

# IBM SPSS Regression 20



注：この情報とサポートされている製品をご使用になる前に、「注意事項」（p. 48）の一般情報をお読みください。

本版は IBM® SPSS® Statistics 20 , および新版で指示されるまで後続するすべてのリリースおよび変更に対して適用されます。

Adobe 製品のスクリーンショットは Adobe Systems Incorporated の許可を得て転載しています。

Microsoft 製品のスクリーンショットは Microsoft 社の許可を得て転載しています。

Licensed Materials - Property of IBM

© Copyright IBM Corporation 1989, 2011.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# はじめに

IBM® SPSS® Statistics は、データ分析の包括的システムです。Regression は、このマニュアルで説明されている追加の分析手法を提供するオプションのアドオン モジュールです。Regression アドオン モジュールは SPSS Statistics Core システムと組み合わせて使用し、Core システムに完全に統合されます。

## IBM Business Analytics について

IBM Business Analytics ソフトウェアは、意思決定者がビジネス パフォーマンスを向上させるために信頼する完全で、一貫した正確な情報を提供します。ビジネス インテリジェンス、予測分析、財務実績および戦略管理、および 分析アプリケーションの包括的なポートフォリオを利用することによって、現在の実績を明確、迅速に理解し、将来の結果を予測することができます。豊富な業界のソリューション、実績ある実例、専門サービスと組み合わせ、さまざまな規模の組織が、高い生産性を実現、意思決定を自信を持って自動化し、より良い決定をもたらします。

このポートフォリオの一部として、IBM SPSS Predictive Analytics ソフトウェアを使用する組織は、将来のイベントを予測し、その洞察に基づいて積極的に行動し、より優れた業績を実現することができます。全世界の企業、政府、学術分野のお客様が IBM SPSS の技術を活用し、不正行為を減少させ、リスクを軽減させながら、顧客の獲得、保持、成長において、競争優位を高めることができます。IBM SPSS ソフトウェアを日々の業務に取り入れることによって、組織は業務目標を達成し、大きな競争的優位を獲得することができるよう、意思決定を方向付け、自動化することができるようになります。お問い合わせは、<http://www.ibm.com/spss> を参照してください。

## テクニカル サポート

テクニカル サポートのサービスをご利用いただけます。IBM Corp. 製品の使用方法や、対応しているハードウェア環境へのインストールに関して問い合わせることもできます。テクニカル サポートの詳細については、IBM Corp. Web サイト (<http://www.ibm.com/support>) を参照してください。連絡の際は、所属団体名、サポート契約などを確認できるよう、あらかじめ手元にご用意ください。

## 学生向けテクニカル サポート

IBM SPSS ソフトウェア製品の Student 版、アカデミック版、Grad パック版を使用している学生の場合、学生用の特別オンライン ページ、[Solutions for Education \(http://www.ibm.com/spss/rd/students/\)](http://www.ibm.com/spss/rd/students/) ページを参照してください。大学提供の IBM SPSS ソフトウェアのコピーを使用している場合、大学の IBM SPSS 製品コーディネータにお問い合わせください。

## カスタマ サービス

配送やアカウントに関するご質問は、お近くの営業所にお問い合わせください。お問い合わせの際には、シリアル番号をご用意ください。

## トレーニング セミナー

IBM Corp. では一般公開およびオンサイトで トレーニング セミナーを実施しています。セミナーでは実践的な講習を行います。セミナーは主要都市で定期的開催されます。セミナーに関する詳細については、<http://www.ibm.com/software/analytics/spss/training> を参照してください。

## 追加の出版物

Marija Noruš による『SPSS Statistics: Guide to Data Analysis』、『SPSS Statistics: Statistical Procedures Companion』、『SPSS Statistics: Advanced Statistical Procedures Companion』が Prentice Hall から出版されました。補助的な資料としてご利用いただけます。これらの出版物には、SPSS Statistics Base モジュール、Advanced Statistics モジュール、Regression モジュールの統計的手続きについて記載されています。初めてデータ分析を行う場合、高度なアプリケーションを使用する場合に応じて、この本は IBM® SPSS® Statistics が提供している機能を効率よく使用するための手助けとなります。出版物の内容、サンプルの図表などの詳細は、作者の Web サイトを参照してください。  
<http://www.norusis.com>

---

# 内容

<b>1</b>	<b>2 項ロジスティック回帰の手続きの選択</b>	<b>1</b>
<b>2</b>	<b>ロジスティック回帰</b>	<b>3</b>
	ロジスティック回帰の規則の設定 . . . . .	5
	ロジスティック回帰の変数選択方法 . . . . .	5
	ロジスティック回帰のカテゴリ変数の定義 . . . . .	6
	ロジスティック回帰の新変数の保存 . . . . .	7
	ロジスティック回帰のオプション . . . . .	9
	LOGISTIC REGRESSION コマンドの追加機能 . . . . .	10
<b>3</b>	<b>多項ロジスティック回帰</b>	<b>11</b>
	多項ロジスティック回帰 . . . . .	13
	項の構築 . . . . .	14
	多項ロジスティック回帰の参照カテゴリ . . . . .	15
	多項ロジスティック回帰統計量 . . . . .	16
	多項ロジスティック回帰の収束基準 . . . . .	18
	多項ロジスティック回帰のオプション . . . . .	19
	多項ロジスティック回帰の保存 . . . . .	20
	NOMREG コマンドの追加機能 . . . . .	21
<b>4</b>	<b>プロビット分析</b>	<b>22</b>
	プロビット分析の範囲の定義 . . . . .	24
	プロビット分析のオプション . . . . .	24
	PROBIT コマンドの追加機能 . . . . .	25
<b>5</b>	<b>非線型回帰分析</b>	<b>26</b>
	条件論理 (非線型回帰分析) . . . . .	28
	非線型回帰分析のパラメータ . . . . .	28

非線型回帰分析の一般モデル	29
非線型回帰分析の損失関数	30
非線型回帰分析のパラメータの制約	31
非線型回帰分析の新変数の保存	32
非線型回帰分析のオプション	33
非線型回帰分析の結果の解釈	34
NLR コマンドの追加機能	34

## 6 重み付け推定 36

重み付け推定のオプション	38
WLS コマンドの追加機能	38

## 7 2 段階最小 2 乗回帰 39

2 段階最小 2 乗回帰のオプション	41
2SLS コマンドの追加機能	41

## 付録

### A カテゴリ変数コード化方式 42

全平均	42
[単純]	43
Helmert	44
差分 (階差)	44
多項式	45
[反復測定]	45
特別 (ピボットテーブル GLM)	46
指標	47

<b>B 注意事項</b>	<b>48</b>
<b>索引</b>	<b>51</b>



## 2 項ロジスティック回帰の手続きの選択

ロジスティック回帰手続きまたは多項ロジスティック回帰手続きのいずれかを使用して、2 項ロジスティック回帰モデルの当てはめを行うことができます。それぞれの手続きには、他の手続きでは使用できないオプションが用意されています。重要な理論上の違いは、ロジスティック回帰手続きでは、データの入力方法に関わらず、また共変量パターンの数がケースの総数よりも少ないかどうかに関わらず、個別のケース レベルのデータを使用してすべての予測、残差、影響力の統計、および適合度検定を出力しますが、多項ロジスティック回帰手続きでは、ケースを内部で集計し、予測変数に対して同一の共変量パターンを持つ部分母集団を構成し、これらの部分母集団に基づいて予測、残差、および適合度検定を出力するという点です。すべての予測変数がカテゴリ変数であるか、または量的予測変数が限られた数の値しか取らない場合、また、そのためにそれぞれの異なる共変量パターンに複数のケースがある場合は、部分母集団アプローチにより有効な適合度検定が得られますが、個別のケース レベル アプローチでは得られません。

**ロジスティック回帰**には、次の独自の機能があります。

- モデルに対する Hosmer-Lemeshow の適合度の検定
- ステップワイズ分析
- モデル パラメータ化を定義する対比
- 分類のための代替分割点
- 分類プロット
- 1 つのケース グループ上でホールドアウト ケース グループに当てはめられたモデル
- 予測、残差、および影響力の統計の保存

**多項ロジスティック回帰**には、次の独自の機能があります。

- モデルの適合度に対する Pearson と逸脱カイ 2 乗検定
- 適合度の検定のためデータをグループ化する部分母集団の指定
- 部分母集団別の度数、予測度数、および残差のリスト表示
- 過剰な散らばりに対する分散推定値の修正
- パラメータ推定値の分散共分散行列
- パラメータの線型結合の検定

- 入れ子になっているモデルの明示的指定
- 差分変数を使用した、1 対 1 一致条件付きロジスティック回帰モデルの当てはめ

# ロジスティック回帰

ロジスティック回帰は、1 組の予測変数の値に基づく特性または結果の真偽の予測に役立ちます。これは線型回帰モデルに似ていますが、従属変数が 2 分変数のモデルに適しています。ロジスティック回帰係数は、モデル中の各独立変数のオッズ比の推定に使用できます。ロジスティック回帰は判別分析よりも広い範囲の調査に適用できます。

**例:**どのような生活様式の特性が、冠状動脈性心臓病 (CHD) の危険因子になるのでしょうか。喫煙状況、食生活、運動、飲酒、および CHD の状態を測定した患者のサンプルがある場合に、4 つの生活様式を変数として使用するモデルを作成し、患者のサンプルについて CHD の有無を予測できます。次に、モデルを使用して、各因子のオッズ比の推定値を求め、たとえば、喫煙者は非喫煙者と比べて、CHD になる確率がどれくらい高いかなどを示すことができます。

**統計量。** 各分析: ケースの総数、選択したケース、有効なケース。各カテゴリ変数: パラメータ コーディング。各ステップ: 投入または除去される変数、反復の記述、-2 対数尤度、適合度、Hosmer-Lemeshow の適合度の統計量、カイ 2 乗モデル、改良カイ 2 乗、分類表、変数の相関、観測グループと予測確率図、残差のカイ 2 乗。等式の各変数: 係数 (B)、B の標準誤差、特典統計量、推定オッズ比 ( $\exp(B)$ )、 $\exp(B)$  の信頼区間、モデルから項が削除された場合は対数尤度。等式にない各変数: スコア統計量。各ケース: 観測グループ、予測確率、予測されたグループ、標準化残差。

**方法。** 変数のブロック エントリ、またはステップワイズ方法、条件付き変数増加法、変数増加法 LR、変数増加法 Wald、条件付き変数減少法、変数減少法 LR、または変数減少法 Wald のいずれかを使用してモデルを推定できます。

**データ。** 従属変数は 2 分変数でなければなりません。独立変数は間隔尺度変数またはカテゴリ変数でなければなりません。カテゴリ変数の場合には、ダミーまたは指標コードの変数でなければなりません (この手続きのオプションで、カテゴリ変数を自動再割り当てにすることができます)。

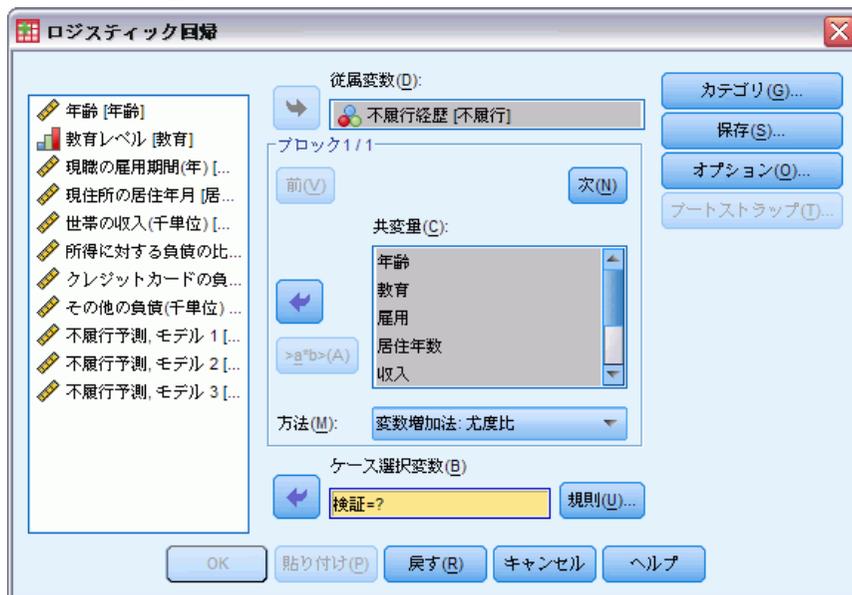
**仮定。** ロジスティック回帰では、分布の仮定を判別分析とは異なる形で利用します。しかし、予測変数が多変量正規分布の場合には、より安定した解が得られます。また、別の回帰形式を使用した場合と同様に、予測変数間の多重共線性により推定値が偏ったり、標準誤差が膨張することがあります。ロジスティック回帰手続きは、所属グループが正確なカテゴリ変数で構成されている場合に最も効果的です。所属グループが連続変数 (たとえば、「高い IQ」対「低い IQ」) の値に基づいている場合には、連続変数自体の豊富な情報量を利用するため、線型回帰を使用することをお勧めします。

**関連手続き。**データの多重共線性の検査には、散布図手続きを使用します。多変量の正規性と等分散共分散行列の仮定が満たされる場合、判別分析手続きを使用すれば、より迅速に解を求めることができます。すべての予測変数がカテゴリ変数である場合には、対数線型手続きも使用できます。従属変数が連続型の場合には、線型回帰手続きを使用します。ROC 曲線手続きを使用して、ロジスティック回帰手続きで保存した確率を作図できます。

### ロジスティック回帰分析を行うには

- ▶ メニューから次の項目を選択します。  
分析(A) > Regression > 二項ロジスティック...

図 2-1  
[ロジスティック回帰] ダイアログ ボックス



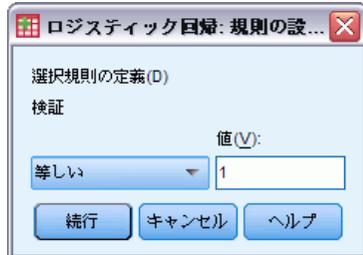
- ▶ 2 分従属変数を 1 つ選択します。この変数は、数値型または文字型のどちらでもかまいません。
- ▶ 1 つ以上の共変量を選択します。交互作用項を含めるには、交互作用に含めるすべての変数を選択し、次に [ $>a*b>$ ] をクリックします。

変数をグループ ([**ブロック**]) に入力するには、ブロックごとに共変量を選択し、[次] をクリックして新しいブロックを指定します。同様に、すべてのブロックを指定します。

オプションとして、分析するケースを選択できます。ケース選択変数を選択し、[規則の設定] をクリックします。

## ロジスティック回帰の規則の設定

図 2-2  
[ロジスティック回帰: 規則の設定] ダイアログ ボックス



選択規則で定義したケースがモデル推定に含まれます。たとえば、1 つの変数と [= (等しい)] を選択し、[値] ボックスで 5 を指定すると、選択された変数の値が 5 のケースだけがモデルの推定で使用されます。

統計および分類の結果は、選択されたケースと選択されていないケースの両方に対して生成されます。この仕組みにより、既存のデータに基づいて新規ケースを分類したり、データを学習用および検証用サブセットに分割して、生成されたモデルの有効性を確認できます。

## ロジスティック回帰の変数選択方法

[方法] ボックスでの選択により、独立変数を分析に投入する方法を指定できます。変数は同じ組み合わせのまま方法を変えることで、さまざまな種類の回帰モデルを求めることができます。

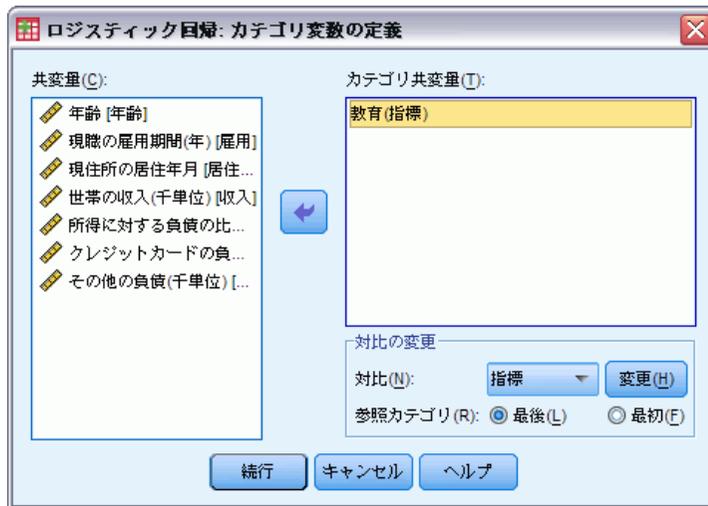
- **Enter.** 変数選択のための手順で、ブロック内のすべての変数は、1 つのステップで投入されます。
- **変数増加法 (条件付).** スコア統計量の有意差に基づくステップワイズ変数増加法。条件付きパラメータ推定値を基本とする尤度比統計量の確率に基づき除去します。
- **変数増加法 (尤度比).** スコア統計量の有意差に基づくステップワイズ変数増加法。最大偏尤度推定値による尤度比統計量の確率に基づき除去を行います。
- **変数増加法 (Wald).** スコア統計量の有意差に基づくステップワイズ変数増加法。Wald 統計量の確率に基づき除去を行います。
- **変数減少法 (条件付).** 変数減少法のステップごとの選択。条件付きパラメータ推定値に基づく尤度比統計量の確率に基づいて、削除検定を行ないます。
- **変数減少法 (尤度比).** 変数減少法のステップごとの選択。最大偏尤度推定値に基づく尤度比統計量の確率に基づいて、削除検定を行ないます。
- **変数減少法 (Wald).** 変数減少法のステップごとの選択。Wald 統計量の確率に基づいて、削除検定を行ないます。

出力内の有意値は、1 つのモデルの適合に基づいたものです。したがって、ステップワイズ法を使用したときには、一般に、有意値は無効になります。

選択されたすべての独立変数が、1 つの回帰モデルに追加されます。しかし、変数の組み合わせを変えて、変数の投入方法を指定することもできます。たとえば、1 番目の変数のブロックをステップワイズ法で回帰モデルに投入し、2 番目のブロックを変数増加法で投入することもできます。2 番目のブロックを回帰モデルに追加するには、[次へ] をクリックします。

## ロジスティック回帰のカテゴリ変数の定義

図 2-3  
[ロジスティック回帰: カテゴリ変数の定義] ダイアログ ボックス



ロジスティック回帰手続きで、カテゴリ変数の処理方法の詳細を指定できます。

**共変量。**メイン ダイアログ ボックスで指定した単独または交互作用の一部としての共変量が、層を問わず、すべて一覧になっています。これらの共変量の中に文字型変数またはカテゴリ変数が含まれている場合は、カテゴリ共変量としてだけ使用できます。

**カテゴリ共変量。**カテゴリ変数として識別されている変数の一覧を表示します。各変数のかっこ内には、使用する対比コードが表示されています。文字型変数（変数名の後に記号 < 付き）は、[カテゴリ共変量] ボックスの一覧に前もって表示されています。[共変量] リストからその他のカテゴリ共変量を選び、それを [カテゴリ共変量] リストに移動します。

**対比の変更。**対比方法は変更できます。次の対比方法を使用できます。

- **指標。**対比は所属カテゴリの有無を示します。[参照カテゴリ] で指定したカテゴリは、ゼロの行として対比行列内に表示されます。

- **単純。** 予測変数（[参照カテゴリ] で指定したカテゴリを除く）の各カテゴリを [参照カテゴリ] で指定したカテゴリと比べます。
- **差分。** 最初のカテゴリを除く予測変数の各カテゴリを、それ以前のカテゴリの平均的な結果と比べます。逆 Helmert 対比とも呼ばれています。
- **Helmert。** 最後のカテゴリを除いた予測変数の各カテゴリを、それ以後のカテゴリの平均的な結果と比べます。
- **反復測定。** 最初のカテゴリを除いた予測変数の各カテゴリを、直前のカテゴリと比べます。
- **多項式。** 直交多項式の対比。カテゴリは、一定の間隔を等しく空けて配置されていると仮定されます。多項式の対比は、数値型変数に限り利用できます。
- **偏差。** [参照カテゴリ] で指定したカテゴリを除く予測変数の各カテゴリを、全体的な結果と比べます。

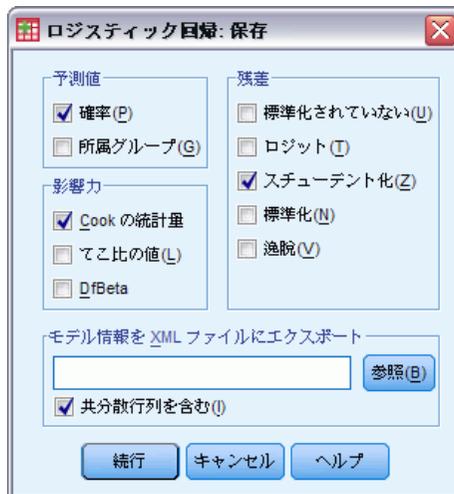
[偏差]、[単純]、または [指標] を選択する場合は、[参照カテゴリ] として、[最初] または [最後] を選択します。[変更] をクリックして初めて、モデルが実際に変更されることに注意してください。

文字型共変量はカテゴリ共変量でなければなりません。[カテゴリ共変量] リストから文字型変数を削除するには、メイン ダイアログ ボックスの [共変量] リストからその変数を含むすべての項を削除する必要があります。

## ロジスティック回帰の新変数の保存

図 2-4

[ロジスティック回帰: 新変数の保存] ダイアログ ボックス



ロジスティック回帰の結果を、新変数としてアクティブ データセットに保存できます。

**予測値。**モデルにより予測した値を保存します。使用できるオプションは [確率] と [所属グループ] です。

- **確率 (ロジスティック回帰 保存).** 各ケースに対して、イベント発生の予測確率を保存します。出力のテーブルは、新しい変数の名前と内容を表示します。「イベント」は値が高い従属変数のカテゴリです。たとえば、従属変数が 0 および 1 である場合、カテゴリ 1 の予測確率が保存されます。
- **予測された所属グループ (判別分析).** 判別得点に基づいた、最大事後確率のあるグループ。所属するケースをモデルが推定するグループ。

**影響力。**ケースの予測値への影響力を測定する統計の値を保存します。使用できるオプションは [Cook の統計量]、[てこ比の値]、および [DfBeta] です。

- **Cook の統計量.** Cook の影響統計量のロジスティック回帰の類似値。特定のケースが回帰係数の計算から除外された場合に、すべてのケースの残差がどのくらい変化するかを示す測定量。
- **てこ比の値.** モデルの適合度に関する各観測の相対的な影響度。
- **DfBeta(s).** ベータ値の差は、特定のケースの除外から発生する回帰係数の変化です。値は、モデル内の各項に対して定数項を含めて計算されます。

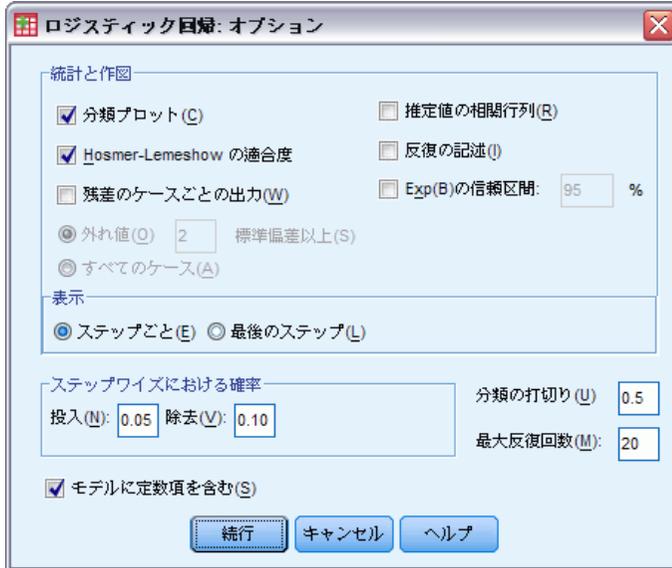
**残差。**残差を保存します。使用できるオプションは [標準化されていない]、[ロジット]、[スチューデント化]、[標準化]、および [逸脱] です。

- **標準化されていない残差.** 観測値およびモデルで予測された値との差。
- **ロジット残差.** ロジット尺度で予測された場合のケースの残差。ロジット残差は、残差を予測される確率で割り、それに 1 から予測される確率を引いた値を掛けたものです。
- **スチューデント化された残差 (LOGREG).** あるケースがモデルから除外された時の、モデルの逸脱の変化量です。
- **標準化残差 (一般的な対数線型分析).** 残差を標準偏差の推定値で割った値。標準化残差は Pearson 残差とも呼ばれ、平均は 0 で、標準偏差 1 になります。
- **逸脱.** モデルの逸脱に基づいた残差。

**モデル情報を XML ファイルにエクスポート。** パラメータ推定とその共分散 (オプション) は指定されたファイルに XML (PMML) 形式でエクスポートされます。このモデル ファイルを使用して、得点付けのために他のデータ ファイルにモデル情報を適用できます。

## ロジスティック回帰のオプション

図 2-5  
[ロジスティック回帰:オプション] ダイアログ ボックス



ロジスティック回帰分析では、次のオプションを指定できます。

**統計と作図。**統計と作図を求めることができます。使用できるオプションは [分類プロット]、[Hosmer-Lemeshow の適合度]、[残差のケースごとの出力]、[推定値の相関行列]、[反復の記述]、および [Exp(B) の信頼区間] です。[表示] の [ステップごと] または最後のモデルのみを表示する [最後のステップ] のいずれかをクリックし、統計と作図を表示します。

- **Hosmer-Lemeshow の適合度統計。**この適合度統計は、特に連続共変量を持つモデルおよび、小さいサンプル サイズを持つ検定に対して、ロジスティック回帰で使用される従来の適合度統計量よりも頑健です。これはケースを危険の十分位数にグループ化することに基づき、各十分位数内で観測確率と期待確率を比較します。

**ステップワイズにおける確率。**変数を式に入力する、または削除する基準を制御できます。[投入] または [除去] の各ボックスに変数の基準を指定します。

- **ステップワイズにおける確率 (ロジスティック回帰 オプション)。**得点統計量の確率が [投入] 値よりも小さい場合、変数はモデルに投入され、確率が [除去] 値よりも大きい場合、変数は除去されます。デフォルト設定を上書きするには、[投入] と [除去] に正の整数を入力します。[投入] 値は [除去] 値よりも小さくしなければなりません。

**分類の打切り。**ケースを分類する分割点を決めることができます。予測値が分類分割点を超えるケースは正に分類され、分割点より小さい予測値を持つケースは負に分類されます。デフォルトを変更するには、0.01 から 0.99 までの値を入力します。

**最大反復回数。**終了のモデルの反復回数の最大値を変更します。

**モデルに定数項を含む。**モデルに定数項を含めるかどうかを指定します。チェック ボックスをオフにした場合は、定数項が 0 になります。

## LOGISTIC REGRESSION コマンドの追加機能

コマンド シンタックス言語を使用して、次のことも実行できます。

- ケースごとの出力を識別するため、変数ラベルや値ラベルを使用。
- 反復報告書の間隔を制御。出力パラメータで各反復後に推定する代わりに、n 番目の反復の後で常に推定するように指定できます。
- 反復の終了と剰余のチェック基準の変更。
- ケースごとの出力で変数一覧を指定。
- 処理中、分割ファイルの各グループのデータを外部作業ファイルで保持することによるメモリー節約。

複雑なシンタックス情報については、「コマンド シンタックス リファレンス」を参照してください。

## 多項ロジスティック回帰

多項ロジスティック回帰は、1組の予測変数の値に基づく被験者の分類に役立ちます。この種類の回帰はロジスティック回帰に似ていますが、従属変数が2つのカテゴリに制限されていないので、より一般的と言えます。

**例:** 映画のマーケティングをより効果的に行うために、映画会社では、映画ファンが見そうな映画の種類を予測したいと考えています。映画会社は、多項ロジスティック回帰を実行することにより、客の年齢、性別、および同伴者の状況が、彼らが好む映画の種類に及ぼす影響の大きさについて調べることができます。こうして、映画会社は、特定の映画に対する宣伝キャンペーンを、その映画を見に来ると思われるグループに合わせるすることができます。

**統計量。** 反復の記述、パラメータ係数、漸近分散共分散行列、モデルおよび部分効果の尤度比検定、 $-2$  対数尤度、Pearson および逸脱のカイ 2 乗適合度、Cox と Snell、Nagelkerke、および McFadden  $R^2$ 。分類: 応答カテゴリ別の観測度数対予測度数。クロス表: 共変量パターンおよび応答カテゴリ別の観測度数、予測度数 (残差あり)、および比率。

**方法。** 多項ロジット モデルは、すべての因子によるモデル、またはユーザー指定モデルに適しています。パラメータ推定は、反復最尤法アルゴリズムを使用して実行されます。

**データ。** 従属変数は、カテゴリ変数でなければなりません。独立変数は、因子または共変量です。一般に、因子はカテゴリ変数でなければならず、共変量は量的変数でなければなりません。

**仮定。** 任意の2つのカテゴリのオッズ比が他のすべての応答カテゴリから独立していると仮定しています。たとえば、新製品を市場に投入する場合、この仮定では、他のすべての製品の市場占有率は同じ比率で影響を受けるとしています。さらに、共変量パターンが与えられたとすると、応答は独立多項分布変数であると仮定されます。

### 多項ロジスティック回帰分析を行うには

- ▶ メニューから次の項目を選択します。  
分析(A) > 回帰 > 多項ロジスティック...

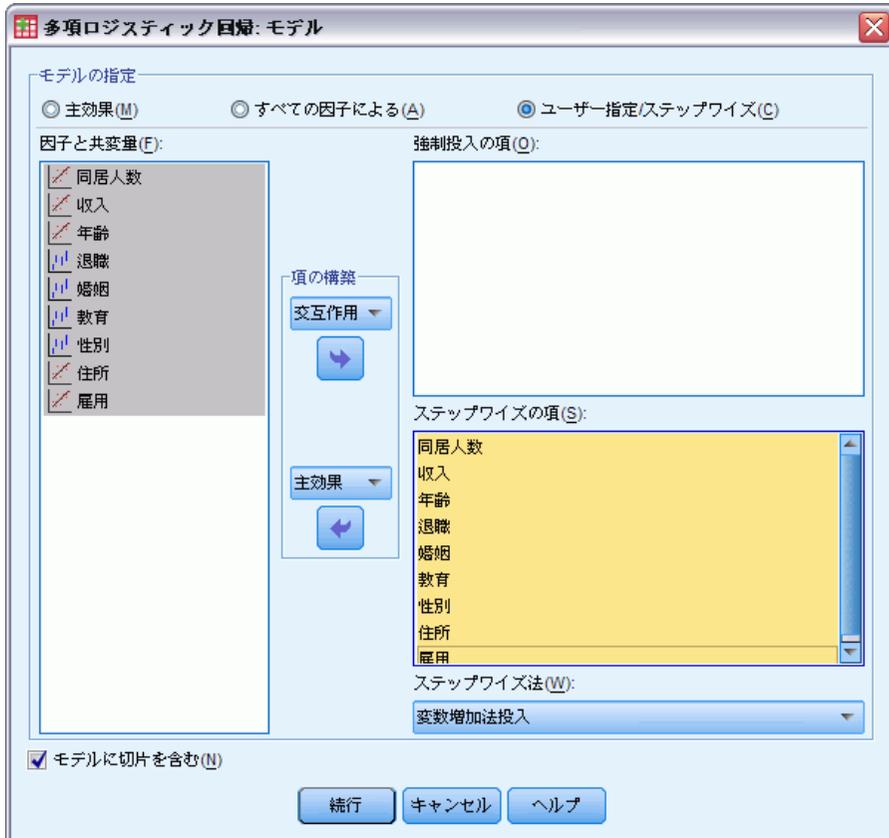
図 3-1  
[多項ロジスティック回帰] ダイアログ ボックス



- ▶ 従属変数を 1 つ選択します。
- ▶ [因子] はオプションであり、数値変数またはカテゴリ変数のいずれかです。
- ▶ [共変量] はオプションですが、指定した場合は、数値変数でなければなりません。

## 多項ロジスティック回帰

図 3-2  
[多項ロジスティック回帰: モデル] ダイアログ ボックス



デフォルトでは、多項ロジスティック回帰手続きは因子および共変量主効果を含むモデルを作成しますが、このダイアログ ボックスでユーザーの指定によるモデルやステップワイズ モデルを選択できます。

**モデルの指定。** 主効果モデルには、共変量および因子主効果が含まれますが、交互作用効果は含まれません。すべての因子によるモデルには、すべての主効果とすべての因子ごとの交互作用が含まれます。共変量の交互作用は含まれません。ユーザーの指定によるモデルを作成して、因子交互作用または共変量交互作用のサブセットを指定したり、モデル項のステップ ワイズ法を要求できます。

**[因子&と共変量]**要素および共変量はリストされます。

**[強制投入の項]** 強制投入リストに追加された項は、常にモデルに含まれます。

**[ステップワイズの項]** ステップワイズ リストに追加された項は、ユーザーが選択した次のいずれかの方法によりモデルに含まれます。

- **[変数増加法投入]** この方法は、モデル内にステップワイズの項が含まれていない状態で開始します。各ステップにおいて、モデルから除去されたステップワイズの項がモデルに追加されても統計的に有意な寄与率を持たなくなるまで、最大有意の項がモデルに追加されます。
- **[変数減少法]** この方法は、ステップワイズ リストの指定された項をすべてモデルに追加することにより開始します。各ステップでは、残りのすべてのステップワイズの項がモデルに対して統計的に有意な寄与率を持つまで、最小有意のステップワイズの項がモデルから除去されます。
- **変数増加ステップワイズ法:** この方法では、変数増加法投入法で選択されるモデルから開始します。アルゴリズムは、モデル内のステップワイズの項に対する変数減少法と、モデルから除去された項に対する変数増加法投入との間で変化します。投入または除去の基準を満たす項がなくなるまで続きます。
- **[変数減少ステップワイズ法]** この方法では、変数減少法により選択されるモデルから開始します。アルゴリズムは、モデルから除去された項に対する変数増加法投入と、モデル内のステップワイズの項に対する変数減少法との間で変化します。投入または除去の基準を満たす項がなくなるまで続きます。

**[モデルに定数項を含む]** モデルに定数項を含めたり、除外したりできます。

## 項の構築

選択した因子や共変量について、次の項を作成できます。

**[交互作用]** 選択した変数のすべてについて、最高水準の交互作用項を作成します。

**[主効果]** 選択した変数のそれぞれに主効果の項目を作成します。

**[2 次まで]** 選択した変数に 2 次までの交互作用を作成します。

**[3 次まで]** 選択した変数に 3 次までの交互作用を作成します。

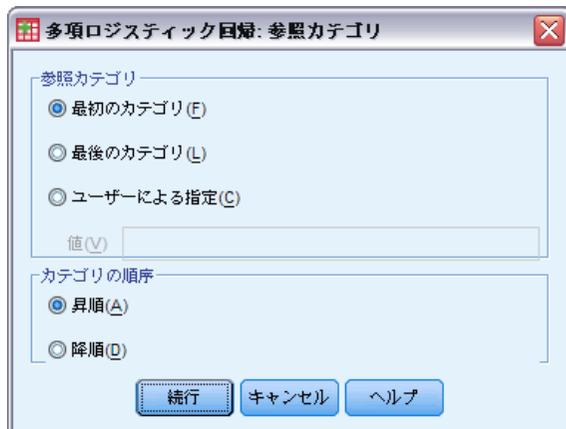
**[4 次まで]** 選択した変数に 4 次までの交互作用を作成します。

**[5 次まで]** 選択した変数に 5 次までの交互作用を作成します。

## 多項ロジスティック回帰の参照カテゴリ

図 3-3

[多項ロジスティック回帰: 参照カテゴリ] ダイアログ ボックス



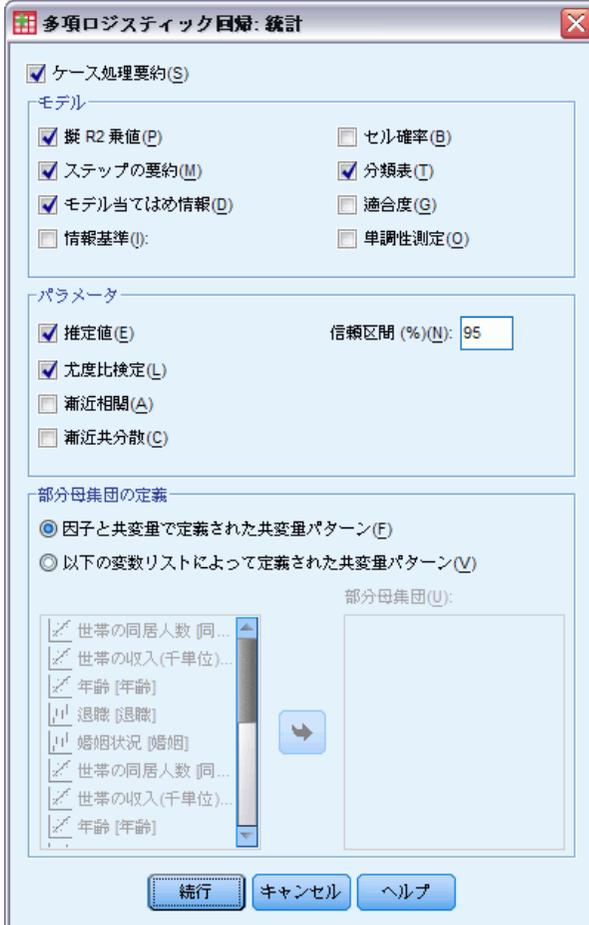
デフォルトでは、多項ロジスティック回帰手続きにより、最後のカテゴリが参照カテゴリになります。このダイアログ ボックスでは、参照カテゴリおよびカテゴリの並び順を制御できます。

**[参照カテゴリ]** 最初のカテゴリ、最後のカテゴリ、またはユーザー指定のカテゴリを指定します。

**[カテゴリ順序]** 昇順では、最低値により最初のカテゴリが定義され、最高値により最後のカテゴリが定義されます。降順では、最高値により最初のカテゴリが定義され、最低値により最後のカテゴリが定義されます。

## 多項ロジスティック回帰統計量

図 3-4  
[多項ロジスティック回帰: 統計] ダイアログ ボックス



多項ロジスティック回帰に対して、次の統計量を指定できます。

**[ケース処理要約]** このテーブルには、指定したカテゴリ変数の情報が含まれています。

**[モデル]** モデル全体の統計量です。

- **擬 R2 乗値。** Cox と Snell、Nagelkerke、および McFadden  $R^2$  統計量 を出力します。
- **[ステップの要約]** このテーブルは、ステップワイズ法により各ステップで投入または除去された効果を集計します。**[モデル]** ダイアログ ボックスでステップワイズ モデルが指定されていない場合は作成されません。

- **[モデル当てはめ情報]** このテーブルは、当てはめられたモデルと、定数項のみのモデルまたは NULL モデルを比較します。
- **[情報量基準]** このテーブルでは、'赤池情報量基準 (AIC) および Schwarz' のベイズ基準 (BIC) を出力します。
- **セルの確率。** 共変量パターンおよび応答カテゴリ別の観測確率、期待確率 (残差あり)、および比率のテーブルを出力します。
- **分類表。** 観測された応答と予測された応答の表を出力します。
- **適合度カイ 2 乗統計量。** Pearson および尤度比カイ 2 乗統計量を出力します。すべての因子および共変量、または因子および共変量のユーザー定義サブセットで決定される共変量パターンについて統計量を計算します。
- **単調性測定方法。** 一致するペア、一致しないペア、および同順位のペアの数を表すテーブルを表示します。このテーブルには、Somers の D、Goodman と Kruskal のガンマ、Kendall のタウ a、および一致指数 C も表示されます。

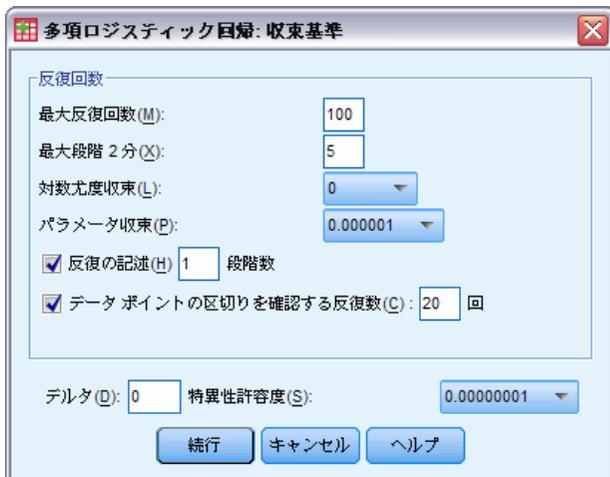
**パラメータ モデル** パラメータに関連する統計量です。

- **推定値。** ユーザー指定の信頼係数で、モデル パラメータの推定値を出力します。
- **尤度比検定。** モデルの部分効果の尤度比検定を出力します。モデル全体の検定が自動的に出力されます。
- **漸近相関。** パラメータ推定相関係数の行列を出力します。
- **漸近共分散。** パラメータ推定分散共分散の行列を出力します。

**部分母集団の定義。** セル確率および適合度検定で使用する共変量パターンを定義するために、因子と共変量のサブセットを選択できます。

## 多項ロジスティック回帰の収束基準

図 3-5  
[多項ロジスティック回帰: 収束基準] ダイアログ ボックス



多項ロジスティック回帰に対して、次の基準を指定できます。

**[反復回数]** アルゴリズムを繰り返し実行する最大回数、段階 2 分の最大段階数、対数尤度とパラメータの変更に対する収束許容度、反復アルゴリズムの進行状況を出力する頻度、手続きでデータの完全な分離またはほぼ完全な分離のチェックを開始する必要がある反復を指定できます。

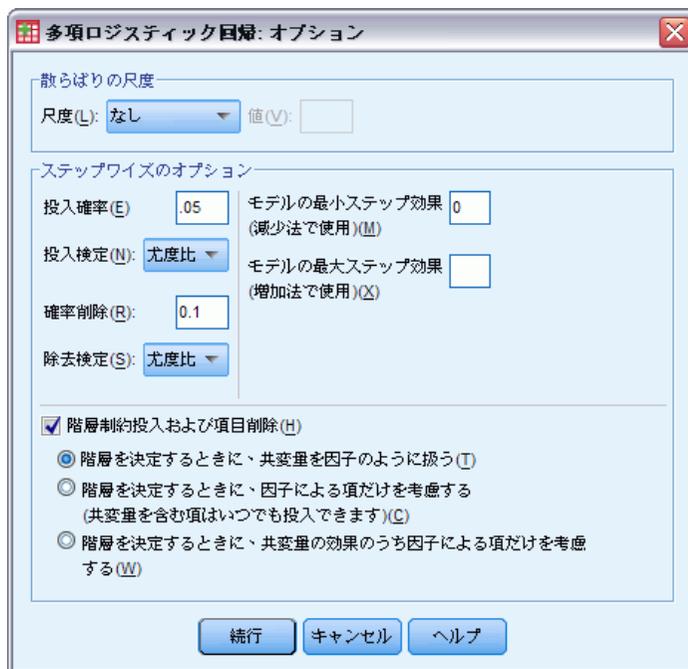
- **対数尤度収束。** 対数尤度関数における絶対変化が指定された値より小さくなると、収束したと考えられます。値が 0 の場合、この基準は使用されません。負でない値を指定してください。
- **[パラメータ収束]** パラメータ推定値における絶対変化がこの値より小さい場合に収束したと考えられます。値が 0 の場合、この基準は使用されません。

**デルタ。** 1 よりも小さい負でない値を指定できます。この値を共変量パターン別の応答カテゴリのクロス表のそれぞれの空セルに加えます。これにより、アルゴリズムを安定させ、推定値に偏りが入り込むのを防ぎます。

**特異性許容度。** 特異性の点検に使用する許容度を指定できます。

## 多項ロジスティック回帰のオプション

図 3-6  
[多項ロジスティック回帰: オプション] ダイアログ ボックス



多項ロジスティック回帰に対して、次のオプションを指定できます。

**[散らばりの尺度]** パラメータ分散共分散行列の推定値を修正するために使用される散らばり尺度値を指定できます。**[逸脱]** では、逸脱関数（尤度比カイ 2 乗）統計量を使用して尺度値を推定します。**[Pearson]** では、Pearson カイ 2 乗統計量を使用して尺度値を推定します。さらに、独自の尺度値を指定することもできます。尺度値は正の数値でなければなりません。

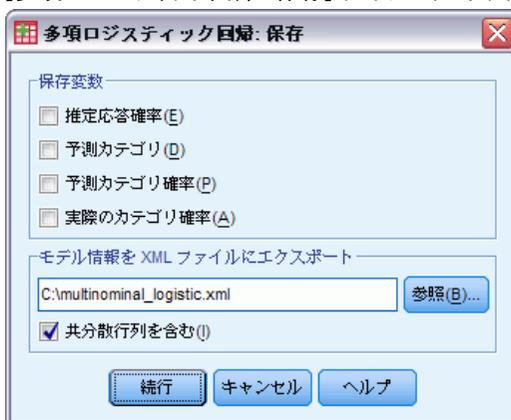
**ステップワイズのオプション。** このオプションにより、ステップワイズ法を使用してモデルを構築する場合の統計基準を制御できます。ただし、**[モデル]** ダイアログ ボックスでステップワイズ モデルが指定されるまで無視されます。

- **投入の確率。** これは、変数の投入の尤度比統計量の確率です。指定された確率が大きくなるほど、変数をモデルに投入しやすくなります。この基準は、変数増加法投入法、変数増加ステップワイズ法、または変数減少ステップワイズ法が選択されていない場合は無視されます。
- **[投入検定]** これは、ステップワイズ法で項目を投入するための方法です。尤度比検定とスコア検定のいずれかを選択します。この基準は、変数増加法投入法、変数増加ステップワイズ法、または変数減少ステップワイズ法が選択されていない場合は無視されます。

- **除去の確率。** これは、変数の除去の尤度比統計量の確率です。指定された確率が大きくなるほど、変数をモデルに保持しやすくなります。この基準は、変数減少法、変数増加ステップワイズ法、または変数減少ステップワイズ法が選択されていない場合は無視されます。
- **[除去検定]** これは、ステップワイズ法で項目を除去するための方法です。尤度比検定と Wald 検定のいずれかを選択します。この基準は、変数減少法、変数増加ステップワイズ法、または変数減少ステップワイズ法が選択されていない場合は無視されます。
- **モデルの最小ステップ効果。** 変数減少法または変数減少ステップワイズ法を使用する場合に、モデルに含める項の最小数を指定します。切片は、モデルの項として数えられません。
- **モデルの最大ステップ効果。** 変数増加法投入法または変数増加ステップワイズ法を使用する場合に、モデルに含める項の最大数を指定します。切片は、モデルの項として数えられません。
- **項の投入および除去に、階層的な制約を加える。** このオプションにより、モデル項を含める場合の制約を定めるかどうかを選択できます。階層では、含める項すべてに対して、項の一部であり順位の高いものをすべて最初にモデルに含めることが必要です。たとえば、階層の必要条件が有効である場合は、Marital status\*Gender 交互作用を追加する前に、因子 Marital status と Gender を両方ともモデルに含めておく必要があります。3 つのラジオ ボタン オプションにより、階層を決定するときの共変量の役割が決まります。

## 多項ロジスティック回帰の保存

図 3-7  
[多項ロジスティック回帰: 保存] ダイアログ ボックス



[保存] ダイアログ ボックスでは、変数を作業ファイルに保存したり、モデル情報を外部ファイルにエクスポートできます。

**保存変数。**

- **[推定応答確率]** 因子/共変量パターンを応答カテゴリに分類する推定確率です。推定確率は応答変数のカテゴリの数だけあり、最大 25 個まで保存されます。
- **予測カテゴリ。** 因子/共変量パターンに対して最大期待確率を持つ応答カテゴリです。
- **予測カテゴリ確率。** 最大の推定応答確率です。
- **実カテゴリ確率。** 因子/共変量パターンを観測カテゴリに分類する推定確率です。

**モデル情報を XML ファイルにエクスポート。** パラメータ推定とその共分散（オプション）は指定されたファイルに XML（PMML）形式でエクスポートされます。このモデル ファイルを使用して、得点付けのために他のデータ ファイルにモデル情報を適用できます。

## NOMREG コマンドの追加機能

コマンド シンタックス言語を使用して、次のことも実行できます。

- 従属変数の参照カテゴリを指定する。
- ケースにユーザー欠損値を含める。
- 帰無仮説をパラメータの線型結合として指定することによって仮説の検定をカスタマイズする。

複雑なシンタックス情報については、「コマンド シンタックス リファレンス」を参照してください。

# プロビット分析

この手続きは、刺激の強さと、刺激に対して一定の反応を示すケースの割合との関係を測定します。独立変数のレベルによって影響を受けている、または起因していると思われる 2 分的な結果がある場合に役立ちます。また、特に実験データに適しています。この手続きでは、たとえばメディアン有効量のように、一定の割合の応答を導くために必要な刺激の強さを推定することができます。

**例。**新しい殺虫剤はアリの駆除にどのくらい効果があるでしょうか。また、どれぐらいの濃度で使用するのが適当でしょうか。アリのサンプルにさまざまな濃度の殺虫剤を散布する実験を行い、次に、駆除されたアリの数と殺虫剤を散布したアリの数を記録します。データにプロビット分析を適用すると、濃度と殺虫力の間の関係の強さを決めることができます。また、たとえば、殺虫剤を散布したアリの 95% を確実に駆除した場合に適当な殺虫剤の濃度を決定できます。

**統計量。**回帰係数と標準誤差、切片と標準誤差、Pearson のカイ 2 乗の適合度、観測度数と期待度数、独立変数の有効レベルの信頼区間。プロット：変換された応答プロット。

この手続きでは、「NPSOL® by Gill, Murray, Saunders & Wright」で提案され、実装されているアルゴリズムを使用して、モデルパラメータを推定します。

**データ。**独立変数のそれぞれの値について（複数の独立変数の値の各対について）、応答変数は対象の応答を示す値を持つケースの数にし、総観測度数変数は独立変数の値を持つケースの数にしてください。因子変数は整数でコード化したカテゴリ変数にしてください。

**仮定。**観測値は独立している必要があります。観測調査のように、観測値の数に関連する独立変数の値が多数ある場合、カイ 2 乗と適合度の統計量が有効でないことがあります。

**関連手続き。**プロビット分析は、ロジスティック回帰に似ています。実際、ロジット変換を選択した場合に、プロビット分析手続きは必ずロジスティック回帰を計算します。一般に、プロビット分析は計画した実験に適しているのに対し、ロジスティック回帰は観測調査に適しています。その違いは出力に顕著に現れます。プロビット分析手続きでは、さまざまな応答の比率の影響値（メディアン影響量を含む）の推定値を出力します。一方、ロジスティック回帰手続きは、独立変数のオッズ比の推定値を出力します。

## プロビット分析を行うには

- ▶ メニューから次の項目を選択します。  
分析 > 回帰 > プロビット...

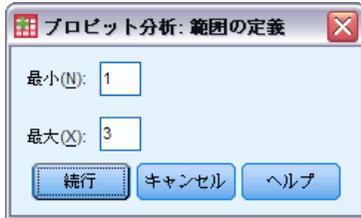
図 4-1  
[プロビット分析] ダイアログ ボックス



- ▶ 応答度数変数を選択します。この変数は検定刺激に回答するケース数を示します。この変数の値は負の値にできません。
- ▶ [総観測度数変数] ボックスに変数を 1 つ選択します。この変数は刺激が適用されるケース数を示します。この変数の値は負の値にできず、また、ケースごとの応答度数変数の値以上である必要があります。  
必要に応じて、[因子] ボックスに変数を 1 つ選択できます。この場合、[範囲の定義] をクリックして、グループを定義します。
- ▶ 1 つ以上の共変量を選択します。この変数には各観測値に適用される刺激のレベルが含まれます。共変量を変換する場合は、[変換] のドロップダウン リストから、適用する変換を選択します。変換を適用せず、対照群がある場合には、対照群が分析に含まれます。
- ▶ [モデル] の [プロビット] と [ロジット] のどちらかをクリックします。
  - **プロビットモデル(プロビット分析)**. プロビット変換 (累積標準正規分布関数の逆関数) を応答率に適用します。
  - **ロジット モデル**. ロジット (対数オッズ) 変換を応答率に適用します。

## プロビット分析の範囲の定義

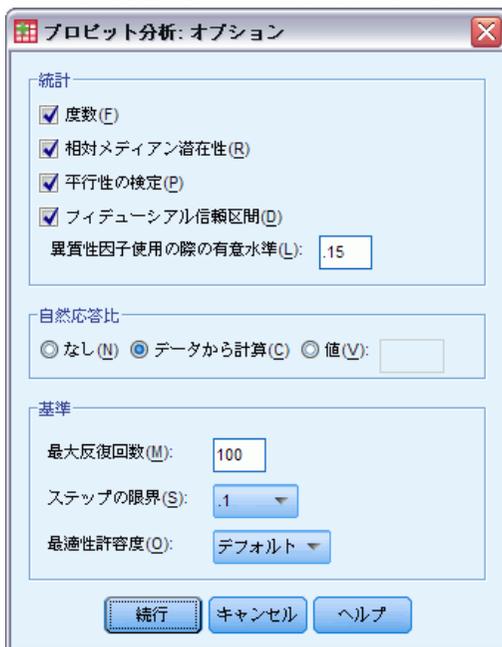
図 4-2  
[プロビット分析: 範囲の定義] ダイアログ ボックス



分析する因子変数のレベルを指定します。因子レベルは連続する整数でコード化する必要があります。指定した範囲内のすべてのレベルを分析します。

## プロビット分析のオプション

図 4-3  
[プロビット分析: オプション] ダイアログ ボックス



プロビット分析では、次のオプションを指定できます。

**統計量。** オプション統計量（[度数]、[相対メディアン潜在性]、[平行性の検定]、および [フィデューシアル信頼区間]）を指定できます。

- **相対メディアン潜在性 (プロビット分析).** 因子レベルの各ペアに対するメディアン潜在性の比を表示します。各相対メディアン潜在性の 95% 信頼限界も表示します。相対メディアン潜在性は、因子変数がない場合または複数の共変量がある場合は使用できません。
- **平行性の検定.** 全ての因子が共通の傾きをもつという帰無仮説を検定します。交互作用が無いということです。
- **フィデューシアル信頼区間.** ある確率で反応を得るために必要な作用物（例えば、薬剤）の適量についての信頼区間です。

複数の共変量を選択した場合、[フィデューシアル信頼区間] と [相対メディアン潜在性] は使用できません。[相対メディアン潜在性]、[平行性の検定] は、[因子] ボックスで変数を選択した場合に限り使用できます。

**自然応答比.** 刺激が存在しない場合でも自然応答比を示すことができます。使用できるオプションは [なし]、[データから計算]、または [値] のいずれか 1 つです。

- **データから計算 (プロビット分析).** サンプル データから自然応答比を推定します。データには、共変量の値が 0 の管理レベルを表すケースがなければなりません。プロビットは、制御レベルの応答比を初期値として使用することによって自然応答比を推定します。
- **値 (ピボットテーブル ノンパラメトリック検定).** モデルの自然応答比を設定します（この項目は、自然応答比を事前に知っているときに選択します）。自然応答比を入力します（比率は 1 より小さくしなければなりません）。たとえば、刺激が 0 のときに時間の 10% で応答が発生したら、0.10 を入力します。

**基準.** 反復パラメータ評価アルゴリズムのパラメータを制御できます。[最大反復回数]、[ステップの限界]、[最適性許容度] の各ボックスでデフォルトを変更できます。

## PROBIT コマンドの追加機能

コマンド シンタックスを使用すると、次の作業も実行できます。

- プロビット モデルとロジット モデルの両方の分析の要求。
- 欠損値の扱いの制御。
- 常用対数または自然対数以外の対数による共変量の変換。

シンタックスの詳細は、『Command Syntax Reference』を参照してください。

# 非線型回帰分析

非線型回帰分析は、従属変数と独立変数グループ間の関係の非線型モデルを見つける手法です。従来の線型回帰は線型モデルの推定だけに制約されますが、非線型回帰分析では従属変数と独立変数に任意の関係があるモデルを推定できます。これは、反復推定アルゴリズムを使用することによって可能になります。この手続きは、 $Y = A + BX^{**2}$  の形式の単純な多項式モデルでは必要ありません。 $W = X^{**2}$  を定義することにより、単純な線型モデル  $Y = A + BW$  が求められます。これは、線型回帰手続きのような従来の手法を使用して推定できます。

**例。**人口を時間に基づいて予測することはできるでしょうか。散布図は、人口と時間の強い関係を示唆しています。しかし、その関係が非線型のため、評価には特殊な予測手法である非線型回帰分析手続きが必要です。ロジスティック人口成長モデルのような適当な方程式を設定することにより、モデルの妥当な推定を得ることができます。この推定により、実際には測定していない、さまざまな時点の人口について予測を行うことができます。

**統計量。**各反復ごとに：パラメータ推定値および残差平方和。各モデルごとに：回帰平方和、残差、無修正総和と修正総和、パラメータ推定値、漸近標準誤差、パラメータ推定値の漸近相関行列。

**注：**制約条件付き非線型回帰分析では、「NPSOL<sup>®</sup> by Gill, Murray, Saunders, and Wright」で提案され、実装されているアルゴリズムを使用してモデルパラメータを推定します。

**データ。**従属変数および独立変数は量的でなければなりません。宗教、専攻、居住地区などのカテゴリ変数は、2値（ダミー）変数またはその他の種類の対比変数に値を再割り当てする必要があります。

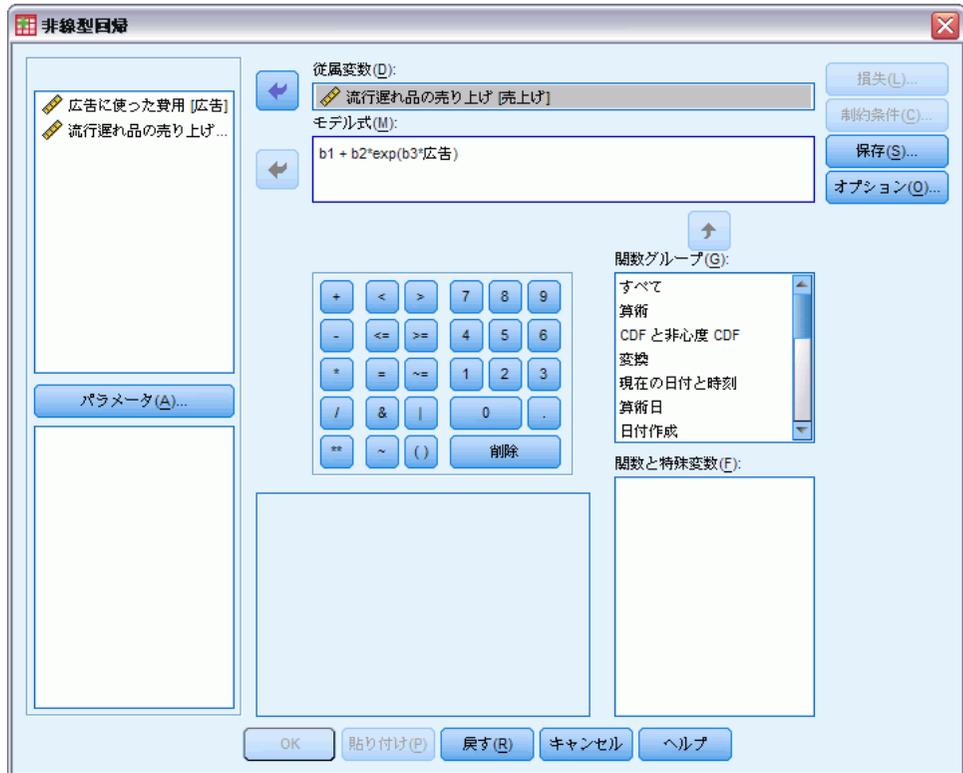
**仮定。**結果は、従属変数と独立変数間の関係を正確に示す関数を指定した場合に限り有効です。また、適切な初期値の選択が非常に重要になります。モデルの関数形式を正しく指定したとしても、不適切な初期値を使用すると、モデルが収束しなかったり、または解がグローバルな最適解ではなくローカルな最適解になることがあります。

**関連手続き。**最初是非線型に見えるさまざまなモデルも線型モデルに変換できるので、これらのモデルを線型回帰手続きを使用して分析できるようになります。適切なモデルがはっきりしない場合には、データの関数関係の識別に、曲線推定手続きが役立ちます。

## 非線型回帰分析を行うには

- ▶ メニューから次の項目を選択します。  
分析 > 回帰 > 非線形...

図 5-1  
[非線型回帰分析] ダイアログ ボックス



- ▶ アクティブ データセットの変数の一覧から、数値型従属変数を 1 つ選択します。
- ▶ モデル式を作成するには、[モデル式] ボックスに式を入力するか、成分 (変数、パラメータ、関数) をボックスに貼り付けます。
- ▶ [パラメータ] をクリックして、モデルのパラメータを識別します。  
分割モデル (ドメインが違うと別の形式になるモデル) は、単一モデル文内の条件論理を使用して指定する必要があります。

## 条件論理（非線型回帰分析）

条件論理を使用すれば、分割モデルを指定できます。モデル式、または損失関数で条件論理を使用するには、条件ごとに項の級数の合計を 1 つ作ります。各項は、論理式（カッコ内）と、論理式が真となる場合に生じる式の積で構成されます。

たとえば、 $X \leq 0$  の場合に 0、 $0 < X < 1$  の場合に  $X$ 、 $X \geq 1$  の場合に 1 となる分割モデルについて考えると、この式は次のようになります。

$$(X \leq 0) * 0 + (X > 0 \ \& \ X < 1) * X + (X \geq 1) * 1.$$

かっこ内の論理式はすべて 1（真）または 0（偽）になります。したがって、

$X \leq 0$  の場合、上記は  $1 * 0 + 0 * X + 0 * 1 = 0$  になります。

$0 < X < 1$  の場合、上記は  $0 * 0 + 1 * X + 0 * 1 = X$  になります。

$X \geq 1$  の場合、上記は  $0 * 0 + 0 * X + 1 * 1 = 1$  になります。

別の論理式と結果式を置き換えることで、その他の複雑な例も簡単に作成できます。 $0 < X < 1$  のような二重不等式は、 $(X > 0 \ \& \ X < 1)$  などの複合式で表す必要があることに注意してください。

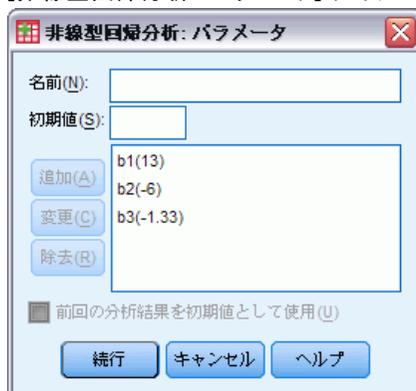
論理式に文字型変数を使用できます。

$$(\text{city} = \text{'New York'}) * \text{costliv} + (\text{city} = \text{'Des Moines'}) * 0.59 * \text{costliv}$$

ここでは、ニューヨークの住人に 1 つの式（変数  $\text{costliv}$  の値）を、デモインの住人にはもう 1 つの式（値の 59%）を使用します。文字型定数は、ここに示しているように、引用符またはアポストロフィーで囲みます。

## 非線型回帰分析のパラメータ

図 5-2  
[非線型回帰分析: パラメータ] ダイアログ ボックス



パラメータは、非線型回帰分析が推定するモデルの一部になります。パラメータには付加定数、乗法的係数、べき指数、または関数評価に使用する値があります。定義したすべてのパラメータが、メイン ダイアログ ボックスのパラメーター一覧に（初期値で）表示されます。

**名前。**各パラメータの名前を指定します。この名前は有効な変数名で、メイン ダイアログ ボックスのモデル式で使った名前でない限りなりません。

**初期値。**パラメータに初期値を指定できます。可能であれば、初期値は期待する最終解にできるだけ近い値を指定します。不適切な値が指定された場合には、収束しない、解の収束が（グローバルではなく）ローカルになる、または収束が物理的にあり得ない解を導くことがあります。

**前回の分析結果を初期値として使用。**このダイアログ ボックスから非線型回帰分析を実行したことがある場合、このチェックボックスをオンにすれば、前回の値からパラメータの初期値を求めることができます。これにより、アルゴリズムが徐々に収束している間に検索を続けることができます。（初期値はメイン ダイアログ ボックスのパラメーター一覧に表示されたままです）。

**注：**セッション中、この選択はダイアログ ボックスで有効のままになります。モデルを変える場合には、必ずこのオプションをオフにします。

## 非線型回帰分析の一般モデル

一般的に知られている非線型回帰分析モデルのモデル シンタックスの例を、次の表に示します。無作為に選択したモデルをデータに適切に適合させることはできません。パラメータには適切な初期値が必要です。また、収束させるには制約が必要なモデルもあります。

テーブル 5-1  
モデル シンタックス例

名前 (スクリプト ウィンドウ、新規手続き)	モデル式
漸近回帰	$b1 + b2 * \exp(b3 * x)$
漸近回帰	$b1 - (b2 * (b3 ** x))$
密度関数	$(b1 + b2 * x) ** (-1 / b3)$
Gauss	$b1 * (1 - b3 * \exp(-b2 * x ** 2))$
Gompertz	$b1 * \exp(-b2 * \exp(-b3 * x))$
Johnson-Schumacher	$b1 * \exp(-b2 / (x + b3))$
対数修正	$(b1 + b3 * x) ** b2$
対数ロジスティック	$b1 - \ln(1 + b2 * \exp(-b3 * x))$
Metzlerlich の収穫逡減の法則	$b1 + b2 * \exp(-b3 * x)$
Michaelis Menten	$b1 * x / (x + b2)$
Morgan-Mercer-Florin	$(b1 * b2 + b3 * x ** b4) / (b2 + x ** b4)$

名前 (スクリプト ウィンドウ、新規手続き)	モデル式
Peal-Reed	$b1 / (1 + b2 * \exp(-(b3 * x + b4 * x ** 2 + b5 * x ** 3)))$
3 次の有理式	$(b1 + b2 * x + b3 * x ** 2 + b4 * x ** 3) / (b5 * x ** 3)$
2 次の有理式	$(b1 + b2 * x + b3 * x ** 2) / (b4 * x ** 2)$
Richards	$b1 / ((1 + b3 * \exp(-b2 * x)) ** (1 / b4))$
Verhulst	$b1 / (1 + b3 * \exp(-b2 * x))$
Von Bertalanffy	$(b1 ** (1 - b4) - b2 * \exp(-b3 * x)) ** (1 / (1 - b4))$
ワイブル (正規確率プロット検定分布)	$b1 - b2 * \exp(-b3 * x ** b4)$
収穫密度	$(b1 + b2 * x + b3 * x ** 2) ** (-1)$

## 非線型回帰分析の損失関数

図 5-3  
[非線型回帰分析: 損失関数] ダイアログ ボックス



非線型回帰分析の**損失関数**は、アルゴリズムで最小化される関数です。残差平方和を最小化するには**[残差平方和]**、その他の関数を最小化するには**[ユーザーの定義した損失関数]**をクリックします。

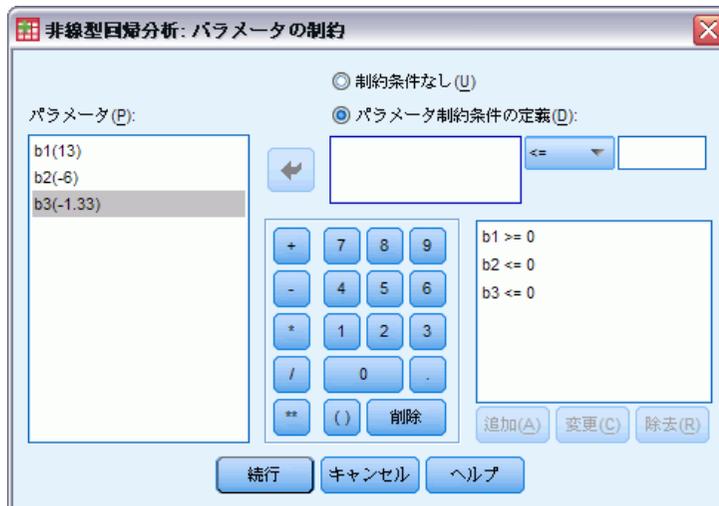
**[ユーザーの定義した損失関数]**を選択した場合には、ボックスで損失関数を定義する必要があります。

- 定義する損失関数の（すべてのケースでの）和は、パラメータ値の選択により最小化される必要があります。ほとんどの損失関数には特殊変数 RESID\_ があります。これは残差を表しています。（デフォルトの残差平方和損失関数を「RESID\_\*\*2」として明示的に入力することもできます）。損失関数に予測値を使用する必要がある場合、これは従属変数から残差を引いた値となります。
- 条件論理を使用して、条件損失関数を指定できます。

**[ユーザーの定義した損失関数]** ボックスに式を入力するか、このボックスに式の要素を貼り付けることができます。文字型定数は引用符、つまりアポストロフで両側を囲みます。また数値型定数は、少数点にドットを使用した形式で入力します。

## 非線型回帰分析のパラメータの制約

図 5-4  
[非線型回帰分析: パラメータの制約] ダイアログ ボックス



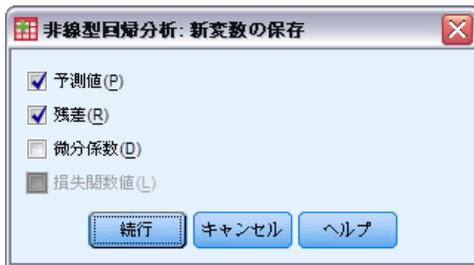
**制約条件**は、解を求めるための反復の際にパラメータが取ることのできる値に対する制限です。線型式はステップの実行前に評価されるので、線型制約を使用してステップのオーバーフローを防止できます。非線型式は、ステップが実行された後に評価されます。

各式には次の要素が必要です。

- モデルに少なくとも 1 つのパラメータを含んでいる式。式はキーボードで入力するか、キーパッドを使用します。必要なパラメータを残りの式と一緒に入力するか、画面左側のパラメーター一覧から貼り付けることができます。制約条件には通常の変数を使用できません。
- 論理演算子。[<=]、[=]、または [>=] の 3 つの演算子の中から 1 つを選択します。
- 数値定数。式は論理演算子を使用して数値定数と比べられます。定数を入力します。数値定数は小数点にドットを使用する形式で入力します。

## 非線型回帰分析の新変数の保存

図 5-5  
[非線型回帰分析: 新変数の保存] ダイアログ ボックス



作業中のデータ ファイルには、多数の新しい変数を保存できます。使用できるオプションは、[予測値]、[残差]、[微分係数]、および [損失関数値] です。この後の分析でこれらの変数を使用して、モデルの適合の検定や問題のあるケースを識別できます。

- **残差**. 残差を変数名 `resid` で保存します。
- **予測値 (非線型回帰 保存)**. 予測値を変数名 `pred_` で保存します。
- **微分係数 (非線型回帰 保存)**. 各モデル パラメータに対して 1 つの微分係数が保存されます。微分係数名は、パラメータ名の最初の 6 文字の前に「d.」を付けて作成します。
- **損失関数値 (非線型回帰 保存)**. このオプションは、ユーザー独自の損失関数を指定した場合に使用することができます。損失関数の値に、変数名 `loss_` が割り当てられます。

## 非線型回帰分析のオプション

図 5-6  
[非線型回帰分析: オプション] ダイアログ ボックス



オプションで非線型回帰分析のさまざまな機能を制御できます。

**ブートストラップ推定値.** 元のデータ セットから反復したサンプルを使うことによって統計量の標準誤差を推定する方法。これは、元のデータ セットと同じサイズのサンプルを（置換も伴って）取り出して行われます。これらのサンプルのそれぞれについて非線型方程式が評価されます。次に、各パラメータ推定値の標準誤差が、ブートストラップ推定値の標準偏差として計算されます。元のデータ セットから取り出されたパラメータ値は、各ブートストラップ サンプルの開始値として使われます。これには、逐次 2 次計画法のアルゴリズムが必要です。

**推定方法.** 可能であれば、推定方法を選択できます。（このダイアログ ボックスまたは別のダイアログ ボックスでの選択内容により、逐次 2 次計画法のアルゴリズムが必要な場合があります）。使用できるオプションは、[逐次 2 次計画法] または [Levenberg-Marquardt 法] のどちらかです。

- **逐次 2 次計画法 (非線型回帰 オプション).** この方法は、制約条件のあるモデルおよび制約条件のないモデルに利用できます。逐次 2 次計画法は、制約条件のあるモデル、ユーザー定義の損失関数、またはブートストラップを指定した場合に自動的に使用されます。[最大反復回数]、[ステップの限界] の各ボックスには新しい値を入力できます。また、[最適性許容度]、[関数精度]、[無限ステップサイズ] の各ボックスの一覧の選択を変更できます。
- **Levenberg-Marquardt 法 (非線型回帰 オプション).** 制約条件のないモデルのデフォルトのアルゴリズムです。Levenberg-Marquardt 法は、制約条件のあるモデル、ユーザー定義の損失関数、またはブートストラップを指定した場合、使用できません。[最大反復回数] ボックスには新しい値

を入力できます。また、[平方和の収束基準]、[パラメータ収束基準]の各ドロップダウン リストの選択を変更できます。

## 非線型回帰分析の結果の解釈

非線型回帰分析では、うまく計算ができないという問題がしばしば起こります。

- パラメータの初期値の選択は収束に影響します。妥当で、期待する最終解にできるだけ近い初期値を選択するようにします。
- 問題によっては、他のアルゴリズムを使用した方がよい結果を得られる場合もあります。他のアルゴリズムが選択できる場合は、[オプション] ダイアログ ボックスで選択します。(損失関数または一定の制約条件を指定する場合には、[Levenberg-Marquardt 法] は指定できません)。
- 反復回数が最大値に達したために反復が停止した場合、「最終」モデルはおそらく適切な解ではありません。[パラメータ] ダイアログ ボックスの [前回の分析結果を初期値として使用] チェック ボックスをオンにして反復を続けるか、さらにより結果を得るには別の初期値を選択します。
- 大きなデータ値のべき乗や大きな値を指数とするモデルは、オーバーフローやアンダーフロー (数値が大きすぎ、または小さすぎてコンピュータが表示できない) を起こすことがあります。場合によっては、初期値を適切に選択するかパラメータに制約条件を設けることにより、このような状態を回避できます。

## NLR コマンドの追加機能

コマンド シンタックスを使用すると、次の作業も実行できます。

- パラメータ推定値の初期値の読み込みに使用するファイルを指定する。
- 複数のモデル文と損失関数を指定する。分割モデルの指定が簡単になります。
- プログラムが計算する微分係数の代わりに独自の微分係数を使用する。
- 生成するブートストラップ数を指定する。
- 微分係数チェック用の臨界値の設定や残差と微分係数間の相関の収束基準の定義などの付加的な反復基準を指定する。

CNLR (制約条件付き非線型回帰分析) コマンド用の付加的な基準を使用すると、次の作業を実行できます。

- 主要な反復の中で行われる副次的な反復の最大値を指定する。
- 微分係数のチェックに使用する臨界値を設定する。

- ステップの限界を設定する。
- 初期値が指定された範囲内にあるか決定するためのクラッシュ許容度を指定する。

シンタックスの詳細は、『Command Syntax Reference』を参照してください。

# 重み付け推定

標準の線型回帰モデルでは、調査中の母集団内で分散が一定していると仮定します。これが当てはまらない場合、たとえば、特定の属性の高いケースがその属性の低いケースよりもさらに変動を示す場合、通常最小 2 乗法を使用する線型回帰 (OLS) では最適なモデル推定ができません。変動の差が別の変数から予測できる場合には、重み付け推定手続きで、重み付き最小 2 乗法 (WLS) を使用して線型回帰モデルの係数を計算できます。したがって、さらに正確な観測値 (つまり、変動が少ない観測値) には回帰係数を決定する際により大きな重みを与えます。重み付け推定手続きは、さまざまな重み変換を検定し、データに最も適合するものを示します。

**例。** 株式価格の変動にインフレーションと失業は影響するのでしょうか。高値の株は、ほとんどの場合、低値の株よりも変動が激しいので、通常最小 2 乗法では最適な推定値を求めることができません。重み付け推定を使用すると、線型モデルの計算で価格変動量の変動に対する株式価格の影響を説明できます。

**[統計量]** 検定された重み付けソース変数の各べき乗ごとの対数尤度、重相関係数、 $R^2$  乗、調整済み  $R^2$  乗、WLS モデルの分散分析表、標準化されていないパラメータ推定値と標準化パラメータ推定値、WLS モデルの対数尤度。

**データ。** 従属変数および独立変数は量的でなければなりません。宗教、専攻、居住地区などのカテゴリ変数は、2 値 (ダミー) 変数またはその他の種類の対比変数に値を再割り当てする必要があります。重み付け変数は量的変数で、従属変数の変動に関連する必要があります。

**仮定。** 独立変数の各値に対して、従属変数の分布は正規分布でなければなりません。従属変数と各独立変数の関係は線型であり、すべての観測が独立している必要があります。従属変数の分散は独立変数のレベルによって異なる場合がありますが、その差は重み付け変数に基づいて予測できなければなりません。

**関連手続き。** 探索的分析手続きをデータの検査に使用できます。[探索的分析] では、正規性と等分散性の検定を、グラフ表示と同時に実行できます。従属変数が独立変数のすべてのレベルで等分散しているような場合には、線型回帰手続きを使用できます。データがある仮定 (たとえば、正規性) に反しているような場合には、そのデータを変換してみます。データに線型関係がなく、変換が役立たない場合は、[曲線推定] 手続きの代替モデルを使用します。従属変数が 2 分変数の場合、たとえば、特定の販売が終了したか否か、または品物が不良であるか否かというような場合は、ロジスティック回帰手続きを使用します。独立変数が打ち切られている場合、たとえば外科手術後の生存時間のような場合は、Advanced

Statistics オプションの生命表分析、Kaplan-Meier、または Cox 回帰分析を使用します。データが独立していない場合、たとえば同じ人をいくつかの条件下で観察する場合などは、Advanced Statistics オプションに含まれる反復測定手続きを使用します。

### 重み付け推定分析を行うには

- ▶ メニューから次の項目を選択します。  
[分析] > [回帰] > [重み付け推定...]

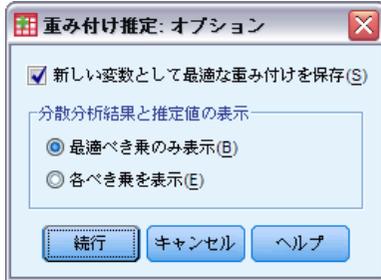
図 6-1  
[重み付け推定] ダイアログ ボックス



- ▶ 従属変数を 1 つ選択します。
- ▶ 1 つ以上の独立変数を選択します。
- ▶ 重み付け変数として不均一分散のソースとなる変数を選択します。
  - **重み付け変数 (WLS).** データは、累乗されたこの変数の逆数で重み付けられます。回帰式は指定したべき乗の範囲に対して計算され、対数尤度関数を最大化するべき乗を示します。
  - **べき乗の範囲 (WLS).** 重み付けを計算するために重み付け変数といっしょに使用します。いくつかの回帰式は、べき乗の範囲内の各値に 1 つずつ適合します。[べき乗の範囲] テキスト ボックスおよび [から] テキスト ボックスに入力する値は、-6.5 から 7.5 の間になければなりません。べき乗値は最低値と最高値の範囲で指定された増分ごとの値を取ります。べき乗の範囲内の値の総数の限界は 150 です。

## 重み付け推定のオプション

図 6-2  
[重み付け推定: オプション] ダイアログ ボックス



重み付け推定分析では、次のオプションを指定できます。

**新しい変数として最適な重み付けを保存。** 重み付け変数を作業中のファイルに追加します。変数は WGT\_n と呼ばれます。ここで、n は変数に一意的な名前を付けるための数字です。

**分散分析結果と推定値の表示。** 統計の出力の表示方法を制御できます。使用できるオプションは、[最適べき乗のみ表示] または [各べき乗を表示] です。

## WLS コマンドの追加機能

コマンド シンタックスを使用すると、次の作業も実行できます。

- べき乗への単一値の割り当て。
- べき乗値の一覧、または値の範囲とべき乗の値の一覧の組み合わせの指定。

シンタックスの詳細は、『Command Syntax Reference』を参照してください。

## 2 段階最小 2 乗回帰

標準の線型回帰モデルでは、従属変数の誤差は独立変数と相関がないと仮定しています。相関がある場合（たとえば、変数の関係が双方向の場合）、通常最小 2 乗法（OLS）を使用した線型回帰では、最適なモデルを推定できません。2 段階最小 2 乗回帰では、誤差項と相関していない操作変数を使用して、問題のある予測変数の推定値を計算します（第 1 段階）。次に、これらの計算値を使用して、従属変数の線型回帰モデルを推定します（第 2 段階）。計算値は、誤差とは相関がない変数に基づいているので、2 段階モデルの結果は最適となります。

**例。**ある商品の需要の大きさは、その価格と消費者の収入に関連しているでしょうか。このモデルの難しさは価格と需要に相互作用があるところです。つまり、価格が需要に影響し、またその反対に、需要が価格に影響することがあり得ます。2 段階最小 2 乗回帰モデルでは、需要の測定誤差との相関がないような価格の代理を計算するのに、消費者の収入ならびにラグのある価格を使用します。この代理は、最初に指定されたモデルの価格自体と置き換えられた後に推定されます。

**統計量。**各モデルごとに、標準回帰係数および標準化されていない回帰係数、重相関係数、 $R^2$ 、調整済み  $R^2$ 、推定値の標準誤差、分散分析表、予測値、および残差。各回帰係数の 95% 信頼区間、パラメータ推定値の相関および分散共分散行列。

**データ。**従属変数および独立変数は量的でなければなりません。宗教、専攻、居住地区などのカテゴリ変数は、2 値（ダミー）変数またはその他の種類の対比変数に値を再割り当てする必要があります。内生説明変数は、（カテゴリ変数ではなく）量的変数でなければなりません。

**仮定。**独立変数の各値に対して、従属変数の分布は正規分布でなければなりません。従属変数の分布の分散は、独立変数のすべての値に対して一定でなければなりません。従属変数と各独立変数の関係は線型でなければなりません。

**関連手続き。**予測変数に従属変数内の誤差と相関しているものがないと思われる場合は、線型回帰手続きを使用できます。データが（正規性または定数分散などの）仮定の 1 つに反している場合は、そのデータを変換します。データに線型関係がなく、変換が役立たない場合は、[曲線推定] 手続きの代替モデルを使用します。従属変数が、特定の販売が終了したか否かというように 2 分される場合、ロジスティック回帰手続きを使用します。データが独立していない場合、たとえば同じ人をいくつかの条件下で観察する場合などは、Advanced Models オプションに含まれる反復測定手続きを使用します。

## 2 段階最小 2 乗回帰分析を行うには

- ▶ メニューから次の項目を選択します。  
分析 > 回帰 > 2 段階最小 2 乗...

図 7-1

[2 段階最小 2 乗回帰] ダイアログ ボックス

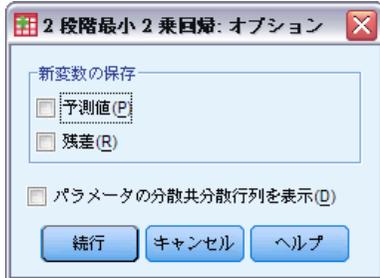


- ▶ 従属変数を 1 つ選択します。
- ▶ [説明変数] ボックスに説明 (予測) 変数を 1 つ以上選択します。
- ▶ [操作変数] ボックスに操作変数を 1 つ以上選択します。
  - **操作変数 (2 段階最小 2 乗)**. 2 段階最小 2 乗分析の最初の段階で内生変数の予測値を計算するとき使用する変数です。同じ変数が、[説明変数] および [操作変数] リスト ボックスの両方に指定されることもあります。操作変数の数は説明変数の数以上である必要があります。説明変数リストと操作変数リストの変数がすべて同じ場合、その結果は線型回帰手続きの結果と同じです。

操作変数として指定されていない説明変数は、内生変数と見なされません。通常、説明変数リスト内のすべての外生変数は、操作変数としても指定されます。

## 2 段階最小 2 乗回帰のオプション

図 7-2  
[2 段階最小 2 乗回帰: オプション] ダイアログ ボックス



分析には、次のオプションを選択できます。

**新変数の保存。**新変数をアクティブなファイルに追加します。使用可能なオプションは [予測値] および [残差] です。

**パラメータの分散共分散行列を表示。**パラメータ推定値の分散共分散行列を出力します。

## 2SLS コマンドの追加機能

コマンド シンタックスを使用すると、同時に複数の方程式を推定できます。シンタックスの詳細は、『Command Syntax Reference』を参照してください。

# カテゴリ変数コード化方式

多くの手続きでは、カテゴリ独立変数から対比変数グループへの自動置換を要求できます。これらの対比変数グループを後で式からブロックとして入力したり削除したりします。対比変数グループをコード化する方法を指定できます。通常は、CONTRAST サブコマンドを使用します。この付録では、CONTRAST サブコマンドで要求したさまざまな種類の対比が実際にどのように機能するかを図解で説明します。

## 全平均

**全平均からの偏差。** 行列の用語でいうと、これらの対比は次の形式です。

$$\begin{array}{l}
 \text{平均値} \quad ( \ 1/k \quad 1/k \quad \cdots \quad 1/k \quad 1/k ) \\
 \text{df (1)} \quad ( \ 1-1/k \quad -1/k \quad \cdots \quad -1/k \quad -1 / k ) \\
 \text{df (2)} \quad ( \ -1/k \quad 1-1/k \quad \cdots \quad -1/k \quad -1 / k ) \\
 \quad \quad \cdot \quad \quad \quad \cdot \\
 \quad \quad \cdot \quad \quad \quad \cdot \\
 \text{df (k-1)} \quad ( \ -1/k \quad -1/k \quad \cdots \quad 1-1/k \quad -1 / k )
 \end{array}$$

ただし、k は独立変数のカテゴリの個数であり、デフォルトでは最後のカテゴリは省略されます。たとえば、カテゴリが 3 つある独立変数の全平均の対比は次のようになります。

$$\begin{array}{l}
 ( \ 1/3 \quad 1/3 \quad 1/3 \ ) \\
 ( \ 2/3 \quad -1/3 \quad -1/3 \ ) \\
 ( \ -1/3 \quad 2/3 \quad -1/3 \ )
 \end{array}$$

最後のカテゴリ以外のカテゴリを除外するには、DEVIATION キーワードの後のカッコ内に除外するカテゴリの個数を指定します。たとえば、次のサブコマンドを実行すると、1 番目と 3 番目のカテゴリの偏差が求められ、2 番目のカテゴリが除外されます。

/CONTRAST (FACTOR) =DEVIATION (2)

factor に 3 つのカテゴリがあると仮定します。得られる対比行列は次のようになります。

$$\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/3 & -1/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}$$

## [単純]

**単純対比。**因子の各水準を最後のものと対比します。一般的な行列形式は次のとおりです。

$$\begin{array}{l} \text{平均値} \\ \text{df(1)} \\ \text{df(2)} \\ \vdots \\ \text{df(k-1)} \end{array} \begin{pmatrix} 1/k & 1/k & \cdots & 1/k & 1/k \\ 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{pmatrix}$$

ただし、k は、独立変数のカテゴリの個数です。たとえば、カテゴリが 4 つある独立変数の単純対比は次のようになります。

$$\begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

最後のカテゴリの代わりに別のカテゴリを参照カテゴリとして使用するには、SIMPLE キーワードの後のカッコ内に、参照カテゴリの番号を指定します。これは、必ずしもそのカテゴリに関連する値ではありません。たとえば、次の CONTRAST サブコマンドを実行すると、2 番目のカテゴリを除去する対比行列が得られます。

```
/CONTRAST(FACTOR) = SIMPLE(2)
```

factor に 4 つのカテゴリがあると仮定します。得られる対比行列は次のようになります。

$$\begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

## Helmert

**Helmert 対比。**独立変数のカテゴリを後続のカテゴリの平均値と比較します。一般的な行列形式は次のとおりです。

$$\begin{array}{l}
 \text{平均値} \quad ( \ 1/k \quad 1/k \quad \cdots \quad 1/k \quad 1/k ) \\
 \text{df(1)} \quad ( \ 1 \quad -1/(k-1) \quad \cdots \quad -1/(k-1) \quad -1/(k-1) ) \\
 \text{df(2)} \quad ( \ 0 \quad 1 \quad \cdots \quad -1/(k-2) \quad -1/(k-2) ) \\
 \quad \cdot \\
 \quad \cdot \\
 \text{df(k-2)} \quad ( \ 0 \quad 0 \quad 1 \quad -1/2 \quad -1/2 ) \\
 \text{df(k-1)} \quad ( \ 0 \quad 0 \quad \cdots \quad 1 \quad -1 )
 \end{array}$$

ただし、k は、独立変数のカテゴリの個数です。たとえば、カテゴリが 4 つある独立変数の Helmert 対比行列の形式は次のとおりです。

$$\begin{array}{l}
 ( \ 1/4 \quad 1/4 \quad 1/4 \quad 1/4 ) \\
 ( \ 1 \quad -1/3 \quad -1/3 \quad -1/3 ) \\
 ( \ 0 \quad 1 \quad -1/2 \quad -1/2 ) \\
 ( \ 0 \quad 0 \quad 1 \quad -1 )
 \end{array}$$

## 差分 (階差)

**逆 Helmert 対比。**独立変数のカテゴリを変数の前のカテゴリの平均値と比較します。一般的な行列形式は次のとおりです。

$$\begin{array}{l}
 \text{平均値} \quad ( \ 1/k \quad 1/k \quad 1/k \quad \cdots \quad 1/k ) \\
 \text{df(1)} \quad ( \ -1 \quad 1 \quad 0 \quad \cdots \quad 0 ) \\
 \text{df(2)} \quad ( \ -1/2 \quad -1/2 \quad 1 \quad \cdots \quad 0 ) \\
 \quad \cdot \\
 \quad \cdot \\
 \text{df(k-1)} \quad ( \ -1/(k-1) \quad -1/(k-1) \quad \cdots \quad 1 ) \\
 \quad \quad -1/(k-1)
 \end{array}$$

ただし、k は、独立変数のカテゴリの個数です。たとえば、カテゴリが 4 つある独立変数の逆 Helmert 対比は次のようになります。

$$\begin{array}{l}
 ( \ 1/4 \quad 1/4 \quad 1/4 \quad 1/4 ) \\
 ( \ -1 \quad 1 \quad 0 \quad 0 ) \\
 ( \ -1/2 \quad -1/2 \quad 1 \quad 0 ) \\
 ( \ -1/3 \quad -1/3 \quad -1/3 \quad 1 )
 \end{array}$$

## 多項式

**直交多項式の対比。**第 1 段階の自由度にはすべてのカテゴリにわたって 1 次効果が含まれ、第 2 段階の自由度には 2 次効果が含まれ、第 3 段階の自由度には 3 次効果が含まれる、というように次数が上がっていきます。

与えられたカテゴリ変数で測定した治療の水準の間に間隔を指定できます。距離を省略した場合のデフォルトである等間隔を 1 から k までの連続整数として指定できます。ただし、k はカテゴリの個数です。変数 DRUG に 3 つのカテゴリがある場合、次のサブコマンドは、

```
/CONTRAST (DRUG)=POLYNOMIAL
```

以下のサブコマンドと同じです。

```
/CONTRAST (DRUG)=POLYNOMIAL (1,2,3)
```

ただし、等間隔がいつでも必要とされるわけではありません。たとえば、DRUG は 3 つのグループに投与されるさまざまな量の薬品を表します。2 番目のグループへの投薬量は 1 番目のグループへの投薬量の 2 倍で、3 番目のグループへの投薬量が 1 番目のグループへの投薬量の 3 倍である場合、治療カテゴリは等間隔であり、この状況の適した距離は次のように連続整数で表されます。

```
/CONTRAST (DRUG)=POLYNOMIAL (1,2,3)
```

ただし、2 番目のグループへの投薬量が 1 番目のグループへの投薬量の 4 倍であり、3 番目のグループへの投薬量が 1 番目のグループへの投薬量の 7 倍である場合、適切な距離は次のとおりです。

```
/CONTRAST (DRUG)=POLYNOMIAL (1,4,7)
```

いずれの場合も、対比を指定した結果、DRUG の第 1 段階の自由度には投薬量水準の線型効果が、第 2 段階の自由度には 2 次効果が含まれます。

多項式の対比は、トレンドの検定を行ったり応答曲面の性質を調べる場合に特に役立ちます。さらに、多項式の対比を使って曲線回帰などの非線型曲線当てはめを実行することもできます。

## [反復測定]

**独立変数の隣接する水準を比較。**一般的な行列形式は次のとおりです。

平均値	(	1/k	1/k	1/k	...	1/k	1/k)
df(1)	(	1	-1	0	...	0	0)
df(2)	(	0	1	-1	...	0	0)

$$df(k-1) \quad ( 0 \quad 0 \quad 0 \quad \dots \quad 1 \quad -1 )$$

ただし、k は、独立変数のカテゴリの個数です。たとえば、カテゴリが 4 つある独立変数の反復は次のようになります。

$$\begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

これらの対比は、プロファイル分析や、偏差得点が必要な場合に利用できます。

## 特別 (ピボットテーブル GLM)

**ユーザー定義の対比。** 与えられた独立変数のカテゴリと同じ数だけ行および列のある正方行列の形式で特別な対比を入力できます。MANOVA と LOGLINEAR では、入力した最初の行は常に平均、つまり一定の効果であって、重みのグループを表し、与えられた変数について他の独立変数があればそれらの平均を求める方法を示します。一般に、この対比は 1 を並べたベクトルです。

行列の残りの行には、変数のカテゴリ間の目的の比較を示す特別な対比が含まれます。通常、直交対比が一番役立ちます。直交対比は統計上独立しており、冗長がありません。対比は、次の場合に直交しています。

- 行ごとに、対比係数の合計が 0 である。
- 非連結な行のすべてのペアについての対応する係数の積和も 0 である。

たとえば、「治療」に水準が 4 つあり、さまざまな水準の治療を相互に比較したいとします。特別な対比で適しているのは次のものです。

$( 1 \quad 1 \quad 1 \quad 1 )$	平均計算のための重み
$( 3 \quad -1 \quad -1 \quad -1 )$	1 番目と 2 ~ 4 番目の比較
$( 0 \quad 2 \quad -1 \quad -1 )$	2 番目と 3 番目および 4 番目の比較
$( 0 \quad 0 \quad 1 \quad -1 )$	3 番目と 4 番目の比較

これは、MANOVA、LOGISTIC REGRESSION、および COXREG に対し次の CONTRAST サブコマンドを使って指定します。

```
/CONTRAST (TREATMNT)=SPECIAL ( 1 1 1 1
                                3 -1 -1 -1
                                0 2 -1 -1
                                0 0 1 -1 )
```

LOGLINEAR については、次のように指定する必要があります。

```
/CONTRAST(TREATMNT)=BASIS SPECIAL( 1  1  1  1
                                     3 -1 -1 -1
                                     0  2 -1 -1
                                     0  0  1 -1 )
```

平均値の行以外の各行の合計は 0 です。非連結行の各ペアの積和も 0 です。

$$\text{行 2 と 3:} \quad (3)(0) + (-1)(2) + (-1)(-1) + (-1)(-1) = 0$$

$$\text{行 2 と 4:} \quad (3)(0) + (-1)(0) + (-1)(1) + (-1)(-1) = 0$$

$$\text{行 3 と 4:} \quad (0)(0) + (2)(0) + (-1)(1) + (-1)(-1) = 0$$

特別な対比は、直交していなくてもかまいません。ただし、相互の線型結合であってはなりません。線形結合である場合、手続きは線型従属であると報告し、処理を停止します。Helmert 対比、逆 Helmert 対比、および多項式対比はすべて直交対比です。

## 指標

**指示変数のコード化。** これは、ダミー コード化とも呼ばれ、LOGLINEAR や MANOVA では使用できません。新しくコード化される変数の数は、 $k-1$  です。参照カテゴリのケースは、すべての  $k-1$  変数で 0 にコード化されます。i 番目のカテゴリのケースは、1 にコード化される i 番目以外のすべての指示変数について 0 にコード化されます。

# 注意事項

この情報は、世界各国で提供される製品およびサービス向けに作成されています。

IBMはこのドキュメントで説明する製品、サービス、機能は他の国では提供していない場合があります。現在お住まいの地域で利用可能な製品、サービス、および、情報については、お近くの IBM の担当者にお問い合わせください。IBM 製品、プログラム、またはサービスに対する参照は、IBM 製品、プログラム、またはサービスのみが使用することができることを説明したり意味するものではありません。IBM の知的所有権を侵害しない機能的に同等の製品、プログラム、またはサービスを代わりに使用することができます。ただし、IBM 以外の製品、プログラム、またはサービスの動作を評価および確認するのはユーザーの責任によるものです。

IBMは、本ドキュメントに記載されている内容に関し、特許または特許出願中の可能性があります。本ドキュメントの提供によって、これらの特許に関するいかなる権利も使用者に付与するものではありません。ライセンスのお問い合わせは、書面にて、下記住所に送ることができます。

IBM Director of Licensing, IBM Corporation, North Castle Drive,  
Armonk, NY 10504-1785, U. S. A.

2 バイト文字セット (DBCS) 情報についてのライセンスに関するお問い合わせは、お住まいの国の IBM Intellectual Property Department に連絡するか、書面にて下記宛先にお送りください。

神奈川県大和市下鶴間1623番14号 日本アイ・ビー・エム株式会社 法務・知的財産 知的財産権ライセンス渉外

**以下の条項は、イギリスまたはこのような条項が法律に反する他の国では適用されません。** International Business Machines は、明示的または黙示的に関わらず、第三者の権利の侵害しない、商品性または特定の目的に対する適合性の暗黙の保証を含むがこれに限定されない、いかなる保証なく、本出版物を「そのまま」提供します一部の州では、特定の取引の明示的または暗示的な保証の免責を許可していないため、この文が適用されない場合があります。

この情報には、技術的に不適切な記述や誤植を含む場合があります。情報については変更が定期的に行われます。これらの変更は本書の新版に追加されます。IBM は、本書に記載されている製品およびプログラムについて、事前の告知なくいつでも改善および変更を行う場合があります。

IBM 以外の Web サイトに対するこの情報内のすべての参照は、便宜上提供されているものであり、決してそれらの Web サイトを推奨するものではありません。これらの Web サイトの資料はこの IBM 製品の資料に含まれるものではなく、これらの Web サイトの使用はお客様の責任によるものとします。

IBM はお客様に対する一切の義務を負うことなく、自ら適切と考える方法で、情報を使用または配布することができるものとします。

本プログラムのライセンス取得者が (i) 別途作成されたプログラムと他のプログラム（本プログラムを含む）との間の情報交換および (ii) 交換された情報の相互利用を目的とした本プログラムに関する情報の所有を希望する場合、下記住所にお問い合わせください。

IBM Software Group, Attention:Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.

上記のような情報は、該当する条項および条件に従い、有料で利用できるものとします。

本ドキュメントに記載されている許可されたプログラムおよびそのプログラムに使用できるすべてのライセンス認証された資料は、IBM Customer Agreement、IBM International Program License Agreement、および当社とかわした同等の契約の条件に基づき、IBM によって提供されます。

IBM 以外の製品に関する情報は、それらの製品の供給業者、公開済みの発表、または公開で使用できるソースから取得しています。IBM は、それらの製品のテストは行っておらず、IBM 以外の製品に関連する性能、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給業者に通知する必要があります。

この情報には、日常の業務処理で用いられるデータや報告書の例が含まれています。できる限り詳細に説明するため、例には、個人、企業、ブランド、製品などの名前が使用されています。これらの名称はすべて架空のものであり、実際の企業で使用される名称および住所とは一切関係ありません。

この情報をソフトコピーでご覧になっている場合は、写真やカラーのイラストが表示されない場合があります。

## 商標

IBM、IBM ロゴ、および [ibm.com](http://www.ibm.com)、SPSS は、世界の多くの国で登録された IBM Corporation の商標です。IBM の商標の現在のリストは、<http://www.ibm.com/legal/copytrade.shtml> を参照してください。

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel、Intel のロゴ、Intel Inside、Intel Inside のロゴ、Intel Centrino、Intel Centrino のロゴ、Celeron、Intel Xeon、Intel SpeedStep、Itanium、および Pentium は、米国およびその他の国の Intel Corporation または関連会社の商標または登録商標です。

Java およびすべての Java ベースの商標およびロゴは、米国およびその他の国の Sun Microsystems, Inc. の商標です。

Linux は、米国およびその他の国における Linus Torvalds の登録商標です。

Microsoft、Windows、Windows NT、および Windows のロゴは、米国およびその他の国における Microsoft 社の商標です。

UNIX は、米国およびその他の国における The Open Group の登録商標です。

この製品は、WinWrap Basic (Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>) を使用します。

その他の製品名およびサービス名等は、IBM または他の会社の商標です。

Adobe 製品のスクリーンショットは Adobe Systems Incorporated の許可を得て転載しています。

Microsoft 製品のスクリーンショットは Microsoft 社の許可を得て転載しています。



# 索引

- 分散共分散行列
    - 多項ロジスティック回帰, 16
  - 変数増加法
    - ロジスティック回帰, 5
  - 変数減少法
    - ロジスティック回帰, 5
  - 信頼区間
    - 多項ロジスティック回帰, 16
  - 反復回数
    - ロジスティック回帰, 9
    - 多項ロジスティック回帰, 18
  - 収束基準
    - 多項ロジスティック回帰, 18
  - 対数尤度
    - 多項ロジスティック回帰, 16
  - 相関行列
    - 多項ロジスティック回帰, 16
  - 逸脱関数
    - 散らばり尺度値の推定用, 19
  - 共変量
    - ロジスティック回帰, 6
  - 分類表
    - 多項ロジスティック回帰, 16
  - 定数項
    - 線型回帰, 9
  - 尤度比
    - 適合度, 16
    - 散らばり尺度値の推定用, 19
  - 特異性
    - 多項ロジスティック回帰, 18
  - 適合度
    - 多項ロジスティック回帰, 16
  - 分離
    - 多項ロジスティック回帰, 18
  - 分類
    - 多項ロジスティック回帰, 11
  - 切片
    - 含まれたケースまたは除かれたケース, 13
  - 商標, 49
  - 対比
    - ロジスティック回帰, 6
  - 段階 2 分
    - 多項ロジスティック回帰, 18
  - 2 項ロジスティック回帰, 1
  - 2 次の有理式モデル
    - 非線型回帰分析, 29
  - 2 段階最小 2 乗回帰, 39
  - コマンドの追加機能, 41
  - 新変数の保存, 41
  - 操作変数, 39
  - 統計量, 39
  - パラメータの共分散, 41
  - 例, 39
- 3 次の有理式モデル
    - 非線型回帰分析, 29
  - Cook の統計量 D
    - ロジスティック回帰, 7
  - Cox と Snell の  $R^2$  乗
    - 多項ロジスティック回帰, 16
  - DfBeta
    - ロジスティック回帰, 7
  - Gauss モデル
    - 非線型回帰分析, 29
  - Gompertz モデル
    - 非線型回帰分析, 29
  - Hosmer-Lemeshow の適合度統計
    - ロジスティック回帰, 9
  - Johnson-Schumacher モデル
    - 非線型回帰分析, 29
  - McFadden の  $R^2$  乗
    - 多項ロジスティック回帰, 16
  - Metzler の収穫逓減の法則
    - 非線型回帰分析, 29
  - Michaelis-Menten モデル
    - 非線型回帰分析, 29
  - Morgan-Mercer-Flourin モデル
    - 非線型回帰分析, 29
  - Nagelkerke の  $R^2$  乗
    - 多項ロジスティック回帰, 16
  - Peal-Reed モデル
    - 非線型回帰分析, 29
  - Pearson のカイ 2 乗
    - 適合度, 16
    - 散らばり尺度値の推定用, 19
  - Richards モデル
    - 非線型回帰分析, 29
  - Verhulst モデル
    - 非線型回帰分析, 29
  - Von Bertalanffy モデル
    - 非線型回帰分析, 29
  - Weibull モデル
    - 非線型回帰分析, 29
  - 重み付け推定, 36
    - 新しい変数として最適な重み付けを保存, 38
    - コマンドの追加機能, 38
    - 対数尤度, 36
    - 統計量, 36
    - 反復の記述, 38
    - 分散分析結果と推定値の表示, 38
    - 例, 36

## 索引

カテゴリ共変量, 6  
 参照カテゴリ  
 多項ロジスティック回帰, 15

収獲密度モデル  
 非線型回帰分析, 29

ステップワイズ法  
 ロジスティック回帰, 5  
 多項ロジスティック回帰, 13  
 スtring共変量  
 ロジスティック回帰, 6  
 すべての因子によるモデル  
 多項ロジスティック回帰, 13

制約条件  
 非線型回帰分析, 31  
 空セル  
 多項ロジスティック回帰, 18  
 セルの確率表  
 多項ロジスティック回帰, 16  
 線型回帰  
 2 段階最小 2 乗回帰, 39  
 重み付け推定, 36  
 漸近回帰  
 非線型回帰分析, 29

相対メディアン潜在性  
 プロビット分析, 24

対数修正モデル  
 非線型回帰分析, 29  
 対数尤度  
 重み付け推定, 36

てこ比の値  
 ロジスティック回帰, 7  
 デルタ  
 空セルの修正, 18

法律に関する注意事項, 48

反復の記述  
 多項ロジスティック回帰, 18

パラメータ制約条件  
 非線型回帰分析, 31  
 パラメータ推定値  
 多項ロジスティック回帰, 16

反復回数  
 プロビット分析, 24

非線型回帰分析, 26  
 Levenberg-Marquardt 法のアルゴリズム, 33  
 一般非線形モデル, 29  
 結果の解釈, 34  
 コマンドの追加機能, 34  
 残差, 32  
 条件論理, 28  
 初期値, 28  
 新変数の保存, 32  
 推定方法, 33  
 損失関数, 30  
 逐次 2 次計画法, 33  
 統計量, 26  
 パラメータ, 28  
 パラメータ制約条件, 31  
 微分係数, 32  
 ブートストラップ推定値, 33  
 分割モデル, 28  
 予測値, 32  
 例, 26  
 非線形モデル  
 非線型回帰分析, 29

フィデューシアル信頼区間  
 プロビット分析, 24  
 プロビット分析  
 基準, 24  
 コマンドの追加機能, 25  
 自然応答比, 24  
 相対メディアン潜在性, 24  
 統計量, 22, 24  
 範囲を定義, 24  
 反復回数, 24  
 フィデューシアル信頼区間, 24  
 平行性の検定, 24  
 例, 22

平行性の検定  
 プロビット分析, 24

密度モデル  
 非線型回帰分析, 29

主効果モデル  
 多項ロジスティック回帰, 13

ユーザーの指定によるモデル  
 多項ロジスティック回帰, 13

## 散らばり尺度値

- 多項ロジスティック回帰, 19

## ロジスティック回帰, 3

- 変数選択方法, 5

- 反復回数, 9

- 予測値, 7

- 定数項, 9

- 統計量, 3

- 係数, 3

- 対比, 6

- 残差, 7

- 例, 3

- Hosmer-Lemeshow の適合度統計, 9

- 表示オプション, 9

- カテゴリ共変量, 6

- コマンドの追加機能, 10

- ステップワイズにおける確率, 9

- ストリング共変量, 6

- 統計量と作図, 9

- 影響力の測定方法, 7

- 選択規則の定義, 5

- 新変数の保存, 7

- 規則の設定, 5

- 分類の打切り, 9

## 多項ロジスティック回帰, 11, 16

- 統計量, 16

- 保存, 20

- 基準, 18

- 参照カテゴリ, 15

- コマンドの追加機能, 21

- モデル, 13

- モデル情報のエクスポート, 20

## ロジスティック回帰

- 2 項, 1