)

2001年1月27日

第6回－高橋セミナーでの追加

テーマ：統計モデルを用いた計数値（発現率）の差の信頼区間の推定

日科技連の実験計画法セミナーで用いている芳賀先生の「計数値」の講義テキストを高橋が全面改訂を行いました。これは、ロジット変換による回帰分析のアプローチを柱にしました。今回は、これに、群間の差の信頼区間を付け加えたテーマとします。

1. ロジット変換とは
2. 2 x 2表でのオッズ比その95%信頼区間
3. 2 x 2表でのロジスティック回帰分析を用いた群間の差の95%信頼区間
4. 共変量を含む2群間の比較とその差の95%信頼区間
5. 3群の場合
6. 順序カテゴリの場合，累積ロジットモデル /* 追加なし */
7. 2因子実験の場合 /* 追加なし */

改訂の記録

2000年10月12日，芳賀 [テキスト1999]，高橋 [JMP 実習資料1999] を元に改訂

2001年1月20日，第87回DEセミナー計数値テキストに加筆

高橋 行雄

ファイル名：Semi_anz_計数値データ.doc

2000年11月

計数値，比率データの解析

目次

1	まえがき	1
2	基礎知識	3
2.1	母集団とサンプル	3
2.2	2項分布	3
2.3	仮説検定	6
2.4	区間推定	9
3	1因子実験（説明変数が量的，反応変数が質的）	13
3.1	データ	13
3.2	破壊率による単純な回帰分析	13
3.3	ロジット変換	14
3.4	経験ロジットを反応変数とする回帰分析	16
3.5	p と z の分散	17
3.6	重みつき最小2乗法	18
3.7	重み付き最小2乗法の逐次計算	19
3.8	JMPによる解析	20
3.9	2次式の当てはめ	23
4	1因子実験（説明変数が質的，反応変数も質的）	25
4.1	実験とデータ	25
4.2	出現率の差の信頼区間	28
4.3	適合度の検定（参考）	32
4.4	カテゴリの反応変数（カテゴリーの順序に意味なし）	34
4.5	尤度比検定（カテゴリーに順序なし，参考）	36
5	2因子実験（説明変数は質と量，反応変数は質的）	41

5.1	データ	41
5.2	ロジット変換の考え方	41
5.3	JMPによる解析	44
5.4	2次曲線のあてはめ (演習)	47
6	累積データの解析 (反応が順序)	51
6.1	1 因子実験	51
6.2	多カテゴリの反応変数 (カテゴリの順序に意味あり)	55
6.3	2 因子実験	58

図表目次

表示 2.1	2項分布	5
表示 2.2	2分布（確率と累積確率）	8
表示 2.3	2項分布の区間推定（1）	10
表示 2.4	2項分布の正確な区間推定（2）	11
表示 3.1	強化ガラスの破壊データ	13
表示 3.2	破壊率に対する回帰直線	14
表示 3.3	シグモイド曲線	14
表示 3.4	p と z の関係	15
表示 3.5	ロジット変換	15
表示 3.6	経験ロジットに対する回帰分析	16
表示 3.7	逆ロジットによる p^* の推定値	16
表示 3.8	$V[p]$ と $V[z]$ の分散	17
表示 3.9	重み付き最小2乗法	19
表示 3.10	JMPでの解析データ	20
表示 3.11	JMPによるロジスティック回帰分析の結果	21
表示 3.12	ロジスティック回帰曲線の当てはめ	22
表示 3.13	p と \hat{p} の散布図	22
表示 3.14	2次式の当てはめ	24
表示 4.1	鋳物の不良率データ	25
表示 4.2	不良率とその正確な95%信頼区間	25
表示 4.3	解析用データ	26
表示 4.4	Fit Modelによるロジスティック回帰分析	27
表示 4.5	A_2 との差の推定と検定	27
表示 4.6	結果のまとめ	27
表示 4.7	自由度無限大での多重比較，両側， $\alpha=0.05$	28
表示 4.8	信頼区間の計算Excelシート	30
表示 4.9	比率の差の95%信頼区間	30
表示 4.10	JMPロジスティック回帰による95%信頼区間	31
表示 4.11	SASによる検証	31
表示 4.12	松田方式による信頼区間，***未検証***	32
表示 4.13	実測度数と期待度数	33
表示 4.14	多特性のデータ	34
表示 4.15	棒グラフ	34
表示 4.16	多特性のロジット	35
表示 4.17	多特性の解析用データ	35
表示 4.18	経験ロジットによる多特性のグラフ	35

表示 4.19	多特性のロジスティック回帰分析	36
表示 4.20	逸脱度の分解.....	38
表示 5.1	2 因子実験の不良率	41
表示 5.2	x と p , および x と z^* の関係	41
表示 5.3	効果の加法性と乗法性	42
表示 5.4	ロジットの構成.....	43
表示 5.5	直線の当てはめ.....	43
表示 5.6	解析用データ.....	44
表示 5.7	交互作用の検討.....	45
表示 5.8	主効果モデル.....	45
表示 5.9	主効果モデルの平行なシグモイド曲線.....	46
表示 5.10	A1 を基準にした差の信頼区間.....	47
表示 5.11	2 次曲線データ (r/n)	47
表示 5.12	2 次曲線の当てはめ	48
表示 5.13	交互作用のためのダミー変数	49
表示 5.14	曲率が同じロジスティック 2 次曲線	50
表示 6.1	順序がある破壊データ	51
表示 6.2	累積比率 P_1 , P_2 と x の関係	52
表示 6.3	1 因子累積データ	53
表示 6.4	2 つの切片をもつ累積ロジットの解	53
表示 6.5	推定値のグラフ	54
表示 6.6	2 群として解析したロジスティック回帰分析.....	55
表示 6.7	臨床試験データ	55
表示 6.8	Excel による帯グラフ	55
表示 6.9	解析用データ.....	56
表示 6.10	1 因子, 累積ロジットの解析結果.....	57
表示 6.11	逆ロジット, 累積有効率 P のとロジット Z の推定.....	57
表示 6.12	2 因子実験のデータ	58
表示 6.13	累積パーセント	59
表示 6.14	交互作用の検定結果.....	59
表示 6.15	主効果モデル.....	60
表示 6.16	効果の推定値.....	60
表示 6.17	推定値「上」のロジットの計算, Excel による確認	61
表示 6.18	JMP の解との比較.....	61

1 まえがき

通常の実験データの解析では、**反応変数**（特性値、目的変数、従属変数ともいう）が**計量値**（量的変数）である場合を取り扱う。計量値とは、強度、純度、売上数量などである。それに対して、良品か不良品かという変数は**質的変数**と呼ばれ、 n 個中の不良品の個数は**計数値**という。計数とは数を数えるという意味である。ただし、たとえば、壊れるまでの衝撃回数を数えて得られるような反応変数は、計量値として扱ってさしつかえない。製品の表面の傷の数（欠点数）も計数値である。

このテキストでは、何個中何個に不良品があるという場合（割合）だけを扱い、欠点数の場合は扱わない。反応変数は不良率であったり、成功率であったり、様々であるが、説明を簡単にするために、不良率という言葉を使う。第2章では、不良率を取り扱う場合に必要な2項分布に関する統計の基礎知識を、Excelのプログラムでの計算事例を示しつつ説明する。

不良率を分散分析的に扱うために、便宜的に、良品を0、不良品を1として計量値とみなした解析が行われてきた。不良率は多くの場合に2項分布に従うことが知られており、その場合に不良率をロジット変換した**ロジスティック回帰分析**を行うのが適切な方法である。ただし、その計算には、重み付き最小二乗法で解が収束するまで繰返すか、または、繰返し計算を前提とした最尤法を用いる必要があり、一般的ではなかった。

最近PCで稼動する統計ソフトが充実し、それらにはロジスティック回帰分析が組み込まれるようになってきた。第3章では、説明変数が量的な場合に、不良率のままの回帰分析、ロジットをそのまま用いた回帰分析を示しつつ、それらも問題点を示し、ロジスティック回帰分析の結果と対比できるようにした。第4章では、説明変数（独立変数、因子などともいう）が質的な1因子実験を扱い、因子にダミー変数を用いたロジスティック回帰分析について説明する。

[良品・不良品]、[非破壊・破壊] というように、1つの製品が2つのカテゴリーのいずれかに分類される場合が最も多いが、[良品・格下げ品・不良品]、[正常・ひび・破壊] のように3つ以上のカテゴリーのいずれかに分類される

場合もある。3つ以上のカテゴリーの場合は、さらに、カテゴリーの順序に意味のある場合とない場合がある。

上に挙げた2つの例は、いずれも順序に意味のある場合である。金属表面の発色の有無で、発色ありの場合、赤くなる場合と黒くなる場合があるとすると、[発色なし・赤く発色・黒く発色]の3つのカテゴリーの順序には意味がない。

[寸法が小さすぎる・ちょうど良い・大きすぎる]という3カテゴリーは、順序に意味があると見ることができる場合と、できない場合がある。具体的には第4章の4.4節で説明する。

順序に意味のある場合は、[良品・(格下げ品+不良品)]と[(良品+格下げ品)・不良品]の2通りの解析をして、その結果を総合化するという方法が考えられる。従来は、その代表的な手法として**田口の累積法**が使われてきた。この方法は、[良品:0・(格下げ品+不良品):1]、[(良品+格下げ品):0・不良品:1]のように計量値とみなした分散分析を行い2つの分散分析表の偏差平方和に重みを付けて併合する簡便的な方法である。第6章で、累積法に代わる方法として**累積ロジット**を用いたロジスティック回帰分析について説明する。

なお、このテキストでは、不良率をロジスティック回帰分析で取扱うことを主体にし、分割表に対する適合度のカイ2乗検定、累積カイ2乗検定、尤度比検定などは、要因配置実験や直交表実験への拡張が困難なので補足的な解説にとどめた。

2 基礎知識

2.1 母集団とサンプル

40 世帯について、あるテレビ番組の視聴を調査し、6 世帯が視聴していたという結果が得られたとする。調査対象の大きさを n 、視聴していた世帯数を r とする。視聴率 p は

$$p = r/n = 6/40 = 0.15$$

となる。同様の調査を 100 世帯に対して実施し、 $r=15$ という結果が得られたとき、 p は同様に 0.15 となる。規模が異なる 2 つの調査の違いは、どのように解析結果に反映したらよいであろうか。真の視聴率が含まれるであろう範囲を $0.15 \pm ?$ とするとき、 $?$ をどのように決めたら良いであろうか。

これまでの視聴率の平均は 25% であった。今回の調査から、「視聴率に変化があった」と言えるであろうか。また、「視聴率が 25% を割ったら放映を中止する」という基準があったとき、どちらの結果ならば中止を決定することができるであろうか。この章では、このような問題に対する解答を導くために、2 項分布、仮説検定、区間推定、などについて説明する。

$n=40$ 、または $n=100$ の調査対象となった世帯（サンプル）は、母集団（調査目的によって、全国の世帯であったり、首都圏の世帯であったりする）の全世帯の視聴率を推定するためにとられたごく一部に過ぎない。調査の目的はサンプルの視聴率ではなく、母集団の視聴率を正しく把握することである。サンプルの視聴率は母集団の視聴率の近傍にあると言える。近傍にあるというのでは曖昧過ぎるので、これを定量的に表すのが区間推定である。

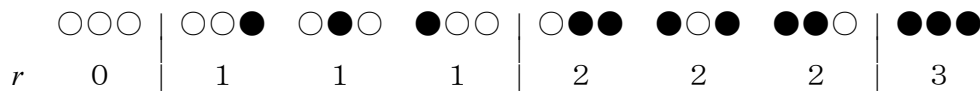
以下に説明する論理が成立するためには、サンプルが母集団からランダム（無作為）に、かつ、互いに独立に抽出されたものであるという前提が必要である。

2.2 2 項分布

白黒の碁石を良く混ぜ合わせて袋に入れる。1 個取り出して、白だったら勝ち、

黒だったら負けという賭を考える。このような賭を繰返したとき、何回負け続けたとき、袋の中身を見せろとクレームをつけるだろうか？ 白黒が半々であっても、3回負け続けることはしばしば起こると思われるので、文句を言えないが、9回負け続けることは極めて起こり難いので、袋の中に細工がしてあるのではないかと疑うのが常識的であろう。それでは、9回の内8回負けたとき、7回負けたときはどう対処したらよいであろうか？ この問題の解答を導くためには、まず2項分布について理解しなければならない。

3回続けて賭をした時の基石の白黒の出方は、次の8通りである。 r は●(黒)の回数である。



袋の中の黒の割合を π 、黒の基石の出る数を r とした場合に、 $r=0, 1, 2, 3$ である確率 p_r は次の式で表される。

$$\begin{aligned}
 p_0 &= (1 - \pi)^3 = 1 \cdot \pi^0 (1 - \pi)^{3-0} \\
 p_1 &= 3 \cdot \pi (1 - \pi)^2 = 3 \cdot \pi^1 (1 - \pi)^{3-1} \\
 p_2 &= 3 \cdot \pi^2 (1 - \pi)^1 = 3 \cdot \pi^2 (1 - \pi)^{3-2} \\
 p_3 &= \pi^3 = 1 \cdot \pi^3 (1 - \pi)^{3-3}
 \end{aligned}$$

係数の 1, 3 は、 n 個から r 個を取り出す場合の数、 $\frac{n!}{r!(n-r)!}$ であって ${}_n C_r$ で表される。上の式は

$$p_r = {}_n C_r \cdot \pi^r (1 - \pi)^{n-r} \tag{2.1}$$

という一般式で表すことができる。確率がこの式で表される分布を**2項分布**と呼ぶ。表示 2.1 は、Excelプログラム「計数值」のシート「2項分布」にあり、 n と π を入力すると、確率 p_r 、下側累積確率 $(\sum_{i=0}^r p_i)$ 、および、上側累積確率 $(\sum_{i=r}^n p_i)$ を計算し、横軸に $p = r/n$ を、縦軸に p_r をとったグラフが描かれる¹。

¹ 確率と累積確率を求める Excel 関数を次に示す。
 確率 = BINOMDIST(r, n, π, FALSE)
 上側累積確率 = 1 - BINOMDIST($n, \pi, r - 1, \text{TRUE}$)

$p = r / n$ は π を中心として分布し、その広がり（標準偏差）は、 π と n によって変化する。この関係は、

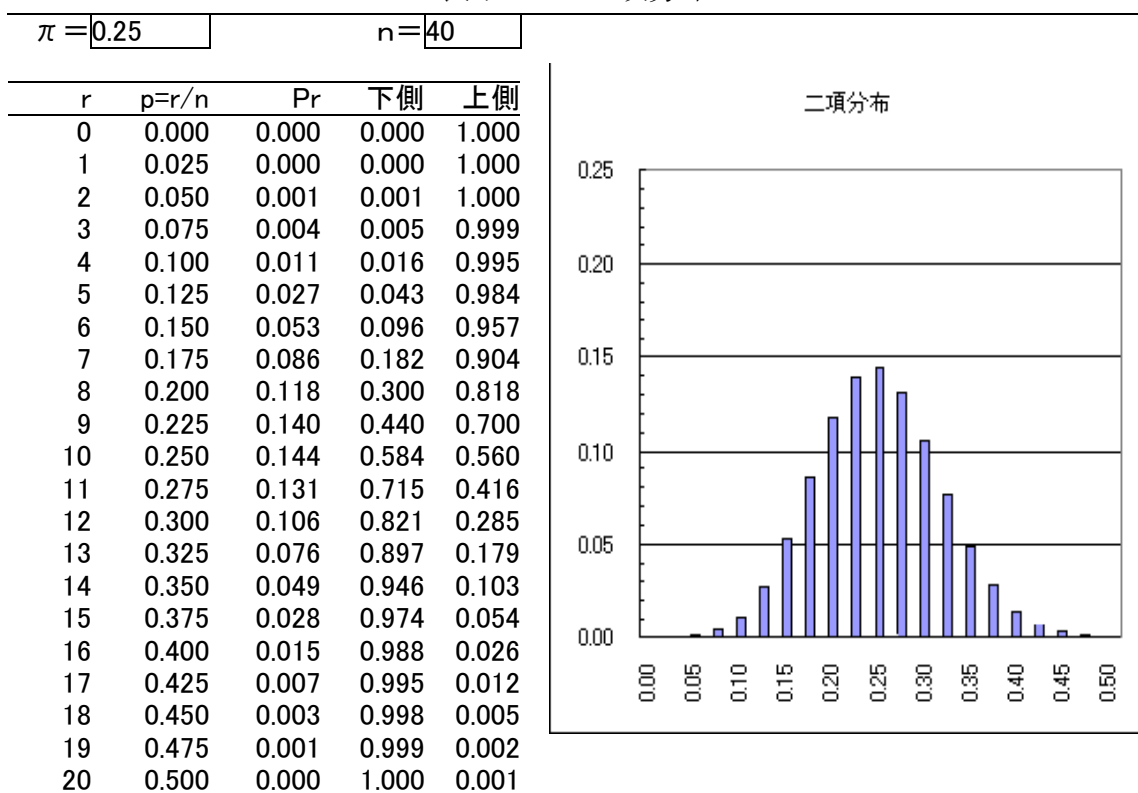
$$E[p] = \pi \tag{2.2}$$

$$D[p] = \sqrt{\frac{\pi(1-\pi)}{n}} \tag{2.3}$$

という式で表わされる。なお、 $n\pi, n(1-\pi)$ が共に大きいとき、2項分布は正規分布で近似できる。

演習 Excelのプログラムで、 n と π を変化させて、上の式(2.2)、(2.3)の意味を確かめよ。

表示 2.1 2項分布



注) 式(2.2)、(2.3)は、 n 回の試行中、反応がある場合を1、反応がない場合を0とした場合に、計量値での平均値と分散(分母が n)と等しくなる。 π の推定値を p としよう。反応の平均値は、反応の合計をデータ数で割ったものであるから、 $p = r/n$ となる。偏

下側累積確率 = BINOMDIST(r, n, π, TRUE)

差平方和は $S = r - r^2 / n$ であるので、 $r = np$ を代入すると、 $S = np - (np)^2 / n$ となり、整理すると $S = np(1 - p)$ となる。これから、分散は S を n で割り $V = p(1 - p)$ が得られる。平均値の標準偏差 (標準誤差) は、分散をデータ数 n で割り $\sqrt{\quad}$ をとったものであり、式 (2.3)に一致する。

2.3 仮説検定

公正な賭 (1回の勝ち負けの確率がそれぞれ 1/2) を 9回した時の負けの回数 r は、 $\pi = 0.5$, $n = 9$ の 2項分布に従うので、式 (2.1) を使って p_r を計算すると

r	0	1	2	3	4	5	6	7	8	9
$512 \times p$	1	9	36	84	126	126	84	36	9	1

となる²。これから全敗する確率は 1/512 (約 0.2%) である。1回しか勝てない確率も 9/512 (約 2%) と小さい。これから、全敗または1回しか勝てないという事は、公正な賭では起こり難いことが起こったと考えられる。すなわち、公正な賭ではないと判断できそうである。このような推論の手順は次のように考えると理解し易い。論理学で次の定理を学んだ。

命題 A が成立すれば命題 B が成立する

が正しいとき、その対偶

命題 B が成立しないときには命題 A が成立しない

も正しい。これを

$$\begin{array}{ll} A \implies B & A \not\Leftarrow B \text{ (逆)} \\ \bar{A} \not\implies \bar{B} \text{ (裏)} & \bar{A} \not\Leftarrow \bar{B} \text{ (対偶)} \end{array}$$

と表すことにする。ここで $\bar{\quad}$ は否定を表す。**B ならば A である**という逆は必

² 本文の確率は、式 (2.1) を使って計算される。この分布の近似値は、Excel でシミュレーションすることによって得ることができる。シート「Simulation」にそのプログラムが記録されている。 $\pi = 0.5$, $n = 9$ で 200 回のサンプリングして、 r を数え、 r の度数分布が求められている。

その手順は、1: 0~1 の一様乱数を =RAND() で求める。2: 乱数が π 以下の個数 r を =COUNTIF(一様乱数,'<0.5') で数える。3: r の度数分布を =FREQUENCY(データの範囲, r の値の区切り) で求める。 π, n , 試行回数 を変えて実行することができる。

ずしも成立しないことは良く知られている。ここで、

命題 A が成立すれば命題 B が成立する

を

命題 A が成立すれば命題 B がほぼ成立する

とすると、その対偶は

命題 B が成立しないときには命題 A がほぼ成立しない

となる。賭の問題に置き換えると

$$\begin{array}{llll} A & \text{賭が公正である} & \pi = 0.5 & \approx \Rightarrow & B: \text{負け数 } r \text{ が } 0 \sim 7 \\ \bar{A} & \text{賭が不公正である} & \pi > 0.5 & \approx \Leftarrow & \bar{B}: \text{負け数 } r \text{ が } 8 \sim 9 \end{array}$$

となる。また、ほぼ の程度は $1 - \frac{1+9}{512} \cong 0.980$ と定量的に表される。この考え方は、

最初の A が正しいことを確認するためのものではなく、A を否定する根拠を求めるためのものである。すなわち、A は否定されるべき仮説という意味で帰無仮説呼ばれ、一般に H_0 という記号で表される。それに対し、帰無仮説が否定されたときに採択される \bar{A} を対立仮説と呼び、 H_1 という記号で表される。帰無仮説が否定される観測値の領域 \bar{B} を棄却域と呼ぶ。

観測値が棄却域 \bar{B} に入ったから必ず \bar{A} が正しいというわけではない。このような論理で結論を出すと、僅かながら誤りを犯す危険がある。この確率は1からほぼ の程度 0.98 を引いた約 2% である。この誤りを**第1種の誤り**と呼び、その確率を α で表す。 α は 5% 以下にとられるのが普通である。このような論理で組立てられた統計手法を**仮説検定**という。

実測値が B の領域に入ったから A が正しいことを保証しているわけではないことは、**逆必ずしも真ならず**から理解できるであろう。すなわち、仮説検定は帰無仮説を証明することに使うことはできない。ある薬には副作用があることを実験で証明することはできても、絶対に副作用がないことを証明することができないことと論理的に同じである。

これまで、 r が 0 (0 敗, 全勝) または 1 (1 敗) のときは問題にしなかった

が、これもおかしいと思われる。このように考えられるときには、

$$\begin{array}{lcl} \text{A} & \pi=0.5 & \approx \Rightarrow \text{B} \quad 2 \leq r \leq 7 \\ \overline{\text{A}} & \pi \neq 0.5 & \approx \Leftarrow \overline{\text{B}} \quad r \leq 1 \text{ または } 8 \leq r \end{array}$$

とすれば良い。この場合は棄却域が r の大きい方と小さい方の両側にとられるので**両側検定**と呼ばれる。検定の目的によって**片側検定**と**両側検定**を使い分ける。

「視聴率が 25%を割ったら放映を中止する」という基準があるとする。 $n=40$ の調査結果が $p=r/n=6/40=0.15$ であったとき、中止するという結論は正しいであろうか。これを検定するためには、まず、

$$\text{帰無仮説 } H_0: \pi = 0.25$$

が正しいとき、 $n=40$ のときの r の分布から、棄却域を決める。この場合は、

$$\text{対立仮説 } H_1: \pi < 0.25$$

であるから、片側検定であり、下側確率が 5% 以下となる領域が棄却域である。表示 2.1 から、 $\pi=0.25$ 、 $n=40$ のとき、下側累積確率が 0.05 を越えない最大の r を探すと、 $r \leq 5$ の確率が 0.043 であることが分かる。これから、棄却域は $r \leq 5$ となる。この調査の結果 $r=6$ は棄却域に含まれないので、これから、真の視聴率が 25%を割ったと結論することはできない。このように、予め棄却域を決めて置かなくても、 $r \leq 6$ となる累積確率を計算し、それが、0.05 以下であれば、観測値が棄却域に入っていると判断することができる。帰無仮説の下で $r \leq$ 観測値 の確率を **p 値**と呼ぶ。

表示 2.2 のExcelプログラムの黒枠の中に、 π, n, r を入力すると、下側と上側の累積確率が簡単に求められる。この場合は、下側の累積確率が 0.0962 で、この値は 0.05 よりも大きい。これから、 $r=6$ は棄却域に含まれないことが分かる。

表示 2.2 2分布（確率と累積確率）

π	n	r	確率		累積確率	
			$p=r/n$	Pr	下側	上側
0.25	40	6	0.15	0.0530	0.0962	0.9567

上の行の π, n, r を修正すると確率と累積確率が求められる。

演習 $n=100$ の世帯について視聴率を調査し、 $r=15$ の結果を得た。これから、サンプルの視聴率を計算すると、 $p=r/n=0.15$ である。この調査結果から、「視聴率が 25%

を割ったら放映を中止する」という基準に照らして、放映を中止できるかを検討せよ。ヒント：表示 2.2 に、 $\pi=0.30$ 、 $n=100$ 、 $r=15$ を入力し、下側累積確率を求める。この演習の結果から、サンプルの視聴率が 0.15 であっても、調査の大きさ n によって、結論が異なることが分かる。

次に、「従来の平均視聴率は 25%であった。今回の調査から、**視聴率に変化があった**と言えるであろうか？」という質問に答えるにはどうしたら良いかを考える。この場合は、サンプルの視聴率が 0.25 より有意に小さくなった場合だけでなく、有意に大きくなった場合も、視聴率に変化があったという結論を出すことになる。したがって、このときは両側検定となる。この場合は、表示 2.2 の上側の累積確率または下側の累積確率のいずれか小さい方が、0.025 以下のとき、帰無仮説が棄却される。

2.4 区間推定

40 人についての調査から、視聴率は $6/40=0.15$ という値が得られたが、これは真の視聴率とは一致しない。しかし、大きく離れていることはないであろう。したがって、真の視聴率は **15%前後である**といえる。100 人についての調査で得られた視聴率 $15/100=0.15$ は、上と同じであり、真の視聴率は 15%前後であるといっても、こちらの方がより真値に近いと考えられる。

この違いを表わすために、**前後**というあいまいな表現ではなく、真の視聴率は「いくら以上である」、または、「いくらからいくら範囲内である」と定量的に表現する方法が、**区間推定**である。これに対して、サンプルの視聴率は 15%であるという推定を**点推定**という。

まず、真の視聴率が「いくらからいくら範囲内である」を求めることにする。区間推定は、仮説検定を逆に考え、

$$\text{帰無仮説 } H_0: \pi = p_0$$

の p_0 が棄却されない範囲を求める。そのための手順は、表示 2.3 に示すように、

- (a) 表示 2.2 の π を試行錯誤的に変化させて、上側累積確率がほぼ 0.025 となる π を求める。 $\pi = 0.06$ のとき上側確率が 0.0309 となった。これを p_1

とする。 p_0 が p_l 以上であれば、帰無仮説 $H_0: \pi = p_0$ は棄却されない。

- (b) 同様に、下側累積確率が 0.025 となる π を求める。 $\pi = 0.30$ のとき上側確率が 0.02376 となった。これを p_u とする。 p_0 が p_u 以下であれば、帰無仮説 $H_0: \pi = p_0$ は棄却されない。

これから、真の視聴率は 0.05~0.30 の範囲内であろうと結論する。真の視聴率がいくら以上であるかを知るためには、上側累積確率が 0.05 となる π を求めれば良い。

表示 2.3 2項分布の区間推定 (1)

π	n	r	$p=r/n$	Pr	下側	上側
0.060	40	6	0.15	0.0218	0.9909	0.0309
0.300	40	6	0.15	0.0151	0.0238	0.9914

両側：上側累積確率と下側累積確率が 0.025 になる π を試行錯誤で求める。

片側：上側累積確率または下側累積確率が 0.05 になる π を試行錯誤で求める。

この方法は繁雑で実用にならない。2項分布の累積確率は F 分布で計算できるので、次の式を使って容易に比率の区間推定値を求めることができる。

$$p_l = \frac{r}{(n-r+1) \cdot F(2(n-r+1), 2r; 0.025) + r} \tag{2.4}$$

$$p_u = 1 - \frac{n-r}{(r+1) \cdot F(2(r+1), 2(n-r); 0.025) + n-r}$$

この式は自由度の大きい F 表が必要であり、計算が面倒であるため、従来は用いられなかったが、Excelでは簡単に計算できる。表示 2.4 の危険率 α を 0.05 とし、 n, r を入力すると、真の割合 π が含まれる確率が 95%である範囲が求められる。

ここに得られた 0.0571 と 0.2984 を表示 2.3 の π に入力すると、累積確率がちょうど 0.025 になることを確かめることができる。この推定値は、母集団の比率 π 信頼率 95%の区間推定といい、得られた区間を信頼区間 または信頼限界と呼ぶ。

区間推定は仮説検定を裏から眺めたものである。 π の上限値が 0.25 以上であれば、真の視聴率が 25%以上である可能性があるため、 $\pi = 0.25$ の帰無仮説を否

定できない。逆に、 π の上限値が 0.25 以下であれば、真の視聴率が 25%以上である可能性が小さいので、 $\pi=0.25$ の帰無仮説は棄却される。

信頼区間を、 $p_l < \pi < p_u$ のように、範囲で表わすとき、両側の信頼区間という。

表示 2.4 2 項分布の正確な区間推定 (2)

$\alpha =$	0.05	
$n =$	40	
$r =$	6	
$p = r/n =$	0.15	
下側		
$df1 =$	$2(n-r+1) =$	70
$df2 =$	$2r =$	12
F	$(df1, df2, \alpha/2) =$	2.831
F	$(df1, df2, \alpha) =$	2.372
	$p_l(0.025) =$	0.0571
	$p_l(0.05) =$	0.0674
上側		
$df1 =$	$2(r+1) =$	14
$df2 =$	$2(n-r) =$	68
F	$(df1, df2, \alpha/2) =$	2.07
F	$(df1, df2, \alpha) =$	1.84
	$p_u(0.025) =$	0.2984
	$p_u(0.05) =$	0.2747

演習 $n = 100$, $r = 15$ のとき、両側の信頼区間を求めよ。 $n = 40$, $r = 6$ の信頼区間と比較して、区間の幅が狭くなることを確かめよ。両方の場合について、真の視聴率がいくら以上であるか（片側の信頼区間）を求めよ。その結果を、仮説検定の結果と比較せよ。

(注) 参考のために、従来使われていた近似式を示す。

不良率を $p = r/n$ とすると、 p の期待値と標準偏差は

$$E[p] = \pi \tag{2.5}$$

$$D[p] = \sqrt{\frac{\pi(1-\pi)}{n}} \tag{2.6}$$

であった。 r および $n-r$ がいずれも 5 以上のとき、正規分布で近似できる。正規近似を用いると、 p が含まれる確率が約 95% の区間は

$$\pi - 1.96\sqrt{\frac{\pi(1-\pi)}{n}} < p < \pi + 1.96\sqrt{\frac{\pi(1-\pi)}{n}} \tag{2.7}$$

となる．式 (2.7) の不等式を移項して， π を挟む不等式

$$p - 1.96\sqrt{\frac{\pi(1-\pi)}{n}} < \pi < p + 1.96\sqrt{\frac{\pi(1-\pi)}{n}} \quad (2.8)$$

に変える． p を移項して，両辺を2乗すると π についての2次方程式

$$\left(1 + \frac{1.96^2}{n}\right)\pi^2 - \left(2p + \frac{1.96^2}{n}\right)\pi + p^2 = 0$$

が得られる．それを解くと， π の信頼区間を求める式

$$\begin{aligned} \pi &\sim \frac{(2p + 1.96^2/n) \pm \sqrt{(2p - 1.96^2/n)^2 - 4(1 + 1.96^2/n)p^2}}{2(1 + 1.96^2/n)} \\ &\sim (0.0705 \sim 0.2907) \end{aligned} \quad (2.9)$$

が導かれる．

もっと簡単には，式 (2.8) の両側の π を p で代用して，次の式が用いられる．

$$\begin{aligned} \pi &\sim p \pm 1.96\sqrt{\frac{p(1-p)}{n}} \\ &\sim (0.0393 \sim 0.2607) \end{aligned} \quad (2.10)$$

式 (2.10) で信頼区間を計算すると， p の両側に同じ距離の区間が得られ，場合によっては， $0 \sim 1$ の範囲外に出る場合がある．また， $r=0$ または $r=n$ のとき区間推定の幅が 0 になってしまうという欠点がある． (注終)

3 1 因子実験 (説明変数が量的, 反応変数が質的)

3.1 データ

表示 3.1 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実験の結果である (仮想例). 計算を簡単にするために衝撃の強さから 30 を引いて 5 で割った値を x とする. 第 i 水準で n_i 個を検査し, 破壊した個数を r_i , 破壊率を p_i とする. 破壊しなかった個数は $n_i - r_i$ 個である.

表示 3.1 強化ガラスの破壊データ

衝撃 kg	変換値 x	破壊数 r	非破壊数 $n-r$	合計 n	比率 p
35	1	0	9	9	0.000
40	2	1	9	10	0.100
45	3	0	8	8	0.000
50	4	3	16	19	0.158
55	5	3	8	11	0.273
60	6	6	8	14	0.429
65	7	14	5	19	0.737

このようなデータは多くの分野で観測される. たとえば, ラットに薬の投与量を変えて実験し, 効果または毒性が現れたラットの数を観測する. このとき, 投与量は 5, 10, 20, ... のように等比級数的に変化させる場合が多い.

3.2 破壊率による単純な回帰分析

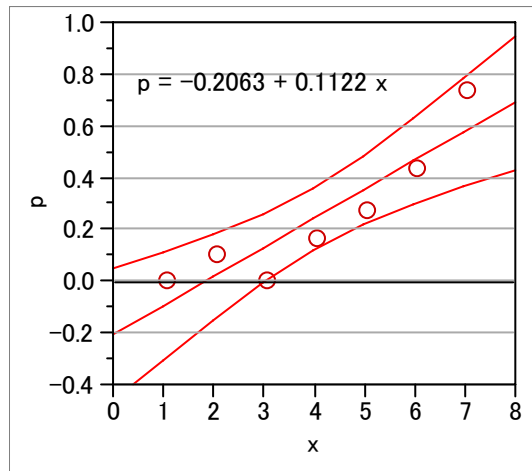
r_i と n_i から, 比率 p_i を計算し, x_i と p_i の散布図を描き, 回帰直線を引くと表示 3.2 が得られる. 回帰直線は

$$\hat{p} = -0.2063 + 0.1122x \quad (3.1)$$

となる. 回帰直線は実測点の変化を表していない. $x=1$ の場合の推定値は -0.1 , 95%信頼区間の下限は -0.3 とグラフから読み取れる. ところが破壊率の物理的な範囲は, 0 から 1 であるのに, 破壊率の推定値がマイナスになっている. さらに, 95%信頼区間の下限は $x=3$ からマイナスになっており, 不合理である.

したがって, 推定値が 0~1 の範囲に入るような変数変換が必要である. そこで, 次節のような曲線を考える.

表示 3.2 破壊率に対する回帰直線



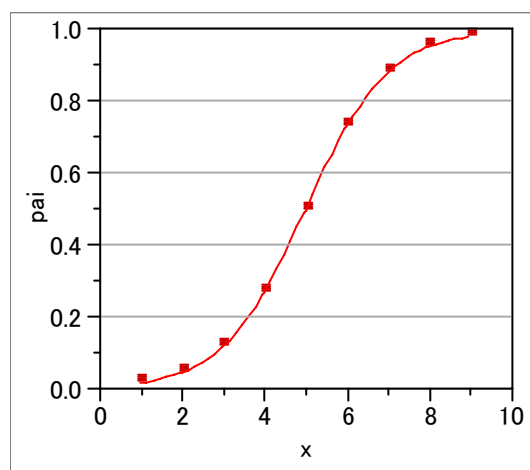
3.3 ロジット変換

表示 3.3 のように x と p の間にS字型の曲線 (シグモイド曲線) 関係が想定されるとき p を適当に変換することにより, 直線関係が成立するようにしたい. そのための変換が**ロジット変換**であり,

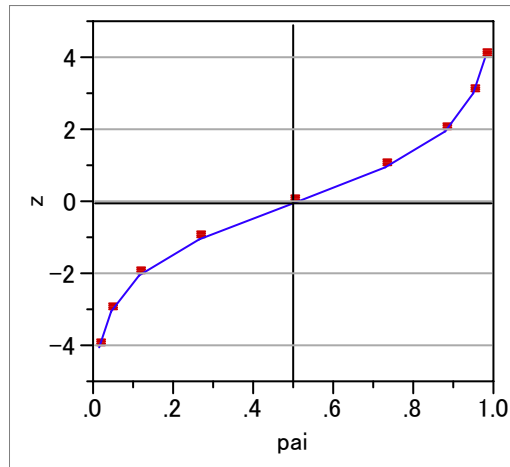
$$z = \ln\left(\frac{p}{1-p}\right) \tag{3.2}$$

で表される. ここに, \ln は自然対数 (e を底とする対数) である. $p = 0.5$ のとき z は 0 で, p が 0 または 1 に近づくとき, z は急激に変化する.

表示 3.3 シグモイド曲線



表示 3.4 p と z の関係



r が 0 または n のとき, すなわち, 一つも破壊しない, または, 全部が破壊しないとき, z の値は $-\infty$ または ∞ になってしまう. それを避けるために n を 1 増やし, r を 0.5 増やして破壊率 p^* を求めロジットを計算し直したものを経験ロジットという.

$$p^* = \frac{r + 0.5}{n + 1} \tag{3.3}$$

$$z^* = \ln \left(\frac{p^*}{1 - p^*} \right) \tag{3.4}$$

この変換は, 実験数が 9 回で, 破壊個数が 0 回の場合, $p^* = 0.05$ となり, いわば, 破壊が 0.5 回あったとみなしている. 表示 3.1 のデータから経験ロジット z^* を計算した結果を表示 3.5 に示す.

表示 3.5 ロジット変換

衝撃 kg	変換値 x	破壊数 r	合計 n	比率 p	比率 p^*	経験ロジット z^*
35	1	0	9	0.000	0.050	-2.944
40	2	1	10	0.100	0.136	-1.846
45	3	0	8	0.000	0.056	-2.833
50	4	3	19	0.158	0.175	-1.551
55	5	3	11	0.273	0.292	-0.887
60	6	6	14	0.429	0.433	-0.268
65	7	14	19	0.737	0.725	0.969

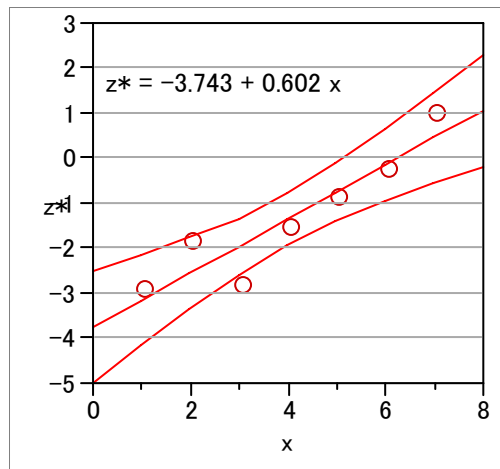
3.4 経験ロジットを反応変数とする回帰分析

x と z^* の散布図を作成し、回帰直線を求めると、観測点がほぼ直線的に並んでいる。 z^* を反応変数として回帰式を求めると

$$\hat{z}^* = -3.743 + 0.602x$$

となり、散布図より、ほぼ直線が当てはまっているとみなせる。

表示 3.6 経験ロジットに対する回帰分析

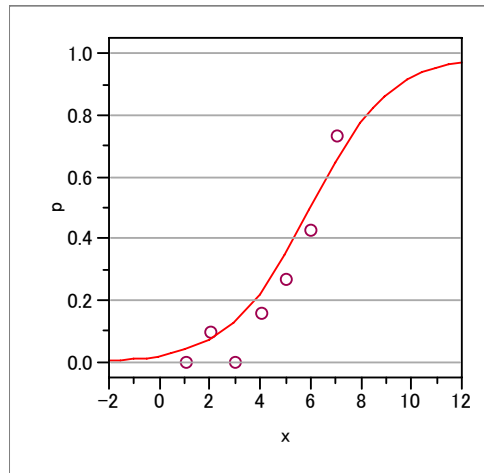


ロジット z から p の逆推定は、式 (3.2) を p について解くと

$$p = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}} \tag{3.5}$$

となる。表示 3.7 は、回帰式 $\hat{z}^* = -3.743 + 0.602x$ で求めた \hat{z}^* を式 (3.4) に代入した得られた結果をプロットしたものである。推定値がS字曲線となっている

表示 3.7 逆ロジットによる p^* の推定値



$$\hat{p}^* = \frac{1}{1+e^{-(-3.743+0.602x)}}$$

ることが確認できる. $x = 1$ の場合に表示 3.6 の散布図上の推定曲線よりも○印が上にあっただのが, 表示 3.7 の逆ロジットの散布図上では下になっている. これは, 経験ロジットによる回帰直線は, その変換の性質により傾きが緩やかになるためである.

3.5 p と z の分散

x_i についての実測値 n_i , r_i から計算した p_i は, 二項分布 $B(\pi_i, n_i)$ に従う. すなわち p_i の分散は,

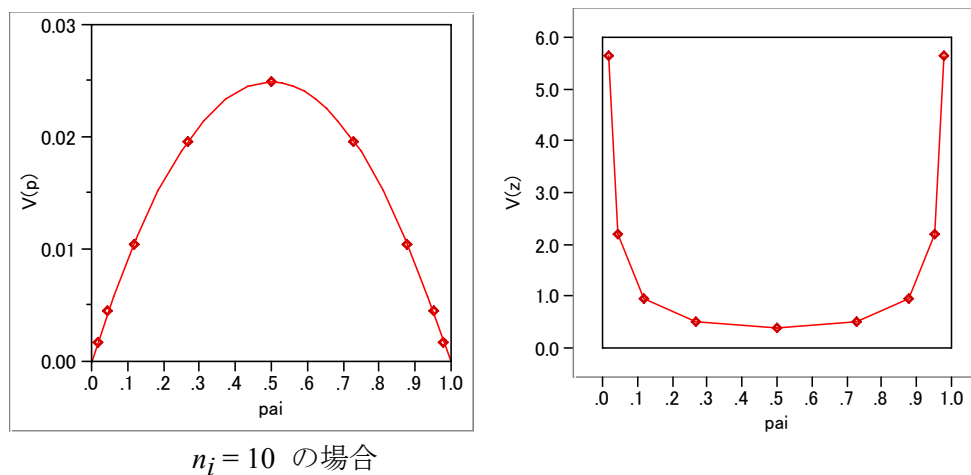
$$V[p_i] = \frac{\pi_i(1-\pi_i)}{n_i} \tag{3.6}$$

となる. ここに π_i は, x_i に対する真の破壊確率である. 表示 3.8 に示すように π_i が 0, 1 に近いところでは, 分散が小さくなる. p_i のロジット z_i の分散は近似的に

$$V(z_i) \approx \frac{1}{\pi_i(1-\pi_i)n_i} \tag{3.7}$$

で表され³ p_i と同様 π_i , n_i の関数である. z_i の分散は, p の分散とは逆に示すように π_i が 0, 1 に近くなると分散が大きくなる.

表示 3.8 $V[p]$ と $V[z]$ の分散



3

$$V[z] \approx V[p] \left(\frac{dz}{dp} \right)^2 = \frac{\pi(1-\pi)}{n} \left(\frac{1}{\pi} + \frac{1}{1-\pi} \right)^2 = \frac{1}{\pi(1-\pi)n}$$

3.6 重みつき最小2乗法

ロジット変換によって直線関係に持って行くことはできたが、回帰分析での等分散性の仮定が満たされていない表示 3.8 から分るように、真の比率が 5%~95%の範囲内であれば、分散の違いは5倍以内である。この範囲内ならば、経験的に等分散みなして分散分析を行うことから、重みなしの回帰分析を用いても差し支えない。

しかし、真の比率が 0 および 1 に近い場合、あるいは、 n_i が i によって大きく違う場合には、分散の違いを考えた重み付き最小2乗法 (Weighted Least Squares Method) を適用しなければならない。重み w として式 (3.7) の逆数をとる。

$$w_i = \pi_i(1 - \pi_i)n_i \quad (3.8)$$

この式は、重みが n_i に比例し p_i が 0.5 付近の値を重視し、0 や 1 に近い値は軽く見ている。重みを計算するためには π_i の値が必要である。これは推定の対象であって、最初から分っているものではない。そこで、その近似値として p_i を使うことが考えられるが、 p_i が 0 または 1 のとき重みが 0 になってしまうので、 $p^* = (r + 0.5)/(n + 1)$ を使って重みを計算する。

$$w_i = p_i^*(1 - p_i^*)n_i \quad (3.9)$$

重み付き最小2乗法は、Excel には含まれていないので、JMP などの統計ソフトウェアを用いて重み付き最小2乗法を試みる。重み付き最小2乗法と通常回帰分析との違いは、回帰による推定値と観測値の残差を最小にする際に、重み w_i を掛けた上で最小化することである。

$$\sum_i w_i (y_i - \hat{y}_i)^2 \Rightarrow \min \quad (3.10)$$

したがって、 p^* が 0 または 1 に近い場合は、重み w_i は小さくなるので、推定しようとする回帰直線から観測点が離れていても相対的に軽く扱われる。

回帰係数の行列演算による正規方程式の解法で、回帰係数のベクトル \mathbf{b} は、

$$\mathbf{b} = (\mathbf{X}^T \mathbf{Y})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.11)$$

と示されているが、重み付き最小2乗法は、 \mathbf{W} を重み w_i を対角行列としたときに

$$\mathbf{b}^* = (\mathbf{X}^T \mathbf{W} \mathbf{Y})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{y} \quad (3.12)$$

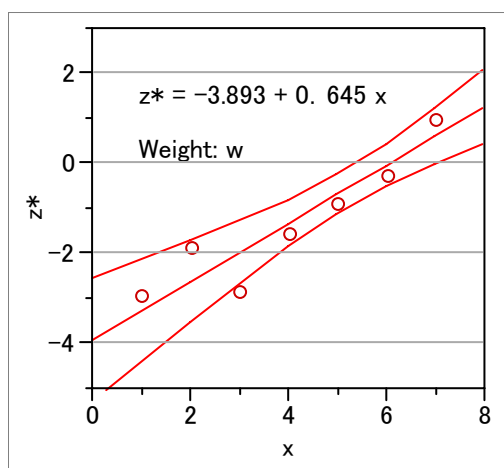
で計算される.

表示 3.9 にJMPによる重み付き最小 2 乗法の結果を示す. これまでと同様な手順で, 回帰直線, 95%信頼区間を描いてみる. 得られたロジスティック回帰式は

$$\hat{z}^* = -3.893 + 0.645x \quad (3.13)$$

となり, 経験ロジットが小さい場合の 95%信頼区間が, 表示 3.6 の標準回帰分析に比べてすそ広がりになっている.

表示 3.9 重み付き最小 2 乗法



3.7 重み付き最小 2 乗法の逐次計算

π_i の値が分からないので, 式 (3.6) の代わりに 式 (3.7) で重みを計算した. π_i により近い値を使って重み w_i を求めるため, 上に求めた回帰式 (3.13) でロジットの推定値 \hat{z}^* を求め, さらに, ロジット z から p の逆推定する式 (3.5)

$$\hat{p}_i = \frac{1}{1 + \exp(-\hat{z}_i)} \quad (3.14)$$

を使って比率の推定値 \hat{p}_i を求める. これを, 式 (3.8) の π_i の代りに用いて重みを計算する. このように, 重みの近似値の精度を向上させつつ, 逐次計算を進める. 重みの近似値を求めるために使った \hat{p}_i と新しく求めた p_i の推定値 \hat{p}_i の差が十分に小さくなったら, 計算を終了する.

なお、最近の統計ソフトは、重み付き最小 2 乗法の繰返しではなく、尤度を最大にする最尤法 (Maximum Likelihood) による逐次計算を行っているが、最終的な結果は一致する。

3.8 JMP による解析

解析用データの形式

JMPによるロジスティック回帰分析を行うためには、表示 3.1 の形式ではなく表示 3.10 のように

- x : 説明変数
- y : 反応 0: 破壊あり
1: 破壊なし
- r : 反応の個数

反応の有無の個数を行方向に入力する。比率 p は必要としない。

表示 3.10 JMP での解析データ

x	y	r
1	0	0
1	1	9
2	0	1
2	1	9
3	0	0
3	1	8
4	0	3
4	1	16
5	0	3
5	1	8
6	0	6
6	1	8
7	0	14
7	1	5

回帰式の推定

JMPによるロジスティック回帰分析の結果を 表示 3.1 に示す。Parameter Estimates のから、回帰式は、

$$\hat{z} = -5.194 + 0.859x$$

となる．経験ロジットで計算した回帰分析の結果は， $\hat{z}^* = -3.743 + 0.602x$ であり，ロジスティック回帰分析の傾斜 b_1 より小さくなる傾向にある．これは， p^* は p より中心 ($p=0.5$) に近づくように変換されているからである．

表示 3.11 JMP によるロジスティック回帰分析の結果

Iteration History				
Iter	LogLikelihood	Step	Delta-Criterion	Obj-Criterion
1	-62.38324625	Initial	2984784204	.
2	-42.44742287	Newton	0.50907309	0.4695486
3	-40.13756376	Newton	0.14213642	0.05753423
4	-39.91512362	Newton	0.01728068	0.00557143
5	-39.91188154	Newton	0.00028025	0.00008121
6	-39.91188068	Newton	0.00000008	0.00000002

Whole Model Test				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	15.065907	1	30.13181	<.0001
Full	39.911881			
Reduced	54.977787			

Parameter Estimates				
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-5.1941151	1.1397411	20.77	<.0001
x	0.8592572	0.2002501	18.41	<.0001

寄与率

通常回帰分析で重要な役割を果たす平方和の代わりにロジスティック回帰分析では Iteration History に示されているように尤度 (likelihood) の対数が用いられる．通常回帰分析の総平方和に相当する対数尤度は，全体の平均破壊比率 $p = 30/90 = 0.300$ から

$$L = p^r (1-p)^{n-r}$$

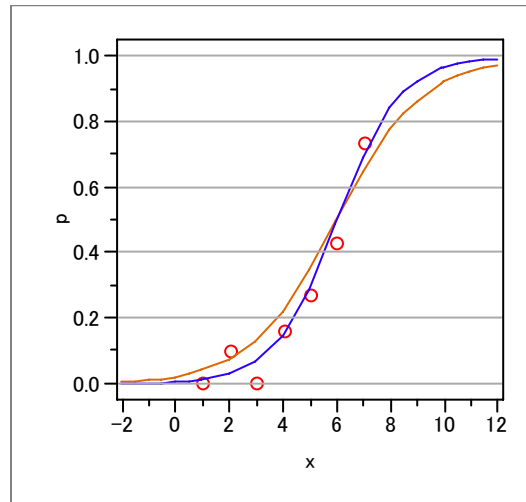
$$-\ln L = -2\{30 \ln p + (90-30) \ln(1-p)\} = 54.978$$

として計算される．これは，表示 3.11 Whole-Model Test の Reduced に示されている $-\text{LogLikelihood}$ に一致する．回帰平方和に相当する対数尤度は，表示 3.11 の Whole-Model Test の Difference に表示されている尤度比検定量の $\text{ChiSquare} = 30.132$ である．寄与率は $-\ln L$ を 2 倍にし，上の 2 つの値の比

$$R^2 = 30.132 / (2 \times 54.978) = 0.274$$

として求められる. この値は 表示 3.12 から予想される値よりもかなり小さい. それは, 総平方和は (p_1, p_2, \dots, p_7) の平方和ではなく, 0 (非破壊) が 63 個, 1 (破壊) が 30 個 というデータについての平方和であるからである.

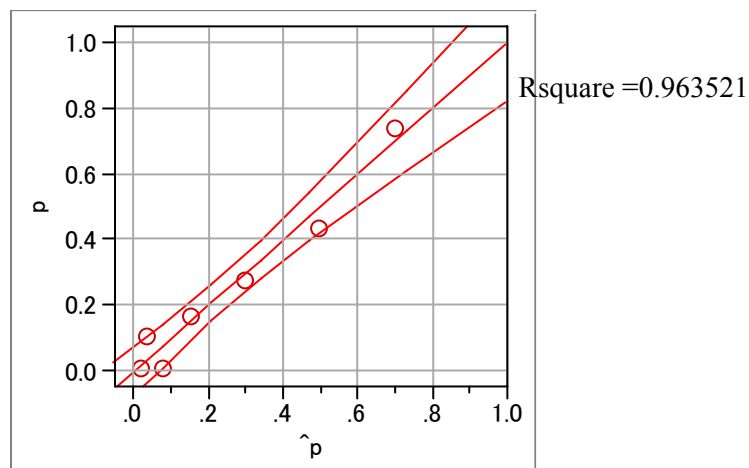
表示 3.12 ロジスティック回帰曲線の当てはめ



--- : p^* の回帰, — : ロジスティック回帰

通常の見方と同様に見たい場合は 表示 3.13 に示すように p と \hat{p} の散布図を求め, その相関係数を見るのがよいであろう. \hat{p} と p は直線の上に乗っており, その相関係数は 0.982 である. これから, ロジスティック回帰モデルが良く当てはまっていると考えられる.

表示 3.13 p と \hat{p} の散布図



回帰係数の区間推定と仮説検定

ロジスティック回帰係数 $b_1 = 0.859$ の標準誤差は 0.200 であるから, β_1 の 95% 信頼区間は,

$$\beta_1 \sim 0.859 \pm 1.96 \cdot 0.200 \sim (0.467, 1.251)$$

となる. 帰無仮説 $H_0: \beta_1 = 0$ の検定には,

$$\chi^2 = (0.859/0.200)^2 = 4.295^2 = 18.4 \quad (3.15)$$

が用いられる. この値は, 表示 3.11 のParameter Estimateのxの行, ChiSquareの列に示されている. 帰無仮説が正しいとき, この値は自由度 1 の χ^2 分布に従う. 18.4 は, $\chi^2(1; 0.01) = 6.63$ より大きいので, 高度に有意である. 表示 3.11 には, 有意性を評価するための値がProb>ChiSqの列に表示される. この値は一般に p 値と呼ばれるものであって, α より小さいとき, 危険率 α で有意と判定する. この例では 0.01 より小さいので高度に有意である.

3.9 2次式の当てはめ

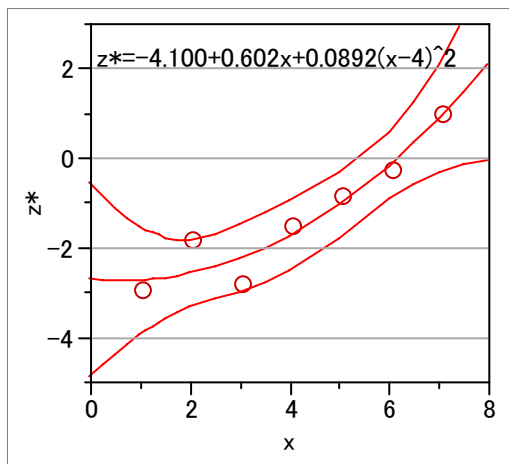
通常 of 回帰分析と同様に,

$$z = f(x)$$

の関数 f は x の線形式である必要はない. p が有効率や不良率であって, それが最大または最小になる x を知りたいというような場合は, $f(x)$ として x の 2 次式を用いることができる. 最尤法によるロジスティック回帰分析も可能であるが, 見とおしを良くするために経験ロジット z^* について普通の 2 次式を当てはめる. さらに, その 95% 信頼区間を作ったのが表示 3.14 である.

2 次の項の p 値は, 0.1762 であり, 明らかに 2 次曲線の当てはまりが良いとは言えそうにない. ここでは, 形式的に 2 次式を当てはめたのであるが, 技術的に x が大きくなるに従って不良率が逆に小さくなるのが推測できる場合には場合には意味がある. この例のように, 不良率が単調に大きくなるような場合には, 誤差的な変動に惑わされてはいけな

表示 3.14 2次式の当てはめ



参考文献

- 1) 丹後俊郎, 山岡和枝, 高木晴良 (1996), ロジスティック回帰分析, 朝倉書店.
統計ソフト SAS の使い方が含まれている.
交互作用についても述べられている.
疫学研究, 臨床研究の事例が豊富である.
- 2) 田中豊, 森川敏彦, 栗原考次訳 (1993), 統計モデル入門, 共立出版.
一般化線形モデルの観点からロジスティック回帰分析を取り扱っている.
最小2乗法, 重み付き最小2乗法, 最尤法について丁寧に解説されている.
第8章が, 「2値変数とロジスティック回帰」となっている.
- 3) D. Collet (1991), Modeling Binary Data, Chapman & Hall.

4 1 因子実験 (説明変数が質的, 反応変数も質的)

4.1 実験とデータ

鋳物に使う砂を 3 種類とりあげ, 鋳物に発生する巣の発生率を比較する実験を実施した. 3 種類の砂を A_1, A_2, A_3 とし, それぞれの砂を使って数十個の鋳物を作成し, 巣の入った鋳物の数を数えた. 現実には同数個のサンプルを作成するのが普通であるが, ここでは, 個数が等しくなくても解析が可能であることを示すために, 不揃いのデータを用いた.

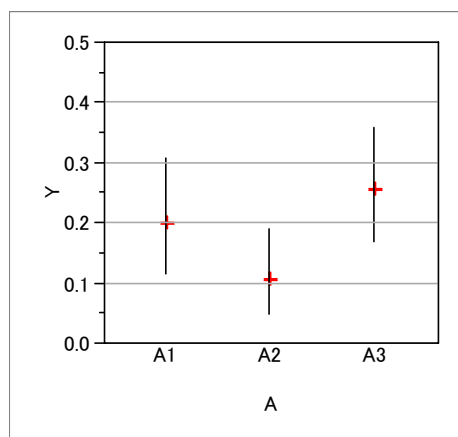
その結果 表示 4.1 のデータが得られた. この結果から, 砂の種類毎に巣の発生率を推定し, 砂の種類によって巣の発生率に差があるかどうかを確かめたい.

表示 4.1 鋳物の不良率データ

	良品	不良品	合計	不良率	95%信頼区間
A_1	60	15 (r_1)	75 (n_1)	0.200 (p_1)	0.116~0.308
A_2	76	9 (r_2)	85 (n_2)	0.106 (p_2)	0.050~0.192
A_3	67	23 (r_3)	90 (n_3)	0.256 (p_3)	0.169~0.358
合計	203	47 (r)	250 (n)	0.188 (p)	0.142~0.242

A_2 の砂の不良率は 0.106 と他の砂より小さく, また, F 分布を用いた正確な計算公式による 95%信頼区間の上限は 0.192 と, 他の水準の不良率より小さい.

表示 4.2 不良率とその正確な 95%信頼区間



砂全体として有意な差があるか確かめ, 次に A_2 と他の水準間に有意な差があるか検討する. 3 種類の砂によって巣の発生率に差がないとすれば, どの砂でも,

全体の不良率 0.188 の割合で巣が発生するはずである.

表示 4.1 は, 反応が 2 値の 1 元配置分散分析として考えられる. 反応が連続量の場合には, 質的な説明因子をダミー変数化した上で, Excel の LINEST 関数を用いた回帰分析によって分散分析を導入し, 水準間の推定も行ってきた. 反応が不良の [有, 無] の 2 値の場合でも, 説明変数が質的な場合には, 同様にダミー変数を生成しロジスティック回帰を適用できる. JMP で解析する場合は, 表示 4.3 に示したように,

- a_1 : A_1 のとき 1, A_1 以外は 0
- a_2 : A_2 のとき 1, A_2 以外は 0
- a_3 : A_3 のとき 1, A_3 以外は 0
- y : 不良有 0, 不良無 1
- r : 反応の数

ダミー変数を生成する. 反応変数 y の 不良・有 を 0 としたのは, JMP のロジットの計算が 0 の率を基準に計算しているためである. どちらを基準にするは統計ソフトによりまちまちなので使用する前に確認する必要がある.

表示 4.3 解析用データ

A	a_1	a_2	a_3	y	r
A_1	1	0	0	1	60
A_1	1	0	0	0	15
A_2	0	1	0	1	76
A_2	0	1	0	0	9
A_3	0	0	1	1	67
A_3	0	0	1	0	23

質的因子をダミー変数とした回帰分析と同じ考え方が, ロジスティック回帰分析についても適用できる. 反応変数を y , a_2 および a_3 を説明変数とし, 反応の数 r を繰返し度数 (Freq) とすると表示 4.4 の結果が得られる.

Whole-Model Test の Difference 結果から因子全体で有意な差 ($p=0.0324$) があることが分かる. Parameter Estimates に説明変数の推定値があり, a_2 の推定値, これは A_2 と A_1 の差であるが, 有意な差ではない. A_3 についても A_1 との間には有意な差はない.

表示 4.4 Fit Model によるロジスティック回帰分析

Whole Model Test				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	3.43060	2	6.861194	0.0324
Full	117.39688			
Reduced	120.82748			

Parameter Estimates				
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.3862944	0.2886751	23.06	<.0001
a2	-0.7472096	0.4556335	2.69	0.1010
a3	0.3170959	0.3764797	0.71	0.3996

表示 4.2 から A_2 と A_3 の間には有意な差がありそうなので、ダミー変数に a_1 と a_3 を選択し直して（これは A_2 との差の検定となる）再度実行し 表示 4.5 の結果を得る。 a_3 の回帰係数は、もちろんロジットの差であるが、1.064 と有意 ($p=0.0128$) となっている。さて、この p 値から A_2 と A_3 に有意な差 ($\alpha=0.05$) があると信じて良いのだろうか。

表示 4.5 A_2 との差の推定と検定

Parameter Estimates				
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-2.1335039	0.3525174	36.63	<.0001
a1	0.7472095	0.4556335	2.69	0.1010
a3	1.0643055	0.4274016	6.20	0.0128

表示 4.6 結果のまとめ

	不良/合計	不良率	ロジット	A_1 との差		A_2 との差	
				ロジット	p 値	ロジット	p 値
A_1	15 / 75	0.200	-1.386	-	-	0.747	0.1010
A_2	9 / 85	0.106	-2.134	-0.7472	0.1010	-	-
A_3	23 / 90	0.256	-1.069	0.3171	0.3996	1.064	0.0123
合計	47 / 250	0.188	-1.463	-	-	-	-

これは、要因実験のテキスト 4.5 節で述べられているように、多水準間の比較を行う場合には、ダネットの多重比較、またはチューキーの多重比較などを考慮する必要がある、しかしながら、これらの多重比較は、反応が量的な 1 元配置実験を前提としており、ロジスティック回帰分析では使うことができない。

要因実験のテキスト p. 67 表示 4.11 に多重比較のための数表がある．この数表の自由度が無限大の t 値を p 値に換算して，多重性を考慮した比較を行ってみよう．3水準間の全ての比較を行いたいのであるから Tukey の t 値，自由度無限大であるから標準正規分布の両側確率を Excel の NORMSDIST 関数を使って計算した p 値を 表示 4.7 に示す， p 値が 0.0191 なので，危険率 5% で， A_2 と A_3 に有意な差あるといえる．

表示 4.7 自由度無限大での多重比較，両側， $\alpha=0.05$

水準数	Dunnet		Tukey	
	t 値	p 値	t 値	p 値
2	1.960	0.0500	1.960	0.0500
3	2.212	0.0270	2.343	0.0191
4	2.349	0.0188	2.569	0.0102
5	2.442	0.0146	2.728	0.0064

4.2 出現率の差の信頼区間

1 群での出現率（不良率）の区間推定については，2.4節で幾つかの推定方式について示した．表示 4.1の各水準の 95%信頼区間は，式(4.1)を用いて計算されたものである． A_1 を現行水準としたときに， A_2 水準との差についての 95%信頼区間を求めたい．

それぞれの不良率を p_1 と p_2 としたときに， $p_1 - p_2$ の分散は，

$$V(p_1 - p_2) = V(p_1) + V(p_2) = \frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2}$$

となり，標準偏差（標準誤差）は，

$$D(p_1 - p_2) = \sqrt{\frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2}}$$

であり， π_1 と π_2 を p_1 と p_2 で代用して，95%信頼区間は，

$$(p_1 - p_2) \pm 1.96 \sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}$$

となる．この式は，2項分布を正規分布で近似したものである．SAS などの統計ソフトでもこの式による信頼区間が計算されている．

さて，ロジスティック回帰が，様々な 2 値データの実験データの解析法として応用できることを示してきた．ロジスティック回帰は，出現率をロジット変

換 $\ln(p/(1-p))$ して, 回帰分析を適用する. ところで, $p/(1-p)$ は, オッズ (odds) として賭け事の世界でよく用いられている. 全体で 25% の人がある馬が勝つことに掛けた場合のオッズは, 負ける人を p とすると $0.75/0.25 = 4$ 倍となる. 50% の人が掛けた場合には, $0.5/0.5 = 1$ 倍となる. 2 つのオッズの比

$$\psi = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$$

は, 疫学研究でオッズ比として常用されている. 出現率の差 $(p_1 - p_2)$ は, $(-1, 1)$ の範囲であるのに対し, オッズ比 ψ は, $(0, \infty)$ の範囲となる.

さて, 2 つのロジット $\ln(p_1/(1-p_1))$ と $\ln(p_2/(1-p_2))$ の差は,

$$\ln \psi = \ln\left(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\right)$$

対数オッズ比に等しくなる. 対数オッズの標準偏差 (標準誤差) は,

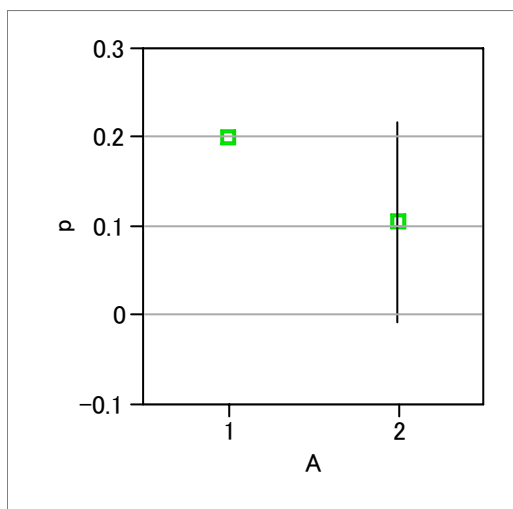
$$D[\ln \psi] \approx \sqrt{\frac{1}{n_1 p_1} + \frac{1}{n_1 (1-p_1)} + \frac{1}{n_2 p_2} + \frac{1}{n_2 (1-p_2)}}$$

となることが知られている. 対数オッズ比の 95% 信頼区間は,

$$\ln \psi \pm 1.96 D[\ln \psi]$$

となる.

Excel による計算結果より, 正規近似の場合 $(p_1 - p_2) = 0.200 - 0.106 = 0.094$, 95% 信頼区間は $(-0.018, 0.206)$ となる. p_1 を基準としたときに p_2 上に差の 95% 信頼区間を描いてみよう. 差の 95% 信頼区間は p_1 を 0 とした相対的な目盛りとなっているので, p_2 上の実目盛りでは $(-0.006, 0.218)$ となる.



表示 4.8 信頼区間の計算 Excel シート

比率の差の95%信頼区間						
2xk表, 頻度, 出現率, SE						正規近似
A	y=1	y=2	n _i	p _{i1}	p _{i2}	V
1	15	60	75	0.200	0.800	0.002133
2	9	76	85	0.106	0.894	0.001114

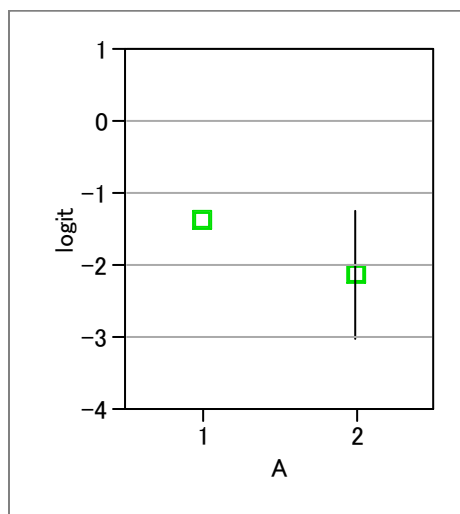
A _i と A ₁ の比率の差と正規近似による95%信頼区間				
A _i -A ₁	p _{i1} -p ₁₁	SE	L95%	U95%
1 - 2	0.094	0.057	-0.018	0.206
2			-0.006	0.218

A	logit 対数odds	logit ln(OR)	logit, ln(OR) D[lnOR]	L95%	U95%
1	-1.386	—	—		
2	-2.134	0.747	0.455634	-3.02655	-1.24047

ロジットの場合 (-3.02655, -1.24047)

$$L95\%(p1)=1/(1+\exp(-z1))$$

比率に直した場合 (0.046241, 0.224355) となり, 負にならない.



表示 4.9 比率の差の 95%信頼区間

表示 4.10 JMP ロジスティック回帰による 95%信頼区間

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	1.388671	1	2.777341	0.0956
Full	66.244784			
Reduced	67.633454			

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiS	Lower 95%	Upper 95%
Intercept	-1.3862944	0.2886751	23.06	<.0001	-1.9885714	-0.8485262
A2	-0.7472096	0.4556335	2.69	0.1010	-1.6753579	0.13039285

表示 4.11 SAS による検証

Parameter Estimates and 95% Confidence Intervals
Profile Likelihood
Confidence Limits

Variable	Parameter Estimate	Lower	Upper
INTERCPT	-1.3863	-1.9886	-0.8485
A2	-0.7472	-1.6754	0.1304

Wald
Confidence Limits

Variable	Parameter Estimate	Lower	Upper
INTERCPT	-1.3863	-1.9521	-0.8205
A2	-0.7472	-1.6402	0.1458

表示 4.12 松田方式による信頼区間, ***未検証***

松田2次形式		オッズ比の信頼区間					
A	y=1	y=2	T				
1	a	b	n	p	106	13	119 0.890756
2	c	d	m	q	100	15	115 0.869565
T	g	h	N		206	28	234
$r_{.1} \leq r \leq r_{.2}$							
$r = (p/(1-p))/(q(1-q))$					0.8579455		
$s.e. = \sqrt{1/a+1/b+1/c+1/d}$					0.4037619		
$r_{1,2} = \exp(\ln(r) \pm 1.96 \text{ s.e.}(r))$					r1	r2	
					0.3888397	1.89299218	
$np + mq = g$					206	206	
$A(r)p^2 + N(r)p + C(r) = 0$							
$A(r_i) = n(r_i - 1)$					-72.72807	106.266069	
$B(r_i) = g - m - r_i(n+g)$					-35.37291	-524.22246	
$C(r_i) = r_i * g$					80.100981	389.956389	
$P(r_i) = \text{2次式解(-)}$					0.8340863	0.91276192	
$Q(r_i) = (g - nP(r))/m$					0.9282063	0.84679418	
$p_{.1,2} = P(r_i) - Q(r_{.1})$					-0.09412	0.06596774	

4.3 適合度の検定 (参考)

表示 4.1 のデータは、一般的には、2 x 3 の分割表に対する適合度の χ^2 検定として解析される。ところが、この方法は、説明変数が 2 因子以上へ拡張がしにくいために、実験計画法の標準的な解析方法として適していない。しかし、基本的な統計手法であるので、その考え方を示す。

表示 4.13 の左半分の実測度数は、表示 4.1 のデータをそのまま転記したものである。この表を分割表 (Contingency Table) という。行の数を m 、列の数を k で表わす。

表示 4.13 実測度数と期待度数

	実測度数			期待度数	
	良品	不良品	合計	良品	不良品
A_1	60 (n_{11})	15 (n_{12})	75 ($n_{1.}$)	60.90 (\hat{n}_{11})	14.10 (\hat{n}_{12})
A_2	76 (n_{21})	9 (n_{22})	85 ($n_{2.}$)	69.02 (\hat{n}_{21})	15.98 (\hat{n}_{22})
A_3	67 (n_{31})	23 (n_{32})	90 ($n_{3.}$)	73.08 (\hat{n}_{31})	16.92 (\hat{n}_{32})
合計	203 ($n_{.1}$)	47 ($n_{.2}$)	250 ($n_{..}$)	203.00 ($n_{.1}$)	47.00 ($n_{.2}$)

3種類の砂によって巣の発生率に差がないとすれば、どの砂でも、全体の不良率 0.188 の割合で巣が発生するはずである。そのように考えて計算した良品と不良品の個数を**期待度数**という。実測度数は整数であるが、期待度数は一般に端数が付く。 i 行 j 列の実測度数を n_{ij} 、期待度数を \hat{n}_{ij} という記号で表すことにする。また、 i 行の合計を $n_{i.}$ 、 j 列の合計を $n_{.j}$ 、総合計を $n_{..}$ とする。期待度数は次の式で計算される。

$$\hat{n}_{ij} = \frac{n_{i.} \cdot n_{.j}}{n_{..}} \quad (4.2)$$

期待度数を 表示 4.13 の右半分に示す。

帰無仮説 H_0 : 3種類の砂によって巣の発生率に差がない

が正しければ、実測度数と期待度数の差は大きくないはずである。分割表の全部の要素について、両者の差を総合化するために、 $n_{ij} - \hat{n}_{ij}$ を2乗して加える。しかし、期待度数が5に対して差が1である場合と、期待度数が10に対して差が1である場合とでは、前者の方が違いが大きい。そこで、(実測度数 - 期待度数)² / 期待度数 = $(n_{ij} - \hat{n}_{ij})^2 / \hat{n}_{ij}$ をすべての要素について加え合わせる。得られた量を χ^2 という記号で表す。

$$\begin{aligned} \chi^2 &= \sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} \quad (4.3) \\ &= \frac{(60 - 60.90)^2}{60.90} + \frac{(15 - 14.10)^2}{14.10} + \dots + \frac{(23 - 16.92)^2}{16.92} \\ &= 6.516 \end{aligned}$$

帰無仮説の下で、 χ^2 は、近似的に自由度が $(m-1)(k-1) = (3-1)(2-1) = 2$ の χ^2 分布に従う。上に求めた χ^2 の値 6.516は、 χ^2 表の上側5%点 (5.99) よりも大きいので、帰無仮説は棄却される。すなわち、砂の種類によって巣の発生率に有意

な差がある.

4.4 カテゴリの反応変数 (カテゴリの順序に意味なし)

表示 4.1 のデータで, 不良品を傷と割れに分けると, 表示 4.14 が得られた. このデータについて, 前節と同様の解析をして, 砂の種類によって不良の出方や不良の種類に違いがあるかを確認したい.

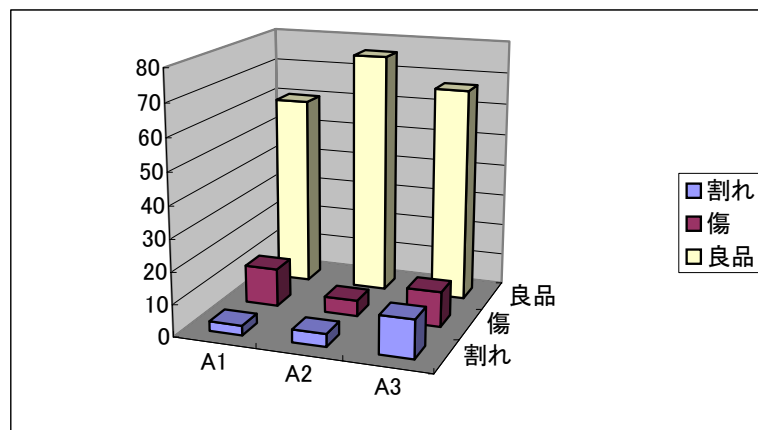
多特性のロジットの定義は, 反応が k 水準の場合にJMPでは, 式 (4.4) で与えられている.

$$z = \ln\left(\frac{P(y = j)}{P(y = k)}\right), \quad j = 1, 2, \dots, k-1 \quad (4.4)$$

表示 4.14 多特性のデータ

	割れ	傷	良品	合計
A_1	3	12	60	75
A_2	4	5	76	85
A_3	12	11	67	90
合計	19	28	203	250

表示 4.15 棒グラフ



表示 4.16 に「良品」を k とした場合の「不良率」 p とロジット z を示す. 傷の率に関しては A_2 が低く, 割れの率に関しては A_1 と A_2 が低い結果となっている. これらを総合して因子 A による不良率に差があるかを検討したい.

表示 4.16 多特性のロジット

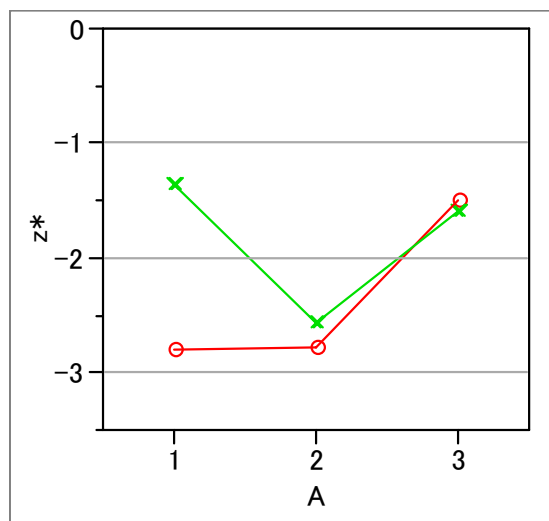
A	割れ (y=1)			傷 (y=2)		
	r_1 / r_3	p_1	z_1	r_2 / r_3	p_2	z_2
A ₁	3 / 60	0.050	-2.944	12 / 60	0.200	-1.386
A ₂	4 / 76	0.053	-2.890	5 / 76	0.066	-2.653
A ₃	12 / 67	0.179	-1.522	11 / 67	0.164	-1.627
合計	19 / 203	0.094	-2.270	28 / 203	0.138	-1.833

JMPによるロジスティック回帰分析を行うために 表示 4.17 に示すように、因子Aについてダミー変数 a_1, a_2, a_3 を作成し、「割れ」を $y=1$ 、「傷」を $y=2$ 、「良品」を $y=3$ とし、率 p とロジット z の計算結果を示す。

表示 4.17 多特性の解析用データ

A	a_1	a_2	a_3	y	r_i	r_3	p	z^*
A ₁	1	0	0	1	3	60	0.050	-2.799
A ₁	1	0	0	2	12	60	0.200	-1.356
A ₁	1	0	0	3	60	60	1.000	4.796
A ₂	0	1	0	1	4	76	0.053	-2.780
A ₂	0	1	0	2	5	76	0.066	-2.565
A ₂	0	1	0	3	76	76	1.000	5.030
A ₃	0	0	1	1	12	67	0.179	-1.491
A ₃	0	0	1	2	11	67	0.164	-1.592
A ₃	0	0	1	3	67	67	1.000	4.905

表示 4.18 経験ロジットによる多特性のグラフ



○ : 割れ (y=1), ×: 傷 (y=2)

A_2 を基準とする差の検討をするために，説明変数として A_1 と A_3 としたロジスティック回帰分析の結果を表示 4.14 に示す．

Whole-Model Test の Difference の p 値が 0.0259 なので，3 種類の砂によって，反応に有意な差がある．Effect Test の結果から， A_1 に比較し A_2 は差がなく， A_3 には $p = 0.0435$ なので，有意な差がある．ただし，多重性を考慮した場合は，有意な差があるとはいえなくなる．Parameter Estimate に切片 Intercept が 2 つあるが，これについては，1 番目の切片が「割れ」に対する推定値であり，2 番目の切片が「傷」に対する推定値に対応する．

表示 4.19 多特性のロジスティック回帰分析

Whole Model Test				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	5.53214	4	11.06428	0.0259
Full	147.00621			
Reduced	152.53835			
Parameter Estimates				
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-2.944438	0.5129889	32.94	<.0001
a1	-0.051173	0.7830192	0.00	0.9479
a3	1.224652	0.6011791	4.15	0.0416
Intercept	-2.7212954	0.4616903	34.74	<.0001
a1	1.1118572	0.5596056	3.95	0.0469
a3	0.9144980	0.5647941	2.62	0.1054
Effect Wald Tests				
Source	Nparm	DF	Wald ChiSquare	Prob>ChiSq
a1	2	2	3.98803104	0.1361
a3	2	2	6.26869189	0.0435

4.5 尤度比検定（カテゴリーに順序なし，参考）

分割表を解析する別の方法について説明する．帰無仮説（2 つの質的変数のカテゴリーは独立である）が正しいとき，

$$\hat{p}_{ij} = p_{i \cdot} \cdot p_{\cdot j}$$

が成立し，そのときの尤度 L_0 とその対数は，2 項係数を除いて，

$$\begin{aligned}
L_0 &= \prod_i \prod_j \hat{p}_{ij}^{n_{ij}} \\
&= \prod_i \prod_j \left(\frac{n_{i.} n_{.j}}{n_{..}^2} \right)^{n_{ij}} \\
\ln L_0 &= \sum_i \sum_j n_{ij} (\ln n_{i.} + \ln n_{.j} - 2 \ln n_{..}) \\
&= \sum_i \sum_j n_{ij} \ln n_{i.} + \sum_i \sum_j n_{ij} \ln n_{.j} - 2 \sum_i \sum_j n_{ij} \ln n_{..} \\
&= \sum_i n_{i.} \ln n_{i.} + \sum_j n_{.j} \ln n_{.j} - 2 n_{..} \ln n_{..}
\end{aligned}$$

となる．対立仮説の下では， $\hat{p}_{ij} = p_{ij}$ であるから，そのときの尤度 L_0 とその対数は，

$$\begin{aligned}
L_1 &= \prod_i \prod_j p_{ij}^{n_{ij}} \\
&= \prod_i \prod_j \left(\frac{n_{ij}}{n_{..}} \right)^{n_{ij}} \\
\ln L_1 &= \sum_i \sum_j n_{ij} (\ln n_{ij} - \ln n_{..}) \\
&= \sum_i \sum_j n_{ij} \ln n_{ij} - \sum_i \sum_j n_{ij} \ln n_{..} \\
&= \sum_i \sum_j n_{ij} \ln n_{ij} - n_{..} \ln n_{..}
\end{aligned}$$

となる．対数尤度の差の 2 倍を**逸脱度 (deviance)** と呼び，帰無仮説が正しいとき，逸脱度は χ^2 分布となることが知られている．ここで，求めた対数尤度は，ロジスティック回帰分析の尤度と一致する．

$$\chi^2 = 2 \left(\sum_i \sum_j n_{ij} \ln n_{ij} - \sum_i n_{i.} \ln n_{i.} - \sum_j n_{.j} \ln n_{.j} + n_{..} \ln n_{..} \right) \quad (4.5)$$

を式 (4.3) の代わりに用いる．この式は，2 因子実験の 2 因子交互作用を求める式

$$S_{A*B} = \sum_i \sum_j x_{ij}^2 - \sum_i \bar{x}_{i.} T_{i.} + \sum_j \bar{x}_{.j} T_{.j} + \bar{x}_{..} T_{..}$$

と類似している．ここに， \bar{x} は平均， T は合計である．

上の検定は，良品，傷，割れの 3 つのカテゴリーの割合の違いを全体として比較したものである．もし，不良の中で，傷と割れの比率が違うかどうかを調

べたい場合は、横のカテゴリーを傷と割れの2つにして、 $m \times 2$ の分割表として解析すれば良い。

表示 4.14 で得られた、良品：不良品の分析と合わせ、3つの方法で解析して比較すると、このように、行や列のカテゴリーが2以上で、カテゴリーに階層構造が考えられる場合には、全体で計算した χ^2 を固有技術的に意味のある階層に分解することにより、有益な情報が得られる可能性がある。これは、 y が計量値の分散分析で、直交する対比を使って平方和を分解することに対応している。

表示 4.20 に逸脱度の分解の計算例を示す。 B_1, B_2, B_3 の違いについて $\chi^2 = 11.06$, 自由度=4 が得られ、これが $(B_1 + B_2), B_3$ の違い $\chi^2 = 6.86$, 自由度=2 と、 B_1, B_2 の違いの $\chi^2 = 4.20$, 自由度=2 に分解されることが分かる。式(4.5)で計算した χ^2 では加法性が完全に成立する⁴。 B_1, B_2, B_3 の違いから得られた $\chi^2 = 11.06$ は、表示 4.19 のWhole Medel Test で得られた χ^2 値に一致していることが確かめられる。

このように、行や列のカテゴリーが3以上で、カテゴリーに階層構造が考えられる場合には、全体で計算した χ^2 を固有技術的に意味のある階層に分解することにより、有益な情報が得られる可能性がある。これは、 y が計量値の分散分析で、直交する対比を使って平方和を分解することに対応している。

表示 4.20 逸脱度の分解

	B_1	B_2	B_3	合計	$B_1 + B_2$	B_3	合計	B_1	B_2	合計	
A_1	3	12	60	75	15	60	75	3	12	15	
A_2	4	5	76	85	9	76	85	4	5	9	
A_3	12	11	67	90	23	67	90	12	11	23	
合計	19	28	203	250		203	250	19	28	47	
			χ^2	11.06			χ^2	5.86		χ^2	4.20
			自由度	4			自由度	2		自由度	2
			p 値	0.026			p 値	0.032		p 値	0.122

この章で述べた手法は、互いに独立かつランダムにとられた n 個の内 r 個が

⁴ 式(4.3)で計算した適合度の検定における χ^2 では加法性が近似的に成立するに過ぎない。いずれの式も近似式である。どちらの式の方が近似が良いかには、ここでは触れないことにする。

不良であった場合に成立するものである。ある条件で n 回独立に試行し, r 回失敗したという場合はこの条件を満たしている。しかし, ある条件でまとめて n 個生産し, その中に r 個の不良品が含まれていたという場合は上の条件を満たさない。もう一度同じ条件で n 個生産したとき, 真の不良率が前と同じであるという保証がないからである。これは, 通常の実験計画法で, 分割実験に相当する。

5 2 因子実験（説明変数は質と量，反応変数は質的）

5.1 データ

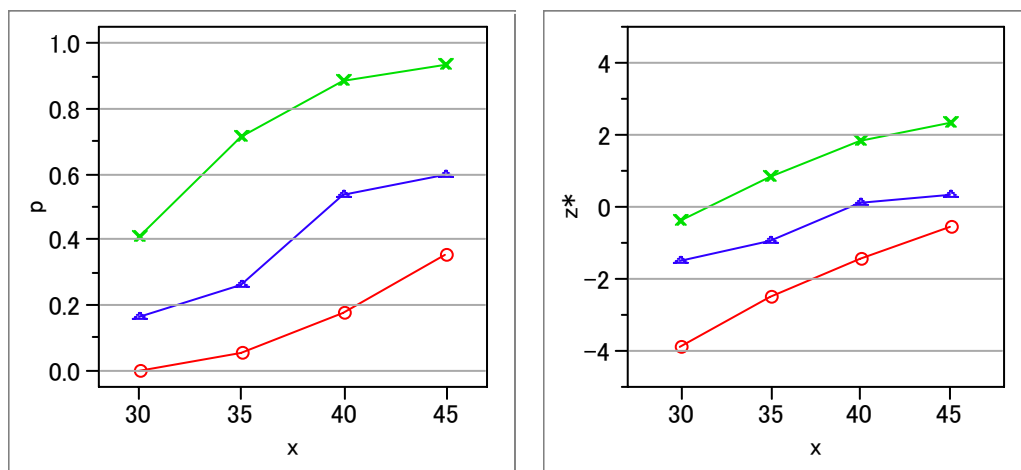
3 水準の質的因子 A （たとえば，ガラスの種類）と 4 水準の量的因子 B （たとえば，衝撃強さ）で $x = (30, 35, 40, 45)$ の 2 元配置の実験を実施した．反応変数は，各水準組合せで作られた製品の個数 n_{ij} と製品中の不良個数 r_{ij} の比率 $p_{ij} = r_{ij} / n_{ij}$ である．表示 5.1 の数字は $r/n, p$ である．

表示 5.1 2 因子実験の不良率

	B_1 (30 kg)	B_2 (35 kg)	B_3 (40 kg)	B_4 (45 kg)
A_1	0/23 0.000	1/18 0.056	4/22 0.182	5/14 0.357
A_2	7/17 0.412	10/14 0.714	16/18 0.889	15/16 0.938
A_3	3/18 0.167	4/15 0.267	7/13 0.538	9/15 0.600

表示 5.1 のデータの比率 p_{ij} を縦軸に，横軸に衝撃強さ x をとり， A の水準毎に線で結ぶと表示 5.2 の左が得られる．この図から， A と B の間には交互作用があるように見える．表示 5.2 右は，縦軸に経験ロジット z^* をとったものである．

表示 5.2 x と p , および x と z^* の関係



5.2 ロジット変換の考え方

説明を簡単にするために， A, B が共に 2 水準の場合を考える， A_1, B_1 は現行水準で，不良率は 0.12 (12%) である．それぞれの改良水準が A_2, B_2 である．

1つの因子ごとに改良をした場合に表示 5.3 の不良率の欄に示すような不良率 0.08 と 0.06 が得られたとする。A₂B₂ の不良率は、いくらになると予想されるであろうか。表示 5.3 の中央に示すように、A₁ を A₂ にすると不良率が 0.06 減少し、B₁ を B₂ にすると不良率が 0.04 減少すると考えると、A₂B₂ の不良率は、

$$p_{22} = 0.12 - 0.06 - 0.04 = 0.02$$

になると予想される。製造現場の現実を知っている人には不良率がこんなに小さくなるとは考えない。

表示 5.3 効果の加法性と乗法性

	不良率		加法モデル			乗法モデル		
	B ₁	B ₂	B ₁	差	B ₂	B ₁	比	B ₂
A ₁	0.12	0.08	0.12	(-0.04)	0.08	0.12	(2/3)	0.08
A ₂	0.06	?	(-0.06)	(-0.04)	(-0.06)	(1/2)	(2/3)	(1/2)

表示 5.3 の乗法モデルの欄に示すように、A₁ を A₂ にすると不良率が 1/2 に減少し、B₁ を B₂ にすると不良率が 2/3 に減少すると考えると、A₂B₂ の不良率は、

$$p_{22} = 0.12 \cdot (1/2) \cdot (2/3) = 0.04$$

になると予想するのが現実的である。このように A, B の効果は乗算的である。

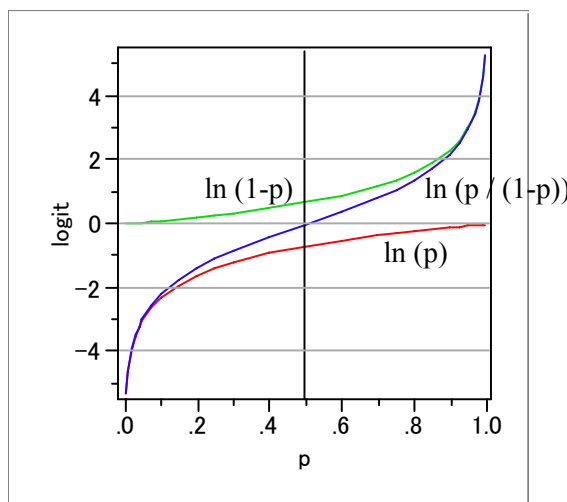
このようなとき、比率 p の対数 ln(p) をとると加法性が成立する。逆に、不良率が 100% に近くなると、-ln(1-p) をとると良い、変換によって大小関係が逆転しないようにするために、- の符号を付ける。比率 p が 0 から 1 まで広く変化するとき、2つの式を合成した

$$\begin{aligned}
 z &= \ln(p) - \ln(1-p) & (5.1) \\
 &= \ln\left(\frac{p}{1-p}\right)
 \end{aligned}$$

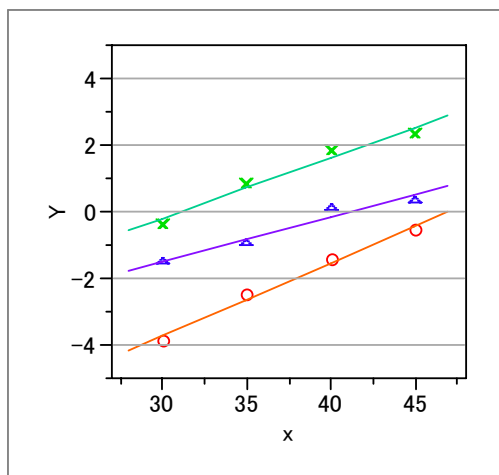
の変換をすると広範囲に加法性が成立することが期待される。合成の経過を表示 5.4 に示す。これはすでに第 3 章で導入したロジット変換に外ならない。これから、ロジット変換は、反応変数の変域を -∞ から +∞ に拡張して、x との直線性を向上するだけでなく、複数の要因の加法性を向上するためにも有効

であることが分る. 表示 5.2 の右の図は, 縦軸を経験ロジット z^* にしたものである. 曲率がやや小さく, さらに交互作用も小さくなっていることが分るであろう.

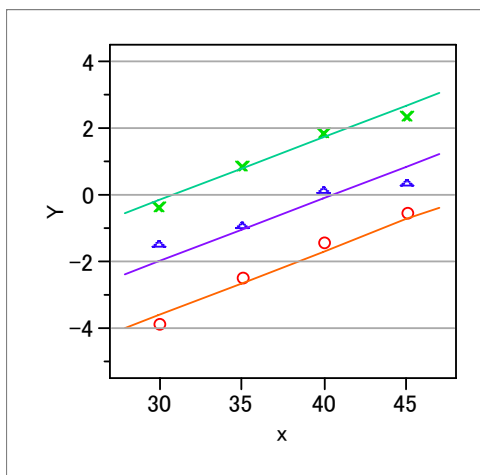
表示 5.4 ロジットの構成



表示 5.5 直線の当てはめ



A1 ○ : $z^* = -10.282 + 0.219 x$
 A2 × : $z^* = -5.6082 + 0.181 x$
 A3 △ : $z^* = -5.489 + 0.1337 x$



A1 : $z^* = -9.314 + 0.191 x$
 A2 : $z^* = -5.882 + 0.191 x$
 A3 : $z^* = -7.726 + 0.191 x$

ロジスティック回帰分析は, ロジット z を用いるが, 解析の手順は一般的な回帰分析と同じである.

- 1) 3種類のガラスそれぞれにロジスティック回帰直線を当てはめる. 表示 5.5

の左の図.

- 2) 共通の傾きを持つが、切片は異なるロジスティック回帰直線を当てはめる.
表示 5.5 の右の図.
- 3) それぞれのモデルの逸脱度の差より交互作用を評価し、どちらのモデルが適切化を判断する.

5.3 JMP による解析

解析用データ

JMPで解析するためのデータを表示 5.6 に示す. 変数 A は質的変数, 変数 x , r は量的変数である. さらに, ダミー変数 a_1, a_2, a_3 を作成しておく. JMPでは, 自動的にダミー変数を作成する機能を持っているが, 他の統計ソフトでは, この機能を持っていないものもあること, 推定値の意味を適切に理解するために, ダミー変数を用いた解析を行う.

表示 5.6 解析用データ

A	a_1	a_2	a_3	B	x	y	r
1	1	0	0	1	30	0	0
1	1	0	0	1	30	1	23
中略							
1	1	0	0	4	45	1	9
2	0	1	0	1	30	0	7
2	0	1	0	1	30	1	10
中略							
2	0	1	0	4	45	1	1
3	0	0	1	1	30	0	3
3	0	0	1	1	30	1	15
中略							
3	0	0	1	4	45	1	6

交互作用の検討

JMP / Fit Model を用いる. 反応変数として y を質的変数 Nominal にし, r を繰返し度数 Freq に指定し, 説明変数として主効果 A , x を, 交互作用として $A * x$ を用いる. 表示 5.7 の Whole Model Test の結果から, モデルを当てはめたときの対数尤度は, $\ln L_{\text{model}} = -46.58$, $-2 \ln L_{\text{model}} = 93.155$ となる. Effect Likelihood Ratio Tests から, 交互作用 $A * x$ の対数尤度は $-2 \ln L_{A*x} = 1.62$ となり, これが自由度 2 の χ^2 分布に従うことから $p = 0.4451$ と有意な差ではない.

交互作用なしの主効果モデル

JMP / Fit Model を再度実行し，説明変数をダミー変数 a_2 と a_3 ，および x にする．Whole Model Test の結果から， $-2 \ln L_{\text{model}} = 91.536$ となる．交互作用を含むモデルとの対数尤度の差 $-91.536 - (-93.155) = 1.619$ が交互作用 $-2 \ln L_{A*x}$ に

表示 5.7 交互作用の検討

Whole Model Test				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	46.57738	5	93.15476	<.0001
Full	89.96248			
Reduced	136.53986			

Effect Likelihood Ratio Tests				
Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
A	2	2	63.3580154	0.0000
x	1	1	34.4203251	0.0000
A*x	2	2	1.6187567	0.4451

一致する．反応変数が連続量の場合での平方和の考え方と同じことが，対数尤度の場合のも当てはまる．Parameter Estimates の χ^2 値は，

$$\chi^2 = (\text{Estimate} / \text{Std Error})^2$$

で計算されたワルド χ^2 である．ワルド χ^2 は，尤度から求めた χ^2 値と完全に一致しないが，計算上の簡便性から一般的に用いられている．

表示 5.8 主効果モデル

Whole Model Test				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	45.76800	3	91.536	<.0001
Full	90.77186			
Reduced	136.53986			

Parameter Estimates						
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq	Lower 95%	Upper 95%
Intercept	-9.3138293	1.539179	36.62	<.0001	-12.5325	-6.46797
a2	3.4314013	0.516391	44.16	<.0001	2.47418	4.50932
a3	1.5877494	0.472946	11.27	0.0008	0.68820	2.55455
x	0.1913376	0.036987	26.76	<.0001	0.12215	0.26791

x の傾きは，0.191， χ^2 検定の結果は $p < 0.0001$ と有意である． A_1 を基準とし

た A_2 との水準間の差は 3.431 で χ^2 検定の結果 $p < 0.0001$ と有意である. 同様に A_1 を基準とした A_3 との水準間のロジットの差は 1.588 で χ^2 検定の結果は $p = 0.0008$ と有意である.

回帰式

x の傾きは, 0.191, χ^2 検定の結果は $p < 0.0001$ と有意である. A_1 を基準とした A_2 との水準間の差は 3.431 で χ^2 検定の結果 $p < 0.0001$ と有意である. 同様に A_1 を基準とした A_3 との水準間のロジットの差は 1.588 で χ^2 検定の結果は $p = 0.0008$ と有意である.

表示 5.8 の Parameter Estimates から, A_1 , A_2 および A_3 の回帰式が次のように得られる.

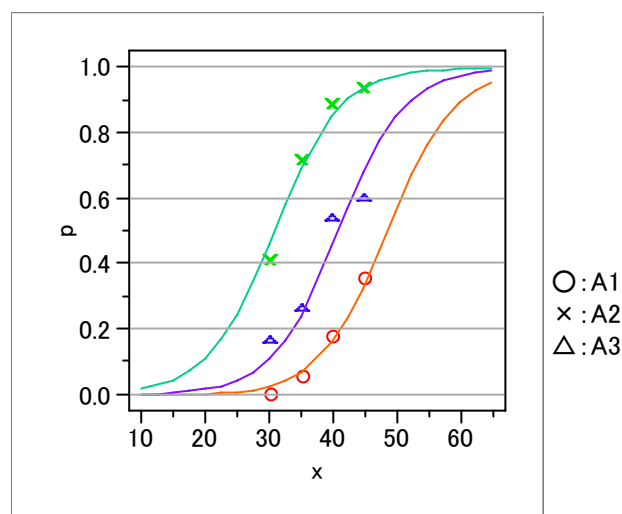
$$\hat{z} = -9.314 + \begin{pmatrix} 0.000 \\ 3.432 \\ 1.588 \end{pmatrix} + 0.191x \quad (5.1)$$

この式から逆ロジットで A_1 の推定値は,

$$\begin{aligned} \hat{p} &= \frac{1}{1 + e^{-\hat{z}}} \\ &= \frac{1}{1 + e^{-(-9.314 + 0.191x)}} \end{aligned} \quad (5.2)$$

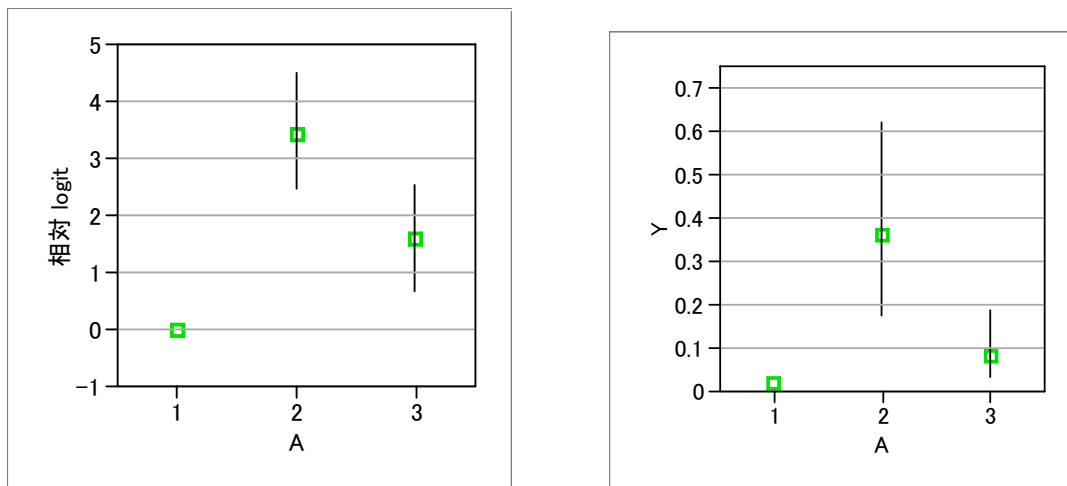
で計算できる. A_2 と A_3 も同様に逆ロジットを計算し 表示 5.9 に示のように図示すると, A_1 , A_2 および A_3 が平行なシグモイド曲線になる.

表示 5.9 主効果モデルの平行なシグモイド曲線



水準間の差の95%信頼区間

A1 水準と A2 水準間の差の95%信頼区間, A1 と A3 の間, これは A1 を 0 としたとき



A	L95logit	m	U95logit	L95_p	m_p	U95_p
1	.	0.000	.	.	0.018	.
2	2.474	3.431	4.509	0.179	0.362	0.625
3	0.688	1.588	2.555	0.035	0.082	0.191

表示 5.10 A1 を基準にした差の信頼区間

5.4 2次曲線のあてはめ (演習)

3種類の原料 (A_1, A_2, A_3) について, 成形圧力を5水準に変化させた実験を行い, 表示 5.11 のデータが得られた. 不良率を最小にする条件を求めたい. ただし, 原料によって最適条件が異なることが考えられる.

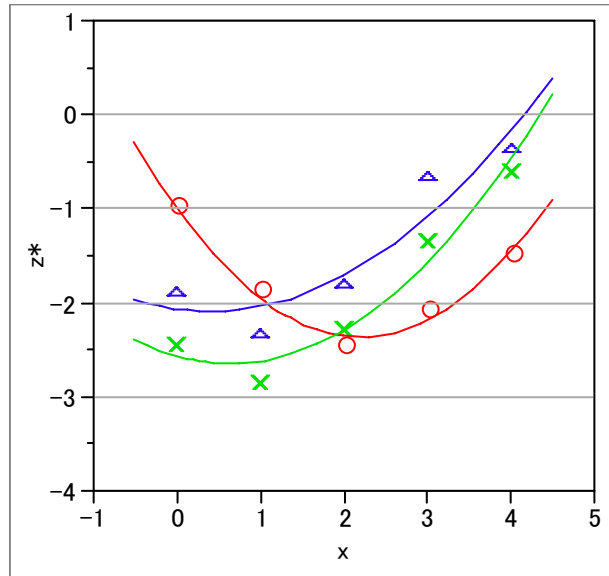
表示 5.11 2次曲線データ (r/n)

	x				
	0	1	2	3	4
A_1	27 / 100	13 / 101	7 / 95	11 / 102	20 / 111
A_2	6 / 80	5 / 100	9 / 101	21 / 103	33 / 93
A_3	13 / 102	8 / 95	14 / 101	33 / 98	40 / 98

表示 5.12 にA の水準ごとに経験ロジットに 2 次曲線を当てはめた結果を示

す. この場合は, A と x に交互作用があるとみなした解析となっている.

表示 5.12 2次曲線の当てはめ



○ : A1 $z^* = -2.098 - 0.122x + 0.281(x-2)^2$
 × : A2 $z^* = -3.327 + 0.520x + 0.192(x-2)^2$
 △ : A3 $z^* = -2.650 + 0.471x + 0.148(x-2)^2$

このデータに,

$$\hat{z}_{(1)} = b_{0(1)} + b_{1(1)}x + b_2x^2 \quad (A1)$$

$$\hat{z}_{(2)} = b_{0(2)} + b_{1(2)}x + b_2x^2 \quad (A2) \tag{5.3}$$

$$\hat{z}_{(3)} = b_{0(3)} + b_{1(3)}x + b_2x^2 \quad (A3)$$

という式を当てはめたい. ここに, b_0 と b_1 は A の水準によって異なるが, b_2 はどの原料でも同じであるとする. この解を求めるには, 説明変数として

質的因子 A のための, 質的変数 A

量的因子 B のための, 量的変数 x と x^2

の他に, x の係数が A の水準によってどれだけ異なるかを求めるための変数 (ダミー変数) が必要である. JMP は質的変数 A のダミー変数, 傾斜の違いのためのダミー変数も生成してくれるが, 他の統計ソフトにこの機能は, あまり見かけないのでないで, 自分で生成する方式も示す.

傾斜 b_1 の違いのためのダミー変数の生成方法を理解するためには, 定数項 b_0 のダミー変数について理解しておかねばならない. 2種類のダミー変数を表

示 5.13に示す. 表示 5.13 を見ると, 傾斜についてのダミー変数は, 定数項のダミー変数 (2 個) と x とを掛けたものになっていることがわかる.

表示 5.13 交互作用のためのダミー変数

No.	A	ダミー変数		x	ダミー変数	
		a_2	a_3		$(ax)_2$	$(ax)_3$
		$b_{0.2}$	$b_{0.3}$		$b_{1.2}$	$b_{1.3}$
1	1	0	0	x_1	0	0
2	1	0	0	x_2	0	0
:						
6	2	1	0	x_1	x_1	0
7	2	1	0	x_2	x_2	0
:						
11	3	0	1	x_0	0	x_1
12	3	0	1	x_2	0	x_2
:						
15	3	0	1	x_5	0	x_5

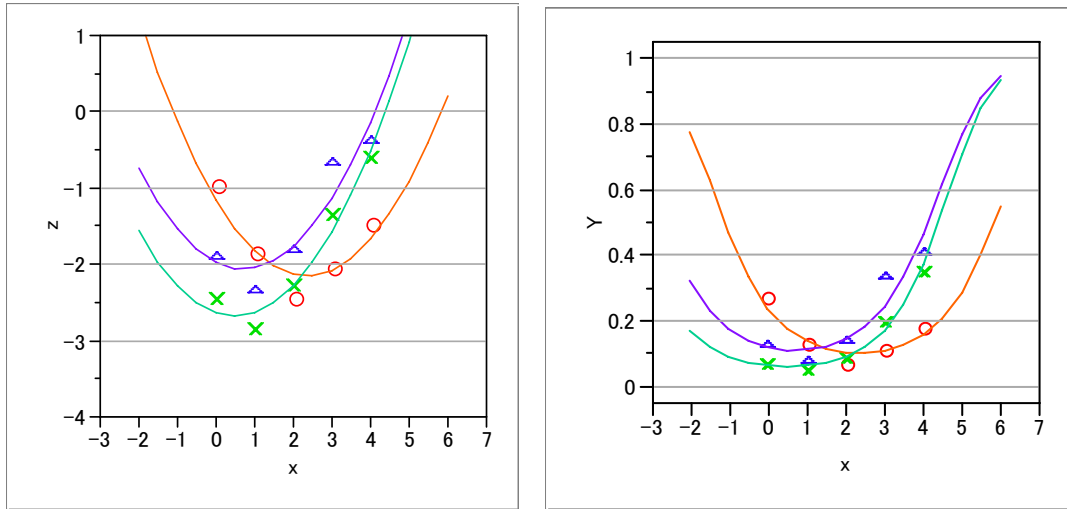
ダミー変数に対する回帰係数を使うと, 式 (5.3) は, 次のように書き直される.

$$\hat{z}_{(1)} = (b_0 + 0) + (b_1 + 0)x + b_2x^2 \quad (A1)$$

$$\hat{z}_{(2)} = (b_0 + b_{0.2}) + (b_1 + b_{1.2})x + b_2x^2 \quad (A2) \quad (5.4)$$

$$\hat{z}_{(3)} = (b_0 + b_{0.3}) + (b_1 + b_{1.3})x + b_2x^2 \quad (A3)$$

$b_{0.2} = b_{0(2)} - b_{0(1)}$, $b_{1.2} = b_{1(2)} - b_{1(1)}$ の関係があり, A_1 を基準として, A_2 の $b_{0(2)}$, $b_{1(2)}$ がどれだけ異なるかを示している. JMPで出力されたロジットの推定値の関数は, 表示 5.14 のようである.



表示 5.14 曲率が同じロジスティック 2 次曲線

- : A1 $\hat{z}_1 = -1.849 - 0.131x + 0.1780(x-2)^2$, $\hat{P}_1 = 1/(1+\exp(-\hat{z}_1))$
- × : A2 $\hat{z}_2 = -3.342 + 0.530x + 0.1780(x-2)^2$, $\hat{P}_2 = 1/(1+\exp(-\hat{z}_2))$
- △ : A3 $\hat{z}_3 = -2.670 + 0.455x + 0.1780(x-2)^2$, $\hat{P}_3 = 1/(1+\exp(-\hat{z}_3))$

6 累積データの解析（反応が順序）

6.1 1 因子実験

累積ロジット

表示 3.1 の強化ガラスの衝撃試験には結果では、破壊と非破壊にデータが分類されていた。実は、破壊数だけでなく「ひび入り」も観察されていた。そのデータを表示 6.1 に示す。一番右の欄に、破壊数 r_1 と ひび入り数 r_2 の合計を示す。これを R_2 とする。また、破壊数 r_1 を改めて R_1 とする。

田口の累積法は、まず R_1 と R_2 を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 6.1 順序がある破壊データ

衝撃 kg	変換値 x	破壊 $r_1=R_1$	ひび入 r_2	非破壊 r_3	合計 n	累積 $r_1+r_2=R_2$
35	1	0	1	8	9	1
40	2	1	2	7	10	3
45	3	0	2	6	8	2
50	4	3	7	9	19	10
55	5	3	6	2	11	9
60	6	6	6	2	14	12
65	7	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_1 と R_2 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、偏りの原因（有意差が出やすくなる）である。田口の累積法は、電卓もない時代に、ソロバンでも解析できる手法として広く用いられてきた。計算手段が進歩した現在では、ロジスティック回帰分析により簡単に解析することができる。

横軸に x を、縦軸に $P_1 = R_1/n$ と $P_2 = R_2/n$ をとったグラフを描くと表示

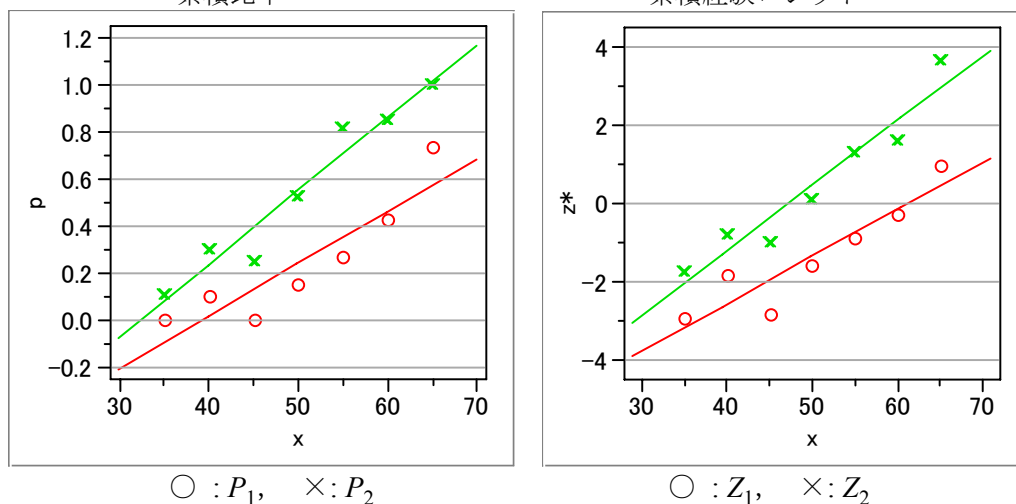
6.2 左の図が得られる. 表示 6.2 の左の図は, 累積経験ロジットで, 累積比率について

$$\begin{aligned}
 P_2^* &= ((R_2 + 0.5) / (n + 1)) \\
 Z_2^* &= \ln((P_2^* / (1 - P_2^*))) \tag{6.1}
 \end{aligned}$$

として計算したものである. $x = 35$ の場合, $R_2 = 1$, $n = 9$, $P_2^* = 0.15$, $Z_2^* = -1.73$ となる. このグラフは, 表示 5.2 の 2 因子実験のグラフと同じような形をしている. 表示 5.2 の 3 本の線は因子 A の 3 つの水準に対応しているが, 表示 6.2 の 2 本の線は 2 つの累積比率に対応している.

2 つの累積不良率を 2 因子実験と場合と同じように 2 本の平行線を当てはめの問題としても考えられる.

表示 6.2 累積比率 P_1 , P_2 と x の関係
 累積比率 累積経験ロジット



累積ロジスティック回帰分析

R_1 と R_2 は 9 個のテストピースから求めたものであって, 独立ではない. 通常のロジスティック回帰分析は観測された比率が 2 項分布に従うという前提で導かれる. 表示 6.1 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,

$$\begin{aligned}
 P_1 &= p_1, \\
 P_2 &= p_1 + p_2 \tag{6.2}
 \end{aligned}$$

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ー数が多く、中間のカテゴリーの比率が小さいとき相関が強くなる。表示 6.2 の場合であれば2つの折れ線が接近しているとき相関が強くなる。

表示 6.3 1 因子累積データ

x	y	r	R	n	P	Z^*
35	0	0	0	9	0.000	-2.94
40	0	1	1	10	0.100	-1.85
中略						
65	0	14	14	19	0.737	0.97
35	1	1	1	9	0.111	-1.73
40	1	2	3	10	0.300	-0.76
中略						
65	1	5	19	19	1.000	3.66
35	2	8	9	9	1.000	2.94
40	2	7	10	10	1.000	3.04
中略						
65	2	0	19	19	1.000	3.66

反応変数に順序がある場合に、JMP / Fit Modelでのロジスティック回帰分析は、内部的に累積ロジットにの計算を行うので、反応変数 y を順序尺度Ordinalとし、 x を説明変数、 r を重みとして解析すればよい。結果を表示 6.4 に示す。

表示 6.4 2つの切片をもつ累積ロジットの解

Ordinal Logistic Fit for y, Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	24.741479	1	49.48296	<.0001
Full	73.705959			
Reduced	98.447438			

Parameter Estimates

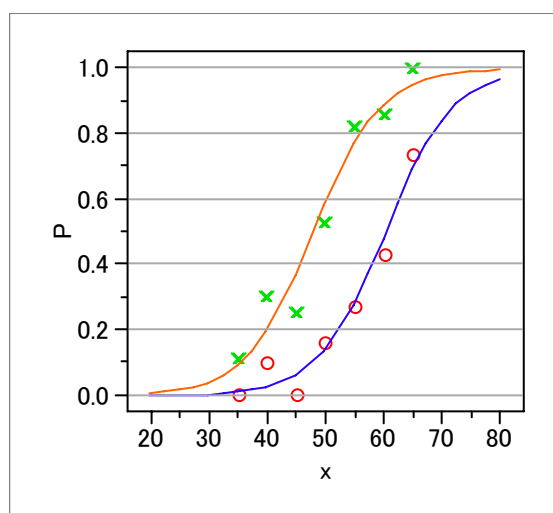
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept[1]	-10.545771	1.7152734	37.80	<.0001
Intercept[2]	-8.404819	1.5410578	29.75	<.0001
x	0.174733	0.0298074	34.36	<.0001

求めるロジスティック回帰直線は、推定結果は、Parameter Estimates から、

$$\begin{aligned}\hat{Z}_1 &= -10.546 + 0.17473 x \\ \hat{Z}_2 &= -8.4048 + 0.17473 x\end{aligned}\quad (6.3)$$

となる. Z の逆ロジットを計算し, 累積比率の推定曲線を示したのが 表示 6.5 であり, シグモイド曲線が横軸方向に平行移動している様子が示されている.

表示 6.5 推定値のグラフ



○ : $P_1=1/(1+\exp(-Z_1))$, × : $P_2=1/(1+\exp(-Z_2))$

P_1 の曲線が $x=50$ の○印の下になっている. 表示 6.2 右では上側であった. これは, 経験ロジット直線の平坦化の現象である.

2 群としての解析

JMP / Fit Model で 2 群として解析した結果を表示 6.6 に示す. Whole Model Testの自由度が 2 となり, 表示 6.4 の結果と異なっていることに注意してもらいたい. χ^2 値の比較は自由度 2 で割った $89.29963 / 2 = 44.64982$ と表示 6.4 の 49.48296 と比べて小さくなっている. これは, 内部の相関を無視したために統計的には当てはまりが悪くなったことの例示である.

推定値は, $y=0$ の場合を群 $G=0$, $y=1$ の場合を群 $G=1$ としたので, 累積比率のロジットの推定値は,

$$\begin{aligned}\hat{Z}_1 &= -10.612 + 0.176x \\ \hat{Z}_2 &= -10.612 + 2.143 + 0.176x = -8.469 + 0.176x\end{aligned}\quad (6.4)$$

となる. 結果は式 (6.3) とほとんど変わらない. 累積ロジットを取扱っていない統計ソフトの場合には, 近似的に多群として解析することも可能である.

表示 6.6 2 群として解析したロジスティック回帰分析

Whole Model Test				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	44.64982	2	89.29963	<.0001
Full	79.57168			
Reduced	124.22150			

Parameter Estimates				
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-10.611778	1.59475	44.28	<.0001
G	2.142775	0.44598	23.08	<.0001
x	0.176398	0.02722	41.99	<.0001

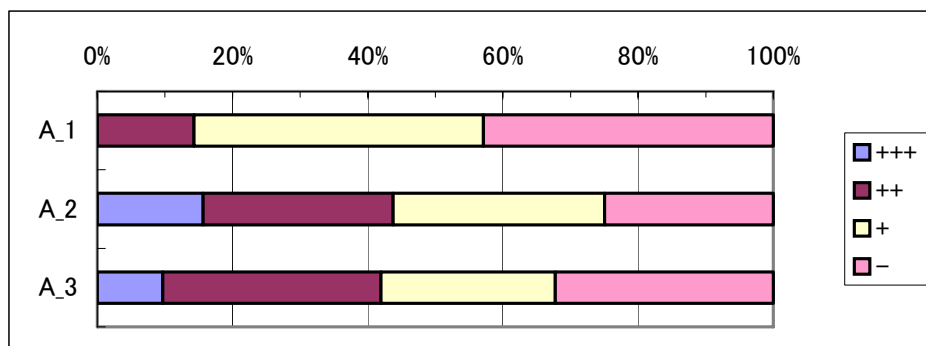
6.2 多カテゴリの反応変数（カテゴリの順序に意味あり）

3 種類の薬を二重盲検比較試験でヒトに投与して、その効果を調べて表示 6.7 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 カテゴリである。カテゴリの順序に意味のある場合には、帯グラフが結果の解釈に役に立つ。

表示 6.7 臨床試験データ

	+++	++	+	-	合計
A ₁	0	4	12	12	28
A ₂	5	9	10	8	32
A ₃	3	10	8	10	31
合計	8	23	30	30	91

表示 6.8 Excel による帯グラフ



このグラフでは、薬による累積比率の違いが明瞭に読み取れる。カテゴリ

の順序に意味のある場合は、帯グラフの3箇所の分割点での比較に意味がある。これは、効果の優れているカテゴリから順番に出現度数を累積し比率を計算したことに対応する。

ロジスティック回帰分析の場合も、反応が順序カテゴリの場合には、累積度数の考えを反映した累積ロジットは、反応が k 水準の場合にJMPでは、式(6.5)で与えられている。

$$z = \ln\left(\frac{P(y \leq i)}{P(y > i)}\right), \quad i = 1, 2, \dots, k \tag{6.5}$$

A_2 水準の場合で累積ロジットの構成は、

$$+++ \text{ vs } (++, +, -): \quad Z_{12} = \ln\left(\frac{5/32}{27/32}\right) = -1.86$$

$$(+++, ++) \text{ vs } (+, -): \quad Z_{22} = \ln\left(\frac{14/32}{18/32}\right) = -0.25$$

$$(+++, ++, +) \text{ vs } (-): \quad Z_{32} = \ln\left(\frac{24/32}{8/32}\right) = -1.10$$

のごとくである。表示 6.9 にJMP / Fit Modelでの解析のためのデータを示す。説明変数として因子 A は a_2, a_3 とダミー変数に展開し、反応変数 y は順序尺度とし、 r を繰返し数として実行する。

表示 6.9 解析用データ

A	y	r	R	n	P	Z^*
1	1:+++	0	0	28	0.00	-4.04
1	2:++	4	4	28	0.14	-1.69
1	3:+	12	16	28	0.57	0.28
1	4:-	12	28	28	1.00	4.04
2	1:+++	5	5	32	0.16	-1.61
2	2:++	9	14	32	0.44	-0.24
2	3:+	10	24	32	0.75	1.06
2	4:-	8	32	32	1.00	4.17
3	1:+++	3	3	31	0.10	-2.10
3	2:++	10	13	31	0.42	-0.32
3	3:+	8	21	31	0.68	0.72
3	4:-	10	31	31	1.00	4.14

表示 6.10 にロジスティック回帰分析の結果を示す。尤度比検定の結果は、因子 A 全体としては、 $\chi^2 = 6.096$ 、これが自由度2の χ^2 分布に従うことから、 $P=0.0475$ と有意である。 A_1 を基準とした場合に、 A_2 との差は、 $P=0.0210$ 、 A_3 との差は、 $P=0.0794$ である。ところで、Parameter Estimatesの結果は何を意味しているのだろうか。3つの切片がそれぞれの累積ロジットに対応していて、 a_2 の推定値は、

A_1 と A_2 の差, a_3 の推定値は, A_1 と A_3 の差となっている. A_3 の $(+++ , ++ , +) vs (-)$ の累積ロジット Z_{33} は,

$$\hat{Z}_{33} = 0.088 + 0.852 = 0.940$$

となる.

表示 6.10 1 因子, 累積ロジットの解析結果

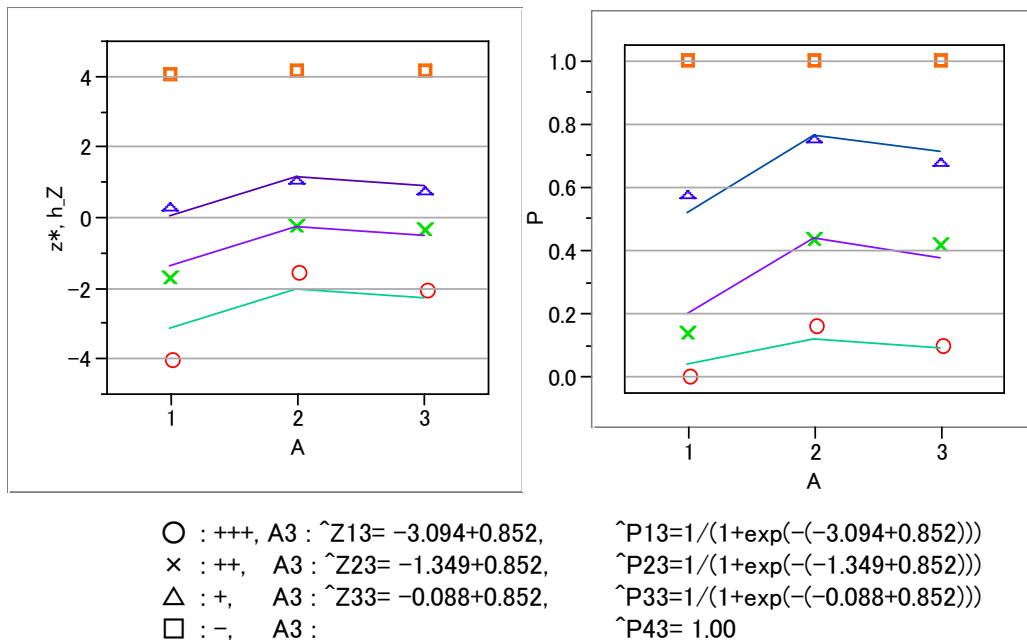
Ordinal Logistic Fit for y, Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	3.04794	2	6.095881	0.0475
Full	114.61653			
Reduced	117.66447			

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept[2:++]	-3.0943319	0.5043942	37.64	<.0001
Intercept[3:+]	-1.3497294	0.3902887	11.96	0.0005
Intercept[4:-]	0.0880739	0.3611391	0.06	0.8073
a2	1.1221722	0.4861639	5.33	0.0210
a3	0.8521236	0.4857452	3.08	0.0794

表示 6.11 逆ロジット, 累積有効率 P のとロジット Z の推定



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、通常のリ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。

6.3 2 因子実験

田口玄一著「実験計画法 第3版 (上)」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元配置地方で法で作った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人づつ112人がランダムに上中下に評価をしたとみなして解析した。その結果を表示6.12に示す。

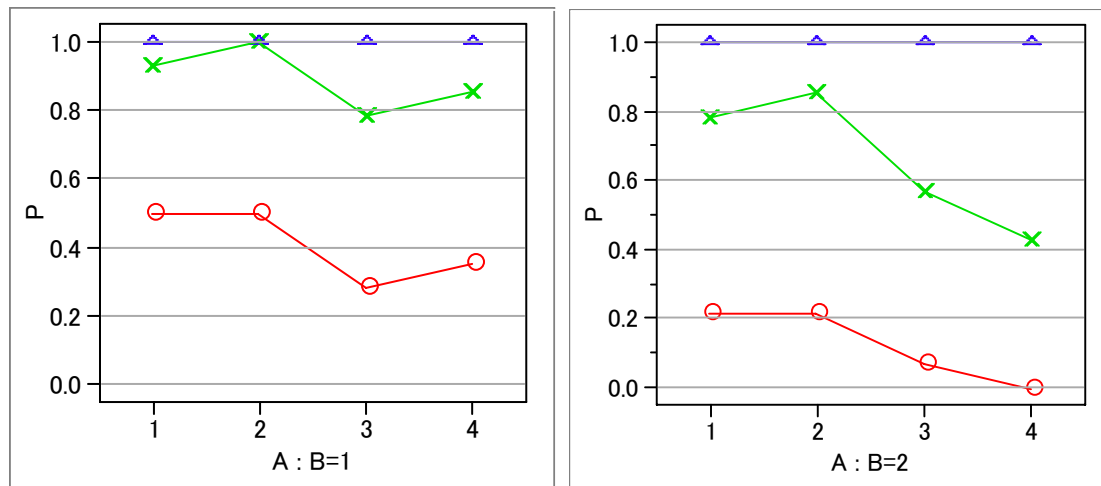
表示 6.12 2 因子実験のデータ

	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

これまでの JMP / Fit Model での解析には、他の統計ソフトのロジスティック回帰分析を利用する場合の配慮と、理解を確実にできるようにとの理由で、質的

変数を取扱うときにダミー変数を用いてきた。2 因子以上の場合には、交互作用

表示 6.13 累積パーセント



○ : 上, × : (上+中)

の検討のためにダミー変数を自ら指定することは煩わしいばかりか間違いの元になりやすい。JMP / Fit Model には、累積ロジットの解析においても自動的にダミー変数を内部で生成してくれる機能が備わっている。ここでは、それを活用する。

解析用のデータは示さないが、前節までと同じマナーで作成する。表示 6.14 に交互作用 $A*B$ の検討結果を示す。p値は 0.6452 と有意ではないので誤差にプールする。

表示 6.14 交互作用の検定結果

Effect Wald Tests				
Source	Nparm	DF	Wald ChiSquare	Prob>ChiSq
B	1	1	15.9963523	0.0001
A	3	3	10.5526337	0.0144
B*A	3	3	1.6628835	0.6452

表示 6.15 の結果から、因子A, 因子B共に高度に有意である。JMP / Fit Model には、推定された回帰式を用いた推定値の計算結果をJMPデータとして書き出してくれるので、それをグラフ化したのが表示 6.16 である。

表示 6.15 主効果モデル

Ordinal Logistic Fit for y Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	13.67420	4	27.3484	<.0001
Full	101.83593			
Reduced	115.51013			

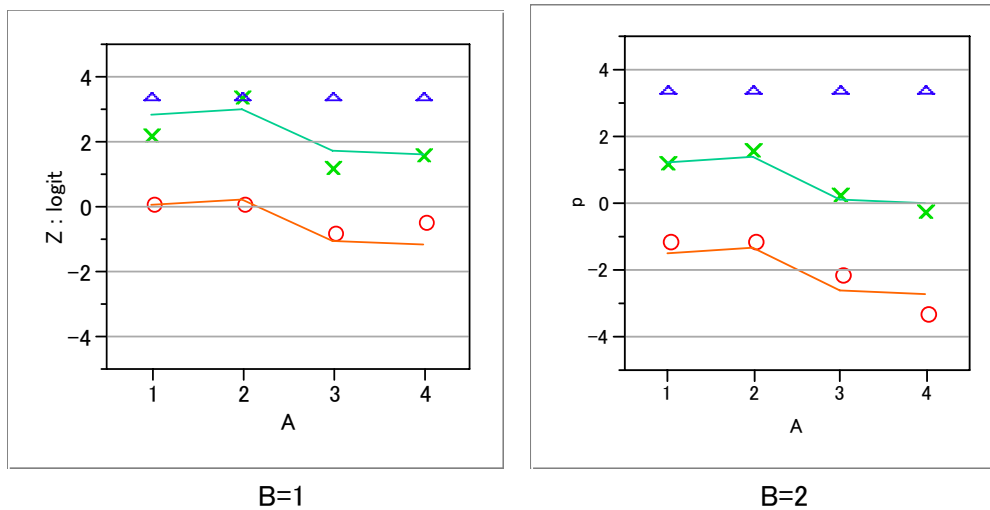
Effect Wald Tests

Source	Nparm	DF	Wald ChiSquare	Prob>ChiSq
B	1	1	15.75601	0.0001
A	3	3	10.47787	0.0149

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept[2]	-1.236728	0.241401	26.25	<.0001
Intercept[3]	1.514974	0.257401	34.64	<.0001
B[1]	0.798035	0.201047	15.76	<.0001
A[1]	0.536346	0.325030	2.72	0.0989
A[2]	0.708652	0.328985	4.64	0.0312
A[3]	-0.566876	0.323807	3.06	0.0800

表示 6.16 効果の推定値



効果の推定は、ロジット軸上で行っている。交互作用をモデルから除いたので4つの推定線は平行している。ロジットは大きい方が望ましい風合いなので、因子Bは、 B_1 が、因子Aは、 A_2 がごくわずかに勝っているが A_1 でも同じ結果とみなせる。

統計ソフトで、ダミー変数を自動的に作成する場合にダミー変数の生成方法

に注意する必要がある。JMP / Fit Model では、最後の水準、因子Aの場合A3であるが $a_1 = -1, a_2 = -1, a_3 = -1$ のようにそれぞれのダミー変数の和が0となるように生成される。これに注意して、Excelによる推定を試みたのが表示 6.17 である。

	a_1	a_2	a_3
A_1	1	0	0
A_2	0	1	0
A_2	0	0	1
A_3	-1	-1	-1

表示 6.17 推定値「上」のロジットの計算，Eexcelによる確認

Intercept[2]	-1.237	B[1]	0.798	A[1]	0.536	計	0.098
Intercept[2]	-1.237	B[1]	0.798	A[2]	0.709	計	0.270
Intercept[2]	-1.237	B[1]	0.798	A[3]	-0.567	計	-1.006
Intercept[2]	-1.237	B[1]	0.798	-A[1]-A[2]-A[3]	-0.678	計	-1.117
Intercept[2]	-1.237	-B[1]	-0.798	A[1]	0.536	計	-1.498
Intercept[2]	-1.237	-B[1]	-0.798	A[2]	0.709	計	-1.326
Intercept[2]	-1.237	-B[1]	-0.798	A[3]	-0.567	計	-2.602
Intercept[2]	-1.237	-B[1]	-0.798	-A[1]-A[2]-A[3]	-0.678	計	-2.713

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 6.18 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分析を利用すると良いであろう。

表示 6.18 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