

IBM SPSS Advanced Statistics
21



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.187）の一般情報をお読みください。

本版は IBM® SPSS® Statistics 21 ,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, 2012.

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。Advanced Statistics 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。Advanced Statistics 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に行われます。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。

内容

| | | |
|----------|--------------------------------|-----------|
| 1 | Advanced Statistics の概要 | 1 |
| 2 | GLM 多変量分析 | 2 |
| | GLM 多変量のモデル | 5 |
| | 項の構築 | 5 |
| | 平方和 | 6 |
| | GLM 多変量の対比 | 7 |
| | 対比の種類 | 8 |
| | GLM 多変量のプロファイル プロット | 8 |
| | GLM 多変量: その後の比較 | 10 |
| | GLM の保存 | 12 |
| | GLM 多変量のオプション | 14 |
| | GLM コマンドの追加機能 | 15 |
| 3 | GLM 反復測定 | 17 |
| | GLM 反復測定の因子の定義 | 21 |
| | GLM 反復測定のモデル | 23 |
| | 項の構築 | 24 |
| | 平方和 | 24 |
| | GLM 反復測定の対比 | 25 |
| | 対比の種類 | 26 |
| | GLM 反復測定のプロファイル プロット | 27 |
| | GLM 反復測定のその後の比較 | 28 |
| | GLM 反復測定の保存 | 30 |
| | GLM 反復測定のオプション | 32 |
| | GLM コマンドの追加機能 | 33 |
| 4 | 分散成分分析 | 35 |
| | 分散成分のモデル | 37 |
| | 項の構築 | 37 |

| | |
|-------------------|----|
| 分散成分のオプション | 38 |
| 平方和(分散成分) | 39 |
| 分散成分の新しいファイルへの保存 | 40 |
| VARCOMP コマンドの追加機能 | 40 |

5 線型混合モデル 42

| | |
|-----------------------|----|
| 線型混合モデルの被験者および反復変数の選択 | 44 |
| 線型混合モデルの固定効果 | 46 |
| 非入れ子の項目の構築 | 46 |
| 入れ子項目の構築 | 47 |
| 平方和 | 47 |
| 線型混合モデルの変量効果 | 48 |
| 線型混合モデルの推定 | 50 |
| 線型混合モデルの統計量 | 51 |
| 線型混合モデルの EM 平均 | 53 |
| 線型混合モデルの保存 | 54 |
| MIXED コマンドの追加機能 | 54 |

6 一般化線型モデル 56

| | |
|------------------|----|
| 一般化線型モデル: 応答 | 61 |
| 一般化線型モデルの参照カテゴリ | 62 |
| 一般化線型モデル: 予測変数 | 63 |
| 一般化線型モデルのオプション | 64 |
| 一般化線型モデル: モデル | 65 |
| 一般化線型モデル: 推定 | 67 |
| 一般化線型モデルの初期値 | 69 |
| 一般化線型モデル: 統計量 | 70 |
| 一般化線型モデル: 推定周辺平均 | 72 |
| 一般化線型モデル: 保存 | 74 |
| 一般化線型モデル: エクスポート | 76 |
| GENLIN コマンドの追加機能 | 78 |

7 一般化推定方程式

79

| | |
|------------------|-----|
| 一般化推定方程式: モデルの種類 | 83 |
| 一般化推定方程式: 応答 | 87 |
| 一般化推定方程式: 参照カテゴリ | 88 |
| 一般化推定方程式: 予測変数 | 89 |
| 一般化推定方程式: オプション | 90 |
| 一般化推定方程式: モデル | 91 |
| 一般化推定方程式: 推定 | 93 |
| 一般化推定方程式: 初期値 | 95 |
| 一般化推定方程式: 統計量 | 97 |
| 一般化推定方程式: 推定周辺平均 | 99 |
| 一般化推定方程式: 保存 | 101 |
| 一般化推定方程式: エクスポート | 103 |
| GENLIN コマンドの追加機能 | 104 |

8 一般化線型混合モデル

106

| | |
|---------------|-----|
| 一般化線型混合モデルの取得 | 108 |
| 対象 | 110 |
| 固定効果 | 114 |
| カスタム項の追加 | 115 |
| 変量効果 | 117 |
| 変量効果ブロック | 118 |
| 重みとオフセット | 120 |
| 作成オプション | 121 |
| 推定平均 | 123 |
| 保存 | 125 |
| モデルビュー | 126 |
| モデル要約 | 127 |
| データ構造 | 128 |
| 予測対観測 | 129 |
| 分類 | 130 |
| 固定効果 | 131 |
| 固定係数 | 133 |
| 変量効果共分散 | 135 |

| | |
|-------------|-----|
| 共分散パラメータ | 136 |
| 推定平均:有意な効果 | 137 |
| 推定平均:カスタム効果 | 137 |

9 モデル選択の対数線型分析 139

| | |
|-----------------------|-----|
| 対数線型分析の範囲の定義 | 140 |
| 対数線型分析のモデル | 141 |
| 項の構築 | 142 |
| モデル選択の対数線型分析のオプション | 142 |
| HILOGLINEAR コマンドの追加機能 | 143 |

10 一般的な対数線型分析 144

| | |
|------------------|-----|
| 一般的な対数線型分析のモデル | 146 |
| 項の構築 | 146 |
| 一般的な対数線型分析のオプション | 147 |
| 一般的な対数線型分析の保存 | 148 |
| GENLOG コマンドの追加機能 | 148 |

11 ロジット対数線型分析 150

| | |
|------------------|-----|
| ロジット対数線型分析のモデル | 152 |
| 項の構築 | 153 |
| ロジット対数線型分析のオプション | 154 |
| ロジット対数線型分析の保存 | 155 |
| GENLOG コマンドの追加機能 | 155 |

12 生命表 (生命表 オプション) 157

| | |
|--------------------|-----|
| 生命表の状態変数の事象の定義 | 159 |
| 生命表の範囲の定義 | 159 |
| 生命表のオプション | 160 |
| SURVIVAL コマンドの追加機能 | 160 |

13 Kaplan–Meier 生存分析 162

| | |
|------------------------------------|-----|
| Kaplan–Meier の状態変数の事象の定義 | 164 |
| Kaplan–Meier の因子レベルの比較 | 164 |
| Kaplan–Meier の新変数の保存 | 165 |
| Kaplan–Meier のオプション | 166 |
| KM コマンドの追加機能 | 167 |

14 Cox 回帰分析 168

| | |
|---------------------------------|-----|
| Cox 回帰分析におけるカテゴリ変数の定義 | 170 |
| Cox 回帰分析の作図 | 171 |
| Cox 回帰分析の新変数の保存 | 172 |
| Cox 回帰分析のオプション | 173 |
| Cox 回帰分析の状態変数の事象の定義 | 174 |
| COXREG コマンドの追加機能 | 174 |

15 時間依存の共変量の計算 175

| | |
|---------------------------------------|-----|
| 時間依存の共変量を計算するには | 176 |
| 時間依存の共変量を使用する Cox 回帰分析の追加機能 | 176 |

付録

A カテゴリ変数コード化方式 177

| | |
|-------------------|-----|
| 全平均 | 177 |
| [単純] | 178 |
| Helmert | 179 |
| 差分 (階差) | 179 |
| 多項式 | 180 |
| [反復測定] | 180 |

| | |
|-------------------------|------------|
| 特別 (ピボットテーブル GLM) | 181 |
| 指標 | 182 |
| B 共分散構造 | 183 |
| C 注意事項 | 187 |
| 索引 | 190 |

Advanced Statistics の概要

Advanced Statistics オプションを使用することにより、Statistics Base オプションから高度なモデリング オプションを適用することができます。

- GLM 多変量とは、GLM 1 変量による一般線型モデルを拡張し、複数の従属変数を扱えるようにしたものです。機能がさらに拡張された GLM 反復測定では、複数の従属変数の反復測定が可能です。
- 分散成分分析とは、1 つの従属変数が示す変動を固定成分とランダム成分に分解する分析方法です。
- 線型混合モデルは、相関関係を持った非定常的な変動が現れるデータも扱えるように、一般線型モデルを拡張したものです。したがって、線型混合モデルは、データの平均値だけでなく、データの分散や共分散もモデリングできる柔軟性を備えています。
- 一般化線型モデル (GZLM) では、誤差項が正規性を持つという前提が緩和されています。このモデルでは、従属変数と予測変数に変換またはリンク関数を介して線型関係にあることだけが必要であり、その他に要件はありません。一般化推定方程式 (GEE) を使用すると、GZLM において反復測定が可能になります。
- [一般的な対数線型分析] では、クロス分類された度数データにモデルを適合させることができます。また、[対数線型分析のモデル選択] では、複数の選択肢からのモデルの選択が可能です。
- ロジット対数線型分析では、対数線型モデルを適合させて、1 つのカテゴリ従属変数と 1 つ以上のカテゴリ予測変数との関係を分析することができます。
- 生存分析は、生命表を基にして、(場合によっては因子変数のレベルごとに) 時間イベント変数の分布を調べるためのものです。たとえば、Kaplan-Meier 生存分析では、因子変数のレベルごとに時間イベント変数の分布を調べることもできるほか、層化変数のレベルごとに分析を作成することもできます。また、Cox 回帰分析では、特定のイベントが発生するまでの経過時間を、与えられた共変量の値に基づいてモデリングすることができます。

GLM 多変量分析

GLM 多変量手続きでは、複数の従属変数の回帰分析と分散分析を 1 つ以上の因子変数か共変量を使って実行します。因子変数により、母集団をいくつかのグループに分けます。この一般線型モデル手続きを使うと、従属変数の同時分布を分けたさまざまなグループの平均値に対する因子変数の効果について、帰無仮説を検定できます。因子間の交互作用と因子ごとの効果の両方を調べることができます。さらに、共変量の効果や共変量と因子の交互作用を含めることができます。回帰分析では、独立（予測）変数は共変量として指定します。

検定は、釣り合い型モデルと不釣り合い型モデルの両方に対して実行できます。モデル内の各セルに含まれているケース数が等しければ、その計画は釣り合っています。多変量モデルの場合は、モデルの効果による平方和と誤差平方和が、1 変量の分析に見られるスカラー形式ではなく、行列形式をとります。この行列は、SSCP（平方和と交差積和）行列と呼ばれます。従属変数を 2 つ以上指定すると、各従属変数についての 1 変量の分散分析だけでなく、Pillai のトレース、Wilks のラムダ、Hotelling のトレース、Roy の最大根基準を近似 F 統計量と併用して、多変量分散分析を行うことができます。GLM 多変量手続きでは、仮説の検定の他に、パラメータの推定値を生成します。

仮説を検定する際は、一般的に使用されている事前対比を使用できます。さらに、全体的な F 検定で有意確率が判明していれば、その後の検定を使用して、特定平均値間の差分を評価できます。推定周辺平均から、モデルに含まれるセルの予測平均値を推定できるとともに、これらの平均値のプロファイル プロット（交互作用プロット）を使用して一部の関係を簡単に視覚化できます。その後の多重比較検定は、それぞれの従属変数ごとに、個別に行われます。

残差、予測値、Cook の距離、てこ比の値は、データ ファイルに新変数として保存し、仮定の確認に利用できます。また、残差の平方和と積和の正方行列である残差 SSCP 行列、残差 SSCP 行列を残差の自由度で割った残差の分散共分散行列、そして残差の分散共分散行列を標準化した形式である残差の相関行列も利用できます。

[WLS 重み] を使うと、重み付き最小 2 乗法 (WLS) 分析の場合、観測値に異なる重みを与えるために使う変数を指定し、それにより異なる測定精度を補足できます。

例: プラスチック製造業者は、プラスチック フィルムの引き裂き抵抗、光沢、不透明度の 3 つのプロパティを測定します。2 種類の比率の押し出しと、量が 2 種類違う添加物を試し、押し出し率と添加物の量のそれぞれについて 3 種類の属性を測定します。製造会社で判明したことは、押し出し

率と添加物の量をもたらす影響は個々には大きいものの、2つの要因の交互作用はそれほど大きくないということです。

方法。異なる仮説を評価する場合は、タイプ I、タイプ II、タイプ III、タイプ IV 平方和を使用できます。デフォルトはタイプ III です。

統計量。その後の範囲検定と多重比較：最小有意差、Bonferroni の方法、Sidak の方法、Scheff の検定、Ryan-Einot-Gabriel-Welsch の多重 F 値、Ryan-Einot-Gabriel-Welsch の多重範囲、Student-Newman-Keuls の検定、Tukey の HSD 検定、Tukey の b 検定、Duncan の方法、Hochberg の GT2、Gabriel の方法、Waller-Duncan の t 検定、Dunnnett の方法（片側と両側）、Tamhane の T2、Dunnnett の T3、Games-Howell の方法、および Dunnnett の C。記述統計：すべてのセルにおけるすべての従属変数の観測平均値、標準偏差および度数、Levene の等分散性検定、従属変数の分散共分散行列の等質性に関する Box の M 検定、および Bartlett の球面性検定。

作図。水準と広がり、残差、およびプロファイル（交互作用）。

データ。従属変数は量的変数でなければなりません。因子はカテゴリ変数で、数値または文字値を持つことができます。共変量は、従属変数に関連する量的変数です。

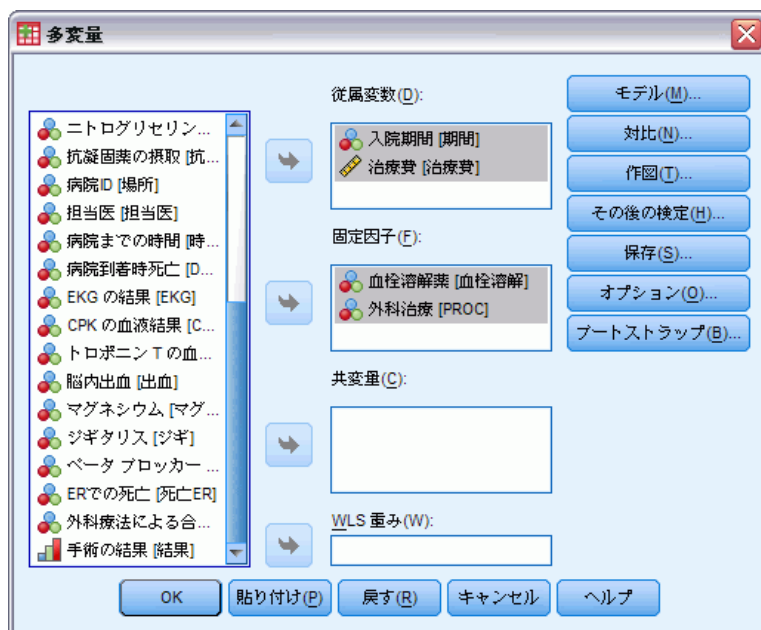
仮定。従属変数については、データは多変量正規母集団から無作為抽出したベクトルです。母集団では、すべてのセルの分散共分散行列は同じです。データは対称であるべきですが、分散分析は正規性からの逸脱に対し頑健です。仮定のチェックには、等分散性の検定（Box の M を含む）、および水準と広がり、残差と残差プロットについて探索的分析を行うこともできます。

関連手続き。探索的分析を使って、分散分析を行う前にデータを調べます。従属変数が単一の場合は、[GLM 1 変量]を使用します。同じ従属変数について、各被験者ごとにいくつかの状況を測定する場合 [GLM 反復測定]を使用します。

GLM 多変量テーブルを取得するには

- ▶ メニューから次の項目を選択します。
分析 > 一般線型モデル > 多変量...

図 2-1
[多変量] ダイアログ ボックス

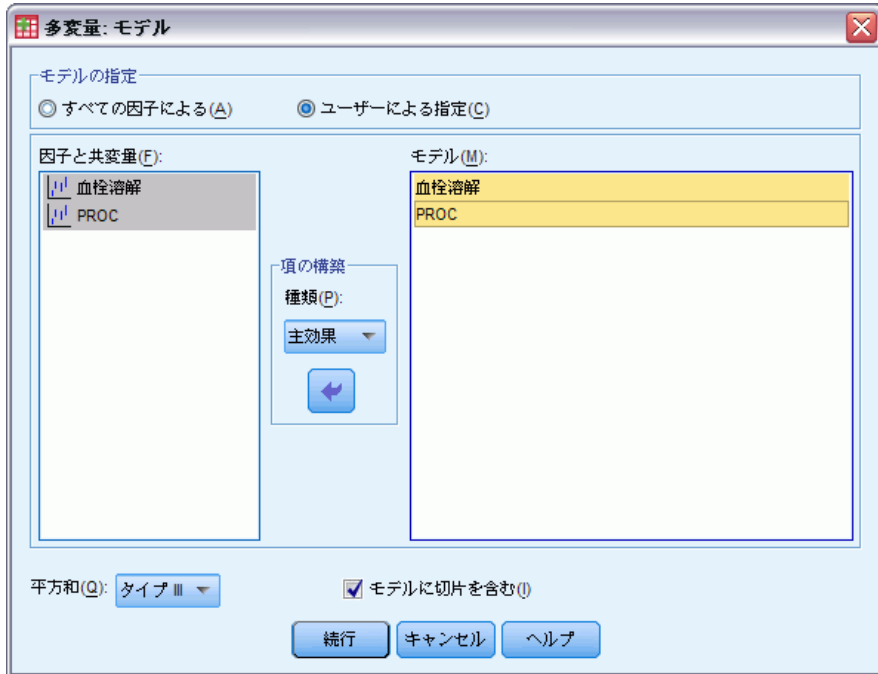


- ▶ 2 つ以上の従属変数を選択します。

オプションとして、[固定因子]、[共変量]、[WLS 重み] を指定することもできます。

GLM 多変量のモデル

図 2-2
[多変量: モデル] ダイアログ ボックス



モデルの指定。 すべての因子によるモデルには、因子の主効果、共変量の主効果、因子間の交互作用のすべてが含まれます。共変量の交互作用は含まれません。交互作用の一部だけを指定する場合や、因子と共変量の交互作用を指定する場合は、[ユーザーの指定] をクリックします。モデルに含める項目はすべて指示する必要があります。

因子と共変量。 要素および共変量はリストされます。

モデル。 モデルは、使用するデータの性質によって異なります。[ユーザーの指定] をクリックすれば、分析対象の主効果と交互作用を選択できます。

平方和。 平方和の計算方法。釣り合い型モデルと不釣り合い型モデルで欠損セルがない場合は、タイプ III の平方和の方法が最もよく利用されます。

モデルに切片を含む。 通常、モデルには定数項が含まれます。データが原点を通ると仮定できる場合は、切片を除外できます。

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。これはデフォルトです。

- 主効果。**選択した変数のそれぞれに主効果の項を作成します。
- 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。

平方和

モデルには、平方和のタイプを選択できます。最も一般的に使われ、デフォルトでもあるのがタイプ III です。

タイプ I。この方法は、平方和の階層的分割法という名でも知られています。各項は、モデル内で先行する項目に対してだけ調整されます。タイプ I の平方和は、一般に次のような場合に使用します。

- 1 次の交互作用の効果より先になんらかの主効果が指定され、2 次の交互作用の効果より先になんらかの 1 次の交互作用の効果が指定されているといった具合の分散分析の釣り合い型モデル。
- 高次の項より先に低次の項目が指定されている多項式回帰モデル。
- 最初に指定された効果が 2 番目に指定された効果の中に入れ子になり、2 番目に指定された効果が 3 番目に指定された効果の中に入れ子になり、という具合に続く純粋な入れ子モデル。(入れ子の形式は、シンタックスを使用しなければ指定できません)。

タイプ II。この方法では、他のすべての「適当な」効果に対して調整されるモデルの効果の平方和が計算されます。適当な効果とは、調査中の効果を含んでいないすべての効果に対応するものです。タイプ II の平方和の方法は、一般に次のような場合に使用します。

- 分散分析の釣り合い型モデル。
- 因子の主効果しか持たないモデル。
- 回帰モデル。
- 純粋に入れ子になっている計画。(入れ子の形式は、シンタックスを使用して指定できます)

タイプ III。デフォルトです。この方法では、計画内にある効果の平方和を、その計画を含まないその他の効果、またはそうした効果に対して直交的な効果を対象に調整した平方和として計算します。タイプ III の平方和には、一般的な推定形式が一定の状態に保たれている限り、セル度数が変わらないという大きな利点があります。したがって、このタイプの平方和は、欠損セルがない不釣り合い型モデルに有用だと考えられます。欠損セルのない多因子計画でこの方法に相当するのが、Yates の平均値の重み付き 2 乗法です。タイプ III の平方和の方法が一般的に使用されるのは次の場合です。

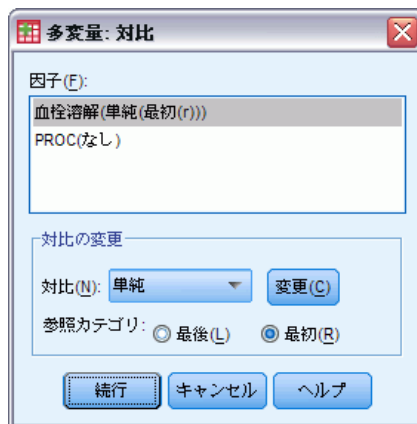
- タイプ I とタイプ II に記載されているモデル。
- 空白セルのない釣り合い型モデルまたは不釣り合い型モデル。

タイプ IV。この方法は、欠損セルが存在する状況を目的としたものです。計画における効果 F に対して、F が他のどの効果にも含まれていないとき、タイプ IV = タイプ III = タイプ II となります。F が他の効果に含まれているとき、タイプ IV は、F におけるパラメータ間で行われている対比を、より高いレベルの効果のすべてに等しく分配します。タイプ IV の平方和の方法は、一般に次のような場合に使用します。

- タイプ I とタイプ II に記載されているモデル。
- 空白セルを伴う釣り合い型モデルまたは不釣り合い型モデル。

GLM 多変量の対比

図 2-3
[多変量: 対比] ダイアログ ボックス



対比は、効果の水準が互いに著しく相異しているかどうかを検定する場合に使用します。対比は、モデルの因子ごとに指定できます。対比は、パラメータの線型結合を表します。

仮説検定は、 $L\mathbf{B} = \mathbf{0}$ という帰無仮説を前提とします。ただし、 L は対比係数行列、 M は単位行列です。単位行列の次元は、従属変数の数に等しくなります。 B はパラメータ ベクトルです。対比を指定すると、因子に対応する列が対比と適合するように L 行列を作成します。残りの列は、 L 行列が推定できるように調整されます。

F 統計量、および Bonferroni の同時信頼区間（すべての従属変数に関する対比の差異についてのスチューデント t 分布に基づく）を使用する 1 変量検定に加え、Pillai のトレース、Wilks のラムダ、Hotelling のトレース、Roy の最大根基準を使用する多変量検定を実行できます。

利用できる対比には、偏差、単純、差分、Helmert、反復測定、多項式があります。偏差対比と単純対比については、参照カテゴリの最初か最終を選択できます。

対比の種類

偏差。各レベルの平均値（参照カテゴリを除く）をすべてのレベルの平均値（全平均）と比較します。因子の水準は任意の順序になります。

単純。各レベルの平均値を特定のレベルの平均値と比較します。この対比の型はコントロール群があるときに有効です。最初または最後のカテゴリを参照として選択できます。

差分。各レベル（最初を除く）の平均値を前のレベルの平均値と比較します（「逆 Helmert 対比」と呼ばれることもあります）。

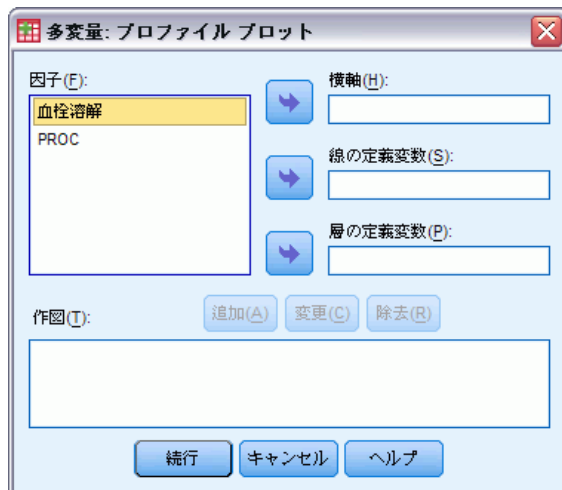
Helmert。因子の各水準（最終を除く）の平均とその後の水準の平均とを比較します。

反復測定。因子の各水準（最終を除く）の平均とその後の水準の平均とを比較します。

多項式。1 次効果、2 次効果、3 次効果等を比較します。第 1 自由度にはすべてのカテゴリを通じての 1 次効果が含まれ、第 2 自由度には 2 次効果が含まれるなど、それぞれが含まれます。このような対比は、多項式のトレンドを推定する場合に頻繁に使用します。

GLM 多変量のプロファイル プロット

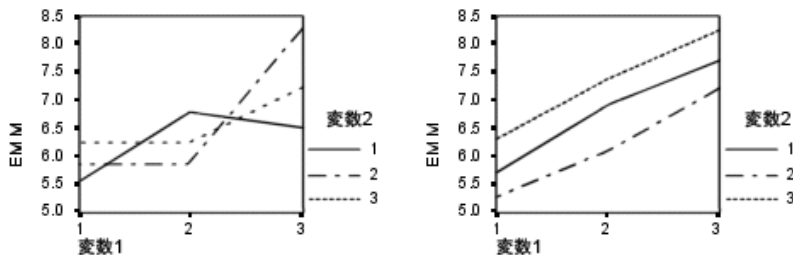
図 2-4
[多変量: プロファイル プロット] ダイアログ ボックス



プロファイル プロット (交互作用プロット) は、モデル中の周辺平均を比較するとき役に立ちます。プロファイル プロットは、点を使って、因子の 1 つの水準で従属変数 (共変量を対象に調整) の推定周辺平均を示す線のプロットです。第 2 因子の水準を使って、個別の線を作成できます。第 3 因子の各水準を使用して、層別のプロットを作成できます。すべての因子が作図で使用できます。プロファイル プロットは各従属変数ごとに作成されます。

1 つの因子についてのプロファイル プロットから、推定周辺平均が水準全体で増えているか減っているかがわかります。2 つ以上の因子では、因子間で交互作用がないと平行線で示されるので、1 つの因子に限ってそのレベルを調べることができます。平行ではない線は、交互作用を示しています。

図 2-5
平行ではないプロット (左) と平行プロット (右)



水平軸の因子を選択してプロットを指定し、オプションとして線の定義変数と層の定義変数を選択したら、そのプロットは [作図] ボックスの一覧に追加しなければなりません。

GLM 多変量: その後の比較

図 2-6
[多変量: 観測平均のその後の多重比較] ダイアログ ボックス



その後の多重比較検定。 平均値の間に差があることが判明した後は、その後の範囲検定とペアごとの多重比較により、どの平均値が相異しているのかを決めることができます。比較は調整済みでない値に基づいて行われます。その後の検定は、各従属変数別に実行します。

多重比較検定では、通常、Bonferroni 検定と Tukey の HSD 検定を使用します。**Bonferroni の検定**では、スチューデントの t 検定統計量に基づいて、多重比較が行われるという事実に対して有意水準が調整されます。**Sidak の t 検定**でも有意水準が調整され、Bonferroni の方法よりも厳しく限定されます。**Tukey の HSD 検定**では、スチューデント化された範囲統計量を使ってすべてのペアごとの比較をグループ間で行い、実験ごとの誤差率をすべてのペアごとの比較の集合の誤差率に設定します。大量の平均値のペアを検定する場合は、Tukey の HSD 検定の方が Bonferroni 検定より有効です。少量のペアの場合は Bonferroni の方法の方が有効です。

Hochberg の GT2 検定は Tukey の HSD 検定と類似していますが、ここで使用するのはスチューデント化された最大法です。一般的には、Tukey の検定の方が有効です。**Gabriel のペアごとの比較検定**も、スチューデント化された最大法を使用しますが、一般的に、セルの大きさが均等ではない

場合は、Hochberg の GT2 より有効です。セルの大きさのばらつきが大きい場合には、Gabriel の検定の方が公平になることがあります。

Dunnett のペアごとの多重比較 t 検定では、処理のグループを単一の対照平均値と比較します。最後のカテゴリは、デフォルトの対照カテゴリです。代わりに、最初のカテゴリを選択できます。両側または片側の検定を選択することもできます。因子の任意のレベルの平均値（対照カテゴリを除く）が対照カテゴリのそれと等しくないことを検定するには、両側の検定を使います。因子の任意のレベルの平均値が対照カテゴリの平均値よりも小さいことを検定するには、**[<対照カテゴリ]**を選択します。同様に、因子の任意のレベルの平均値が対照カテゴリの平均値より大きいことを検定するには、**[>対照カテゴリ]**を選択します。

Ryan と Einot、Gabriel、Welsch (R-E-G-W) は、2 種類のステップダウン多重範囲検定を開発しました。ステップダウン多重手続きでは、最初にすべての平均値が等しいかどうかを検定します。すべての平均値が等しくなければ、サブグループの平均値が等しいかどうかを検定します。**R-E-G-W** の F 値は F 検定に基づき、**R-E-G-W** の Q 値はスチューデント化した範囲に基づきます。この検定は、Duncan の多重範囲検定や Student-Newman-Keuls の検定（これもステップダウン多重手続き）より有効ですが、セルの大きさが等しくない場合はお勧めできません。

分散が等しくない場合は、**Tamhane の T2** (t 検定に基づくペアごとの控えめな比較)、**Dunnett の T3** (スチューデント化した最大偏差に基づくペアごとの比較検定)、**Games-Howell の ペアごとの比較検定** (公平な場合もある)、または **Dunnett の C** (スチューデント化した範囲に基づくペアごとの比較検定) を使用します。

Duncan の多重範囲検定、Student-Newman-Keuls (**S-N-K**) の方法、および **Tukey の b 検定**はグループ平均を順位付け、範囲の値を計算する範囲検定です。これらの検定は先に述べた検定ほど頻繁には使われていません。

Waller-Duncan の t 検定では、Bayesian のアプローチが使われています。この範囲検定では、サンプル サイズが等しくない場合にサンプルサイズの調和平均が使われます。

Scheffé の検定の有意水準は、この機能で利用できるペアごとの比較だけにとどまらず、グループ平均に可能なすべての線型結合を検定できる設計になっています。結果的に、Scheffé の検定は他の検定より控えめになってしまうことが多いため、有意確率を求める場合は、平均値間の差が大きい必要があります。

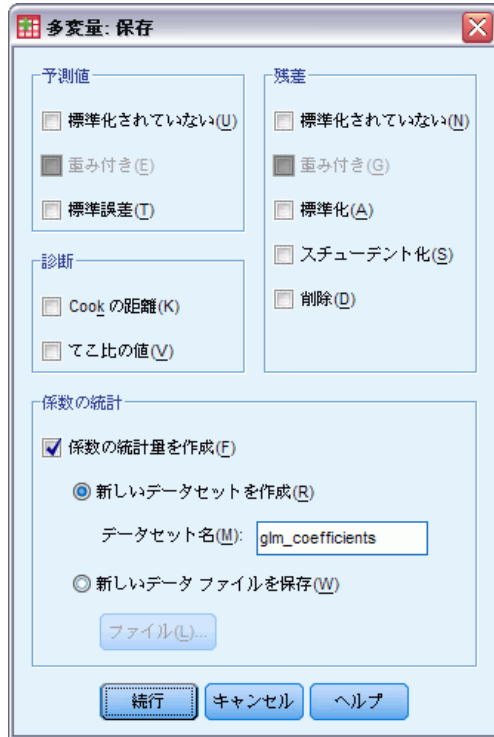
最小有意差 (**LSD**) のペアごとの多重比較検定は、グループのすべてのペア間ごとの多重 T 検定に相当します。この検定の欠点は、観測された有意水準を多重比較向けに調整する試みが一切行われなことです。

表示される検定。 ペアごとの比較は、LSD、Sidak の方法、Bonferroni の方法、Games と Howell の方法、Tamhane の T2 と T3、Dunnett の C と Dunnett の T3 で使用できます。範囲検定の等質サブグループは、S-N-K、Tukey の b、Duncan、R-E-G-W の F、R-E-G-W の Q および Waller の方法

で使用できます。Tukey の HSD 検定、Hochberg の GT2、Gabriel の検定、および Scheffé の検定は、多重比較検定と範囲検定の両方です。

GLM の保存

図 2-7
[保存] ダイアログ ボックス



モデルが予測した値、残差、関連測定値は、データ エディタで新変数として保存できます。これらの変数の多くは、データの仮定を調べるために利用できます。値を保存して別の IBM® SPSS® Statistics セッションで利用するためには、現在のデータ ファイルを保存する必要があります。

予測値。 モデルがケースごとに予測する値。

- **標準化されていない (判別分析).** 従属変数を予測するモデルの値。
- **重み付き.** 重み付きの標準化されていない予測値。WLS 変数がすでに選択されている場合にのみ使用することができます。
- **標準誤差.** 独立変数の同じ値を持つケースに対する従属変数の平均値の標準偏差の推定値。

診断。 独立変数の一般的でない値の組み合わせのケースとモデルに大きな影響を及ぼすケースを識別するための測定方法。

- **Cook の距離.** 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。Cook の距離が大きいつきは、回帰統計量の計算からケースを除外したことが係数を実質的に変化させたことを示しています。
- **てこ比の値.** 非心てこ比の値。モデルの適合度に関する各観測の相対的な影響度。

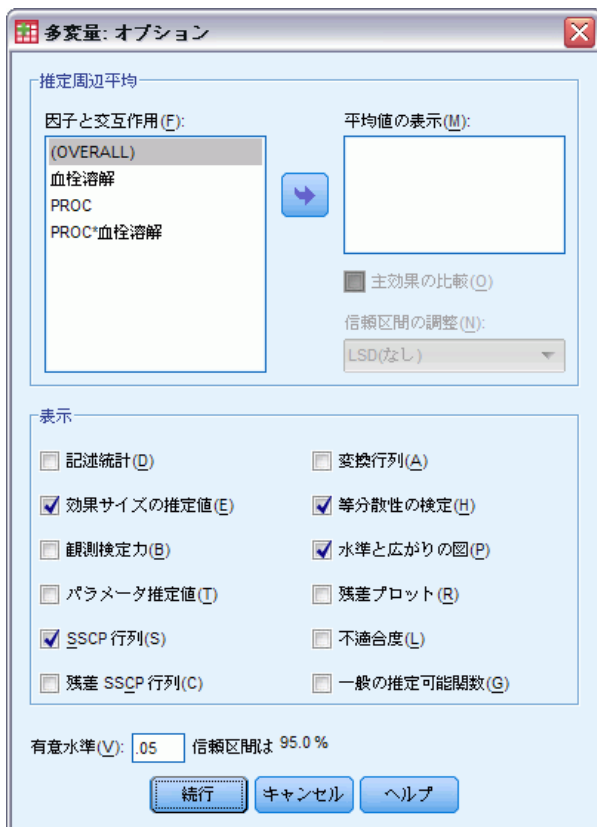
残差. 標準化されていない残差は、従属変数の実際の値からモデルにより予測される値を引いたものです。標準化された残差、スチューデント化された残差、削除された残差も使用できます。WLS 変数を選択すると、重み付けした標準化されていない残差が使用できます。

- **標準化されていない (判別分析).** 観測値およびモデルで予測された値との差。
- **重み付き.** 重み付きの標準化されていない残差。WLS 変数がすでに選択されている場合にのみ使用することができます。
- **標準化残差.** 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。
- **スチューデント化された残差.** 残差を、独立変数の平均値からの独立変数の各ケース値の距離に依存して、ケースごとに違う標準偏差の推定量で割ったもの。
- **削除.** あるケースが回帰係数の計算から除外されたときのケースの残差。従属変数と調整済み予測値の間の差です。

係数統計量. モデルにおけるパラメータ推定値の分散共分散行列を、現在のセッションにある新しいデータセット、または SPSS Statistics 形式の外部データ ファイルに出力します。また、それぞれの従属変数に対しても、パラメータ推定値の行、パラメータ推定値に対応する t 統計量の有意確率の行、および残差自由度の行が存在します。多変量モデルの場合は、各従属変数に同様の行があります。行列ファイルを読み込む別の手続きで、この行列ファイルを使用できます。

GLM 多変量のオプション

図 2-8
[多変量: オプション] ダイアログ ボックス



このダイアログ ボックスでオプションの統計を利用できます。統計量は、固定効果モデルを使用して計算されます。

推定周辺平均。セルにおける母周辺平均値について推定したい因子と交互作用を選択します。共変量が存在する場合、これらの平均値は、共変量に対して調整されます。交互作用は、ユーザーの指定によるモデルを指定している場合にだけ使用できます。

- **主効果の比較。** 被験者間と被験者内因子の両方について、モデル内の主効果に対する推定周辺平均値間で、ペアごとに無補正の比較を行います。この項目は、[平均値の表示] リストで主効果を選択した場合にだけ選択できます。
- **信頼区間の調整。** 最小有意差 (LSD)、Bonferroni の方法、Sidak の方法の信頼区間と有意水準に対する調整を選択します。この項目は、[主効果の比較] を選択している場合にだけ選択できます。

[表示][記述統計量] を選択すると、すべてのセルにおけるすべての従属変数の観測平均値、標準偏差、および度数を求めることができます。**[効果サイズの推定値]** は、偏相関のイータの 2 乗の値をそれぞれの効果とそれぞれのパラメータ推定値に与えるものです。イータの 2 乗統計量は、因子に起因する総変動の比率を記述するものです。観測した値に基づいて対立仮説を立てるときに検定の検出力を得るには、**[観測検定力]** を選択します。パラメータ推定値、標準誤差、T 検定、信頼区間、およびそれぞれの検定での観測検定力を創出するには、**[パラメータ推定値]** を選択します。残差分散共分散行列についての仮説、誤差 SSCP 行列、残差 SSCP 行列プラス Bartlett の球面性の検定を表示できます。

等分散性の検定は、被験者間因子の場合にだけ、各従属変数の分散の等質性分析について、被験者間因子のすべての水準の組み合わせを通じて Levene の検定を創出するものです。等分散性の検定は、被験者間因子のすべての水準の組み合わせを通じて、従属変数の分散共分散行列の等質性分析についての Box の M 検定をも含んでいます。水準と広がり の図および残差プロットのオプションは、データに関する仮定のチェックに役立ちます。この項目は、因子が存在しないときは無効になります。それぞれの従属変数について、標準化された予測による観測残差プロットを作成するには、**[残差プロット]** を選択します。このようなプロットは、分散が等しいという仮定の検証に役立ちます。従属変数と独立変数の間の関係がモデルで正しく記述されているかどうかをチェックするには、**[不適合度]** を選択します。一般の推定可能関数を使うと、一般推定可能関数に基づいてユーザー指定の仮説の検定を構築できます。任意の対比係数行列における行は、一般推定可能関数の一次結合です。

有意水準。 信頼区間を構築するために、その後の検定や信頼係数で使われる有意水準を調整したい場合があります。指定された値は、検定の観測検定力の計算に使用できます。有意水準を指定するときは、信頼区間の関連水準がダイアログ ボックスに表示されます。

GLM コマンドの追加機能

この機能は、1 変量、多変量、反復測定分析に適用できます。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。

- 計画中の入れ子効果の指定 (DESIGN サブコマンドを使用)。
- 効果と効果や値の線型組み合わせ検定の指定 (TEST サブコマンドを使用)。
- 多重対比の指定 (CONTRAST サブコマンドを使用)。
- ユーザー欠損値の包含 (MISSING サブコマンドを使用)。
- EPS 基準の指定 (CRITERIA サブコマンドを使用)。
- ユーザー指定の L 行列、M 行列、K 行列の作成 (LMATRIX、MMATRIX、または KMATRIX サブコマンドを使用)。

- 全平均対比または単純対比での、中間参照カテゴリの指定 (CONTRAST サブコマンドを使用)。
- 多項式対比の計量の指定 (CONTRAST サブコマンドを使用)。
- その後の比較の誤差項の指定 (POSTHOC サブコマンドを使用)。
- 因子または因子の一覧にある因子間の交互作用の推定周辺平均値の計算 (EMMEANS サブコマンドを使用)。
- 仮の変数の名前の指定 (SAVE サブコマンドを使用)。
- 相関行列のデータ ファイルの作成 (OUTFILE サブコマンドを使用)。
- 被験者間因子の分散分析表の統計量を含む行列データ ファイルの作成 (OUTFILE サブコマンドを使用)。
- 計画行列の新しいデータ ファイルへの保存 (OUTFILE サブコマンドを使用)。

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。

GLM 反復測定

[GLM 反復測定] 手続きでは、それぞれの被験者、つまり、ケースに同じ測定を繰り返す場合に分散分析を実施します。被験者間因子を指定すると、母集団がいくつかのグループに分割されます。この一般線型モデル手続きを使うと、被験者間因子と被験者内因子の両方の効果について、帰無仮説を検定できます。因子間の交互作用と因子ごとの効果の両方を調べることができます。その他に、被験者間因子の定共変量の効果と共変量の交互作用を含めることができます。

2 重多変量の反復測定計画の場合、従属変数は、被験者内因子のさまざまなレベルに対して変数が複数の測定を表します。たとえば、被験者ごとに時間を変えて 3 回、脈と呼吸の両方を測定するなどができます。

[GLM 反復測定] 手続きでは、反復測定データに対する 1 変量と多変量の両方の分析を行います。検定は、釣り合い型モデルと不釣り合い型モデルの両方に対して実行できます。モデル内の各セルに含まれているケース数が等しければ、その計画は釣り合っています。多変量モデルの場合は、モデルの効果による平方和と誤差平方和が、1 変量の分析に見られるスカラー形式ではなく、行列形式をとります。この行列は、SSCP (平方和と交差積和) 行列と呼ばれます。[GLM 反復測定] では、仮説の検定の他に、パラメータの推定値を生成します。

被験者間因子についての仮説を検定するときには、一般的に使われている事前対比を利用できます。さらに、全体的な F 検定で有意確率が判明していれば、その後の検定を使用して、特定平均値間の差分を評価できます。推定周辺平均から、モデルに含まれるセルの予測平均値を推定できるとともに、これらの平均値のプロファイル プロット (交互作用プロット) を使用して一部の関係を簡単に視覚化できます。

残差、予測値、Cook の距離、てこ比の値は、データ ファイルに新変数として保存し、仮定の確認に利用できます。また、残差の平方和と積和の正方行列である残差 SSCP 行列、残差 SSCP 行列を残差の自由度で割った残差の分散共分散行列、そして残差の分散共分散行列を標準化した形式である残差の相関行列も利用できます。

[WLS 重み] を使うと、重み付き最小 2 乗法 (WLS) 分析の場合、観測値に異なる重みを与えるために使う変数を指定し、それにより異なる測定精度を補足できます。

例。 12人の生徒が不安度の検定における得点に基づいて、高い不安度グループから低い不安度グループまでに分けられています。不安度は、被験者をグループ分けすることから、「被験者間因子」と呼ばれます。生徒達には、各自学習に関する 4 種類の試行が与えられ、各試行ごとのエラーの数が記録されます。各試行のエラーは別々の変数に記録され、被

験者内因子（試行）は 4 種類の試行に対して 4 つのレベルで定義されます。試行の効果は有意であることがわかりますが、不安度別試行の交互作用は有意ではありません。

方法。異なる仮説を評価する場合は、タイプ I、タイプ II、タイプ III、タイプ IV 平方和を使用できます。デフォルトはタイプ III です。

統計量。その後の範囲検定と多重比較（被験者間因子について）：最小有意差、Bonferroni の方法、Sidak の方法、Scheffé の検定、Ryan-Einot-Gabriel-Welsch の多重 F 値、Ryan-Einot-Gabriel-Welsch の多重範囲、Student-Newman-Keuls の検定、Tukey の HSD 検定、Tukey の b 検定、Duncan の方法、Hochberg の GT2、Gabriel の方法、Waller Duncan の T 検定、Dunnett の方法（片面と両面）、Tamhane の T2、Dunnett の T3、Games-Howell の方法、および Dunnett's の C。記述統計：すべてのセルにおけるすべての従属変数の観測平均値、標準偏差および度数、Levene の等分散性検定、Box の M、および Mauchly の球面性検定。

作図。水準と広がり、残差、およびプロファイル（交互作用）。

データ。従属変数は量的変数でなければなりません。被験者間因子では、たとえば男性と女性といった具合に、サンプルが個別のサブグループに分けられます。因子はカテゴリ変数で、数値または文字値を持つことができます。被験者内因子は、[反復測定]の因子の定義ダイアログボックスで定義されます。共変量は、従属変数に関連する量的変数です。反復測定分析の場合、これらは被験者内変数のそれぞれのレベルで常に一定でなければなりません。

データ ファイルには、被験者の測定グループごとの変数セットが含まれている必要があります。このセットは、グループ内の反復測定 1 回に対して 1 つの変数を持ちます。被験者内因子は、そのグループを対象に、反復回数に等しいレベル数で定義されます。たとえば、何日かに分けて体重測定を行うなどができます。同じ特性の測定値を 5 日に分けて採取したとすると、被験者内因子は 5 つのレベルを伴う「日数」で指定できます。

被験者内因子が複数ある場合は、被験者ごとの測定回数が因子ごとのレベル数の積と等しくなります。たとえば、1 日 3 回の測定を 4 日行くと、合計測定数は、被験者ごとに 12 回となります。被験者内因子は、「日数 (4)」と「回数 (3)」で指定できます。

仮定。反復測定分析には、1 変量と多変量の 2 通りの方法があります。

1 変量のアプローチ（分割プロットによるアプローチ、または混合モデルによるアプローチという名でも知られています）は、従属変数を被験者内因子のレベルに対する応答数と見なします。被験者の測定値は、多変量の正規分布からのサンプルでなければならず、分散共分散行列は、被験者間効果によって形成されるセル全体で同じです。従属変数の分散共分散行列で、特定の仮定が立てられます。1 変量のアプローチに使用された F 値の統計量は、分散共分散行列が循環形式をとっていれば、必ず有効になります (Huynh と Mandeville、1979 年)。

この仮定を検定するには、Mauchly の球面性の検定が使用できます。この検定では、正規直交型の変換従属変数の分散共分散行列について球面性の検定を行います。反復測定分析の場合は、自動的に Mauchly の検定が表示されます。サンプル サイズが小さい場合には、この検定はそれほど有効ではありません。サンプル サイズが大きいと、この検定は、結果に対する逸脱の影響が小さくても有意な場合があります。検定の有意確率が大きい場合は、球面性の仮説を立てることができます。しかし、有意確率が小さく、球面性の仮定に反しているような場合には、分子と分母の自由度を調整して、1 変量 F 統計量の有効性を確認できます。[GLM 反復測定] 手続きでは、**イプシロン**という、この調整の 3 つの推定値を利用できます。自由度の分子と分母のどちらにもイプシロンを掛ける必要があります。F 比の有意確率は、その新たな自由度を使用して評価する必要があります。

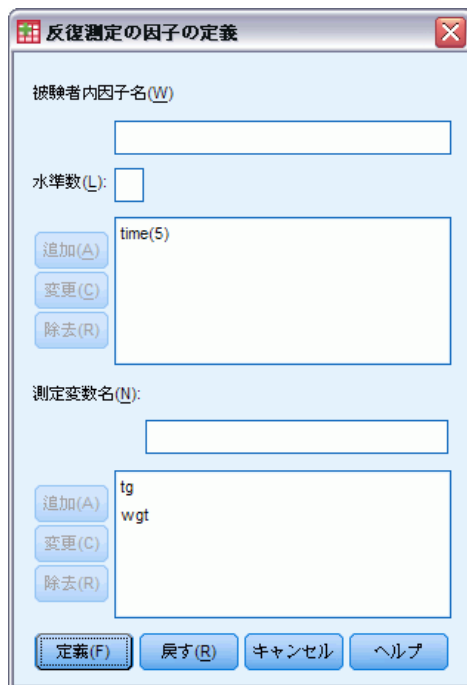
多変量のアプローチは、被験者の測定値を多変量の正規分布のサンプルと見なし、分散共分散行列は、被験者間効果によって形成されるセル全体で同じです。すべてのセルで分散共分散行列が同じかどうかを検定するには、Box の M 検定を利用できます。

関連手続き。 探索的分析を使って、分散分析を行う前にデータを調べます。それぞれの被験者に関する反復測定が存在しない場合は、[GLM 1 変量] または [GLM 多変量] を使用します。各被験者に対する測定が 2 種類（たとえば、事前検定や事後検定）しかなく、しかも被験者間因子が存在しない場合は、[対応のあるサンプルの t 検定] 手続きを使用できます。

GLM 反復測定を実行するには

- ▶ メニューから次の項目を選択します。
[分析] > [一般線型モデル] > [反復測定...]

図 3-1
[反復測定 of 因子の定義] ダイアログ ボックス



- ▶ [被験者内因子名] と [水準数] を入力します。
- ▶ [追加] をクリックします。
- ▶ 被験者内因子ごとにこの手順を繰り返します。
二重多変量の反復測定計画の測定因子を定義するには、
- ▶ [測定変数名] に入力します。
- ▶ [追加] をクリックします。
因子と測定変数をすべて定義したら、
- ▶ [定義] をクリックします。

図 3-2
[反復測定] ダイアログ ボックス



- ▶ [被験者内変数] ボックスの一覧内の被験者内因子（およびオプションとして測定方法）の各組み合わせに対応する従属変数を選択します。

変数の位置を変更するには、上または下向きの矢印をクリックします。

被験者内因子に変更を加える場合、メイン ダイアログ ボックスを閉じないで、[反復測定の因子の定義] ダイアログ ボックスを開くことができます。オプションとして、被験者間因子と共変量を指定できます。

GLM 反復測定の因子の定義

GLM 反復測定では、同じ属性のさまざまな測定値を表す関連従属変数のグループを分析します。[因子の定義] ダイアログ ボックスでは、GLM 反復測定で利用する被験者内因子を 1 つ以上定義できます。図 3-1 p. 20 を参照してください。被験者内因子を指定するときの次数は重要なので、注意してください。各因子が、その前の因子のレベルを構成します。

反復測定を使用する場合は、データを正しく設定する必要があります。このダイアログ ボックスで被験者内因子を定義する必要があります。これらの因子は、データ内の既存の変数ではなく、ここで定義する因子であることに注意してください。

例。 体重減少に関する研究で、数人の体重を 5 週間にわたって毎週測定すると仮定します。データ ファイルでは、それぞれの人が被験者またはケースとなります。数週間にわたる体重は変数 `weight1`、`weight2` などに記録します。別の変数には各人の性別を記録します。体重は、各被験者ごとに反復して測定し、被験者内因子を定義することでグループ分けできます。因子は `week` という名前にして、5 つのレベルがあると定義します。メインのダイアログ ボックスでは、変数 `weight1 ... weight5` を使って `week` の 5 つのレベルを割り当てます。男性と女性にグループ分けされたデータ ファイル内の変数 (`gender`) は、男性と女性による相違点を研究するための被験者間因子として指定できます。

測定方法。 被験者を毎回複数の測定方法で検定した場合には、測定方法を定義します。たとえば、脈拍と呼吸数をそれぞれの被験者ごとに 1 週間測定するとします。これらの測定方法は、データ ファイルには変数として存在しませんが、ここで定義できます。1 種類以上の測定方法を持つモデルを、2 重多変量反復測定モデルと呼ぶこともあります。

GLM 反復測定モデル

図 3-3
[反復測定: モデル] ダイアログ ボックス



モデルの指定。 すべての因子によるモデルには、因子の主効果、共変量の主効果、因子間の交互作用のすべてが含まれます。共変量の交互作用は含まれません。交互作用の一部だけを指定する場合や、因子と共変量の交互作用を指定する場合は、[ユーザーの指定] をクリックします。モデルに含める項目はすべて指示する必要があります。

被験者間。 要素および共変量はリストされます。

モデル。 モデルは、使用するデータの性質によって異なります。[ユーザーの指定による] を選択すると、分析したい被験者内効果と交互作用、および被験者間効果と交互作用を選択できます。

平方和。 被験者間モデルの場合の平方和の計算方法。欠損セルのある釣り合い型または不釣り合い型被験者間モデルの場合、最も一般的に使われる手法はタイプ III 平方和です。

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。これはデフォルトです。

主効果。選択した変数のそれぞれに主効果の項を作成します。

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。

平方和

モデルには、平方和のタイプを選択できます。最も一般的に使われ、デフォルトでもあるのがタイプ III です。

タイプ I。この方法は、平方和の階層的分割法という名でも知られています。各項は、モデル内で先行する項目に対してだけ調整されます。タイプ I の平方和は、一般に次のような場合に使用します。

- 1 次の交互作用の効果より先になんらかの主効果が指定され、2 次の交互作用の効果より先になんらかの 1 次の交互作用の効果が指定されているといった具合の分散分析の釣り合い型モデル。
- 高次の項より先に低次の項目が指定されている多項式回帰モデル。
- 最初に指定された効果が 2 番目に指定された効果の中に入れ子になり、2 番目に指定された効果が 3 番目に指定された効果の中に入れ子になり、という具合に続く純粋な入れ子モデル。(入れ子の形式は、シンタックスを使用しなければ指定できません)。

タイプ II。この方法では、他のすべての「適当な」効果に対して調整されるモデルの効果の平方和が計算されます。適当な効果とは、調査中の効果を含んでいないすべての効果に対応するものです。タイプ II の平方和の方法は、一般に次のような場合に使用します。

- 分散分析の釣り合い型モデル。
- 因子の主効果しか持たないモデル。
- 回帰モデル。
- 純粋に入れ子になっている計画。(入れ子の形式は、シンタックスを使用して指定できます)

タイプ III。デフォルトです。この方法では、計画内にある効果の平方和を、その計画を含まないその他の効果、またはそうした効果に対して直交的な効果を対象に調整した平方和として計算します。タイプ III の平方和には、一

一般的な推定形式が一定の状態に保たれている限り、セル度数が変わらないという大きな利点があります。したがって、このタイプの平方和は、欠損セルがない不釣り合い型モデルに有用だと考えられます。欠損セルのない多因子計画でこの方法に相当するのが、Yates の平均値の重み付き 2 乗法です。タイプ III の平方和の方法が一般的に使用されるのは次の場合です。

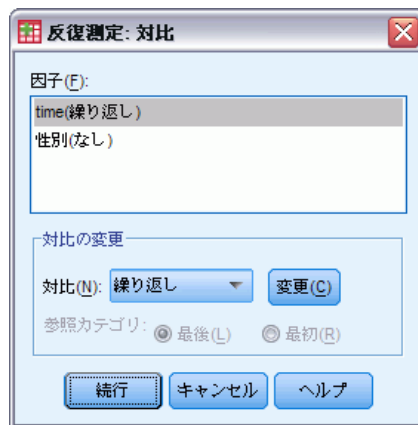
- タイプ I とタイプ II に記載されているモデル。
- 空白セルのない釣り合い型モデルまたは不釣り合い型モデル。

タイプ IV。この方法は、欠損セルが存在する状況を目的としたものです。計画における効果 F に対して、F が他のどの効果にも含まれていないとき、タイプ IV = タイプ III = タイプ II となります。F が他の効果に含まれているとき、タイプ IV は、F におけるパラメータ間で行われている対比を、より高いレベルの効果のすべてに等しく分配します。タイプ IV の平方和の方法は、一般に次のような場合に使用します。

- タイプ I とタイプ II に記載されているモデル。
- 空白セルを伴う釣り合い型モデルまたは不釣り合い型モデル。

GLM 反復測定の対比

図 3-4
[反復測定: 対比] ダイアログ ボックス



対比は、被験者間因子の水準に差異があるかどうかを検定するために使用します。対比は、モデルの被験者間因子ごとに指定できます。対比は、パラメータの線型結合を表します。

仮説の検定は帰無仮説 $LBM = 0$ に基づいています。ここで L は対比係数行列であり、 B はパラメータのベクトル、 M は従属変数の平均変換に対応する平均行列を表しています。[反復測定: オプション] ダイアログ ボックスで [変換行列] を選択すると、変換行列を表示できます。たとえば、4 つの従属変数、4 レベルの被験者内因子があり、多項式の対比 (デフォルト) が被験者内因子で使用されている場合、 M 行列は $(0.5 \ 0.5 \ 0.5 \ 0.5)'$ にな

ります。対比を指定すると、因子に対応する列が対比と適合するように **L** 行列を作成します。残りの列は、**L** 行列が推定できるように調整されます。

利用できる対比には、偏差、単純、差分、Helmert、反復測定、多項式があります。偏差対比と単純対比については、参照カテゴリの最初か最終を選択できます。

[なし]以外の対比は被験者内因子のために必ず選択します。

対比の種類

全平均。 各レベルの平均値（参照カテゴリを除く）をすべてのレベルの平均値（全平均）と比較します。因子の水準は任意の順序になります。

単純。 各レベルの平均値を特定のレベルの平均値と比較します。この対比の型はコントロール群があるときに有効です。最初または最後のカテゴリを参照として選択できます。

差分。 各レベル（最初を除く）の平均値を前のレベルの平均値と比較します（「逆 Helmert 対比」と呼ばれることもあります）。

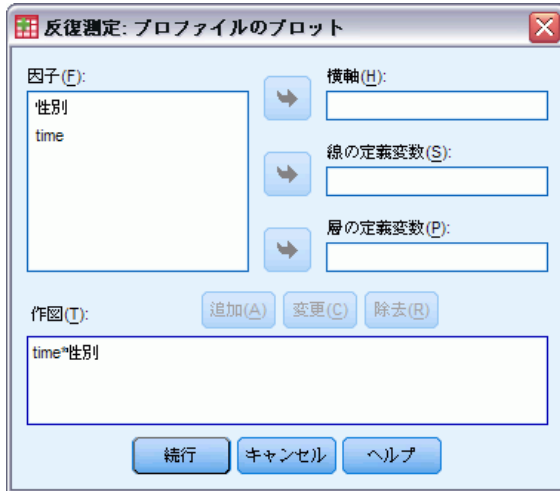
Helmert。 因子の各水準（最終を除く）の平均とその後の水準の平均とを比較します。

反復測定。 因子の各水準（最終を除く）の平均とその後の水準の平均とを比較します。

多項式。 1 次効果、2 次効果、3 次効果等を比較します。第 1 自由度にはすべてのカテゴリを通じての 1 次効果が含まれ、第 2 自由度には 2 次効果が含まれるなど、それぞれが含まれます。このような対比は、多項式のトレンドを推定する場合に頻繁に使用します。

GLM 反復測定のプロフィール プロット

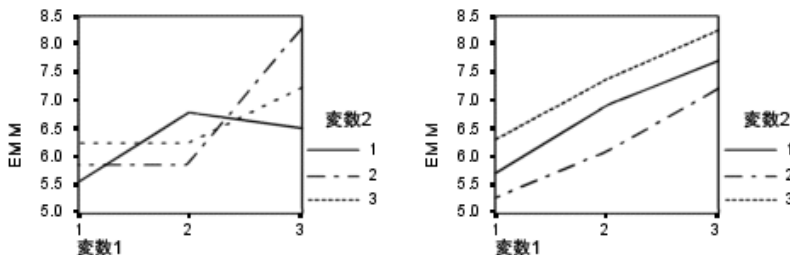
図 3-5
[反復測定: プロファイル プロット] ダイアログ ボックス



プロフィール プロット（交互作用プロット）は、モデル中の周辺平均を比較するとき役に立ちます。プロフィール プロットは、点を使って、因子の 1 つの水準で従属変数（共変量を対象に調整）の推定周辺平均を示す線のプロットです。第 2 因子の水準を使って、個別の線を作成できます。第 3 因子の各水準を使用して、層別のプロットを作成できます。すべての因子が作図で使用できます。プロフィール プロットは各従属変数ごとに作成されます。プロフィール プロットには、被験者間因子と被験者内因子の両方が利用できます。

1 つの因子についてのプロフィール プロットから、推定周辺平均が水準全体で増えているか減っているかがわかります。2 つ以上の因子では、因子間で交互作用がないと平行線で示されるので、1 つの因子に限ってそのレベルを調べることができます。平行ではない線は、交互作用を示しています。

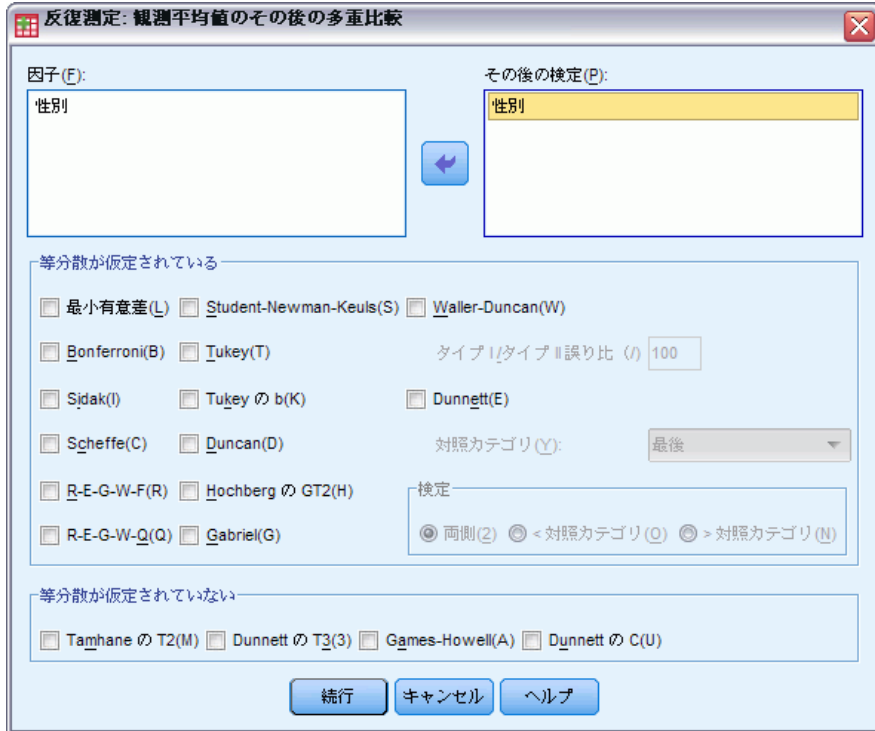
図 3-6
平行ではないプロット（左）と平行プロット（右）



水平軸の因子を選択してプロットを指定し、オプションとして線の定義変数と層の定義変数を選択したら、そのプロットは [作図] ボックスの一覧に追加しなければなりません。

GLM 反復測定のためのその後の比較

図 3-7
[反復測定: 観測平均値のその後の多重比較] ダイアログ ボックス



その後の多重比較検定。 平均値の間に差があることが判明した後は、その後の範囲検定とペアごとの多重比較により、どの平均値が相異しているのかを決めることができます。比較は調整済みでない値に基づいて行われます。ただし、この検定は、被験者間因子がなければ利用できません。また、その後の多重比較検定は、被験者内因子の水準全体の平均に対して実行します。

多重比較検定では、通常、Bonferroni 検定と Tukey の HSD 検定を使用します。**Bonferroni の検定**では、スチューデントの t 検定統計量に基づいて、多重比較が行われるという事実に対して有意水準が調整されます。**Sidak の t 検定**でも有意水準が調整され、Bonferroni の方法よりも厳しく限定されます。**Tukey の HSD 検定**では、スチューデント化された範囲統計量を使ってすべてのペアごとの比較をグループ間で行い、実験ごとの誤差率をすべてのペアごとの比較の集合の誤差率に設定します。大量の平均値のペアを検定する場合は、Tukey の HSD 検定の方が Bonferroni 検定より有効です。少量のペアの場合は Bonferroni の方法の方が有効です。

Hochberg の GT2 検定は Tukey の HSD 検定と類似していますが、ここで使用するのはスチューデント化された最大法です。一般的には、Tukey の検定の方が有効です。**Gabriel のペアごとの比較検定**も、スチューデント化された最大法を使用しますが、一般的に、セルの大きさが均等ではない

場合は、Hochberg の GT2 より有効です。セルの大きさのばらつきが大きい場合には、Gabriel の検定の方が公平になることがあります。

Dunnett のペアごとの多重比較 t 検定では、処理のグループを単一の対照平均値と比較します。最後のカテゴリは、デフォルトの対照カテゴリです。代わりに、最初のカテゴリを選択できます。両側または片側の検定を選択することもできます。因子の任意のレベルの平均値（対照カテゴリを除く）が対照カテゴリのそれと等しくないことを検定するには、両側の検定を使います。因子の任意のレベルの平均値が対照カテゴリの平均値よりも小さいことを検定するには、`<対照カテゴリ`を選択します。同様に、因子の任意のレベルの平均値が対照カテゴリの平均値より大きいことを検定するには、`>[対照カテゴリ]`を選択します。

Ryan と Einot、Gabriel、Welsch (R-E-G-W) は、2 種類のステップダウン多重範囲検定を開発しました。ステップダウン多重手続きでは、最初にすべての平均値が等しいかどうかを検定します。すべての平均値が等しくなければ、サブグループの平均値が等しいかどうかを検定します。**R-E-G-W の F 値**は F 検定に基づき、**R-E-G-W の Q 値**はスチューデント化した範囲に基づきます。この検定は、Duncan の多重範囲検定や Student-Newman-Keuls の検定（これもステップダウン多重手続き）より有効ですが、セルの大きさが等しくない場合はお勧めできません。

分散が等しくない場合は、**Tamhane の T2** (t 検定に基づくペアごとの控えめな比較)、**Dunnett の T3** (スチューデント化した最大偏差に基づくペアごとの比較検定)、**Games-Howell の ペアごとの比較検定** (公平な場合もある)、または **Dunnett の C** (スチューデント化した範囲に基づくペアごとの比較検定) を使用します。

Duncan の多重範囲検定、Student-Newman-Keuls (**S-N-K**) の方法、および **Tukey の b 検定**はグループ平均を順位付け、範囲の値を計算する範囲検定です。これらの検定は先に述べた検定ほど頻繁には使われていません。

Waller-Duncan の t 検定では、Bayesian のアプローチが使われています。この範囲検定では、サンプル サイズが等しくない場合にサンプルサイズの調和平均が使われます。

Scheffé の検定の有意水準は、この機能で利用できるペアごとの比較だけにとどまらず、グループ平均に可能なすべての線型結合を検定できる設計になっています。結果的に、Scheffé の検定は他の検定より控えめになってしまうことが多いため、有意確率を求める場合は、平均値間の差が大きい必要があります。

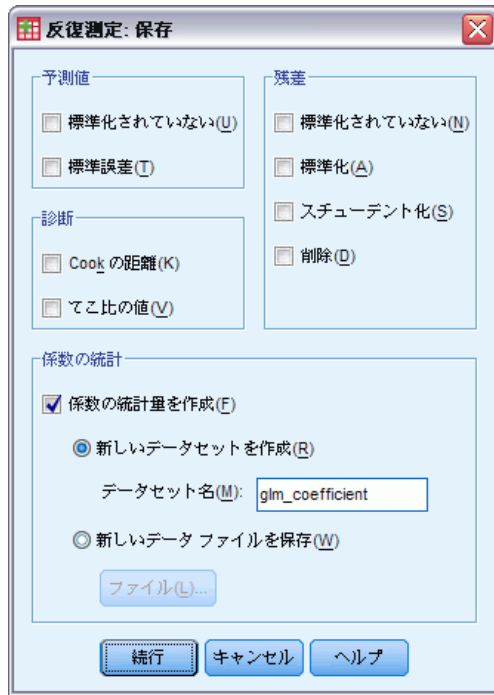
最小有意差 (**LSD**) のペアごとの多重比較検定は、グループのすべてのペア間ごとの多重 T 検定に相当します。この検定の欠点は、観測された有意水準を多重比較向けに調整する試みが一切行われなことです。

表示される検定。 ペアごとの比較は、LSD、Sidak の方法、Bonferroni の方法、Games と Howell の方法、Tamhane の T2 と T3、Dunnett の C と Dunnett の T3 で使用できます。範囲検定の等質サブグループは、S-N-K、Tukey の b、Duncan、R-E-G-W の F、R-E-G-W の Q および Waller の方法

で使用できます。Tukey の HSD 検定、Hochberg の GT2、Gabriel の検定、および Scheffé の検定は、多重比較検定と範囲検定の両方です。

GLM 反復測定 of 保存

図 3-8
[反復測定: 保存] ダイアログ ボックス



モデルが予測した値、残差、関連測定値は、データ エディタで新変数として保存できます。これらの変数の多くは、データの仮定を調べるために利用できます。値を保存して別の IBM® SPSS® Statistics セッションで利用するためには、現在のデータ ファイルを保存する必要があります。

予測値。 モデルがケースごとに予測する値。

- **[標準化されていない]** 従属変数を予測するモデルの値。
- **標準誤差。** 独立変数の同じ値を持つケースに対する従属変数の平均値の標準偏差の推定値。

診断。 独立変数の一般的でない値の組み合わせのケースとモデルに大きな影響を及ぼすケースを識別するための測定方法。Cook の距離と非心てこ比の値が使用できます。

- **Cook の距離.** 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。Cook の距離が大きいときは、回帰統計量の計算からケースを除外したことが係数を実質的に変化させたことを示しています。
- **てこ比の値.** 非心てこ比の値。モデルの適合度に関する各観測の相対的な影響度。

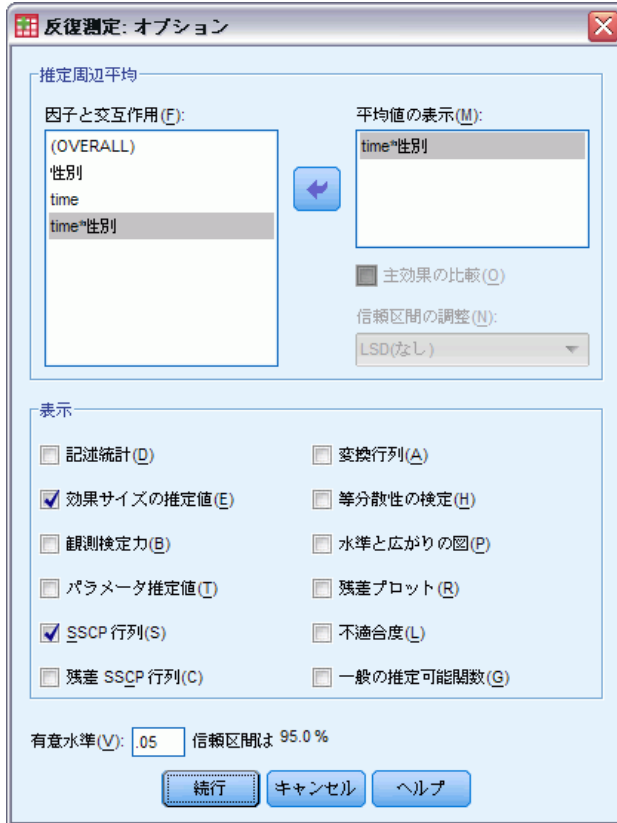
残差. 標準化されていない残差は、従属変数の実際の値からモデルにより予測される値を引いたものです。標準化された残差、スチューデント化された残差、削除された残差も使用できます。

- **[標準化されていない].** 観測値およびモデルで予測された値との差。
- **標準化残差.** 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。
- **スチューデント化された残差.** 残差を、独立変数の平均値からの独立変数の各ケース値の距離に依存して、ケースごとに違う標準偏差の推定量で割ったもの。
- **削除.** あるケースが回帰係数の計算から除外されたときのケースの残差。従属変数と調整済み予測値の間の差です。

係数統計量. パラメータ推定値の分散共分散行列をデータセットやデータファイルに保存します。また、それぞれの従属変数に対しても、パラメータ推定値の行、パラメータ推定値に対応する t 統計量の有意確率の行、および残差自由度の行が存在します。多変量モデルの場合は、各従属変数に同様の行があります。行列ファイルを読み込む別の手続きで、この行列データを使用できます。データセットは、同じセッションの今後で利用可能ですが、セッション終了前に明示的に保存しない限り、保存されません。データセット名は、変数命名規則に従う必要があります。

GLM 反復測定オプション

図 3-9
[反復測定: オプション] ダイアログ ボックス



このダイアログ ボックスでオプションの統計を利用できます。統計量は、固定効果モデルを使用して計算されます。

推定周辺平均。セルにおける母周辺平均値について推定したい因子と交互作用を選択します。共変量が存在する場合、これらの平均値は、共変量に対して調整されます。被験者内と被験者間因子の両方を選択できます。

- **主効果の比較。**被験者間と被験者内因子の両方について、モデル内の主効果に対する推定周辺平均値間で、ペアごとに無補正の比較を行います。この項目は、[平均値の表示] リストで主効果を選択した場合にだけ選択できます。
- **信頼区間の調整。**最小有意差 (LSD)、Bonferroni の方法、Sidak の方法の信頼区間と有意水準に対する調整を選択します。この項目は、[主効果の比較] を選択している場合にだけ選択できます。

[表示][記述統計量] を選択すると、すべてのセルにおけるすべての従属変数の観測平均値、標準偏差、および度数を求めることができます。**[効果サイズの推定値]** は、偏相関のイータの 2 乗の値をそれぞれの効果とそれぞれのパラメータ推定値に与えるものです。イータの 2 乗統計量は、因子に起因する総変動の比率を記述するものです。観測した値に基づいて対立仮説を立てるときに検定の検出力を得るには、**[観測検定力]** を選択します。パラメータ推定値、標準誤差、T 検定、信頼区間、およびそれぞれの検定での観測検定力を創出するには、**[パラメータ推定値]** を選択します。残差分散共分散行列についての仮説、誤差 SSCP 行列、残差 SSCP 行列プラス Bartlett の球面性の検定を表示できます。

等分散性の検定は、被験者間因子の場合にだけ、各従属変数の分散の等質性分析について、被験者間因子のすべての水準の組み合わせを通じて Levene の検定を創出するものです。等分散性の検定は、被験者間因子のすべての水準の組み合わせを通じて、従属変数の分散共分散行列の等質性分析についての Box の M 検定をも含んでいます。水準と広がり の 図および残差プロットのオプションは、データに関する仮定のチェックに役立ちます。この項目は、因子が存在しないときは無効になります。それぞれの従属変数について、標準化された予測による観測残差プロットを作成するには、**[残差プロット]** を選択します。このようなプロットは、分散が等しいという仮定の検証に役立ちます。従属変数と独立変数の間の関係がモデルで正しく記述されているかどうかをチェックするには、**[不適合度]** を選択します。一般の推定可能関数を使うと、一般推定可能関数に基づいてユーザー指定の仮説の検定を構築できます。任意の対比係数行列における行は、一般推定可能関数の一次結合です。

有意水準。 信頼区間を構築するために、その後の検定や信頼係数で使われる有意水準を調整したい場合があります。指定された値は、検定の観測検定力の計算に使用できます。有意水準を指定するときは、信頼区間の関連水準がダイアログ ボックスに表示されます。

GLM コマンドの追加機能

この機能は、1 変量、多変量、反復測定分析に適用できます。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。

- 計画中の入れ子効果の指定 (DESIGN サブコマンドを使用)。
- 効果と効果や値の線型組み合わせ検定の指定 (TEST サブコマンドを使用)。
- 多重対比の指定 (CONTRAST サブコマンドを使用)。
- ユーザー欠損値の包含 (MISSING サブコマンドを使用)。
- EPS 基準の指定 (CRITERIA サブコマンドを使用)。
- ユーザー定義の L 行列、M 行列、または K 行列の作成 (LMATRIX、MMATRIX、および KMATRIX サブコマンドを使用)。

- 全平均対比または単純対比での、中間参照カテゴリの指定 (CONTRAST サブコマンドを使用)。
- 多項式対比の計量の指定 (CONTRAST サブコマンドを使用)。
- その後の比較の誤差項の指定 (POSTHOC サブコマンドを使用)。
- 因子または因子の一覧にある因子間の交互作用の推定周辺平均値の計算 (EMMEANS サブコマンドを使用)。
- 仮の変数の名前の指定 (SAVE サブコマンドを使用)。
- 相関行列のデータ ファイルの作成 (OUTFILE サブコマンドを使用)。
- 被験者間因子の分散分析表の統計量を含む行列データ ファイルの作成 (OUTFILE サブコマンドを使用)。
- 計画行列の新しいデータ ファイルへの保存 (OUTFILE サブコマンドを使用)。

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

分散成分分析

混合効果モデルの分散成分手続きは、従属変数の分散に対する変量効果ごとの寄与率を推定します。この手続きは、分割プロット、1 変量の反復測定、任意のブロック計画など混合モデルの分析を行う場合に特に役立ちます。分散成分を計算すると、どこに焦点を絞れば分散を減らすことができるかを確定できます。

分散成分の推定には、最小ノルム 2 次形式不偏推定量 (MINQUE)、分散分析 (ANOVA)、最尤法 (ML)、制限された最尤法 (REML) の 4 つの方法が利用できます。これらの方法で、各種の指定を利用できます。

すべての方法のデフォルトの出力に、分散成分推定値が含まれます。ML 法か REML 法を使用して、分散共分散行列の漸近表も表示されます。この他に利用できる出力には、ANOVA 表と ANOVA 法の期待平均平方、および ML 法と REML 法の反復の記述などがあります。分散成分手続きは、「GLM 1 変量」手続きと完全に適合します。

[WLS 重み] で、測定方法ごとに異なる精度を補正するなどの目的で、観測値に重み付き分析のためのさまざまな重みを付けるのに使用する変数を指定できます。

例。農業学校で、6 匹の母豚から生まれた子豚の 1 か月後の体重増を測定します。同腹子 (同じ親から生まれた子) 変数は、6 つの水準を持つ変量因子です。(研究対象の 6 匹の母豚は、さらに大きな、母豚の母集団からの任意のサンプルです)。ここで、研究者は、同じ母豚から生まれた子豚の間の差よりも、異なる同腹子間の差の方が、体重増の分散をもたらすはるかに大きな原因であることがわかります。

データ。従属変数は量的変数です。因子も同じくカテゴリ型です。数値または最高 8 バイトまでの文字値を持つことができます。因子は、最低 1 つは変量でなければなりません。つまり、因子の水準は、可能な水準の任意のサンプルでなければなりません。共変量は、従属変数に関連する量的変数です。

仮定。どの方法においても、変量効果のモデル パラメータの平均値は 0 で、有限定数分散があり、相関はないものと仮定します。異なる変量効果のモデル パラメータも相関はありません。

残差の項も平均値が 0 で、有限定数分散があります。どの変量効果のモデル パラメータとも相関はありません。異なる観測値からの残差項は、相関がないと仮定されます。

以上の仮定に基づいて、変量因子の同一水準からの観測値に相関を持たせません。この点から、分散成分モデルは一般線型モデルと区別されます。

分散分析と MINQUE は、正規性の仮定を必要としません。どちらも、正規性の仮定からの平均的な逸脱については、許容範囲としています。

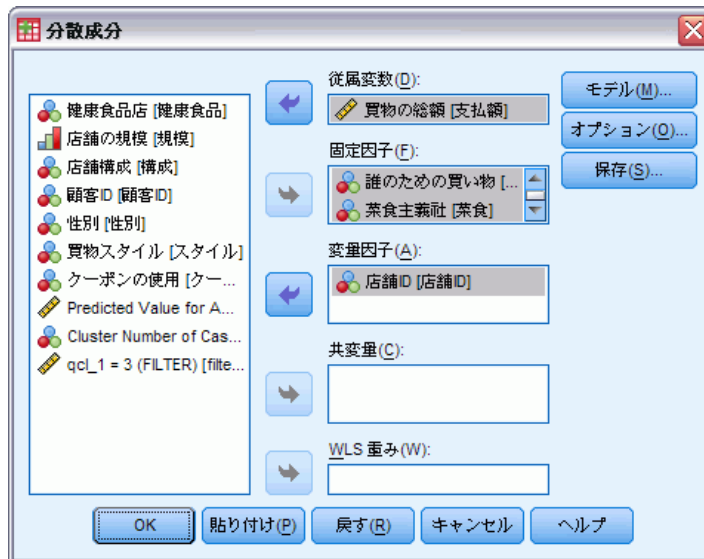
ML と REML では、モデルのパラメータと残差項が正規分布している必要があります。

関連手続き。分散成分分析の実行前にデータを調べる場合は、探索的分析手続きを使用します。仮説の検定には、GLM 1 変量、GLM 多変量、GLM 反復測定を使用します。

分散成分表を作成するには

- ▶ メニューから次の項目を選択します。
分析 > 一般線型モデル > 分散成分...

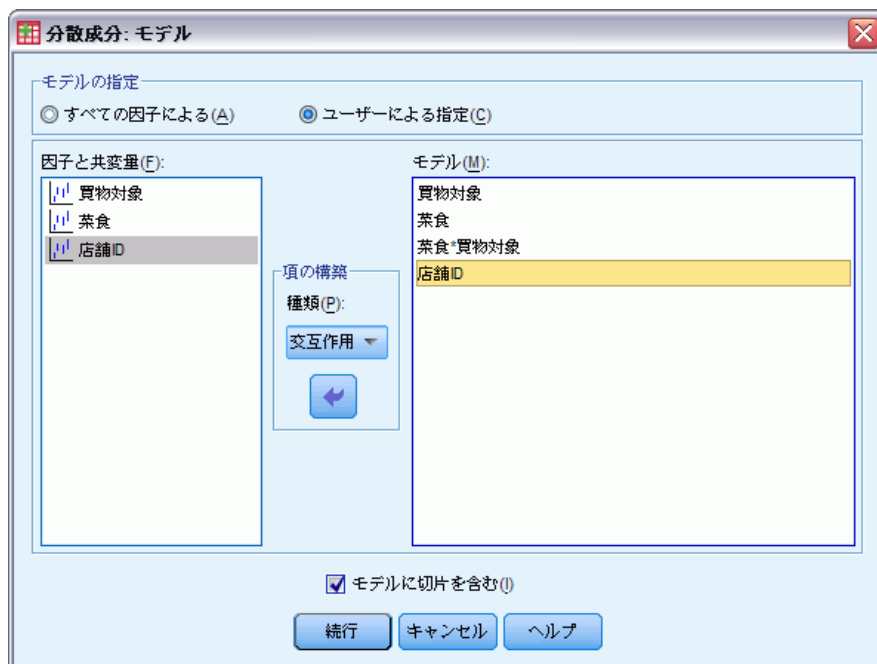
図 4-1
[分散成分] ダイアログ ボックス



- ▶ [従属変数] ボックスに従属変数を選択します。
- ▶ データに合わせて、[固定因子]、[変量因子]、[共変量] の変数を選択します。重み付き変数は、[WLS 重み] ボックスで指定します。

分散成分のモデル

図 4-2
[分散成分: モデル] ダイアログ ボックス



モデルの指定。すべての因子によるモデルには、因子の主効果、共変量の主効果、因子間の交互作用のすべてが含まれます。共変量の交互作用は含まれません。交互作用の一部だけを指定する場合や、因子と共変量の交互作用を指定する場合は、[ユーザーの指定] をクリックします。モデルに含める項目はすべて指示する必要があります。

因子と共変量。要素および共変量はリストされます。

モデル。モデルは、使用するデータの性質によって異なります。[ユーザーの指定] をクリックすれば、分析対象の主効果と交互作用を選択できます。モデルには、変量因子が含まれている必要があります。

モデルに切片を含む。通常、モデルには切片が含まれます。データが原点を通ると仮定できる場合は、切片を除外できます。

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。

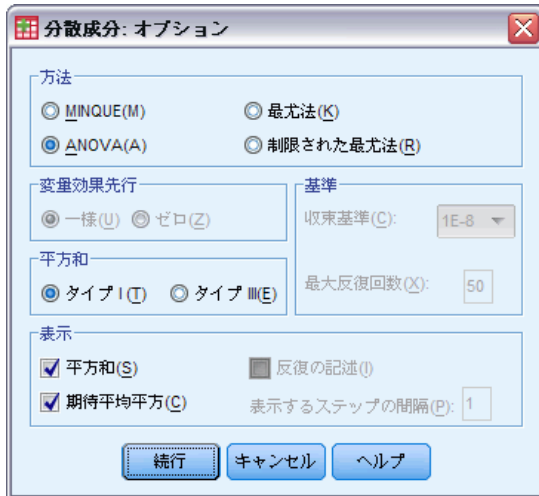
交互作用。選択した変数のすべてについて、最高水準の交互作用項を作成します。これはデフォルトです。

主効果。選択した変数のそれぞれに主効果の項を作成します。

- 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。

分散成分のオプション

図 4-3
[分散成分: オプション] ダイアログ ボックス



方法。4 つの方法から 1 つを選んで、分散成分の推定を行います。

- [MINQUE] (最小ノルム 2 次形式不偏推定値) は、固定効果が不変の推定値を生成します。データの分布が正規で、推定値が正しければ、この方法は、すべての不偏推定値の中で最小分散を生成します。変数効果前の重みの方法を選択できます。
- [分散分析] (ANOVA) 効果ごとにタイプ I とタイプ III の平方和のどちらか一方を使用して、不偏推定値を計算します。ANOVA 法は、分散に負の推定値を生成する場合がありますが、間違ったモデル、不適切な推定方法、あるいはデータの不足を示している可能性があります。
- [最尤法] (ML) 手続きは、反復を使用するので、実際に観測されたデータに最も一致する推定値を生成します。この推定値は偏向する可能性があります。この方法は、漸近的に正規です。ML と REML の推定値は、解釈においては不変です。この方法は、固定効果の推定に用いられる自由度を考慮しません。
- [制限された最尤法] (REML) の推定値は、平行データの (すべてではないが) 多くのケースに対する ANOVA 推定値を減じます。この方法は固定効果に合わせて調整されるため、ML 法の場合より標準誤差が小さ

くなければなりません。この方法は、固定効果の推定に用いられる自由度を考慮します。

変量効果先行。[一様] とは、すべての変量効果と残差の項が観測値に同等の影響をもたらすことを意味しています。[ゼロ] 設定は、変量効果分散に 0 を仮定するのと同じです。これを利用できるのは、MINQUE 法だけです。

平方和。[タイプ I] の平方和は、階層モデルで使用されますが、分散成分の文献では、このモデルが頻繁に使われます。GLM のデフォルトである [タイプ III] を選択すると、GLM 1 変量において分散の推定値を、タイプ III の平方和を使用した仮説の検定に利用できます。これを利用できるのは、ANOVA 法だけです。

基準。[収束基準] と [最大反復回数] を指定できます。これを利用できるのは、ML 法または REML 法だけです。

表示。ANOVA 法の場合は、[平方和] と [期待平均平方] の表示を選択できます。[最尤法] か [制限された最尤法] を選択すると、[反復の記述] を表示できます。

平方和 (分散成分)

モデルには、平方和のタイプを選択できます。最も一般的に使われ、デフォルトでもあるのがタイプ III です。

タイプ I。この方法は、平方和の階層的分割法という名でも知られています。各項は、モデル内で先行する項目に対してだけ調整されます。タイプ I の平方和の方法は、一般に次のような場合に使用します。

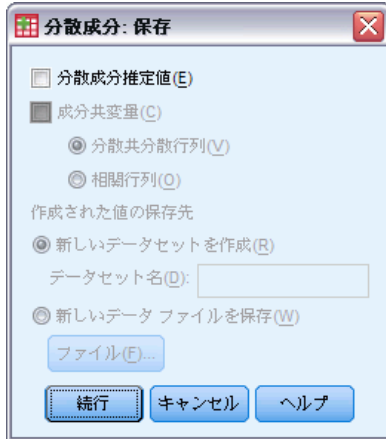
- 1 次の交互作用の効果より先になんらかの主効果が指定され、2 次の交互作用の効果より先になんらかの 1 次の交互作用の効果が指定されているといった具合の分散分析の釣り合い型モデル。
- 高次の項より先に低次の項目が指定されている多項式回帰モデル。
- 最初に指定された効果が 2 番目に指定された効果の中に入れ子になり、2 番目に指定された効果が 3 番目に指定された効果の中に入れ子になり、という具合に続く純粋な入れ子モデル。(入れ子の形式は、シNTAXS を使用しなければ指定できません)。

タイプ III。デフォルトです。この方法では、計画内にある効果の平方和を、その計画を含まないその他の効果、またはそうした効果に対して直交的な効果を対象に調整した平方和として計算します。タイプ III の平方和には、一般的な推定形式が一定の状態に保たれている限り、セル度数が変わらないという大きな利点があります。したがって、このタイプの平方和は、欠損セルがない不釣り合い型モデルに有用だと考えられます。欠損セルのない多因子計画でこの方法に相当するのが、Yates の平均値の重み付き 2 乗法です。タイプ III の平方和の方法が一般的に使用されるのは次の場合です。

- タイプ I に記載されているモデル。
- 空白セルのない釣り合い型モデルまたは不釣り合い型モデル。

分散成分の新しいファイルへの保存

図 4-4
[分散成分: 新しいファイルへ保存] ダイアログ ボックス



この手続きの結果の一部を新しい IBM® SPSS® Statistics データ ファイルに保存できます。

分散成分推定値。分散成分の推定値と推定値ラベルをデータ ファイルやデータセットに保存します。さらに多くの統計量を計算したり、GLM の手続きで踏み込んだ分析を行う場合に使用できます。たとえば、信頼区間の計算や仮説の検定などに使用できます。

成分共変量。分散共分散行列や相関行列をデータ ファイルやデータセットに保存します。利用できるのは、[最尤法] か [制限された最尤法] が指定されている場合だけです。

作成された値の保存先。分散成分の推定値や行列が保存されたファイルのデータセット名や外部ファイル名を指定できます。データセットは、同じセッションの今後で利用可能ですが、セッション終了前に明示的に保存しない限り、保存されません。データセット名は、変数命名規則に従う必要があります。

MATRIX コマンドを使うと、データ ファイルから必要なデータを抽出でき、信頼区間の計算や検定を行うことができます。

VARCOMP 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。

- 計画中の入れ子効果の指定 (DESIGN サブコマンドを使用)。
- ユーザー欠損値の包含 (MISSING サブコマンドを使用)。
- EPS 基準の指定 (CRITERIA サブコマンドを使用)。

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

線型混合モデル

線型混合モデル手続きは、相関関係を持った非定常的な変動が現れるデータも扱えるように、一般線型モデルを拡張したものです。線形混合モデルは、データの平均値だけでなく、データの分散および共分散をモデリングできる柔軟性を備えています。

また、線型混合モデル手続きは、線型混合モデルとして公式化できる他のモデルを当てはめるための柔軟なツールでもあります。それらのモデルには、マルチレベルモデル、階層線型モデル、および変量係数モデルがあります。

例。ある食料雑貨店チェーンが、さまざまなクーポンが顧客支出に対して与える影響に関心を持っていると仮定します。常連となっている顧客からサンプルを無作為に選択し、10 週間にわたって各顧客の支出を調べます。1 週間ごとに違う種類のクーポンが顧客に送付されます。線型混合モデルを使用して、10 週間にわたって各被験者から得られる反復観測値に基づく相関に合わせて調整しながら、支出に対するさまざまなクーポンの影響を推定します。

方法。最尤法 (ML) および制限された最尤法 (REML) の推定。

統計量。記述統計：従属変数のサンプル サイズ、平均値、および標準偏差と、因子の異なるそれぞれのレベルの組み合わせの共変量。因子レベル情報：各因子のレベルを並べ替えた値と、それらの度数分布表。また、固定効果のパラメータ推定値および信頼区間、分散共分散行列のパラメータの Wald 検定および信頼区間。異なる仮説を評価するには、[タイプ I] および [タイプ III] 平方和を使用できます。デフォルトは タイプ III です。

データ。従属変数は量的変数でなければなりません。因子はカテゴリ変数でなければならず、数値または文字値を持つことができます。共変量および重み付け変数は量的でなければなりません。被験者および反復変数はどのような種類でもかまいません。

仮定。従属変数は固定因子、変量因子、および共変量と線型関係があると仮定されます。固定効果は従属変数の平均値をモデリングします。変量効果は従属変数の分散共分散構造をモデリングします。多重変量効果は互いに独立していると思われ、個別の分散共分散行列はそれぞれ計算されます。ただし、同じ変量効果で指定したモデル項は、相関している場合があります。反復測定は、残差分散共分散構造をモデリングします。従属変数は、正規分布から抽出されると仮定されます。

関連手続き。分析の実行前にデータを調べる場合は、[探索的分析] 手続きを使用します。相関している変動、または非定数の変動が存在することを疑わない場合は、[1 変量] または [反復測定] 手続きを使用できます。変

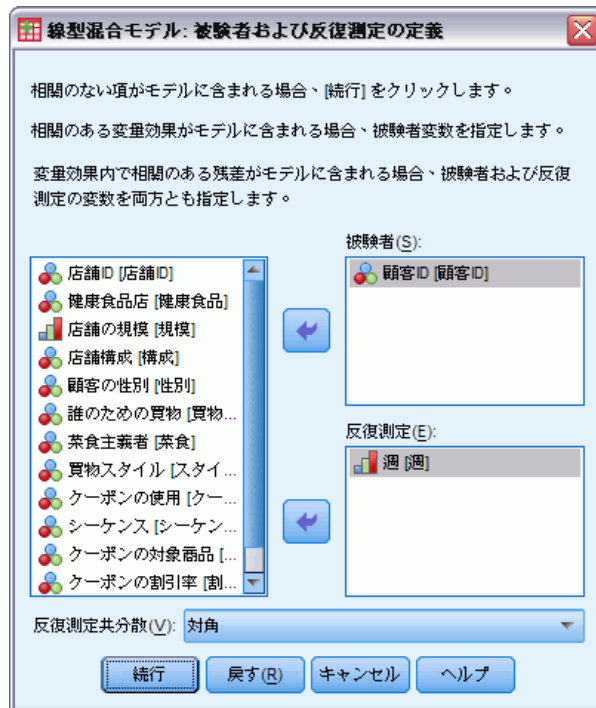
量効果が分散成分の分散共分散構造を持ち、また反復測定がない場合は、代わりに [分散成分分析] 手続きを使用できます。

線型混合モデル分析を行うには

- ▶ メニューから次の項目を選択します。
分析 > 混合モデル > 線型...

図 5-1

[線型混合モデル: 被験者および反復測定の定義] ダイアログ ボックス



- ▶ 必要に応じて、1 つ以上の被験者変数を選択します。
- ▶ 必要に応じて、1 つ以上の反復変数を選択します。
- ▶ 必要に応じて、残差分散共分散構造を選択します。
- ▶ [続行] をクリックします。

図 5-2
[線型混合モデル] ダイアログ ボックス



- ▶ [従属変数] ボックスに従属変数を選択します。
 - ▶ 少なくとも 1 つの因子または共変量を選択します。
 - ▶ [固定] または [変量] をクリックし、固定効果モデルまたは変量効果モデルのどちらか一方を指定します。
- 必要に応じて、残差重みを選択します。

線型混合モデルの被験者および反復変数の選択

このダイアログ ボックスで、被験者および反復観測値を定義する変数を選択し、残差分散共分散構造を選択します。図 5-1 p. 43 を参照してください。

被験者。被験者は観測単位で、他の被験者とは独立していると思なされます。たとえば医学統計では、ある患者の血圧測定値は、他の測定値とは独立していると思なすことができます。被験者ごとに反復測定を行い、その観測値の相関をモデリングする場合、被験者の定義は特に重要です。たとえば、1 人の患者が医者に連続して診断に行く間に、血圧測定値が相関していると予想できます。

被験者は、複数の変数の因子レベルの組み合わせによっても定義できます。たとえば、「性別」および「年齢」カテゴリを被験者変数として指定すると、[65 歳を超える男性] は互いに同類であるものの、[65 歳以下の男性] および [女性] からは独立しているという仮定をモデリングできます。

[被験者] リストで指定された変数はすべて、残差共分散構造の被験者を定義するために使用されます。一部またはすべての変数を、変量効果の分散共分散構造を分析する変数を定義するために使用できます。

反復測定。このリストで指定された変数は、反復観測値を識別するために使用されます。医学統計を例にとれば、単一変数「週」を使用して 10 週間の観測値を識別できます。また、「月」および「日」を同時に使用すれば、毎日の観測値を 1 年を通して識別できます。

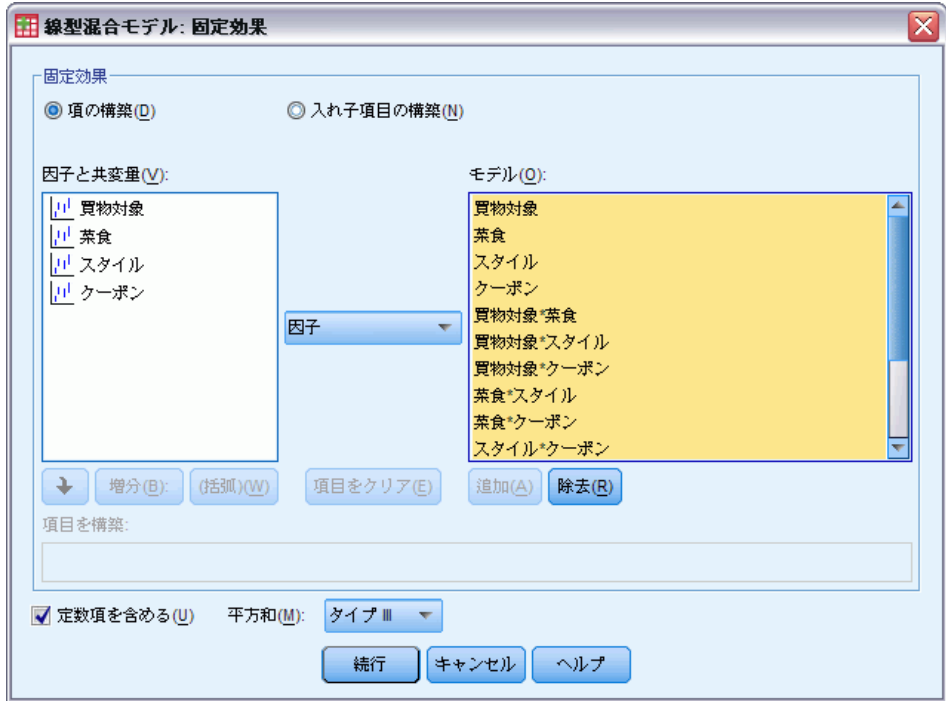
反復共分散型。ここで、残差分散共分散構造を指定します。利用できる構造には、次のものがあります。

- 前従属： 1 次
- AR(1)
- AR(1)： 不均質
- ARMA(1, 1)
- 複合シンメトリ
- 複合シンメトリ： 相関行列
- 複合シンメトリ： 不均質
- 対角
- 因子解析的： 1 次
- 因子解析的： 1 次、不均質
- Huynh-Feldt
- 計測された単位
- Toeplitz
- Toeplitz： 不均質
- 無構造
- 無構造： 相関行列

詳細は、B 付録 p.183 共分散構造 を参照してください。

線型混合モデルの固定効果

図 5-3
[線型混合モデル: 固定効果] ダイアログ ボックス



固定効果。デフォルトのモデルはありません。このため固定効果を明示的に指定する必要があります。入れ子または非入れ子の項を構築することができます。

定数項を含める。通常、モデルには定数項が含まれます。データが原点を通ると仮定できる場合は、切片を除外できます。

平方和。平方和の計算方法。欠損セルのないモデルの場合、最も一般的に使われる手法は [タイプ III] です。

非入れ子の項目の構築

選択した因子や共変量について、次の項を作成できます。

因子。選択した変数に交互作用および主効果を作成します。これはデフォルトです。

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。

入れ子項目の構築

この手続きでは、モデルに入れ子の項を構築できます。入れ子項目を構築すると、別の因子のレベルと相互作用しない因子または共変量の効果をモデリングできます。たとえば、ある食料雑貨店チェーンが、いくつかの店舗の場所ごとの顧客支出を調査すると仮定します。1 人の顧客が頻繁に通うのはこれらの店舗の 1 つだけだするため、「顧客」の効果は、「店舗の場所」の効果の入れ子になっていると言えます。

さらに、入れ子の項に交互作用効果を含めたり、複数のレベルの入れ子を追加することもできます。

制限。入れ子項目には、次の制限があります。

- 1 つの交互作用内のすべての因子は一意でなければなりません。したがって、A が因子である場合、 $A*A$ の指定は無効です。
- 1 つの入れ子効果内のすべての因子は一意でなければなりません。したがって、A が因子である場合、 $A(A)$ の指定は無効です。
- 効果は共変量の入れ子にできません。したがって、A が因子で X が共変量である場合、 $A(X)$ の指定は無効です。

平方和

モデルには、平方和のタイプを選択できます。最も一般的に使われ、デフォルトでもあるのがタイプ III です。

タイプ I。この方法は、平方和の階層的分割法という名でも知られています。各項は、モデル内で先行する項の場合にだけ調整されます。タイプ I の平方和は、一般に次のような場合に使用します。

- 1 次の交互作用の効果より先になんらかの主効果が指定され、2 次の交互作用の効果より先になんらかの 1 次の交互作用の効果が指定されているといった具合の分散分析の釣り合い型モデル。
- 高次の項より先に低次の項目が指定されている多項式回帰モデル。
- 最初に指定された効果が 2 番目に指定された効果の中に入れ子になり、2 番目に指定された効果が 3 番目に指定された効果の中に入れ子になり、という具合に続く純粋な入れ子モデル。(入れ子の形式は、シンタックスを使用しなければ指定できません)。

タイプ III。デフォルトです。この方法では、計画内にある効果の平方和を、その計画を含まないその他の効果、またはそうした効果に対して直交的な効果を対象に調整した平方和として計算します。タイプ III の平方和には、一般的な推定形式が一定の状態に保たれている限り、セル度数が変わらないという大きな利点があります。したがって、このタイプの平方和は、欠損セルがない不釣り合い型モデルに有用だと考えられます。欠損セルのない多因子計画でこの方法に相当するのが、Yates の平均値の重み付き 2 乗法です。タイプ III の平方和の方法が一般的に使用されるのは次の場合です。

- タイプ I に記載されているモデル。
- 空白セルのない釣り合い型モデルまたは不釣り合い型モデル。

線型混合モデルの変量効果

図 5-4
[線型混合モデル: 変量効果] ダイアログ ボックス

線型混合モデル: 変量効果

変量効果 1 の 1

前へ(P) 次(N)

共分散タイプ(V) 分散成分

変量効果

項の構築(O) 入れ子項目の構築(N) 定数項を含める(U)

因子および共変量(F):

- 買物対象
- 茶食
- スタイル
- クーポン

因子

モデル(M):

- 買物対象

↓ 増分: (括弧)(W) 項目をクリア(E) 追加(A) 除去(R)

項目を構築:

被験者のグループ化:

被験者(S):

- 顧客ID [顧客ID]

組み合わせ(O):

- 顧客ID [顧客ID]

続行 キャンセル ヘルプ

共分散型。変量効果モデルに分散共分散構造を指定します。各変量効果に個別の分散共分散行列が推定されます。利用できる構造には、次のものがあります。

- 前従属: 1次
- AR(1)
- AR(1): 不均質
- ARMA(1, 1)
- 複合シンメトリ
- 複合シンメトリ: 相関行列
- 複合シンメトリ: 不均質
- 対角
- 因子解析的: 1次
- 因子解析的: 1次、不均質
- Huynh-Feldt
- 計測された単位
- Toeplitz
- Toeplitz: 不均質
- 無構造
- 無構造: 相関行列
- 分散成分

詳細は、B 付録 p.183 共分散構造 を参照してください。

変量効果。デフォルトのモデルはありません。このため変量効果を明示的に指定する必要があります。入れ子または非入れ子の項を構築することができます。変量効果モデルに定数項を含めることもできます。

多重変量効果を指定できます。最初のモデルの構築後、次のモデルを構築するには [次] をクリックします。[前] をクリックすると、スクロールして既存のモデルに戻ることができます。各変量効果モデルは、他のすべての変量効果モデルから独立している、つまり、各モデルに対し、個別に分散共分散行列独立が計算されるものと仮定します。同じ変量効果モデルで指定した項は、相関している場合があります。

被験者のグループ化。[被験者および反復測定の設定] ダイアログ ボックスに、被験者変数として選択した変数が一覧で表示されます。リストの一部またはすべての変数を選択して、変量効果モデルの被験者を定義します。

線型混合モデルの推定

図 5-5
[線形混合モデル: 推定] ダイアログ ボックス

方法。[最尤法] または [制限された最尤法] 推定を選択します。

反復回数。

- **最大反復回数。**負でない整数を指定してください。
- **最大段階 2 分。**対数尤度が増加するか、最大段階 2 分に達するまで、反復ごとにステップ サイズが 0.5 の因子ずつ減少します。正の整数を指定します。
- **すべての反復の記述。**対数尤度関数値、および 0 回目（初期推定値）から始まる反復の n 回ごとの反復パラメータ推定値を含むテーブルを表示します。反復の記述の出力を選択した場合、n の値にかかわらず、最後の反復は常に出力されます。

対数尤度収束。対数尤度関数の絶対変化または相対変化が、指定した負でない値未満である場合に収束します。指定した値が 0 の場合、この収束基準は使用されません。

パラメータ収束。パラメータ推定値の絶対変化または相対変化の最大値が、指定した負でない値以下である場合に収束します。指定した値が 0 の場合、この収束基準は使用されません。

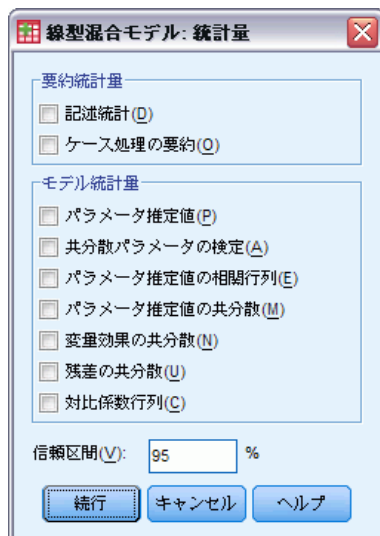
Hessian 収束。[絶対値] を指定した場合は、Hessian に基づく統計量が、指定した値よりも小さい場合に収束するとみなされます。[相対値] を指定した場合は、統計量が、指定した値と対数尤度の絶対値の積より小さい場合に収束するとみなされます。指定した値が 0 の場合、この収束基準は使用されません。

最大スコア ステップ数。反復回数が n に達するまで、Fisher スコア法のアルゴリズムを使用するよう要求します。正の整数を指定します。

特異性許容度。特異性の点検のときに許容度として使用される値です。正の値を指定します。

線型混合モデルの統計量

図 5-6
[線型混合モデル: 統計量] ダイアログ ボックス



要約統計量。次のテーブルが作成されます。

- **記述統計。**従属変数のサンプル サイズ、平均値、および標準偏差と、共変量を表示します（指定した場合）。これらの統計は、因子レベルの組み合わせごとに表示されます。
- **ケース処理の要約。**因子の並べ替えた値、反復測定変数、反復測定の被験者、および変量効果の被験者と、それらの度数分布表を表示します。

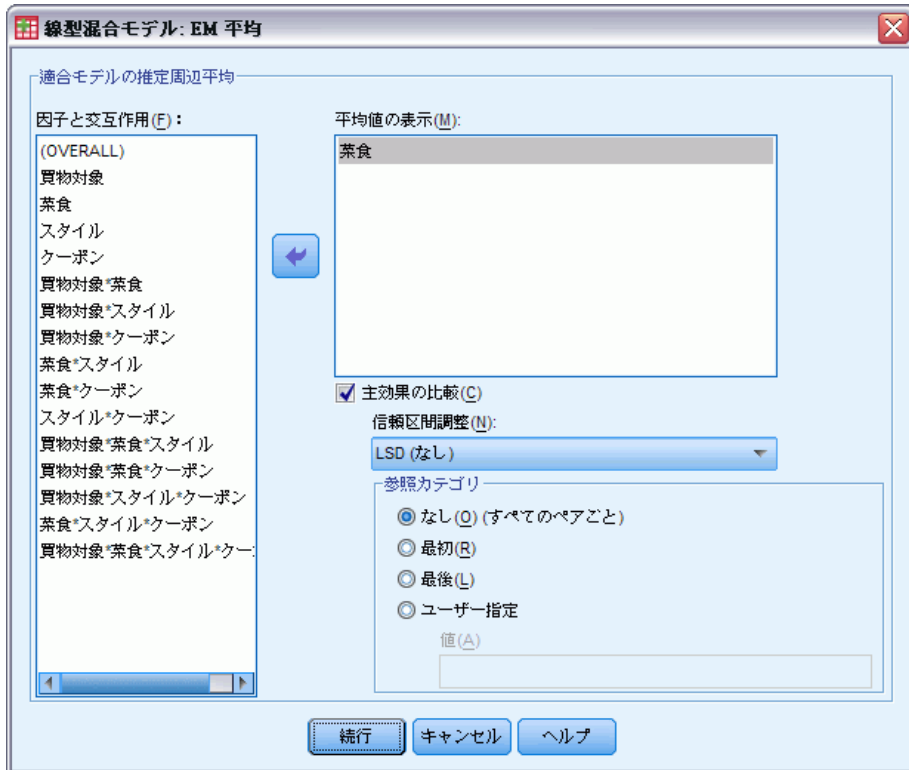
モデル統計量。次のテーブルが作成されます。

- **パラメータ推定値。**固定効果および変量効果のパラメータ推定値と、それらの標準誤差の近似を表示します。
- **共分散パラメータの検定。**分散共分散パラメータの漸近標準誤差および Wald 検定を表示します。
- **パラメータ推定値の相関。**固定効果のパラメータ推定値の漸近相関行列を表示します。
- **パラメータ推定値の共分散。**固定効果のパラメータ推定値の漸近分散共分散行列を表示します。
- **変量効果の共分散。**変量効果の推定分散共分散行列を表示します。このオプションは、1 つ以上の変量効果が指定された場合に限り使用できます。変量効果に被験者変数が指定された場合、共通ブロックが表示されます。
- **残差の共分散。**推定残差分散共分散行列を表示します。このオプションは、反復変数が指定されている場合に限り使用できます。被験者変数が指定されている場合、共通ブロックが表示されます。
- **対比係数行列。**このオプションは、固定効果およびユーザー指定の仮説の検定に使用する推定可能関数を表示します。

信頼区間。信頼区間を構築する場合は、常にこの値を使用します。0 以上、100 未満の値を指定します。デフォルト値は 95 です。

線型混合モデルの EM 平均

図 5-7
[線型混合モデル: 推定周辺平均] ダイアログ ボックス



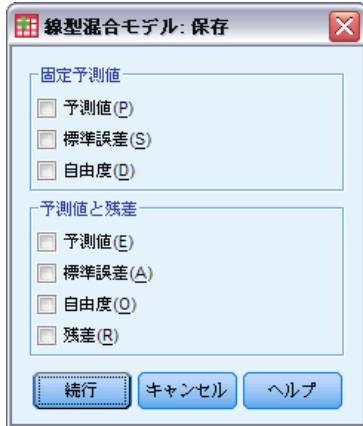
適合モデルの推定周辺平均。このグループでは、セル内の従属変数のモデル予測推定周辺平均と、指定した因子に対するそれらの標準誤差を求めることができます。さらに、主効果の因子の水準を比較することもできます。

- **因子と交互作用。**このリストには、[固定] ダイアログ ボックスで指定した因子と交互作用、および OVERALL 項が含まれています。共変量から作成したモデル項は、このリストから除外されます。
- **平均値の表示。**この手順で、このリストに選択した因子と交互作用の推定周辺平均を計算します。OVERALL を選択した場合は、すべての因子について、従属変数の推定周辺平均が表示されます。ただし、選択した因子または交互作用は、関連する変数がメイン ダイアログ ボックスの [因子] リストから除外されるまで、選択されたままの状態になります。
- **主効果の比較。**このオプションを使用して、選択した主効果の水準をペアごとに比較できます。[信頼区間の調整] では、多重比較に適用する信頼区間および有意確率を調整できます。利用できる方法は次のとおりです。[LSD (なし)]、[Bonferroni]、および [Sidak]。最後に、各因子に対して、比較の対象となる参照カテゴリを選択できます。参

照カテゴリを選択しない場合は、すべてのペアごとの比較が実行されます。参照カテゴリのオプションは、[最初]、[最後]、または [カスタム]（参照カテゴリの値を入力）です。

線型混合モデルの保存

図 5-8
[線型混合モデル: 保存] ダイアログ ボックス



このダイアログ ボックスを使用して、モデルの結果を作業ファイルに保存します。

固定予測値。効果のない回帰平均値と関連する変数を保存します。

- **予測値。**変量効果のない回帰平均値。
- **標準誤差。**推定値の標準誤差。
- **自由度。**推定値に関連する自由度。

予測値と残差。値を当てはめられたモデルに関連する変数を保存します。

- **予測値。**値を当てはめられたモデル。
- **標準誤差。**推定値の標準誤差。
- **自由度。**推定値に関連する自由度。
- **残差。**データ値から予測値を引いた値。

MIXED 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。

- 効果と効果や値の線型組み合わせ検定の指定 (TEST サブコマンドを使用)。
- ユーザー欠損値の包含 (MISSING サブコマンドを使用)。

- 指定した共変量の値の推定周辺平均の計算 (EMMEANS サブコマンドの WITH キーワードを使用)。
- 交互作用の単純な主効果の比較 (EMMEANS サブコマンドを使用)。

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

一般化線型モデル

一般化線型モデルは、指定したリンク関数によって従属変数が因子および共変量と線型関係になるよう、一般線型モデルを拡張したものです。さらにこのモデルでは、非正規分布の従属変数を使用することができます。一般化線型モデルは、定式化の一般性が非常に高いため、正規分布する応答に関する線型回帰、2 値データに関するロジスティック モデル、度数データに関する対数線型モデル、区間打ち切りの生存データに関する補ログ マイナス ログ モデルなど、一般によく使用される統計モデルだけでなく、その他のさまざまな統計モデルを網羅しています。

例。海運会社では一般化線型モデルを使用して、異なる期間に建設された複数の種類の船の損害数にポアソン回帰を当てはめることができ、構築されたモデルによって損害を受けやすい船の種類を判断することができます。

自動車保険会社では一般化線型モデルを使用して、自動車に対する損害請求にガンマ回帰を当てはめることができ、構築されたモデルによって請求に最も寄与する因子を判断することができます。

医療研究者は、一般化線型モデルを使用して区間打ち切り生存率データに補ログ マイナス ログを当てはめ、病状が再発する時間を予測します。

データ。応答には、尺度、度数、2 値、または試行におけるイベントがあります。因子は、カテゴリ型であるとしします。また、共変量、尺度重み付け、およびオフセットは、尺度であるとしします。

仮定。ケースは、独立した観測値と仮定します。

一般化線型モデルの構成

メニューから次の項目を選択します。

分析(A) > 一般化線型モデル > 一般化線型モデル...

図 6-1
[一般化線型モデル] の [モデルの種類] タブ



- ▶ 分布およびリンク関数を指定します（オプションの詳細は以下を参照してください）。
- ▶ [応答] タブで、従属変数を選択します。
- ▶ [予測変数] タブで、従属変数の予測に使用する因子および共変量を選択します。
- ▶ [モデル] タブで、選択した因子および共変量を使用してモデル効果を指定します。

[モデルの種類] タブで、モデルに対する分布およびリンク関数を指定し、応答の種類によって分類されているいくつかの一般モデルに対しショートカットを提供します。

モデルの種類

尺度の応答。

- **1次**。分布として [正規分布]、リンク関数として [同一] を指定します。
- **ログリンクを持つガンマ**。分布として [ガンマ]、リンク関数として [ログ] を指定します。

順位データ応答。

- **順位データ ロジスティック**。分布として多項順序を指定し、リンク関数として累積ロジットを指定します。
- **順位データ プロビット**。分布として多項順序を指定し、リンク関数として累積プロビットを指定します。

度数。

- **ポアソン対数線型**。分布として [ポアソン分布]、リンク関数として [ログ] を指定します。
- **ログリンクのある負の2項分布**。分布として [負の2項] (補助パラメータの値が 1)、リンク関数として [ログ] を指定します。手続きで補助パラメータの値を推定するには、負の2項分布を持つカスタム モデルを指定し、[パラメータ] グループで、[推定値] を選択します。

2 値応答またはイベント/試行データ。

- **2項ロジスティック**。分布として [2項]、リンク関数として [ロジット] を指定します。
- **2項プロビット**。分布として [2項]、リンク関数として [プロビット] を指定します。
- **区間打ち切り生存率**。分布として [2項]、リンク関数として [補ログ・マイナス・ログ] を指定します。

組み合わせ。

- **ログリンクの Tweedie**。分布として [Tweedie 分布]、リンク関数として [ログ] を指定します。
- **同一リンクの Tweedie**。分布として [Tweedie 分布]、リンク関数として [同一] を指定します。

ユーザー指定。分布とリンク関数の独自の組み合わせを指定します。

分布

次の選択項目を使用することにより、従属変数の分布を指定することができます。非正規分布や非恒等リンク関数を指定できるようになったことは、一般線型モデルに対する一般化線型モデルの大きな改善点です。分布とリンク関数とは数多くの組み合わせが可能であり、与えられたデー

タセットに適した組み合わせが複数ある場合もあります。そのため、事前に理論的な考察を行ったり、最適と思われる組み合わせを検討した上で、項目を選択します。

- **2 項**。この分布は、変数が 2 値反応またはイベント数のどちらかを表す場合に適合します。
- **ガンマ**。この分布は、値が正方向に大きくなるにつれて歪みを示す正のスケール値を持った変数に適合します。データ値が 0 以下または欠損値の場合、対応するケースは分析には使用されません。
- **逆ガウス分布**。この分布は、値が正方向に大きくなるにつれて歪みを示す正のスケール値を持った変数に適合します。データ値が 0 以下または欠損値の場合、対応するケースは分析には使用されません。
- **負の 2 項分布**。この分布は、k 回の成功を観測するまでに必要な試行の回数を表すと考えられ、非負整数値を取る変数に適合します。データ値が非整数、0 未満、または欠損値の場合、対応するケースは分析には使用されません。負の 2 項分布の補助パラメータ固定値は、0 以上の任意の数です。固定値に設定したり、手順によって推定することもできます。補助パラメータが 0 に設定されている場合、この分布はポアソン分布に一致します。
- **正規分布**。この分布は、スケール変数の取る値が中心値（平均値）に関して対称な釣鐘型に分布する場合に適合します。従属変数は、数値であることが必要です。
- **ポアソン分布**。この分布は、ある一定期間に目的のイベントが発生する回数を表すと考えられ、非負整数値を取る変数に適合します。データ値が非整数、0 未満、または欠損値の場合、対応するケースは分析には使用されません。
- **Tweedie 分布**。この分布は、ガンマ分布のポアソン混合によって表すことができる変数に適合します。この分布は、連続分布（負ではない実数値を取る）および離散分布（単一の値、0 の正の確率のマス（質量））を組み合わせるという点で「混合」です。従属変数は、数値型で 0 以上のデータ値を持つ必要があります。データ値が 0 未満、または欠損値の場合、対応するケースは分析には使用されません。Tweedie 分布のパラメータの固定値は、1 より大きく 2 より小さい任意の数です。
- **多項分布**。この分布は、順序応答を表す変数に適合します。従属変数は、数値または文字列のどちらかにすることができ、2 つ以上の有効なデータ値が含まれる必要があります。

リンク関数

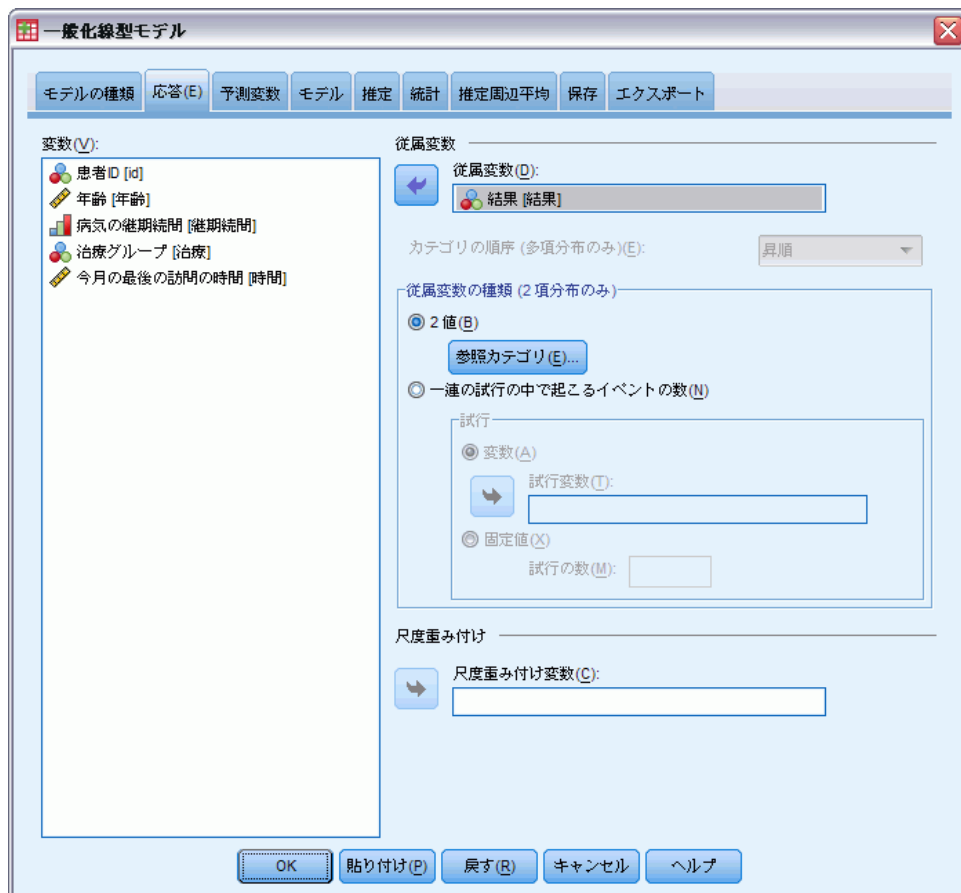
リンク関数とは、モデルの推定を行えるように従属変数を変換するためのものです。使用できる関数は次のとおりです。

- **同一**。f(x)=x。従属変数は変換されません。このリンク関数は、任意の分布で使用できます。

- **補ログ・マイナス・ログ**。 $f(x)=\log(-\log(1-x))$ 。このリンク関数は、2 項分布でだけ使用します。
- **累積コーチット**。 $f(x) = \tan(\pi (x - 0.5))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積補ログ マイナス ログ**。 $f(x)=\ln(-\ln(1-x))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積ロジット**。 $f(x)=\ln(x / (1-x))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積負ログ マイナス ログ**。 $f(x)=-\ln(-\ln(x))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積プロビット**。 $f(x)=\Phi^{-1}(x)$ 。応答の各カテゴリの累積確率に適用します。 Φ^{-1} は累積標準正規分布関数の逆関数です。このリンク関数は、多項分布でだけ使用します。
- **対数**。 $f(x)=\log(x)$ 。このリンク関数は、任意の分布で使用できます。
- **ログ補完**。 $f(x)=\log(1-x)$ 。このリンク関数は、2 項分布でだけ使用します。
- **ロジット**。 $f(x)=\log(x / (1-x))$ 。このリンク関数は、2 項分布でだけ使用します。
- **負の 2 項分布**。 $f(x)=\log(x / (x+k^{-1}))$ 。k は負の 2 項分布の補助パラメータです。このリンク関数は、負の 2 項分布でだけ使用します。
- **負ログ・マイナス・ログ**。 $f(x)=-\log(-\log(x))$ 。このリンク関数は、2 項分布でだけ使用します。
- **オッズべき乗**。 $\alpha \neq 0$ の場合は、 $f(x)=[(x/(1-x))^\alpha - 1] / \alpha$ 。 $\alpha=0$ の場合は、 $f(x)=\log(x)$ 。 α には常に、いずれかの実数を指定する必要があります。このリンク関数は、2 項分布でだけ使用します。
- **プロビット**。 $f(x)=\Phi^{-1}(x)$ 。ただし、 Φ^{-1} は累積標準正規分布関数の逆関数です。このリンク関数は、2 項分布でだけ使用します。
- **べき乗**。 $f(x)=x^\alpha$ ($\alpha \neq 0$ の場合)。 $f(x)=\log(x)$ ($\alpha=0$ の場合)。 α には常に、いずれかの実数を指定する必要があります。このリンク関数は、任意の分布で使用できます。

一般化線型モデル: 応答

図 6-2
[一般化線型モデル] ダイアログ ボックス



多くの場合、従属変数は、指定するだけで他に特別な操作は必要ありません。ただし、2つの値だけを取る変数や、試行で発生するイベントを記録する応答については、特別な注意が必要です。

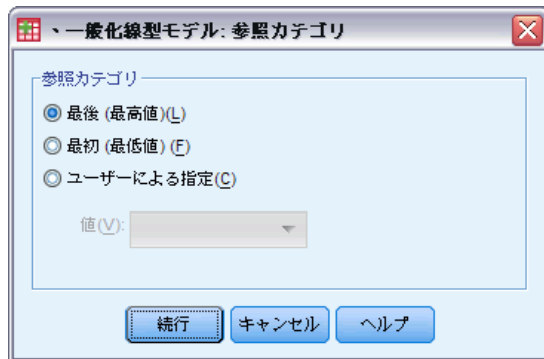
- **2値。**従属変数の取る値が2つだけの場合、パラメータ推定に対して **参照カテゴリ** を指定できます。2値反応変数には、文字または数値を指定できます。
- **一連の試行の中で起こるイベントの数。** 応答が、一連の試行を1回行うごとに発生するイベントの数である場合、その従属変数はイベントの数を表し、追加変数を選択して試行の数を表すことができます。ただし、試行数がすべての被験者に対して同じである場合は、固定値を使用して試行を指定することができます。試行数は、各ケースについて、イベント数以上である必要があります。また、イベント数は非負整数、試行数は正の整数である必要があります。

通常の多項モデルの場合、応答のカテゴリ順を、昇順、降順、またはデータ順の中から指定できます（データ順では、データで最初に出現した値が最初のカテゴリを定義し、最後の値が最後のカテゴリを定義します）。

尺度重み付け。 スケール パラメータは応答の分散に関する推定されたモデルパラメータです。尺度重み付けは、観測ごとに異なる「既知の」値です。尺度重み付け変数が指定された場合、応答の分散と関連性を持つ尺度パラメータは、各観測ごとに尺度重み付け変数によって分割されます。尺度重み付け値が 0 以下または欠損値の場合、対応するケースは分析には使用されません。

一般化線型モデルの参照カテゴリ

図 6-3
[一般化線型モデル: 参照カテゴリ] ダイアログ ボックス

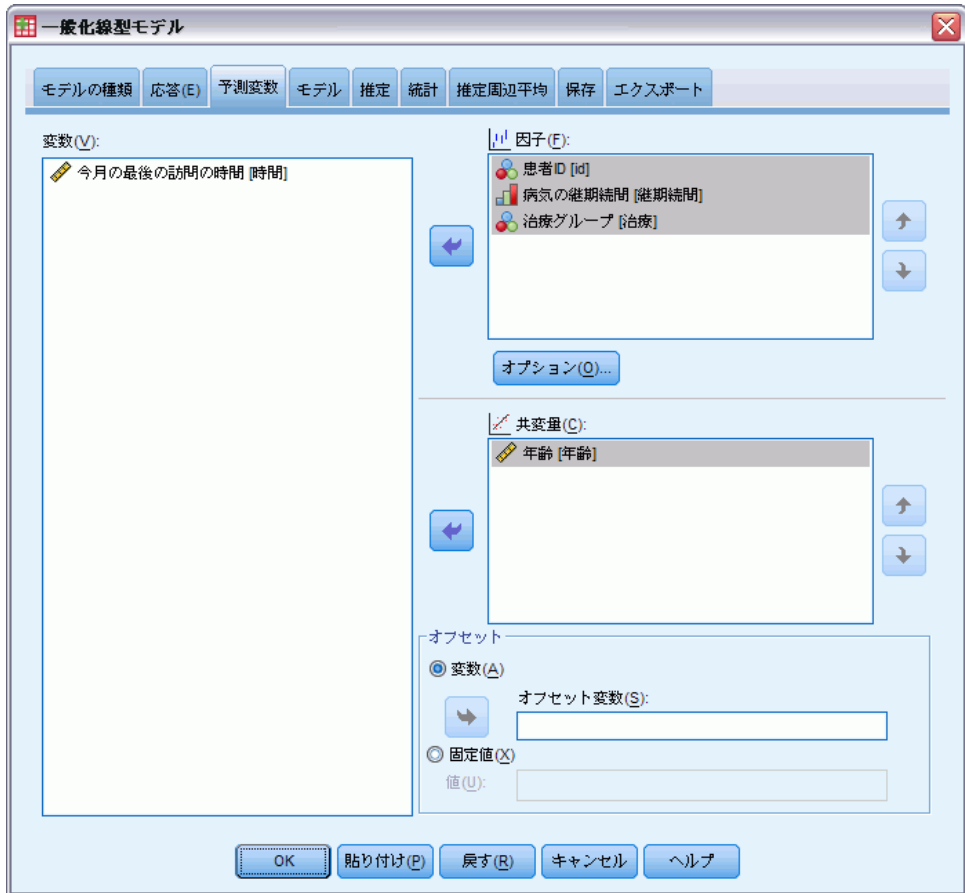


2 値反応では、従属変数に対して参照カテゴリを指定できます。参照カテゴリを指定すると、特定の出力（パラメータ推定値や保存値など）に影響が出ますが、モデルの適合度は変更されません。たとえば、2 値反応で値 0 と値 1 を取る場合は以下ようになります。

- デフォルトでは、手続きにより、最後（最高値）のカテゴリまたは 1（参照カテゴリ）になります。この状況では、与えられたケースが値 0 を取る確率がモデルで保存された確率によって予測されるため、パラメータ推定値は値 0 の尤度に関連しているものとして解釈されます。
- 最初（最低値）のカテゴリまたは 0（参照カテゴリ）を指定した場合は、与えられたケースが値 1 を取る確率が、モデルで保存された確率によって予測されます。
- ユーザー指定のカテゴリを指定した場合に、変数にラベルが定義されているときは、リストから値を選択して参照カテゴリを設定できます。この方法は、モデルを指定中に、特定の変数をコード化した方法を思い出せないときに役立ちます。

一般化線型モデル: 予測変数

図 6-4
Generalized Linear Models:[予測変数] タブ



[予測変数] タブでは、モデル効果の構成や任意オフセットの指定に使用する因子および共変量を指定できます。

因子。因子はカテゴリ予測変数です。因子には数値か文字を指定できます。

共変量。共変量は尺度予測変数で、数値を指定することが必要です。

注: 応答が 2 値形式の 2 項分布に従う場合、この手続きでは、選択した因子および共変量の観測値のクロス分類に基づく部分母集団ごとに、逸脱およびカイ 2 乗の適合度統計量が計算されます。部分母集団の数に一貫性を持たせるため、手続きを複数回実行する間、同じ予測変数を保持することが必要です。

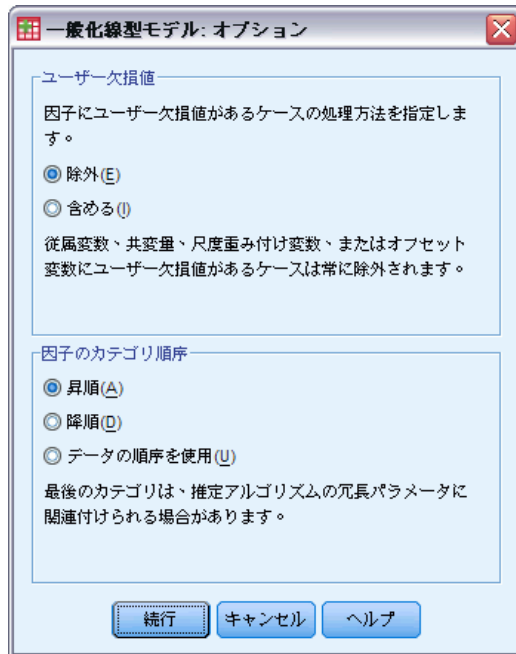
オフセット。オフセット項は、「構造的な」予測変数です。その係数は、モデルによって推定されるのではなく、常に 1 であると仮定されます。したがって、オフセットの値は単に、目標の線型予測変数に追加されるだけです。これは、ケースごとに目的のイベントが発生するレベルが異なるようなポアソン回帰モデルにおいて有用です。

たとえば、個人運転手の事故率をモデリングする場合、過失のある事故を起こした経験が 3 年の経歴の中で 1 度ある運転手と、25 年の経歴の中で 1 度しかない運転手とでは大きな違いがあります。運転手の経験をオフセット項として加味する場合、事故の発生数は対数リンクを持つポアソン応答または負の 2 項応答としてモデル化できます。

分布およびリンクの種類その他の組み合わせには、オフセット変数のその他の変換が必要です。

一般化線型モデルのオプション

図 6-5
[一般化線型モデル: オプション] ダイアログ ボックス



これらのオプションは、[予測変数] タブで指定されたすべての因子に適用されます。

ユーザー欠損値。因子は、分析の対象となるケースに対して有効な値を取る必要があります。このオプションを使用すると、ユーザー欠損値を因子変数で有効な値として扱うかどうかを決定できます。

カテゴリ順序。因子の最後のレベルを指定します。最後のレベルは、推定アルゴリズムの冗長パラメータに関連付けることができます。カテゴリの並び順を変更すると、因子レベル効果の値も変更されることがあります。これは、因子レベル効果の値が、「最後の」レベルに応じて計算されるパラメータ推定値であるためです。因子は、最小値から最大値への昇順、最大値から最小値への降順、または「データ順」で並び替えることができます。つまり、データ内で最初に検出された最初の値が最初のカテゴリを定義し、検出された最後の一意の値が最後のカテゴリを定義します。

一般化線型モデル: モデル

図 6-6
Generalized Linear Models:[モデル] タブ



モデル効果を指定。 デフォルトのモデルは定数項だけのモデルです。このため、その他のモデル効果は明示的に指定する必要があります。入れ子または非入れ子の項を構築することができます。

非入れ子項目

選択した因子や共変量について、次の項を作成できます。

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。

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。

因子。 選択した変数に交互作用および主効果を作成します。

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。

入れ子項目

この手続きでは、モデルに入れ子の項を構築できます。入れ子項目を構築すると、別の因子のレベルと相互作用しない因子または共変量の効果をモデリングできます。たとえば、ある食料雑貨店チェーンが、いくつかの店舗で顧客の消費傾向を調査すると仮定します。1 人の顧客が頻繁に通うのはこれらの店舗の 1 つだけであるため、「顧客」の効果は、「店舗の場所」の効果の**入れ子**になっていると言えます。

また、入れ子の項には、同一の共変量に関する多項式項などの交互作用効果を含めたり、複数レベルで入れ子になった項を追加したりできます。

制限。 入れ子項目には、次の制限があります。

- 1 つの交互作用内のすべての因子は一意でなければなりません。したがって、A が因子である場合、 $A*A$ の指定は無効です。
- 1 つの入れ子効果内のすべての因子は一意でなければなりません。したがって、A が因子である場合、 $A(A)$ の指定は無効です。
- 効果は共変量の入れ子にできません。したがって、A が因子で X が共変量である場合、 $A(X)$ の指定は無効です。

定数項。 通常、モデルには定数項が含まれます。データが原点を通ると仮定できる場合は、切片を除外できます。

多項順序分布のモデルには単一の定数項はありませんが、隣接するカテゴリの間の転移点を定義するしきい値パラメータがあります。しきい値は、常にモデルに含まれています。

一般化線型モデル: 推定

図 6-7
Generalized Linear Models:[推定] タブ

一般化線型モデル

モデルの種類 応答(E) 予測変数 モデル 推定 統計 推定周辺平均 保存 エクスポート

パラメータ推定

方法(M): Hybrid

Fisher スコアリングの最大反復回数(F): 1

尺度パラメータ法(C): 固定値

値(V): 1

分散共分散行列

モデルに基づく推定量(Q)

頑健推定量(B)

パラメータ推定値の初期値をデータセットから取得(G)

初期値(I)...

反復回数

最大反復回数(X): 100

最大段階 2分(U): 5

開始している反復(T): 20

データ ポイントの区切りを確認(K)

収束基準

0 より大きい値を最小値として、収束基準を少なくとも 1 つ指定する必要があります。

最小: 種類:

パラメータ推定値の変化(A) 1E-006 絶対値

対数尤度の変化(E) 絶対値

Hessian 収束(H) 絶対値

特異性許容範囲(S): 1E-012

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

パラメータ推定。このグループの項目を使用することにより、推定方法を指定したり、パラメータ推定値に初期値を与えたりすることができます。

- **方法。**パラメータ推定の方法を選択できます。Newton-Raphson 法、Fisher スコア法、および HYBRID 法（Fisher スコア法を何回か反復実行した後 Newton-Raphson 法に切り替える方法）のいずれかを選択できます。HYBRID 法における Fisher スコア法を実行中、その最大反復回数に到達する前に収束した場合も、Newton-Raphson 法のアルゴリズムは続行されます。
- **尺度パラメータ法。**尺度パラメータ推定の方法を選択できます。最尤法では、尺度パラメータとモデル効果が同時に推定されます。ただしこのオプションは、応答が 負の 2 項分布、ポワソン分布、2 項分布、または多項分布に従う場合は無効です。[逸脱] オプションまたは [Pearson のカイ 2 乗] オプションでは、これらの統計量の値に基づいて、尺

度パラメータの推定が行われます。ただし、尺度パラメータに固定値を指定することもできます。

- **初期値。**この手続きにより、パラメータに対する初期値が自動的に計算されます。ただし、パラメータ推定値に **初期値** を指定することもできます。
- **分散共分散行列** モデルに基づく推定量は、Hessian 行列の一般化逆行列に -1 を乗じたものです。頑健推定量 (Huber/White/サンドウィッチ推定量とも呼ばれる) は「修正された」モデルに基づく推定量で、分散やリンク関数の指定が不適切な場合でも、精度の高い共分散の推定を行うことができます。

反復回数。

- **最大反復回数。** アルゴリズムで実行される反復の最大回数です。負でない整数を指定してください。
- **最大段階 2 分。** 対数尤度が増加するか、最大段階 2 分に達するまで、反復ごとにステップ サイズが 0.5 の因子ずつ減少します。正の整数を指定します。
- **データポイントの区切りを確認。** 選択した場合、パラメータ推定値が一意な値になるようにアルゴリズムによって検定が行われます。区切りは、この手続きによって各ケースが正しく分類されるモデルを作成できる場合に存在します。このオプションは、2 値形式の多項反応および 2 項反応でのみ利用できます。

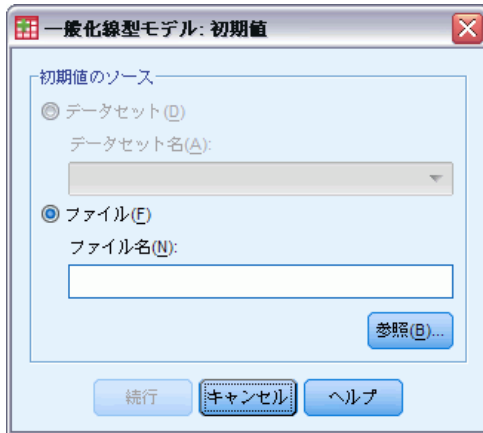
収束基準。

- **パラメータ推定値の変化。** 選択すると、パラメータ推定値の絶対変化または相対変化が指定した値よりも小さくなる反復が発生した後に、アルゴリズムが停止します。指定する値は正の数であることが必要です。
- **対数尤度収束。** 選択した場合、対数尤度関数の絶対変化または相対変化が指定した値よりも小さくなる反復が発生した後に、アルゴリズムが停止します。指定する値は正の数であることが必要です。
- **Hessian 収束。** [絶対値] を指定した場合は、Hessian 収束に基づく統計量が、指定した正の値よりも小さい場合に収束するとみなされます。[相対値] を指定した場合は、統計量が、指定した正の値と対数尤度の絶対値の積より小さい場合に収束するとみなされます。

特異性許容度。 特異 (または不可逆) 行列は、2 つ以上の列が線型従属になっており、推定アルゴリズムにとって大きな不都合となることがあります。準特異行列についても、十分な結果が得られない場合があります。そのためこの手続きでは、行列式が許容値未満である行列は、特異行列とみなされます。正の値を指定します。

一般化線型モデルの初期値

図 6-8
[一般化線型モデル: 初期値] ダイアログ ボックス



初期値を指定する場合は、モデル内のすべてのパラメータ（冗長パラメータも含む）に対して指定する必要があります。データセットでは、変数の順序は左から右へ、RowType_、VarName_、P1、P2 となる必要があります。この場合、RowType_ および VarName_ は文字型変数、P1、P2 はパラメータの順序づけられたリストに対応した数値型変数です。

- 初期値は、変数 RowType_ の値が EST であるレコードに指定されます。ただし実際の初期値は、変数 P1、P2… で与えられます。手続きでは、最初に現れた RowType_ の値が EST になっているレコードよりも後のレコードが無視されます。また、RowType_ の値が EST 以外のレコードもすべて無視されます。
- モデルまたはしきい値パラメータ内に定数項が含まれる場合、応答に多項分布が含まれる場合、定数項がリストされる最初の初期値である必要があります。
- 応答に負の 2 項分布が含まれている場合、尺度パラメータおよび負の 2 項パラメータは、指定される最後の初期値である必要があります。
- [分割ファイル] が有効になっている場合、変数の配列は、分割ファイル作成時に指定した順序で並んでいる分割ファイル変数で始まります。それに続く部分は、上記のとおり、RowType_、VarName_、P1、P2、… になります。指定したデータセット内の分割ファイルの順序は、元のデータセットでの順序と同じにします。

注: 変数名は必ずしも、P1、P2、… という形式である必要はありません。変数は、変数の名前ではなく位置に基づいてパラメータにマッピングされるため、手続きでは、パラメータに対して有効なものであれば任意の変数名を使用できます。最後のパラメータよりも後にある変数はすべて無視されます。

初期値に対するファイル構造は、モデルをデータとしてエクスポートする際に使用するものと同じです。したがって、一度手続きを実行して取得した最後の値を、それ以降に実行する手続きの入力として使用することができます。

一般化線型モデル: 統計量

図 6-9
Generalized Linear Models:[統計量] タブ

モデル効果。

- **分析の種類。** 作成する分析の種類を選択します。タイプ I 分析は一般に、理論的な動機に基づいてモデル内の予測変数を順序付ける場合に適しています。これに対し、タイプ III は、より広い範囲に適用できます。Wald または尤度比統計量は、カイ 2 乗統計量グループ内の選択に基づいて計算されます。

- **信頼区間。** 50 より大きく 100 より小さい信頼水準を指定します。Wald 区間は、パラメータが漸近正規分布に従うという仮定に基づくものです。プロファイル尤度区間はより正確ですが、計算が効率的でない場合があります。プロファイル尤度区間に関する許容水準は、区間の計算に使用される反復アルゴリズムを停止させるために使用する基準となります。
- **対数尤度関数。** 対数尤度関数の表示形式を制御します。完全形の関数には、パラメータ推定値に関して一定の追加項が含まれます。この項は、パラメータ推定に影響を及ぼすことはなく、ソフトウェア製品によっては表示対象から除外されます。

印刷。 オプションの出力は次のとおりです。

- **ケース処理の要約** 分析の対象となるケースおよび分析の対象から除外されるケースのそれぞれの数と割合、および [相関データの集計] 表が表示されます。
- **記述統計。** 記述統計量に加え、従属変数、共変量、および因子に関する要約情報が表示されます。
- **モデル情報。** データセット名、従属変数またはイベント/試行変数、オフセット変数、尺度重み付け変数、確率分布、およびリンク関数が表示されます。
- **適合度統計量。** 逸脱と尺度逸脱、Pearson のカイ 2 乗と尺度付き Pearson カイ 2 乗、対数尤度、赤池情報量基準 (AIC)、有限サンプル相関 AIC (AICC)、ベイズ情報量基準 (BIC)、一致 AIC (CAIC) が表示されます。
- **モデル要約統計量。** モデル適合度のオムニバス検定に関する対数尤度比統計量や、効果ごとのタイプ I またはタイプ III の対比に関する統計量を含むモデル適合度検定が表示されます。
- **パラメータ推定値。** パラメータ推定値およびそれに対応する検定統計量と信頼区間が表示されます。必要があれば、元のパラメータ推定値に加えて、指数化されたパラメータ推定値を表示することもできます。
- **パラメータ推定値の分散共分散行列。** 推定パラメータ分散共分散行列が表示されます。
- **パラメータ推定値の相関行列。** 推定パラメータ相関行列が表示されます。
- **対比係数 (L) 行列。** デフォルトの効果の対比係数が表示されます。また、[推定周辺平均] タブで指定されている場合は、推定周辺平均の対比係数も表示されます。
- **一般の推定可能関数。** 対比係数 (L) 行列を生成するための行列が表示されます。
- **反復の記述。** パラメータ推定値および対数尤度に関する反復の記述が表示されます。また、勾配ベクトルおよび Hessian 行列の最新の評価が出力されます。[反復の記述] 表には、0 回目 (初期推定値) から数えて n 回反復するごとに、パラメータ推定値が表示されます。ただし n

は、印刷間隔を示す値です。反復の記述を指定した場合、 n の値にかかわらず、最後の反復は常に出力されます。

- **LaGrange 乗数検定。** 尺度パラメータの妥当性を評価するための LaGrange 乗数検定統計量が表示されます。正規分布、ガンマ分布、逆ガウス分布、Tweedie 分布の場合、この尺度パラメータは、固定値に設定されるか、逸脱または Pearson カイ 2 乗を使用して計算されます。負の 2 項分布の場合は、固定値の補助パラメータが検定対象となります。

一般化線型モデル: 推定周辺平均

図 6-10
Generalized Linear Models:[EM 平均値] タブ

一般化線型モデル

モデルの種類 応答(E) 予測変数 モデル 推定 統計 推定周辺平均 保存 エクスポート

因子と交互作用(F):

| M | 項目 |
|-------------------------------------|------|
| <input checked="" type="checkbox"/> | 継期続間 |
| <input checked="" type="checkbox"/> | 治療 |
| <input type="checkbox"/> | id |

平均値の表示(O):

| 項目 | 対比 | 参照カテゴリ |
|---------|----|--------|
| 継期続間 | 単純 | 2 |
| 治療 | なし | |
| 継期続間*治療 | なし | |

乗算*(B)

スケール

応答の平均値を計算(C)

線型予測変数の平均値を計算(O)

多重比較の調整(A):

最小有意差

全体の推定平均を表示(S)

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

このタブでは、因子と交互作用のレベルについての推定周辺平均値を表示することができます。また、全体の推定平均を表示することもできます。推定周辺平均は、順序多項モデルには適合しません。

因子と交互作用。 このリストには、[予測変数] タブで指定した因子、および [モデル] タブで指定した交互作用が表示されます。共変量は、このリストから除外されます。このリストから項目を直接選択することができます。また、[乗算*] ボタンを使用し、項目を組み合わせることで交互作用を作成することもできます。

平均値の表示。 推定平均は、選択した因子と交互作用に対して計算されます。対比により、どのように仮説の検定を設定して推定平均を計算するのかを決定します。単純対比では、参照カテゴリ、つまり、比較対象とする因子レベルを指定する必要があります。

- **ペアごと。** ペアごとの比較は、明示的または暗黙的に指定された因子のあらゆるレベルでの組み合わせに対して実行されます。これは、交互作用に対して行える唯一の対比です。
- **単純(S).** 各水準の平均と指定された水準の平均を比較します。この対比の型はコントロール群があるときに有効です。
- **偏差。** 因子の各レベルを全平均値と比較します。偏差（全平均）対比は直交対比ではありません。
- **差分。** 各レベル（最初は除く）の平均値を前のレベルの平均値と比較します。（「逆 Helmert 対比」と呼ばれることもあります）。
- **Helmert.** 因子の各水準（最終を除く）の平均をその後の水準の平均と比較します。
- **反復測定.** 各水準（最終を除く）の平均をそのすぐ後の水準の平均と比較します。
- **多項式.** 1 次効果、2 次効果、3 次効果などを比較します。第 1 自由度にはすべてのカテゴリを通じての 1 次効果が含まれ、第 2 自由度には 2 次効果が含まれるなど、それぞれが含まれます。このような対比は、多項式のトレンドを推定する場合に頻繁に使用します。

スケール。 従属変数の元のスケールに基づいて、応答に対する推定周辺平均を計算できます。また、リンク関数で変換された従属変数に基づいて、線型予測変数に対する推定周辺平均を計算できます。

多重比較の調整。 多重対比を用いた仮説検定を行う場合、各対比に関する有意水準を基に、全体の有意水準を調整できます。このグループでは、調整方法を選択できます。

- **最小有異差。** この方法では、帰無仮説の値と異なる線型の対比が存在するという仮説を棄却する確率の全体的な制御を行いません。
- **Bonferroni の方法。** この方法は、多重比較の検定がなされるときに、有意確率を修正します。
- **逐次 Bonferroni.** 個々の仮説を棄却しますが、すべて同じ有意水準を維持することでは、それほど控えめでもない逐次ステップダウン棄却 Bonferroni 手続きです。

- **Sidak.** この方法は Bonferroni の方法よりもっと狭い限界を提供します。
- **逐次 Sidak.** これは、すべて同じ有意水準を維持したままで個々の仮説が棄却される保守的でない逐次ステップダウン相対 Sidak プロシジャです。

一般化線型モデル: 保存

図 6-11
Generalized Linear Models:[保存] タブ



確認済みの項が、指定した名前でも保存されます。その際、新しい変数と同じ名前を持つ既存の変数を上書きするか、名前の競合を避けるため新しい変数に接尾辞を追加して一意の名前を付けるかを選択することができます。

- **応答の平均値の予測値。** ケースごとにモデルで予測された値を、元の応答の計量値として保存します。応答分布が 2 項分布で従属変数が 2 変数の場合、この手続きは予測確率を保存します。応答分布は多項分布の場合、項目ラベルは 累積予測確率となり、指定されたカテゴリ数を

最大として保存する最後のカテゴリを除き、この手続きでは応答の各カテゴリに対し累積予測確率を保存します。

- **応答の平均値に対する信頼区間の下限。** 応答の平均に対する信頼区間の上限および下限を保存します。応答分布は多項分布の場合、項目ラベルは累積予測確率に対する信頼区間の下限となり、指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、この手続きでは応答の各カテゴリに対し下限を保存します。
- **応答の平均値に対する信頼区間の上限。** 応答の平均に対する信頼区間の上限および上限を保存します。応答分布は多項分布の場合、項目ラベルは累積予測確率に対する信頼区間の上限となり、指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、この手続きでは応答の各カテゴリに対し上限を保存します。
- **予測カテゴリ。** 2 項分布および 2 値従属変数、または多項分布を含むモデルに対し、各ケースの予測応答カテゴリを保存します。このオプションは、その他の応答分布には適用できません。
- **線型予測の予測値。** ケースごとにモデルで予測された値を、線型予測変数（指定したリンク関数で変換された応答）の計量値として保存します。応答分布が多項分布の場合、この手続きは指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、応答の各カテゴリに対して予測値を保存します。
- **線型予測変数の予測値に関する推定標準誤差。** 応答分布が多項分布の場合、この手続きは指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、応答の各カテゴリに対して推定標準誤差を保存します。

次の項目は、応答分布が多項分布の場合適用できません。

- **Cook の距離。** 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。Cook の距離が大きいときは、回帰統計量の計算からケースを除外したことが係数を実質的に変化させたことを示しています。
- **てこ比の値。** 回帰の適合性に対する 1 つの点の影響度を測定します。中心化てこ比の範囲は、0（適合性への影響なし）から $(N-1)/N$ までです。
- **生の残差。** 観測値およびモデルで予測された値との差。
- **Pearson 残差。** Pearson カイ 2 乗統計量に対するケースの寄与率の平方根。符号は、生の残差に従います。
- **標準化 Pearson 残差。** 尺度パラメータと「1 - ケースのてこ比」との積の逆数の平方根を Pearson 残差に乗じた値。
- **最大対数尤度比残差。** 最大対数尤度比統計量に対するケースの寄与率の平方根。符号は、生の残差に従います。

- **標準化最大対数尤度比残差。** 尺度パラメータと「1 - ケースのてこ比」との積の逆数の平方根を最大対数尤度比残差に乗じた値。
- **尤度残差。** 標準化 Pearson 残差の 2 乗と標準化最大対数尤度比残差の 2 乗の（ケースのてこ比に基づく）重み付け平均の平方根。

一般化線型モデル: エクスポート

図 6-12
Generalized Linear Models:[エクスポート] タブ



モデルをデータとしてエクスポート。 パラメータ相関行列またはパラメータ共分散行列から構成される IBM® SPSS® Statistics データセットを出力します。これらの行列には、パラメータ推定値、標準誤差、有意確率、および自由度が含まれています。行列ファイルの変数の順序は次のとおりです。

- **分割変数。** この変数が使用されている場合、いかなる変数も分割を定義します。

- **RowType_**. COV (共分散)、CORR (相関)、EST (パラメータ推定値)、SE (標準誤差)、SIG (有意水準)、および DF (抽出計画の自由度) の値 (および値ラベル) を取ります。各モデル パラメータに行型 COV (または CORR) を持つ個別のケースだけでなく、各モデル パラメータにその他の行型を持つ個別のケースがあります。
- **VarName_**。すべての推定モデル パラメータ (尺度パラメータまたは負の 2 項パラメータを除く) の順序づけられたリストに応じて、行型 COV または CORR に対し値 P1、P2、... を取り、パラメータ推定値テーブルに表示されたパラメータ文字列に従って値ラベルを付けます。その他の行型については、セルが空白になります。
- **P1、P2、...** これらの変数は、すべてのモデル パラメータ (必要に応じて尺度パラメータおよび負の 2 項パラメータを含む) の順序づけられたリストに対応しており、パラメータ推定テーブルに表示されたパラメータ文字列に従った変数ラベルを持ち、行型に従って値を取ります。

冗長パラメータの場合、すべての共分散は 0 に設定され、相関はシステム欠損値に設定されます。また、すべてのパラメータ推定値は 0 に設定され、すべての標準誤差、有意水準、残差自由度はシステム欠損値に設定されます。

尺度パラメータの場合、共分散、相関、有意水準および残差自由度はシステム欠損値に設定されます。最尤法を使用して尺度パラメータが推定された場合、標準誤差が指定されます。そうでない場合は、システム欠損値に設定されます。

負の 2 項パラメータの場合、共分散、相関、有意水準および残差自由度はシステム欠損値に設定されます。最尤法を使用して負の 2 項パラメータが推定された場合、標準誤差が指定されます。そうでない場合は、システム欠損値に設定されます。

分割がある場合、パラメータのリストはすべての分割全体で累積される必要があります。指定された分割内では、無関係のパラメータが存在する場合がありますが、これは冗長とは異なります。無関係なパラメータの場合、すべての共分散または相関、パラメータ推定値、標準誤差、有意水準、および残差自由度はシステム欠損値に設定されます。

この行列ファイルを初期値として使用し、より詳細なモデル推定を行うことができます。その他の手続きではここでエクスポートされたすべての行型が承認されない場合、行列ファイルを読み込むこれらの手続きでこのファイルをすぐに使用してより詳細な分析を行うことはできません。そのような場合でも、この行列ファイルのすべてのパラメータがファイルを読み込む手続きに対して同じ意味を持つことを確認する必要があります。

モデルを XML としてエクスポート。 選択した場合、パラメータ推定値とパラメータ分散共分散行列を XML (PMML) 形式で保存します。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。

GENLIN 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。

- パラメータ推定値に対する初期値を数値のリストとして指定 (CRITERIA サブコマンドを使用)。
- 推定周辺平均の計算時に、平均値ではなく値で共変量を固定 (EMMEANS サブコマンドを使用)。
- 推定周辺平均に対してユーザー指定の多項式対比を指定 (EMMEANS サブコマンドを使用)。
- 推定周辺平均を表示し指定したタイプの対比に基づいて比較を行う因子を指定 (EMMEANS サブコマンドの TABLES キーワードおよび COMPARE キーワードを使用)。

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。

一般化推定方程式

一般化推定方程式の手続きは、反復測定の実験や、クラスターデータなど
関連のある観測値の分析を行えるように、一般化線型モデルを拡張した
ものです。

例: 公衆衛生当局では、一般化推定方程式を使用して反復測定ロジス
ティック回帰を当てはめ、子供に与える大気汚染の影響を調査するこ
とができます。

データ。 応答には、尺度、度数、2 値、または試行におけるイベントがあり
ます。因子は、カテゴリ型であるとし、また、共変量、尺度重み付
け、およびオフセットは、尺度であるとし、また、被験者反復測定または被
験者内反復測定の定義に使用する変数は、応答の定義に使用すること
はできませんが、モデル内のその他の役割を担うことはできます。

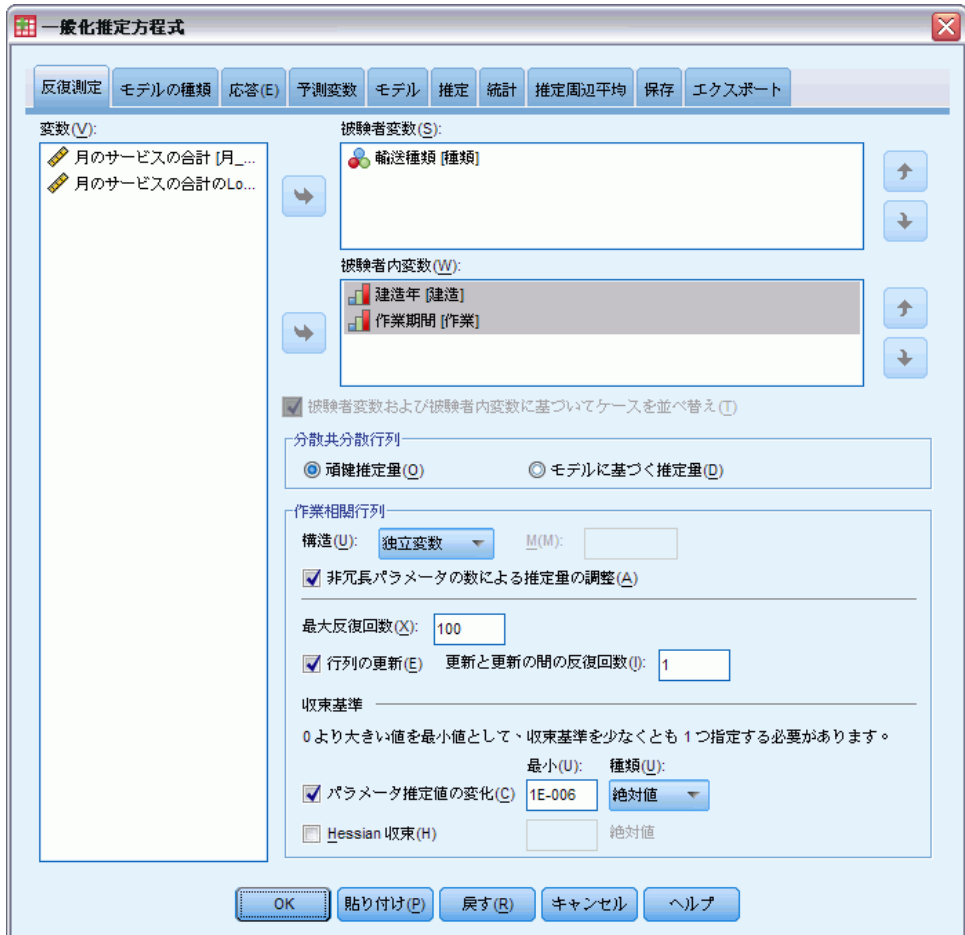
仮定。 各ケースは、被験者内では従属関係にあり、被験者間では相互に
独立していると仮定します。被験者内の従属関係を表す相関行列は、モ
デルの一部として推定されます。

一般化推定方程式を構成するには

メニューから次の項目を選択します。

分析(A) > Generalized Linear Models > 一般化推定方程式...

図 7-1
一般化推定方程式: [反復] タブ



- ▶ 被験者変数を 1 つ以上選択します（オプションについては下記参照）。

指定した変数の値の組み合わせにより、データセット内の **被験者** が一意に定義されることが必要です。たとえば、1 つの病院内の被験者を定義するには、患者 ID 変数が 1 つあれば十分ですが、複数の病院間で患者の ID 番号が重複する場合は、病院 ID と患者 ID を組み合わせて使用することが必要になります。反復測定では、被験者ごとに複数の観測値が記録されるため、各被験者がデータセット内の複数のケースを占めることがあります。

- ▶ [モデルの種類] タブで、分布およびリンク関数を指定します。
- ▶ [応答] タブで、従属変数を選択します。
- ▶ [予測変数] タブで、従属変数の予測に使用する因子および共変量を選択します。

- ▶ **[モデル]** タブで、選択した因子および共変量を使用してモデル効果を指定します。

必要であれば、**[反復測定]** タブで以下のオプションを指定できます。

被験者内変数。 被験者内変数の値の組み合わせにより、被験者内の測定順序が定義されます。そのため、被験者内変数および被験者変数の組み合わせにより、各測定が一意に定義されます。期間、病院 ID、および患者 ID の組み合わせにより、ケースごとに、特定の病院内の特定の患者に対する特定の外来診療が定義されます。

データセットの並べ替えがすでに行われており、各被験者の反復測定が、ケースの連続するブロックに正しい順序で現れる場合は、必ずしも被験者内変数を指定する必要はなく、**[被験者変数および被験者内変数に基づいてケースを並べ替え]** の選択を解除し、(一時的な) 並べ替えの実行に必要な処理時間を節約することができます。ただし通常は、測定の正しい順序が保証されるように、被験者内変数を利用することを推奨します。

被験者変数および被験者内変数は、応答の定義に使用することはできませんが、これらの変数は、モデル内の他の機能を実行することができます。たとえば、病院 ID をモデル内の因子として使用することができます。

分散共分散行列。 モデルに基づく推定量は、Hessian 行列の一般化逆行列に -1 を乗じたものです。頑健推定量 (Huber/White/サンドウィッチ推定量とも呼ばれる) は「補正された」モデルに基づく推定量で、作業相関行列の指定が不適切な場合でも、精度の高い共分散の推定を行うことができます。このオプションは、一般化推定方程式の線型モデル部分に含まれるパラメータに適用されますが、**[推定]** タブで指定した場合は、最初の一般化線型モデルにだけ適用されます。

作業相関行列。 この相関行列は、被験者内の従属関係を表します。この行列のサイズは、測定の数および被験者内変数の値の組み合わせによって決まります。構造は次のいずれかを選択できます。

- **独立。** 各反復測定は、相関関係を持ちません。
- **AR(1)。** 各反復測定は、1 次の自己回帰関係を持ちません。2 つの要素間の相関は、両要素が隣接する場合は ρ に等しく、1 要素分隔たりがある場合は ρ^2 となり、以下同様に続きます。 ρ は $-1 < \rho < 1$ となるよう制約されます。
- **交換。** この構造では、要素間に均質な相関があります。この構造は、複合シンメトリ構造とも呼ばれます。
- **M 従属。** 連続する測定間の相関係数はすべて同じです。また 1 つ分隔たりがある測定間の相関係数もすべて同じです。同様に、 $m-1$ 個分隔たりがある測定間の相関係数もすべて同じです。それ以上の隔たりがある

2 つの測定は、無相関であるとみなされます。この構造を選択した場合は、 m として作業相関行列の次数より小さい値を指定してください。

- **無構造。** これは一般の分散共分散行列です。

この手続きでは、非冗長パラメータの数による相関推定値の調整がデフォルトで行われます。データにおける被験者レベルでの複製変更に対して推定値を不変にするのであれば、この調整を行わない方が適切な場合もあります。

- **最大反復回数。** 一般化推定方程式アルゴリズムの実行を反復する最大回数です。負でない整数を指定してください。このオプションは、一般化推定方程式の線型モデル部分に含まれるパラメータに適用されますが、**[推定]** タブで指定した場合は、最初の一般化線型モデルにだけ適用されます。
- **行列の更新。** 作業相関行列の各要素は、アルゴリズムの実行が反復されるたびに更新されるパラメータ推定値に基づいて推定されます。作業相関行列がまったく更新されない場合は、推定プロセス全体を通して、最初の作業相関行列が使用されます。この行列が更新される場合は、作業相関行列の各要素が更新される間隔を、反復回数で指定することができます。1 より大きい値を指定すると、処理時間が短縮されることがあります。

収束基準。 これらの指定は、一般化推定方程式の線型モデル部分に含まれるパラメータに適用されますが、**[推定]** タブで指定した場合は、最初の一般化線型モデルにだけ適用されます。

- **パラメータ推定値の変化。** 選択すると、パラメータ推定値の絶対変化または相対変化が指定した値よりも小さくなる反復が発生した後に、アルゴリズムが停止します。指定する値は正の数であることが必要です。
- **Hessian 収束。** Hessian に基づく統計量が、指定した正の値よりも小さい場合に、収束するとみなされます。

一般化推定方程式: モデルの種類

図 7-2
一般化推定方程式: [モデルの種類] タブ

一般化推定方程式

反復測定 モデルの種類 応答(E) 予測変数 モデル 推定 統計 推定周辺平均 保存 エクスポート

分布およびリンク関数の組み合わせを指定したり、以下にリストされたモデルの種類のひとつを選択します。

尺度の応答

- 1次(L)
- ログリンクを持つガンマ(G)

順位データ応答

- 順位データロジスティック(O)
- 順位データプロビット(D)

度敷

- ポアソン対数線型(S)
- ログリンクのある負の2項分布(N)

2値応答またはイベント/試行データ

- 2値ロジスティック(B)
- 2値プロビット(A)
- 調査された生存推定値間隔(I)

組み合わせ

- ログリンクの Tweedie(T)
- 同一リンクの Tweedie(W)

ユーザー指定(C)

- ユーザー指定(C)

分布(U): 正規分布 リンク関数(E): 同一

べき乗(E):

パラメータ

- 値を指定(Y)
- 推定値(M)

値(V): 1

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

[モデルの種類] タブで、モデルに対する分布およびリンク関数を指定し、応答の種類によって分類されているいくつかの一般モデルに対しショートカットを提供します。

モデルの種類

尺度の応答。

- **1次**。分布として [正規分布]、リンク関数として [同一] を指定します。
- **ログリンクを持つガンマ**。分布として [ガンマ]、リンク関数として [ログ] を指定します。

順位データ応答。

- **順位データ ロジスティック**。分布として多項順序を指定し、リンク関数として累積ロジットを指定します。
- **順位データ プロビット**。分布として多項順序を指定し、リンク関数として累積プロビットを指定します。

度数。

- **ポアソン対数線型**。分布として [ポアソン分布]、リンク関数として [ログ] を指定します。
- **ログリンクのある負の 2 項分布**。分布として [負の 2 項] (補助パラメータの値が 1)、リンク関数として [ログ] を指定します。手続きで補助パラメータの値を推定するには、負の 2 項分布を持つカスタム モデルを指定し、[パラメータ] グループで、[推定値] を選択します。

2 値応答またはイベント/試行データ。

- **2 項ロジスティック**。分布として [2 項]、リンク関数として [ロジット] を指定します。
- **2 項プロビット**。分布として [2 項]、リンク関数として [プロビット] を指定します。
- **区間打ち切り生存率**。分布として [2 項]、リンク関数として [補ログ・マイナス・ログ] を指定します。

組み合わせ。

- **ログリンクの Tweedie**。分布として [Tweedie 分布]、リンク関数として [ログ] を指定します。
- **同一リンクの Tweedie**。分布として [Tweedie 分布]、リンク関数として [同一] を指定します。

ユーザー指定。分布とリンク関数の独自の組み合わせを指定します。

分布

次の選択項目を使用することにより、従属変数の分布を指定することができます。非正規分布や非恒等リンク関数を指定できるようになったことは、一般線型モデルに対する一般化線型モデルの大きな改善点です。分布とリンク関数とは数多くの組み合わせが可能であり、与えられたデータセットに適した組み合わせが複数ある場合もあります。そのため、事前に理論的な考察を行ったり、最適と思われる組み合わせを検討した上で、項目を選択します。

- **2 項**。この分布は、変数が 2 値反応またはイベント数のどちらかを表す場合に適合します。
- **ガンマ**。この分布は、値が正方向に大きくなるにつれて歪みを示す正のスケール値を持った変数に適合します。データ値が 0 以下または欠損値の場合、対応するケースは分析には使用されません。

- **逆ガウス分布。** この分布は、値が正方向に大きくなるにつれて歪みを示す正のスケール値を持った変数に適合します。データ値が 0 以下または欠損値の場合、対応するケースは分析には使用されません。
- **負の 2 項分布。** この分布は、k 回の成功を観測するまでに必要な試行の回数を表すと考えられ、非負整数値を取る変数に適合します。データ値が非整数、0 未満、または欠損値の場合、対応するケースは分析には使用されません。負の 2 項分布の補助パラメータ固定値は、0 以上の任意の数です。固定値に設定したり、手順によって推定することもできます。補助パラメータが 0 に設定されている場合、この分布はポアソン分布に一致します。
- **正規分布。** この分布は、スケール変数の取る値が中心値（平均値）に関して対称な釣鐘型に分布する場合に適合します。従属変数は、数値であることが必要です。
- **ポアソン分布。** この分布は、ある一定期間に目的のイベントが発生する回数を表すと考えられ、非負整数値を取る変数に適合します。データ値が非整数、0 未満、または欠損値の場合、対応するケースは分析には使用されません。
- **Tweedie 分布。** この分布は、ガンマ分布のポアソン混合によって表すことができる変数に適合します。この分布は、連続分布（負ではない実数値を取る）および離散分布（単一の値、0 の正の確率のマス（質量））を組み合わせるという点で「混合」です。従属変数は、数値型で 0 以上のデータ値を持つ必要があります。データ値が 0 未満、または欠損値の場合、対応するケースは分析には使用されません。Tweedie 分布のパラメータの固定値は、1 より大きく 2 より小さい任意の数です。
- **多項分布。** この分布は、順序応答を表す変数に適合します。従属変数は、数値または文字列のどちらかにすることができ、2 つ以上の有効なデータ値が含まれる必要があります。

リンク関数

リンク関数とは、モデルの推定を行えるように従属変数を変換するためのものです。使用できる関数は次のとおりです。

- **同一。** $f(x)=x$ 。従属変数は変換されません。このリンク関数は、任意の分布で使用できます。
- **補ログ・マイナス・ログ。** $f(x)=\log(-\log(1-x))$ 。このリンク関数は、2 項分布でだけ使用します。
- **累積コーチット。** $f(x) = \tan(\pi (x - 0.5))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積補ログ マイナス ログ。** $f(x)=\ln(-\ln(1-x))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積ロジット。** $f(x)=\ln(x / (1-x))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。

- **累積負ログ マイナス ログ。** $f(x)=-\ln(-\ln(x))$ 。応答の各カテゴリの累積確率に適用します。このリンク関数は、多項分布でだけ使用します。
- **累積プロビット。** $f(x)=\Phi^{-1}(x)$ 。応答の各カテゴリの累積確率に適用します。 Φ^{-1} は累積標準正規分布関数の逆関数です。このリンク関数は、多項分布でだけ使用します。
- **対数。** $f(x)=\log(x)$ 。このリンク関数は、任意の分布で使用できます。
- **ログ補完。** $f(x)=\log(1-x)$ 。このリンク関数は、2 項分布でだけ使用します。
- **ロジット。** $f(x)=\log(x / (1-x))$ 。このリンク関数は、2 項分布でだけ使用します。
- **負の 2 項分布。** $f(x)=\log(x / (x+k^{-1}))$ 。k は負の 2 項分布の補助パラメータです。このリンク関数は、負の 2 項分布でだけ使用します。
- **負ログ・マイナス・ログ。** $f(x)=-\log(-\log(x))$ 。このリンク関数は、2 項分布でだけ使用します。
- **オッズべき乗。** $\alpha \neq 0$ の場合は、 $f(x)=[(x/(1-x))^{\alpha}-1]/\alpha$ 。 $\alpha=0$ の場合は、 $f(x)=\log(x)$ 。 α には常に、いずれかの実数を指定する必要があります。このリンク関数は、2 項分布でだけ使用します。
- **プロビット。** $f(x)=\Phi^{-1}(x)$ 。ただし、 Φ^{-1} は累積標準正規分布関数の逆関数です。このリンク関数は、2 項分布でだけ使用します。
- **べき乗。** $f(x)=x^{\alpha}$ ($\alpha \neq 0$ の場合)。 $f(x)=\log(x)$ ($\alpha=0$ の場合)。 α には常に、いずれかの実数を指定する必要があります。このリンク関数は、任意の分布で使用できます。

一般化推定方程式: 応答

図 7-3
一般化推定方程式: [応答] タブ

一般化推定方程式

反復測定 モデルの種類 応答(E) 予測変数 モデル 推定 統計 推定周辺平均 保存 エクスポート

変数(V):
月のサービスの合計 [月_サービス]

従属変数
従属変数(D):
損傷件数 [損傷_件数]

カテゴリの順序 (多項分布のみ)(E):
昇順

従属変数の種類 (2項分布のみ)
 2値(B)
参照カテゴリ(E)...
 一連の試行の中で起こるイベントの数(N)
 試行
 変数(A)
 試行変数(I):

 固定値(X)
 試行の数(M):

尺度重み付け
尺度重み付け変数(C):

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

多くの場合、従属変数は、指定するだけで他に特別な操作は必要ありません。ただし、2つの値だけを取る変数や、試行で発生するイベントを記録する応答については、特別な注意が必要です。

- **2値。**従属変数の取る値が2つだけの場合、パラメータ推定に対して **参照カテゴリ** を指定できます。2値反応変数には、文字または数値を指定できます。
- **一連の試行の中で起こるイベントの数。**応答が、一連の試行を1回行うごとに発生するイベントの数である場合、その従属変数はイベントの数を表し、追加変数を選択して試行の数を表すことができます。ただし、試行数がすべての被験者に対して同じである場合は、固定値を使用して試行を指定することができます。試行数は、各ケースについて、イベント

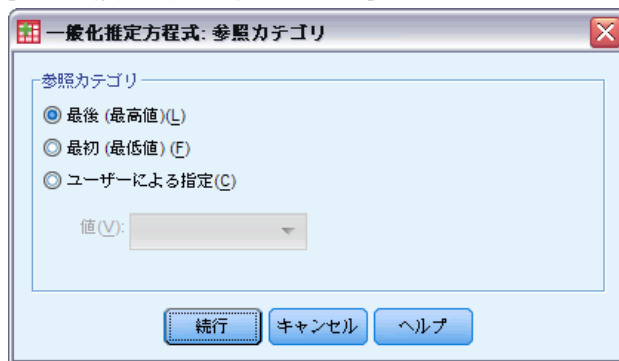
数以上であることが必要です。また、イベント数は非負整数、試行数は正の整数であることが必要です。

通常の多項モデルの場合、応答のカテゴリ順を、昇順、降順、またはデータ順の中から指定できます（データ順では、データで最初に出現した値が最初のカテゴリを定義し、最後の値が最後のカテゴリを定義します）。

尺度重み付け。 スケール パラメータは応答の分散に関する推定されたモデルパラメータです。尺度重み付けは、観測ごとに異なる「既知の」値です。尺度重み付け変数が指定された場合、応答の分散と関連性を持つ尺度パラメータは、各観測ごとに尺度重み付け変数によって分割されます。尺度重み付け値が 0 以下または欠損値の場合、対応するケースは分析には使用されません。

一般化推定方程式: 参照カテゴリ

図 7-4
[一般化推定方程式: 参照カテゴリ] ダイアログ ボックス

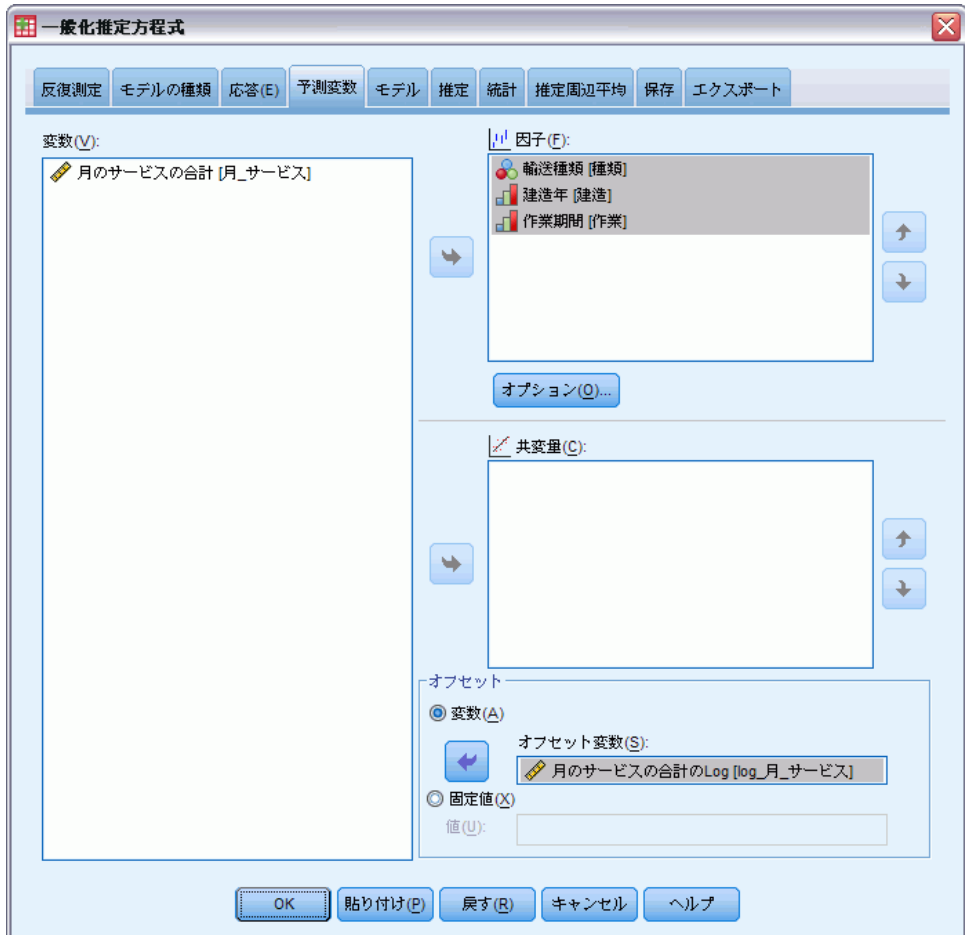


2 値反応では、従属変数に対して参照カテゴリを指定できます。参照カテゴリを指定すると、特定の出力（パラメータ推定値や保存値など）に影響が出ますが、モデルの適合度は変更されません。たとえば、2 値反応で値 0 と値 1 を取る場合は以下のようになります。

- デフォルトでは、手続きにより、最後（最高値）のカテゴリまたは 1（参照カテゴリ）になります。この状況では、与えられたケースが値 0 を取る確率がモデルで保存された確率によって予測されるため、パラメータ推定値は値 0 の尤度に関連しているものとして解釈されます。
- 最初（最低値）のカテゴリまたは 0（参照カテゴリ）を指定した場合は、与えられたケースが値 1 を取る確率が、モデルで保存された確率によって予測されます。
- ユーザー指定のカテゴリを指定した場合に、変数にラベルが定義されているときは、リストから値を選択して参照カテゴリを設定できます。この方法は、モデルを指定中に、特定の変数をコード化した方法を思い出せないときに役立ちます。

一般化推定方程式: 予測変数

図 7-5
一般化推定方程式: [予測値] タブ



[予測変数] タブでは、モデル効果の構成や任意オフセットの指定に使用する因子および共変量を指定できます。

因子。因子はカテゴリ予測変数です。因子には数値か文字を指定できます。

共変量。共変量は尺度予測変数で、数値を指定することが必要です。

注: 応答が 2 値形式の 2 項分布に従う場合、この手続きでは、選択した因子および共変量の観測値のクロス分類に基づく部分母集団ごとに、逸脱およびカイ 2 乗の適合度統計量が計算されます。部分母集団の数に一貫性を持たせるため、手続きを複数回実行する間、同じ予測変数を保持することが必要です。

オフセット。オフセット項は、「構造的な」予測変数です。その係数は、モデルによって推定されるのではなく、常に 1 であると仮定されます。したがって、オフセットの値は単に、目標の線型予測変数に追加されるだけです。これは、ケースごとに目的のイベントが発生するレベルが異なるようなポアソン回帰モデルにおいて有用です。

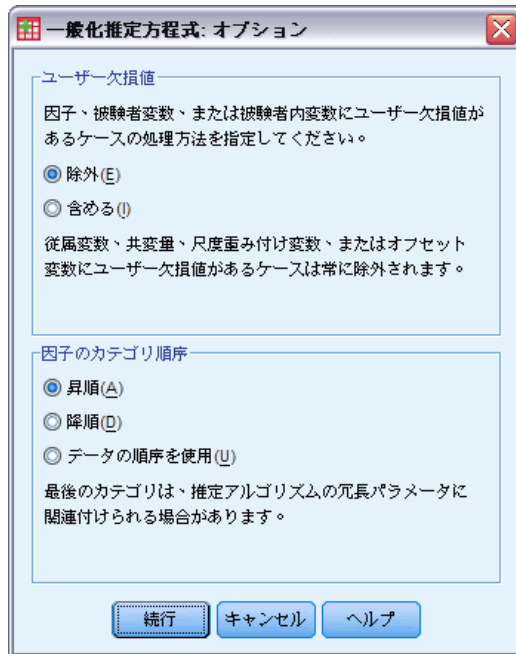
たとえば、個人運転手の事故率をモデリングする場合、過失のある事故を起こした経験が 3 年の経歴の中で 1 度ある運転手と、25 年の経歴の中で 1 度しかない運転手とでは大きな違いがあります。運転手の経験をオフセット項として加味する場合、事故の発生数は対数リンクを持つポアソン応答または負の 2 項応答としてモデル化できます。

分布およびリンクの種類のその他の組み合わせには、オフセット変数のその他の変換が必要です。

一般化推定方程式: オプション

図 7-6

[一般化推定方程式: オプション] ダイアログ ボックス



これらのオプションは、[予測変数] タブで指定されたすべての因子に適用されます。

ユーザー欠損値。因子は、分析の対象となるケースに対して有効な値を取る必要があります。このオプションを使用すると、ユーザー欠損値を因子変数で有効な値として扱うかどうかを決定できます。

カテゴリ順序。因子の最後のレベルを指定します。最後のレベルは、推定アルゴリズムの冗長パラメータに関連付けることができます。カテゴリの並び順を変更すると、因子レベル効果の値も変更されることがあります。これは、因子レベル効果の値が、「最後の」レベルに応じて計算されるパラメータ推定値であるためです。因子は、最小値から最大値への昇順、最大値から最小値への降順、または「データ順」で並べ替えることができます。つまり、データ内で最初に検出された最初の値が最初のカテゴリを定義し、検出された最後の一意の値が最後のカテゴリを定義します。

一般化推定方程式: モデル

図 7-7

一般化推定方程式: [モデル] タブ



モデル効果を指定。 デフォルトのモデルは定数項だけのモデルです。このため、その他のモデル効果は明示的に指定する必要があります。入れ子または非入れ子の項を構築することができます。

非入れ子項目

選択した因子や共変量について、次の項を作成できます。

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。

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。

因子。 選択した変数に交互作用および主効果を作成します。

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。

入れ子項目

この手続きでは、モデルに入れ子の項を構築できます。入れ子項目を構築すると、別の因子のレベルと相互作用しない因子または共変量の効果をモデリングできます。たとえば、ある食料雑貨店チェーンが、いくつかの店舗で顧客の消費傾向を調査すると仮定します。1 人の顧客が頻繁に通うのはこれらの店舗の 1 つだけであるため、「顧客」の効果は、「店舗の場所」の効果の**入れ子**になっていると言えます。

さらに、入れ子の項に交互作用効果を含めたり、複数のレベルの入れ子を追加することもできます。

制限。 入れ子項目には、次の制限があります。

- 1 つの交互作用内のすべての因子は一意でなければなりません。したがって、A が因子である場合、 $A*A$ の指定は無効です。
- 1 つの入れ子効果内のすべての因子は一意でなければなりません。したがって、A が因子である場合、 $A(A)$ の指定は無効です。
- 効果は共変量の入れ子にできません。したがって、A が因子で X が共変量である場合、 $A(X)$ の指定は無効です。

定数項。 通常、モデルには定数項が含まれます。データが原点を通ると仮定できる場合は、切片を除外できます。

多項順序分布のモデルには単一の定数項はありませんが、隣接するカテゴリの間の転移点を定義するしきい値パラメータがあります。しきい値は、常にモデルに含まれています。

一般化推定方程式: 推定

図 7-8
一般化推定方程式: [推定] タブ

一般化推定方程式

反復測定 モデルの種類 応答(E) 予測変数 モデル 推定 統計 推定周辺平均 保存 エクスポート

パラメータ推定

方法(M): Hybrid

Fisher スコアリングの最大反復回数(F): 1

尺度パラメータ法(C): 固定値 パラメータ推定値の初期値をデータセットから取得(G)

値(V): 1 初期値(I)...

反復回数

最大反復回数(X): 100 データポイントの区切りを確認(K)

最大段階 2分(U): 5 開始している反復(T): 20

収束基準

0より大きい値を最小値として、収束基準を少なくとも1つ指定する必要があります。

最小: 種類:

パラメータ推定値の変化(A) 1E-006 絶対値

対数尤度の変化(E) 絶対値

Hessian 収束(H) 絶対値

特異性許容範囲(S): 1E-012

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

パラメータ推定。 このグループの項目を使用することにより、推定方法を指定したり、パラメータ推定値に初期値を与えたりすることができます。

- **方法。** パラメータ推定の方法は選択することができます。Newton-Raphson 法、Fisher スコア法、および Fisher スコア法を何回か反復実行した後 Newton-Raphson 法に切り替える HYBRID 法の中からいずれかを選択できます。HYBRID 法における Fisher スコア法を実行中、その最大反復回数に到達する前に収束した場合も、Newton-Raphson 法のアルゴリズムは続行されます。
- **尺度パラメータ法。** 尺度パラメータ推定の方法を選択できます。

最尤法では、尺度パラメータとモデル効果が同時に推定されます。ただしこのオプションは、応答が負の 2 項分布、ポワソン分布、または 2 項分布に従う場合は無効です。一般化推定方程式には、尤度という概念はないため、ここで指定した内容は、最初の一般化線型モデルにだけ適用されます。この尺度パラメータ推定値は一般化推定方程式に渡されず。一般化推定方程式では、Pearson カイ 2 乗をその自由度で割った値により、尺度パラメータが更新されます。

[逸脱] オプションまたは [Pearson のカイ 2 乗] オプションでは、最初の一般化線型モデルにおけるこれらの統計量の値に基づいて、尺度パラメータの推定が行われます。この尺度パラメータ推定値は一般化推定方程式に渡されます。一般化推定方程式では、この尺度パラメータ推定値が固定値として扱われます。

または、尺度パラメータに固定値を指定することもできます。この尺度パラメータは、最初の一般化線型モデルおよび一般化推定方程式の推定では、固定値として扱われます。

- **初期値。** この手続きにより、パラメータに対する初期値が自動的に計算されます。ただし、パラメータ推定値に **初期値** を指定することもできます。

このタブで指定された反復回数および収束基準が適用されるのは、最初の一般化線型モデルだけです。一般化推定方程式を適合する際に使用する推定基準については、**[反復測定]** タブの項を参照してください。

反復回数。

- **最大反復回数。** アルゴリズムで実行される反復の最大回数です。負でない整数を指定してください。
- **最大段階 2 分。** 対数尤度が増加するか、最大段階 2 分に達するまで、反復ごとにステップ サイズが 0.5 の因子ずつ減少します。正の整数を指定します。
- **データポイントの区切りを確認。** 選択した場合、パラメータ推定値が一意な値になるようにアルゴリズムによって検定が行われます。区切りは、この手続きによって各ケースが正しく分類されるモデルを作成できる場合に存在します。このオプションは、2 値形式の多項反応および 2 項反応でのみ利用できます。

収束基準。

- **パラメータ推定値の変化。** 選択すると、パラメータ推定値の絶対変化または相対変化が指定した値よりも小さくなる反復が発生した後に、アルゴリズムが停止します。指定する値は正の数であることが必要です。

- **対数尤度収束。** 選択した場合、対数尤度関数の絶対変化または相対変化が指定した値よりも小さくなる反復が発生した後に、アルゴリズムが停止します。指定する値は正の数である必要があります。
- **Hessian 収束。** [絶対値] を指定した場合は、Hessian 収束に基づく統計量が、指定した正の値よりも小さい場合に収束するとみなされます。[相対値] を指定した場合は、統計量が、指定した正の値と対数尤度の絶対値の積より小さい場合に収束するとみなされます。

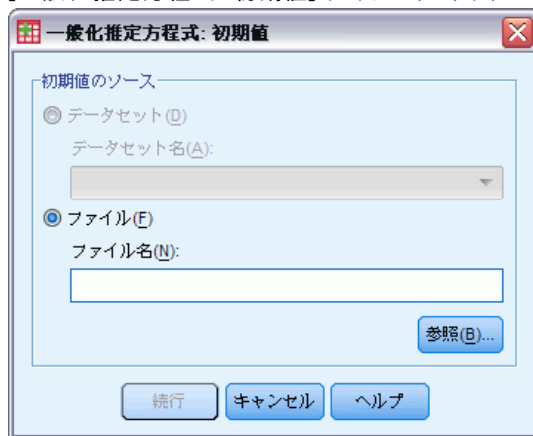
特異性許容度。 特異（または不可逆）行列は、2 つ以上の列が線型従属になっており、推定アルゴリズムにとって大きな不都合となることがあります。準特異行列についても、十分な結果が得られない場合があります。そのためこの手続きでは、行列式が許容値未満である行列は、特異行列とみなされます。正の値を指定します。

一般化推定方程式: 初期値

この手続きでは、一般化線型モデルの推定が行われます。このモデルの推定値は、一般化推定方程式の線型モデル部分におけるパラメータ推定値の初期値として使用されます。行列の要素はパラメータ推定値に基づいて算出されるため、作業相関行列に対する初期値は必要ありません。[推定] タブの [最大反復回数] の値が 0 に設定されていない場合、このダイアログボックスで指定された初期値は、一般化推定方程式ではなく、一般化線型モデルに対する出発点として使用されます。

図 7-9

[一般化推定方程式: 初期値] ダイアログ ボックス



初期値を指定する場合は、モデル内のすべてのパラメータ（冗長パラメータも含む）に対して指定する必要があります。データセットでは、変数の順序は左から右へ、RowType_、VarName_、P1、P2 となる必要があります。この場合、RowType_ および VarName_ は文字型変数、P1、P2 はパラメータの順序づけられたリストに対応した数値型変数です。

- 初期値は、変数 RowType_ の値が EST であるレコードに指定されます。ただし実際の初期値は、変数 P1、P2… で与えられます。手続きでは、最初に現れた RowType_ の値が EST になっているレコードよりも後のレコードが無視されます。また、RowType_ の値が EST 以外のレコードもすべて無視されます。
- モデルまたはしきい値パラメータ内に定数項が含まれる場合、応答に多項分布が含まれる場合、定数項がリストされる最初の初期値である必要があります。
- 応答に負の 2 項分布が含まれている場合、尺度パラメータおよび負の 2 項パラメータは、指定される最後の初期値である必要があります。
- [分割ファイル] が有効になっている場合、変数の配列は、分割ファイル作成時に指定した順序で並んでいる分割ファイル変数で始まります。それに続く部分は、上記のとおり、RowType_、VarName_、P1、P2、… になります。指定したデータセット内の分割ファイルの順序は、元のデータセットでの順序と同じにします。

注: 変数名は必ずしも、P1、P2、… という形式である必要はありません。変数は、変数の名前ではなく位置に基づいてパラメータにマッピングされるため、手続きでは、パラメータに対して有効なものであれば任意の変数名を使用できます。最後のパラメータよりも後にある変数はすべて無視されます。

初期値に対するファイル構造は、モデルをデータとしてエクスポートする際に使用するものと同じです。したがって、一度手続きを実行して取得した最後の値を、それ以降に実行する手続きの入力として使用することができます。

一般化推定方程式: 統計量

図 7-10
一般化推定方程式: [統計量] タブ

一般化推定方程式

反復測定 モデルの種類 応答(E) 予測変数 モデル 推定 **統計** 推定周辺平均 保存 エクスポート

モデル効果

分析の種類(A): タイプ III 信頼区間レベル(%) (V): 95

カイ 2 乗統計

Wald (W)
 一般化得点 (Z)

quasi-likelihood function ログ(L): カーネル

プリント

ケース処理要約(C) 対比係数 行列(T)
 記述統計(S) 一般の推定可能関数(U)
 モデル情報(M) 反復の記述(I)
 適合度統計量(G) 印刷の間隔(I): 1
 モデル要約統計量(Y)
 指数パラメータ推定値を含む(O)
 パラメータ推定値の分散共分散行列(X)
 パラメータ推定値の相関行列(N)
 作業相関行列(W)

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

モデル効果。

- **分析の種類。** 作成する分析の種類を選択し、モデルの効果を検定します。タイプ I 分析は一般に、理論的な動機に基づいてモデル内の予測変数を順序付ける場合に適しています。これに対し、タイプ III は、より広い範囲に適用できます。Wald または一般化スコア統計量は、カイ 2 乗統計量グループ内の選択に基づいて計算されます。

- **信頼区間。**50 より大きく 100 未満の信頼区間を指定します。Wald 区間は常に選択されたカイ 2 乗統計量の種類にかかわらず生成され、パラメータは漸近正規分布に従うという仮定に基づきます。
- **対数尤度関数。**対数尤度関数の表示形式を制御します。完全形の関数には、パラメータ推定値に関して一定の追加項が含まれます。この項は、パラメータ推定に影響を及ぼすことはなく、ソフトウェア製品によっては表示対象から除外されます。

印刷。出力できる内容は次のとおりです。

- **ケース処理の要約** 分析の対象となるケースおよび分析の対象から除外されるケースのそれぞれの数と割合、および [相関データの集計] 表が表示されます。
- **記述統計。**記述統計量に加え、従属変数、共変量、および因子に関する要約情報が表示されます。
- **モデル情報。**データセット名、従属変数またはイベント/試行変数、オフセット変数、尺度重み付け変数、確率分布、およびリンク関数が表示されます。
- **適合度統計量。**モデル選択に対する赤池情報量基準の 2 つの拡張を示します。独立モデルの準尤度基準 (QIC) で最適な相関構造を選択し、もう 1 つの QIC 方で予測の最適サブセットを選択します。
- **モデル要約統計量。**モデル適合度のオムニバス検定に関する対数尤度比統計量や、効果ごとのタイプ I またはタイプ III の対比に関する統計量を含むモデル適合度検定が表示されます。
- **パラメータ推定値。**パラメータ推定値およびそれに対応する検定統計量と信頼区間が表示されます。必要があれば、元のパラメータ推定値に加えて、指数化されたパラメータ推定値を表示することもできます。
- **パラメータ推定値の分散共分散行列。**推定パラメータ分散共分散行列が表示されます。
- **パラメータ推定値の相関行列。**推定パラメータ相関行列が表示されます。
- **対比係数 (L) 行列。**デフォルトの効果の対比係数が表示されます。また、[推定周辺平均] タブで指定されている場合は、推定周辺平均の対比係数も表示されます。
- **一般の推定可能関数。**対比係数 (L) 行列を生成するための行列が表示されます。
- **反復の記述。**パラメータ推定値および対数尤度に関する反復の記述が表示されます。また、勾配ベクトルおよび Hessian 行列の最新の評価が出力されます。[反復の記述] 表には、0 回目 (初期推定値) から数えて n 回反復するごとに、パラメータ推定値が表示されます。ただし n

は、印刷間隔を示す値です。反復の記述を指定した場合、n の値にかかわらず、最後の反復は常に出力されます。

- **作業相関行列。** 被験者内の従属関係を表す行列の値を表示します。その構造は、[反復測定] タブでの指定によって決まります。

一般化推定方程式: 推定周辺平均

図 7-11
一般化推定方程式: [推定周辺平均] タブ

The screenshot shows the 'Generalized Estimation Equation' dialog box with the 'Estimated Marginal Means' tab selected. The interface includes several sections:

- 反復測定**, **モデルの種類**, **応答(E)**, **予測変数**, **モデル**, **推定**, **統計**, **推定周辺平均**, **保存**, **エクスポート** (Navigation tabs)
- 因子と交互作用(F):** A table with columns 'M' and '項目'. Three items are listed: '種類' (checked), '建造' (checked), and '作業' (checked).
- 平均値の表示(O):** A table with columns '項目', '対比', and '参照カテゴリ'. One row is shown: '種類' (対比: 単純, 参照カテゴリ: 1).
- 乗算*(B)** (Calculation button)
- スケール:** Two radio buttons: '● 応答の平均値を計算(C)' (selected) and '○ 線型予測変数の平均値を計算(O)'.
- 多重比較の調整(A):** A dropdown menu set to '最小有意差'.
- 全体の推定平均を表示(S)** (Show overall estimated average checkbox)
- OK**, **貼り付け(P)**, **戻す(R)**, **キャンセル**, **ヘルプ** (Action buttons)

このタブでは、因子と交互作用のレベルについての推定周辺平均値を表示することができます。また、全体の推定平均を表示することもできます。推定周辺平均は、順序多項モデルには適合しません。

因子と交互作用。 このリストには、[予測変数] タブで指定した因子、および [モデル] タブで指定した交互作用が表示されます。共変量は、このリストから除外されます。このリストから項目を直接選択することがで

きます。また、[乗算*] ボタンを使用し、項目を組み合わせで交互作用を作成することもできます。

平均値の表示。 推定平均は、選択した因子と交互作用に対して計算されます。対比により、どのように仮説の検定を設定して推定平均を計算するのかを決定します。単純対比では、参照カテゴリ、つまり、比較対象とする因子レベルを指定する必要があります。

- **ペアごと。** ペアごとの比較は、明示的または暗黙的に指定された因子のあらゆるレベルでの組み合わせに対して実行されます。これは、交互作用に対して行える唯一の対比です。
- **単純(S).** 各水準の平均と指定された水準の平均を比較します。この対比の型はコントロール群があるときに有効です。
- **偏差。** 因子の各レベルを全平均値と比較します。偏差（全平均）対比は直交対比ではありません。
- **差分。** 各レベル（最初は除く）の平均値を前のレベルの平均値と比較します。（「逆 Helmert 対比」と呼ばれることもあります）。
- **Helmert.** 因子の各水準（最終を除く）の平均をその後の水準の平均と比較します。
- **反復測定.** 各水準（最終を除く）の平均をそのすぐ後の水準の平均と比較します。
- **多項式.** 1 次効果、2 次効果、3 次効果などを比較します。第 1 自由度にはすべてのカテゴリを通じての 1 次効果が含まれ、第 2 自由度には 2 次効果が含まれるなど、それぞれが含まれます。このような対比は、多項式のトレンドを推定する場合に頻繁に使用します。

スケール。 従属変数の元のスケールに基づいて、応答に対する推定周辺平均を計算できます。また、リンク関数で変換された従属変数に基づいて、線型予測変数に対する推定周辺平均を計算できます。

多重比較の調整。 多重対比を用いた仮説検定を行う場合、各対比に関する有意水準を基に、全体の有意水準を調整できます。このグループでは、調整方法を選択できます。

- **最小有異差。** この方法では、帰無仮説の値と異なる線型の対比が存在するという仮説を棄却する確率の全体的な制御を行いません。
- **Bonferroni の方法.** この方法は、多重比較の検定がなされるときに、有意確率を修正します。
- **逐次 Bonferroni.** 個々の仮説を棄却しますが、すべて同じ有意水準を維持することでは、それほど控えめでもない逐次ステップダウン棄却 Bonferroni 手続きです。

- **Sidak.** この方法は Bonferroni の方法よりもっと狭い限界を提供します。
- **逐次 Sidak.** これは、すべて同じ有意水準を維持したままで個々の仮説が棄却される保守的でない逐次ステップダウン相対 Sidak プロシジャです。

一般化推定方程式: 保存

図 7-12
一般化推定方程式: [保存] タブ



確認済みの項が、指定した名前でも保存されます。その際、新しい変数と同じ名前を持つ既存の変数を上書きするか、名前の競合を避けるため新しい変数に接尾辞を追加して一意の名前を付けるかを選択することができます。

- **応答の平均値の予測値。** ケースごとにモデルで予測された値を、元の応答の計量値として保存します。応答分布が 2 項分布で従属変数が 2 変数の場合、この手続きは予測確率を保存します。応答分布は多項分布

の場合、項目ラベルは 累積予測確率となり、指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、この手続きでは応答の各カテゴリに対し累積予測確率を保存します。

- **応答の平均値に対する信頼区間の下限。** 応答の平均に対する信頼区間の上限および下限を保存します。応答分布は多項分布の場合、項目ラベルは累積予測確率に対する信頼区間の下限となり、指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、この手続きでは応答の各カテゴリに対し下限を保存します。
- **応答の平均値に対する信頼区間の上限。** 応答の平均に対する信頼区間の上限および上限を保存します。応答分布は多項分布の場合、項目ラベルは累積予測確率に対する信頼区間の上限となり、指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、この手続きでは応答の各カテゴリに対し上限を保存します。
- **予測カテゴリ。** 2 項分布および 2 値従属変数、または多項分布を含むモデルに対し、各ケースの予測応答カテゴリを保存します。このオプションは、その他の応答分布には適用できません。
- **線型予測の予測値。** ケースごとにモデルで予測された値を、線型予測変数（指定したリンク関数で変換された応答）の計量値として保存します。応答分布が多項分布の場合、この手続きは指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、応答の各カテゴリに対して予測値を保存します。
- **線型予測変数の予測値に関する推定標準誤差。** 応答分布が多項分布の場合、この手続きは指定されたカテゴリ数を最大として保存する最後のカテゴリを除き、応答の各カテゴリに対して推定標準誤差を保存します。

次の項目は、応答分布が多項分布の場合適用できません。

- **生の残差。** 観測値およびモデルで予測された値との差。
- **Pearson 残差。** Pearson カイ 2 乗統計量に対するケースの寄与率の平方根。符号は、生の残差に従います。

一般化推定方程式: エクスポート

図 7-13
一般化推定方程式: [エクスポート] タブ



モデルをデータとしてエクスポート。 パラメータ相関行列またはパラメータ共分散行列から構成される IBM® SPSS® Statistics データセットを出力します。これらの行列には、パラメータ推定値、標準誤差、有意確率、および自由度が含まれています。行列ファイルの変数の順序は次のとおりです。

- **分割変数。** この変数が使用されている場合、いかなる変数も分割を定義します。
- **RowType_。** COV (共分散)、CORR (相関)、EST (パラメータ推定値)、SE (標準誤差)、SIG (有意水準)、および DF (抽出計画の自由度) の値 (および値ラベル) を取ります。各モデル パラメータに行型 COV (または CORR) を持つ個別のケースだけでなく、各モデル パラメータにその他の行型を持つ個別のケースがあります。

■ **VarName_**。すべての推定モデル パラメータ（尺度パラメータまたは負の 2 項パラメータを除く）の順序づけられたリストに応じて、行型 COV または CORR に対し値 P1、P2、... を取り、パラメータ推定値テーブルに表示されたパラメータ文字列に従って値ラベルを付けます。その他の行型については、セルが空白になります。

■ **P1、P2、...** これらの変数は、すべてのモデル パラメータ（必要に応じて尺度パラメータおよび負の 2 項パラメータを含む）の順序づけられたリストに対応しており、パラメータ推定テーブルに表示されたパラメータ文字列に従った変数ラベルを持ち、行型に従って値を取ります。

冗長パラメータの場合、すべての共分散は 0 に設定され、相関はシステム欠損値に設定されます。また、すべてのパラメータ推定値は 0 に設定され、すべての標準誤差、有意水準、残差自由度はシステム欠損値に設定されます。

尺度パラメータの場合、共分散、相関、有意水準および残差自由度はシステム欠損値に設定されます。最尤法を使用して尺度パラメータが推定された場合、標準誤差が指定されます。そうでない場合は、システム欠損値に設定されます。

負の 2 項パラメータの場合、共分散、相関、有意水準および残差自由度はシステム欠損値に設定されます。最尤法を使用して負の 2 項パラメータが推定された場合、標準誤差が指定されます。そうでない場合は、システム欠損値に設定されます。

分割がある場合、パラメータのリストはすべての分割全体で累積される必要があります。指定された分割内では、無関係のパラメータが存在する場合がありますが、これは冗長とは異なります。無関係なパラメータの場合、すべての共分散または相関、パラメータ推定値、標準誤差、有意水準、および残差自由度はシステム欠損値に設定されます。

この行列ファイルを初期値として使用し、より詳細なモデル推定を行うことができます。その他の手続きではここでエクスポートされたすべての行型が承認されない場合、行列ファイルを読み込むこれらの手続きでこのファイルをすぐに使用してより詳細な分析を行うことはできません。そのような場合でも、この行列ファイルのすべてのパラメータがファイルを読み込む手続きに対して同じ意味を持つことを確認する必要があります。

モデルを XML としてエクスポート。 選択した場合、パラメータ推定値とパラメータ分散共分散行列を XML (PMML) 形式で保存します。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。

GENLIN 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。

- パラメータ推定値に対する初期値を数値のリストとして指定 (CRITERIA サブコマンドを使用)。
- 固定作業相関行列の指定 (REPEATED サブコマンドを使用)。
- 推定周辺平均の計算時に、平均値ではなく値で共変量を固定 (EMMEANS サブコマンドを使用)。
- 推定周辺平均に対してユーザー指定の多項式対比を指定 (EMMEANS サブコマンドを使用)。
- 推定周辺平均を表示し指定したタイプの対比に基づいて比較を行う因子を指定 (EMMEANS サブコマンドの TABLES キーワードおよび COMPARE キーワードを使用)。

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。

一般化線型混合モデル

一般化線型混合モデルが線型モデルを拡張すると、次のようになります。

- 目標は指定されたリンク関数を介して因子および共変量に線型的に関連します。
- 目標の分布は非正規分布です。
- 観測は関連します。

一般化線形混合モデルは、単純な線形回帰から非正常な時系列データの複雑なマルチレベルに至るまで、さまざまなモデルをカバーします。

例。 地区の教育委員会は、実験的な教授法が数学の点数を向上させる効果的であるかどうかを判断するために一般化線形混合モデルを使用することができます。生徒は同じ教師によって教えられているので、同じ教室の生徒を関連させる必要があります、同じ学校内の教室にも関連がある場合があるため、変動の異なるソースを考慮して学校やクラスのレベルで変量効果を含めることができます。

医学研究者は、新しい抗てんかん薬が、てんかん発作の患者の割合を減らすことができるかどうかを判断するために一般化線形混合モデルを使用することができます。同一患者から繰り返し測定を行う場合通常、正の相関関係があるため、いくつかの変量効果を持つ混合モデルが適切となります。対象フィールドである発作の数は、正の整数値をとるため、ポアソン分布と対数リンクを持つ一般化線形混合モデルが適切となる場合があります。

テレビ、電話、およびインターネットサービスのケーブル プロバイダーの経営陣は、潜在的な顧客についての詳細を知るために一般化線形混合モデルを使用することができます。考えられる回答は名義型尺度であるため、同社のアナリストは、所定の調査応答者の回答の中のサービスの種類（テレビ、電話、インターネット）全体のサービスの使い方についての質問への回答との相関関係をキャプチャするために、ランダム切片と一般化ロジット混合モデルを使用します。

図 8-1
[データ構造] タブ

データ構造 フィールドと効果 作成オプション モデル オプション

データ構造は?
この手順では、複数のレコードが単一の被験者の反復した測定を示すと想定します。

フィールド: 並べ替え: なし(N)

- Center ID
- Gender
- Date of birth
- Treatment received

キャンパス:

| | | | | |
|-------------|------------------------|------------|------|------|
| 被験者 | | | 反復測定 | |
| Center size | Attending Physician ID | Patient ID | Week | |
| | | Patient ID | Week | |
| | Attending Physician ID | Patient ID | Week | |
| | | Patient ID | Week | |
| | | | | Week |
| | | | | Week |

すべて(A) [Icons]

その他

共分散グループを定義: [Empty Box] [Icons]

反復測定共分散: 1次自己回帰 (AR1)

[データ構造] タブは、観測値が相関しているときに、データセット内のレコード間の構造的な関係を指定することができます。データセット内のレコードが独立した観察を表している場合、このタブでは何も指定する必要はありません。

被験者: 指定したカテゴリ型フィールドの組み合わせにより、データセット内の被験者が一意に定義されることが必要です。たとえば、1つの病院内の被験者を定義するには、患者 ID 変数が 1 つあれば十分ですが、複数の病院間で患者の ID 番号が重複する場合は、病院 ID と患者 ID を組み合わせて使用することが必要になります。反復測定では、被験者ごとに複数の観測値が記録されるため、各被験者がデータセット内の複数のレコードを占めることがあります。

被験者は、その他の被験者から独立している見なすことができる観測単位です。たとえば、医学研究では患者の血圧の測定値は、他の患者の測定値から独立していると見なすことができます。被験者の定義は、被験者ごとに測定を繰り返す場合、これらの観測間の相関関係をモデル化したい場

合に重要になります。たとえば、病院に連続して訪問する際に一人の患者の血圧測定値は相関していると期待できる場合があります。

[データ構造] タブの [被験者] として指定されたすべてのフィールドを使用して、残差共分散構造の被験者を定義し、[変量効果ブロック](#)の変量効果共分散構造の被験者を定義するフィールドのリストを提供します。

反復測定: ここで指定されたフィールドは、繰り返し観測を識別するために使用されます。たとえば、変数 [週] は医療研究の 10 週間にわたる観測を指定、また [月] および [日] を同時に使用して 1 年間にわたって行われる日常の観測を示します。

共分散グループを定義: ここで指定されたフィールドは、反復効果の共分散パラメータの独立したセットを定義します。カテゴリごとに 1 つずつ、グループ化フィールドのクロス分類によって定義されます。すべての被験者は、同じ共分散のタイプです。同じ共分散グループ内の被験者は、パラメータに同じ値を持つことになります。

反復共分散タイプ: 残差の共分散構造を指定します。使用できる構造は次のとおりです。

- 一次の自己回帰 (AR1)
- 自己回帰移動平均 (1, 1) (ARMA11)
- 複合対称性
- 対角性
- 尺度化された識別
- Toeplitz
- 無構造
- 分散成分

詳細は、[B 付録 p. 183 共分散構造](#) を参照してください。

一般化線型混合モデルの取得

この機能を使うには、[高度な統計] オプションが必要です。

メニューから次の項目を選択します。

分析 > 混合モデル > 一般化線型モデル

- ▶ [データ構造] タブでデータセットの構造を定義します。
- ▶ [フィールドと効果] タブには、任意の尺度、またはイベントとトライアルの仕様（その場合はイベントとトライアルの使用が連続型でなければなりません）を持つことができる単一の対象が必要です。オプションで、その分布

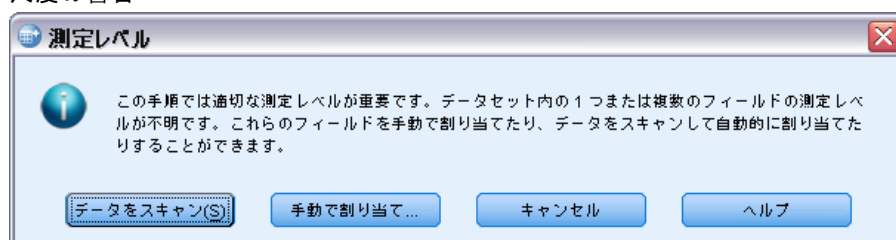
とのリンク関数、固定効果、および任意の変数効果ブロック、オフセット、または分析の重みを指定します。

- ▶ [作成オプション] をクリックして、オプションの作成設定を指定します。
- ▶ [モデル オプション] をクリックして、有効なデータセットにスコアを保存し、モデルを外部ファイルにエクスポートします。
- ▶ [実行] をクリックして、手順を実行し、モデルオブジェクトを作成します。

不明な尺度のフィールド

データセットの 1 つまたは複数の変数（フィールド）の尺度が不明な場合、尺度の警告が表示されます。尺度はこの手順の結果の計算に影響を与えるため、すべての変数に尺度を定義する必要があります。

図 8-2
尺度の警告



- **データをスキャン。** アクティブ データセットのデータを読み込み、デフォルトの尺度を尺度が現在不明なフィールドに割り当てます。データセットが大きい場合は時間がかかります。
- **手動で割り当てる。** 不明な尺度のフィールドをすべて表示するダイアログが開きます。このダイアログを使用して、尺度をこれらのフィールドに割り当てることができます。データ エディタの [変数ビュー] でも、尺度を割り当てることができます。

尺度がこの手順で重要であるため、すべてのフィールドに尺度が定義されるまで、ダイアログにアクセスしてこの手順を実行することはできません。

対象

図 8-3
目標設定

これらの設定は、リンク関数を介してターゲット、その分布、および予測因子との関係を定義します。

対象：対称は必須です。これは、任意の尺度を持つことができ、対象の尺度は適切な分布とリンク関数を制限します。

- **試行数を分母として使用する：**対象レスポンスが一定の試行回数のセット内で発生したイベント数の場合、対象フィールドにはイベント数が含まれ、この試行回数を含んだ追加のフィールドを選択できます。たとえば、新しい農薬をテストするときは、さまざまな濃度の農薬をアリのサンプルに噴霧して死んだアリの数と各サンプルのアリの数を記録します。この場合、死んだアリの数を記録するフィールドは、対象（イベント）フィールドとして指定されなければならないが、各サンプル中のアリの数を記録するフィールドは、試行フィールドとして指定する必要があります。

ます。アリの数は、各サンプルに対して同じである場合、試行回数は、固定値を使用して指定することができます。

試行回数は、各レコードのイベント数以上である必要があります。また、イベント数は非負整数、試行数は正の整数であることが必要です。

- **参照カテゴリのカスタマイズ:** カテゴリ型対象の場合、参照カテゴリを選択できます。このことでパラメータ推定値などの一定の出力に影響を与えることができますが、モデルの適合度を変更してはなりません。たとえば、対象の値がデフォルトで0、1、および2となる場合、手順では最後の（最も高い値を持つ）カテゴリ、または2を参照のカテゴリにします。この場合、パラメータ推定はカテゴリ2の尤度に相対してカテゴリ0または1に関連すると解釈されます。カスタムカテゴリを指定して、対象がラベルを定義している場合、リストから値を選択して基準のカテゴリを設定することができます。これは、モデル指定の途中で特定のフィールドがどのようにコーディングされたか正確にわからないときに便利です。

目標分布および線型モデルとの関係 (リンク): 予測変数の値を考えると、モデルはターゲットの値の分布が、指定された形状に従うことを期待し、そして目標値については、指定したリンク機能を通じて、予測に直線的に関連することを期待します。共通モデルへのショートカットが提供され、または、特定の分布とリンク関数の組み合わせがあり、その短いリスト上に収まらない場合、[カスタム] 設定を選択します。

- **線型モデル:** 対象が線形回帰やANOVAモデルを用いて予測できる場合に有用な、識別リンクの正規分布を指定します。
- **ガンマ回帰:** 対象がすべて正の値が格納されていると大きな値に向かって偏っている場合に使用される対数リンクのガンマ分布を指定します。
- **対数線型:** 対象が一定期間における発生回数を表すときに使用される対数リンクのポアソン分布を指定します。
- **負の二項回帰:** 対象と分母が k の成功を観測するために必要な試行回数を表す場合に使用される対数リンクの負の二項分布を指定します。
- **多項ロジスティック回帰:** 対象がマルチカテゴリ レスポンスである場合に使用される多項分布を指定します。累積ロジットリンク（順序型結果）または一般化ロジット リンク（マルチカテゴリ名義型レスポンス）を使用します。
- **二項ロジスティック回帰:** 対象がロジスティック回帰モデルで予測される二項レスポンスである場合に使用されるロジットリンクの二項分布を指定します。
- **二項プロビット:** 対象が基礎となる正規分布の二項レスポンスである場合に使用されるプロビットリンクの二項分布を指定します。
- **区間打ち切り生存:** いくつかの観測がなく終了イベントを持たない場合生存期間分析で有用な補ログ・マイナス・ログの二項分布を指定します。

分布

このセクションで、対象の分布を指定します。非正規分布と非恒等式リンク関数を指定する機能は、線形混合モデルを超える、一般化線型混合モデルの本質的な機能の向上です。分布とリンク関数には多くの組み合わせの可能性があります、その中のいくつかは特定のデータセットに適切な場合があるので、この選択が先験的な理論考察または一番適合するよう見える組み合わせによって導き出される可能性があります。

- **2 項**。この分布は、二者択一のレスポンスまたはイベント数を表す対象に対してのみ、適しています。
- **ガンマ**。この分布は、より大きな正数値の方向へ歪められる正のスケール値を持つ対象に適しています。データの値が 0 以下または欠損している場合は、対応するケースが分析に使用されません。
- **逆ガウス分布**。この分布は、より大きな正数値の方向へ歪められる正のスケール値を持つ対象に適しています。データの値が 0 以下または欠損している場合は、対応するケースが分析に使用されません。
- **多項分布**。この分布は、マルチカテゴリ レスポンスを表す対象に対してのみ、適しています。モデルの形式は、対象の尺度によって異なります。

名義型対象は、モデル パラメータの異なるセットが、（参照カテゴリを除く）対象のカテゴリごとに推定される名義型多項モデルとなります。特定の予測のためのパラメータ推定値は、参照カテゴリに相対的に、予測値と対象の各カテゴリの尤度との関連性を示します。

順序型対象は、従来の切片項が、対象カテゴリの累積確率に関連する**閾値**パラメータのセットと置き換えられる順序型多項モデルとなります。

- **負の 2 項分布**。負の二項回帰は、ターゲットが高い分散の発生回数を表すときに使用される対数リンクの負の二項分布を使用しています。
- **正規**。これは、値が中央（平均値）の値から対称的に釣鐘型に分布する連続型対象に適しています。
- **ポアソン**。この分布は、長さが固定している期間内に発生した興味あるイベントの発生回数として考えられ、負でない整数を持つ変数に適しています。データの値が非正数、0 より小さい、または欠損している場合は、対応するケースが分析に使用されません。

リンク関数

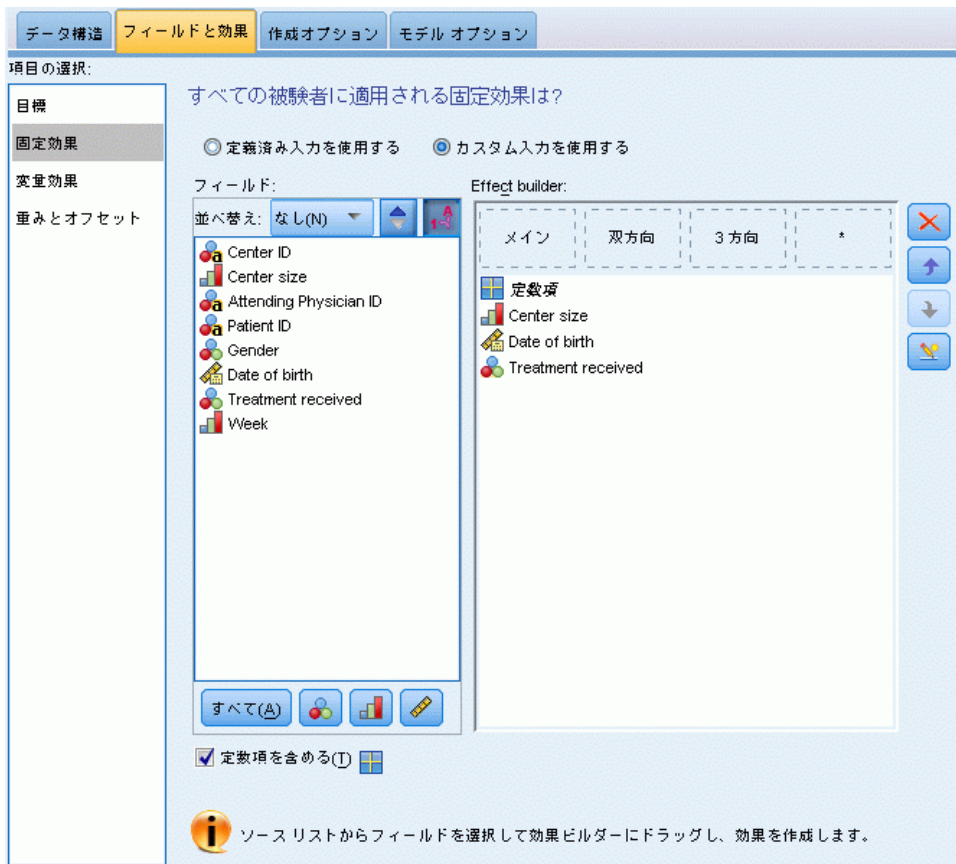
リンク関数は、モデルを推定できるようにする対象の変形です。使用できる関数は次のとおりです。

- **等式**。 $f(x)=x$ 。対象は変換されません。このリンクは、多項分布を除き、どの分布でも共に使用できます。
- **補ログ・マイナス・ログ**。 $f(x)=\log(-\log(1-x))$ 。これは、二項分布とまたは多項分布とのみ使用するのに適しています。

- **コーチット:** $f(x) = \tan(\pi(x - 0.5))$ 。これは、二項分布とまたは多項分布とのみ使用するのが適しています。
- **対数:** $f(x) = \log(x)$ 。このリンクは、多項分布を除き、どの分布でも共に使用できます。
- **ログ補完:** $f(x) = \log(1-x)$ 。これは、二項分布とのみ使用するのが適しています。
- **ロジット:** $f(x) = \log(x / (1-x))$ 。これは、二項分布とまたは多項分布とのみ使用するのが適しています。
- **負ログ・マイナス・ログ:** $f(x) = -\log(-\log(x))$ 。これは、二項分布とまたは多項分布とのみ使用するのが適しています。
- **プロビット:** $f(x) = \Phi^{-1}(x)$ 。 Φ^{-1} は累積標準正規分布関数の逆関数です。これは、二項分布とまたは多項分布とのみ使用するのが適しています。
- **べき乗:** $f(x) = x^\alpha$ ($\alpha \neq 0$ の場合)。 $f(x) = \log(x)$ ($\alpha = 0$ の場合)。 α には常に、いずれかの実数を指定する必要があります。このリンクは、多項分布を除き、どの分布でも共に使用できます。

固定効果

図 8-4
固定効果設定



固定効果の因子は、一般的に、その関心の値がすべてのデータセットで表現されるフィールドとして考えられ、スコアリングに使用することができます。デフォルトでは、ダイアログ内の他の場所で指定されていない定義済みの入力の役割を持つフィールドは、モデルの固定効果部分に入力されます。カテゴリ型（名義型、順序型）フィールドは、モデルでの因子として使用され、連続型フィールドは共変量として使用されます。

ソースリスト内の1つ以上のフィールドを選択し、効果リストにドラッグして、効果をモデルに入力します。作成する効果の種類は、選択項目をドロップするホットスポットによって異なります。

- **主相互作用:** ドロップされたフィールドは、効果リストの一番下にある別の主効果として表示されます。
- **2方向:** ドロップされたフィールドの可能なすべてのペアは、効果リストの下部に2方向の相互作用として表示されます。

- **3 方向:** ドロップされたフィールドの可能なすべての三重項は、効果リストの下部に 3 方向の相互作用として表示されます。
- *****: ドロップされたすべてのフィールドの組み合わせは、効果リストの一番下にある単一の相互作用として表示されます。

効果ビルダーの右側にあるボタンを使用すると、以下のことが可能です。



削除したい条件を選択し、削除ボタンをクリックして、固定効果モデルから用語を削除します。



順序を変更する条件を選択し、上向きまたは下向きの矢印をクリックして、固定効果モデル内の項目を並べ替えます。

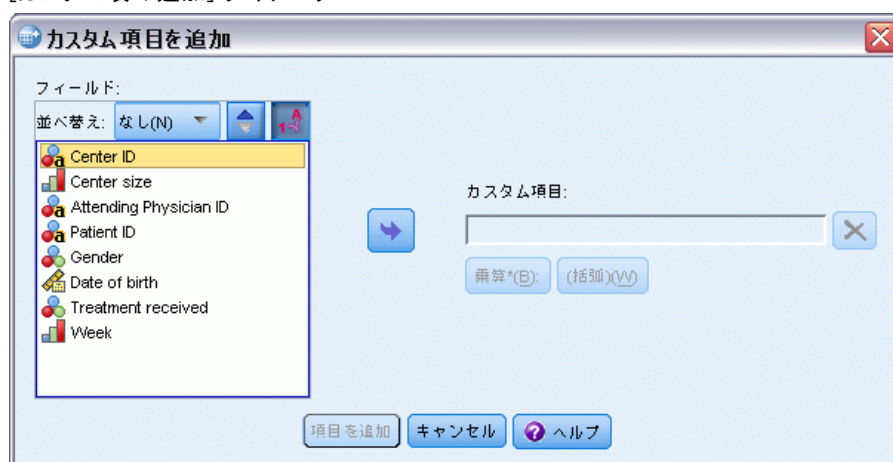


[カスタム項の追加] ボタンをクリックし、[カスタム項の追加](#) ダイアログを使用して、入れ子になった項をモデルに追加します。

定数項を含める: 通常、モデルには定数項が含まれます。データが原点を通ると仮定できる場合は、切片を除外できます。

カスタム項の追加

図 8-5
[カスタム項の追加] ダイアログ



この手順で使用するモデルの入れ子になった項を構築することができます。入れ子になった項は、値が別の因子の水準と相互作用しない因子または共変量の影響をモデル化するのに便利です。たとえば、食料品店チェーンは、複数の店舗の場所で、顧客の支出の習慣に従う場合があります。各

顧客はそれぞれ、これらの場所の 1 つにだけ頻繁に訪れるため、[顧客] 効果は [店舗の場所] 効果内で**入れ子にすることはできません**。

また、そのような同一の共変量を含む多項式の項のような相互作用の効果を含める、または入れ子になった項に複数レベルの入れ子を追加することができます。

制限: 入れ子になった項には以下の制限事項があります。

- 相互作用内のすべての因子は一意でなければなりませんA が因子である場合、 $A*A$ を指定しても無効になります。
- 入れ子になった効果内のすべての因子は一意でなければなりませんA が因子である場合、 $A(A)$ を指定しても無効になります。
- 共変量内に効果を入れ子にすることはできません。A が因子であり、X が共変量である場合、 $A(X)$ を指定しても無効になります。

入れ子になった項の構築

- ▶ もう一つの因子内に入れ子になっている要因および共変量を選択し、矢印ボタンをクリックします。
- ▶ [[内]] をクリックします。
- ▶ 前の因子または共変量が入れ子になっている因子を選択し、矢印のボタンをクリックします。
- ▶ [項の追加] をクリックします。

また、相互作用の効果を含める、または入れ子になった項に複数レベルの入れ子を追加することができます。

変数効果

図 8-6
変数効果設定

変数効果とは?
各被験者の組み合わせまたはグループでブロック内の項目の推定が異なる場合があります。

変数効果ブロック:

| 被験者 | 項目 | 定数項 |
|------------------------|----|-----|
| center_size | なし | はい |
| center_size*attphys_id | なし | はい |

ブロックを追加... ブロックを編集...

ランダム効果の因子は、値がデータファイル内の値のより大きな母集団から無作為標本を検討することができるフィールドです。これらは、対象の過剰な変動を説明するのに便利です。デフォルトでは、[データ構造] タブで複数の被験者を選択した場合、変数効果ブロックが、最も内側の被験者を超えて、被験者ごとに作成されます。たとえば、[データ構造] タブで学校、クラス、生徒を選択した場合、以下の変数効果が自動的に作成されます。

- 変数効果 1: 被験者は学校です (効果なし、定数項のみ)
- 変数効果 2: 被験者は学校 * クラスです (効果なし、定数項のみ)

以下の方法で変数効果ブロックの作業が可能です。

- ▶ 新しいブロックを追加するには、[ブロックの追加...] をクリックします。[変数効果ブロック] ダイアログが開きます。

- ▶ 既存のブロックを編集するには、編集するブロックを選択して、[ブロックの編集] をクリックします。[変数効果ブロック] ダイアログが開きます。
- ▶ ブロックを削除するには、削除ボタンを削除し、クリックしたいブロックを選択します。

変数効果ブロック

図 8-7
[変数効果ブロック] ダイアログ



ソース リスト内の1つ以上のフィールドを選択し、効果リストにドラッグして、効果をモデルに入力します。作成する効果の種類は、選択項目をドロップするホットスポットによって異なります。カテゴリ型（名義型、順序型）フィールドは、モデルでの因子として使用され、連続型フィールドは共変量として使用されます。

- **主相互作用:** ドロップされたフィールドは、効果リストの一番下にある別の主効果として表示されます。
- **2 方向:** ドロップされたフィールドの可能なすべてのペアは、効果リストの下部に 2 方向の相互作用として表示されます。

- **3 方向:** ドロップされたフィールドの可能なすべての三重項は、効果リストの下部に 3 方向の相互作用として表示されます。
- *****: ドロップされたすべてのフィールドの組み合わせは、効果リストの一番下にある単一の相互作用として表示されます。

効果ビルダーの右側にあるボタンを使用すると、以下のことが可能です。



削除したい条件を選択し、削除ボタンをクリックして、固定効果モデルから用語を削除します。



順序を変更する条件を選択し、上向きまたは下向きの矢印をクリックして、固定効果モデル内の項目を並べ替えます。



[カスタム項の追加] ボタンをクリックし、[カスタム項の追加](#) ダイアログを使用して、入れ子になった項をモデルに追加します。

定数項を含める: デフォルトでは、定数項は変量効果モデルには含まれていません。データが原点を通ると仮定できる場合は、切片を除外できます。

共分散グループを定義: ここで指定されたフィールドは、変量効果の共分散パラメータの独立したセットを定義します。カテゴリごとに 1 つずつ、グループ化フィールドのクロス分類によって定義されます。グループ化フィールドの異なるセットを各変量効果ブロックに指定することができます。すべての被験者は、同じ共分散のタイプです。同じ共分散グループ内の被験者は、パラメータに同じ値を持つこととなります。

被験者の組み合わせ: [データ構造] タブから被験者のプリセットの組み合わせから変量効果の被験者を指定することができます。たとえば、学校、クラス、生徒が [データ構造] タブで被験者として指定されている場合、被験者の組合せのドロップダウン リストには、[なし]、[学校]、[学校 * クラス] および [学校 * クラス * 生徒] がオプションとして表示されます。

変量効果共分散タイプ: 残差の共分散構造を指定します。使用できる構造は次のとおりです。

- 一次の自己回帰 (AR1)
- 自己回帰移動平均 (1, 1) (ARMA11)
- 複合対称性
- 対角性
- 尺度化された識別
- Toeplitz

- 無構造
- 分散成分

詳細は、B 付録 p.183 共分散構造 を参照してください。

重みとオフセット

図 8-8
重みとオフセットの設定

データ構造 フィールドと効果 作成オプション モデルオプション

項目の選択:

目標
固定効果
変量効果
重みとオフセット

分析の重み付け:
(なし)

オフセット

このモデルはオフセットである必要はありません

オフセット値を使用
オフセット値: 0.0

オフセットフィールドを使用
オフセットフィールド:
(なし)

分析の重み付け: 尺度パラメータは、レスポンスの分散に関連する推定モデルパラメータです。分析の重み付けは、観測ごとに異なる「既知の」値です。分析の重み付けフィールドが指定された場合、応答の分散と関連性を持つ尺度パラメータは、各観測ごとに尺度重み付け変数によって分割されます。分析の重み値が 0 以下または欠損値のレコードは、分析に使用されません。

オフセット。オフセット項は、「構造的な」予測変数です。その係数は、モデルによって推定されるのではなく、常に 1 であると仮定されます。したがって、オフセットの値は単に、目標の線型予測変数に追加されるだけです。これは、ケースごとに目的のイベントが発生するレベルが異なりうるポアソン回帰モデルにおいて有用です。

たとえば、個人運転手の事故率をモデリングする場合、過失のある事故を起こした経験が 3 年の経歴の中で 1 度ある運転手と、25 年の経歴の中で 1 度しかない運転手とでは大きな違いがあります。運転手の経験をオフセット項として加味する場合、事故の発生数は対数リンクを持つポアソン応答または負の 2 項応答としてモデル化できます。

分布およびリンクの種類のその他の組み合わせには、オフセット変数のその他の変換が必要です。

作成オプション

図 8-9
作成オプションの設定

データ構造 フィールドと効果 **作成オプション** モデルオプション

ソート順

カテゴリ型目標のソート順: 昇順

カテゴリ型予測のソート順: 昇順

停止規則

最大反復回数: 100

事後推定設定

信頼係数(%): 95.0

自由度

- すべての検定に固定 (残差方法)
サンプル サイズが大きいくほど役立ちます。またはデータが均衡している場合、より単純な共分散タイプを使用する場合に役立ちます (計測された単位または対角など)
- 検定全体で異なる (Satterthwaite 近似値)
サンプル サイズが小さいほど役立ちます。あるいはデータが不均衡な場合または複雑な共分散タイプの場合 (無構造など) に役立ちます

固定効果および係数の検定

- モデルの推定が正しいと推定する
(モデルベースの共分散)
- 頑健な推定を使用してモデル推定の違反を処理する
(頑健共分散)

これらの選択は、モデルの作成に使用されるより高度な条件を指定します。

ソート順: これらのコントロールは、「最後」のカテゴリを決定するための対象と因子（カテゴリ型入力）のカテゴリの順序を決定します。対象がカテゴリ型でない場合、またはカスタム参照カテゴリが **[対象]** 設定で指定されている場合、対象のソート順設定は無視されます。

停止規則: アルゴリズムが実行できる反復の最大回数を指定できます。負でない整数を指定してください。デフォルトは 100 です。

推定後設定: これらの設定は、モデルの出力の一部を表示するために計算する方法を指定します。

- **確信度レベル:** モデル係数の間隔の推定値を計算するために使用する確信度のレベルです。0 より大きく、100 より小さいの値を指定します。デフォルトは 95 です。
- **自由度:** 自由度が有意性検定に計算される方法を指定します。サンプルサイズが十分大きい場合、またはデータが均衡である場合、またはモデルが尺度化識別または対角性など単純な共変量タイプを使用する場合、**[すべての検定に固定(残差法)]** を指定します。これはデフォルトです。サンプルサイズが小さい場合、またはデータが不均衡である場合、またはモデルが非構造化など複雑な共変量タイプを使用する場合、**[検定間で変化(Satterthwaiteの近似)]** を指定します。
- **固定効果のおよび係数の検定:** これは、パラメータ推定値の共分散行列を計算する方法です。モデルの想定に反していると考えられる場合、堅牢な推定量を選択してください。

推定平均

図 8-10
推定平均の設定

データ構造
フィールドと効果
作成オプション
モデル オプション

項目の選択:

推定平均

フィールド保存

目標を推定しますか?

カスタム推定平均および対比を指定します。

項目:

| カテゴリベースの項目: | 推定平均 | 対比タイプ | 対比フィールド |
|-------------|-------------------------------------|-------|-------------|
| center_size | <input checked="" type="checkbox"/> | なし | center_size |
| Gender | <input type="checkbox"/> | なし | Gender |
| treatment | <input type="checkbox"/> | なし | treatment |
| Week | <input type="checkbox"/> | なし | Week |

目標推定時に連続型フィールドは定数項となります。

フィールド:

| 連続型フィールド | 定数項 | 値 |
|----------|-----|---|
| dob | 平均値 | |

次についての推定平均を表示する:

元の目標スケール

リンク関数の変換

次を使用して複数比較の調整:

最小有意差

このタブでは、因子と因子の相互作用のレベルの推定周辺平均を表示することができます。推定周辺平均は、多項モデルには使用できません。

項: 完全にカテゴリ型フィールドで構成される固定効果でのモデルの項は、ここに記載されています。モデルに推定周辺平均を生成させる各項を確認します。

- 対比の種類:** 対比のフィールドのレベルに使用する対比のタイプを指定します。[なし] を選択すると、対比は生成されません。[ペアごと] を選択すると、指定の因子の全レベルの組み合わせについてペアごとに比較します。これは、交互作用に対して行える唯一の対比です。[偏差] を指定すると、因子の各レベルを全体の平均と比較します。[単純] を指定すると、因子の各レベル（最後のレベルを除く）を最後のレベルと比較します。「最後の」レベルは、[作成] オプションで指定された因

子のソート順によって決まります。これらの対比の種類の手べてが直交であるわけではありません。

- **対比フィールド:** 因子、つまり選択した対比の種類を使用して比較するレベルを指定します。[なし] が対比の種類に選択されている場合、対比フィールドは選択できません (または必要ありません)。

連続型フィールド: 表示されている連続型フィールドは、連続型フィールドを使用する固定効果の項から抽出されます。推定周辺平均を計算する場合、共変量は指定された値に固定されます。平均値を選択するか、カスタム値を指定します。

次に関する推定平均を表示: 対象の元の尺度に基づいて、またはリンク関数変換に基づいて推定周辺平均を計算するかどうかを指定します。[元の対象の尺度] を指定すると、対象の推定周辺平均を計算します。対象がイベント/試行のオプションを使用して指定された場合、イベントの数ではなくイベント/試行の割合の推定周辺平均を計算します。[リンク関数の閉館] を指定すると、線形予測値の推定周辺平均を計算します。

次を使用して複数の比較を調整: 複数の対比の仮説検定を行うときは、全体の有意水準は含まれている対比の有意水準から調整できます。これにより、調整方法を選択できます。

- **最小有意差:** この方法では、いくつかの線形対比が帰無仮説の値とは異なるという仮説を拒否する全体的な確率を制御しません。
- **逐次 Bonferroni.** 個々の仮説を棄却しますが、すべて同じ有意水準を維持することでは、それほど控えめでもない逐次ステップダウン棄却 Bonferroni 手続きです。
- **逐次 Sidak.** これは、すべて同じ有意水準を維持したままで個々の仮説が棄却される保守的でない逐次ステップダウン相対 Sidak プロシジャです。

最小有意差法は、sequential Bonferroni 法より保守的でない sequential Sidak 法より保守的ではありません。つまり、最小有意差法は、少なくとも sequential Bonferroni と同じ仮設数を拒否する sequential Sidak と同じ仮設数を拒否します。

保存

図 8-11
保存設定

チェックされた項は、指定された名前でも保存されます。既存のフィールド名と競合することはできません。

予測値:対象の予測値を保存します。デフォルトのフィールド名は、PredictedValue です。

カテゴリ型対象の予測確率:対象がカテゴリ型の場合、このキーワードは、最初の n 件のカテゴリの予測確率を、[保存する最大カテゴリ数] で指定した値を上限として保存します。計算される値は、順序型対象の累積確率です。デフォルトのルート名は、PredictedProbability です。予測カテゴリの予測確率を保存するには、確信度を保存します（下記参照）。

信頼区間: 予測値のまたは予測確率の信頼区間の上限と下限を保存します。多項分布以外のすべての分布について、2 つの変数を作成し、デフォルトのルート名は、CI、設備士として `_Lower` および `_Upper` が付きます。

多項分布と名義型対象の場合、従属変数のカテゴリごとにフィールドが 1 つずつ作成されます。最初の n 件のカテゴリの予測確率の上限と下限を、**[保存する最大カテゴリ数]** で指定した値を上限として保存します。デフォルトのルート名は CI、デフォルトのフィールド名は CI_Lower_1、CI_Upper_1、CI_Lower_2、CI_Upper_2、となり、対象カテゴリの順序に対応しています。

多項分布と順序型対象の場合、最後を除く従属変数のカテゴリごとにフィールドが 1 つずつ作成されます（[詳細は、p. 121 作成オプションを参照してください。](#)）。最初の n 件のカテゴリの予測確率の上限と下限を、**[保存する最大カテゴリ数]** で指定した値を上限として保存します。デフォルトのルート名は CI、デフォルトのフィールド名は CI_Lower_1、CI_Upper_1、CI_Lower_2、CI_Upper_2、となり、対象カテゴリの順序に対応しています。

Pearson の残差: モデル適合の推定後診断に使用できる、各レコードの Pearson 残差を保存します。デフォルトのフィールド名は、PearsonResidual です。

確信度: カテゴリ型対象の予測値の確信度を保存します。計算される確信度は、予測値の確率（最も高い予測確率）または最も高い予測確率と 2 番目に高い予測確率との差を基準とする場合があります。デフォルトのフィールド名は、Confidence です。

モデルのエクスポート: モデルを外部の .zip ファイルに書き込みます。このモデル ファイルを使用して、モデル情報を他のデータ ファイルに適用してスコアリングできます。一意で有効なファイル名を指定します。ファイルの指定が既存のファイルを示す場合、ファイルが上書きされます。

モデルビュー

ビューアでモデル オブジェクトを作成します。このオブジェクトを有効化（ダブルクリック）すると、モデルの双方向ビューを取得します。

デフォルトでは、**[モデル要約]** ビューが表示されます。別のモデル ビューを表示するには、ビューのサムネイルから選択します。

モデル オブジェクトの代わりとして、**[オプション]** ダイアログ（**[編集]** → **[オプション]**）の **[出力]** タブの **[出力表示]** グループで **[ピボットテーブルとグラフ]** を選択し、ピボット テーブルとグラフを生成できます。次のトピックでモデル オブジェクトについて説明します。

モデル要約

図 8-12
[モデルの要約] ビュー

モデル要約

目標: Number of convulsions

| | |
|-------|-----------------------|
| 目標 | Number of convulsions |
| 確率分布 | ポアソン |
| リンク関数 | ログ |
| 補正赤池 | 6,983.683 |
| 情報基準 | |
| ベイズ | 7,008.180 |

情報基準は -2 対数擬尤度 (6,975.671) に基づき、モデルの比較に使用されます。情報基準値の小さいモデルの適合度はより高くなります。擬尤度値を使用してモデルを比較する場合、異なるデータ変換がモデル間で使用されるため注意が必要です。

このビューはスナップショットで、モデルとその適合度についての要約が一目でわかります。

テーブル: 表は、**対象設定**で指定された対象、確率分布、およびリンク関数を示します。対象が、イベントや試行によって定義されている場合、セルはイベントのフィールドと試行のフィールド試行の固定数を表示できるよう分割されます。また、有限のサンプルの赤池情報量補正基準 (AICC) およびベイズの情報基準 (BIC) が表示されます。

- **赤池補正.** -2 (制限) 対数尤度に基づいて混合モデルを選択し、比較する測度です。値が小さいほどモデルが良好であることを示します。AICC は小さなサンプルサイズに対して AIC を「修正」します。サンプルサイズが大きくなるに従い、AICC は AIC に収束します。
- **ベイズアン.** -2 対数尤度に基づいてモデルを選択し、比較する測度です。値が小さいほどモデルが良好であることを示します。BIC もパラメータが多すぎるモデルにペナルティを課しますが、AIC よりも厳密に課します。

グラフ: 対象がカテゴリ型の場合、グラフは、正しい分類の割合である最終的なモデルの精度を表示します。

データ構造

図 8-13
[データ構造] ビュー

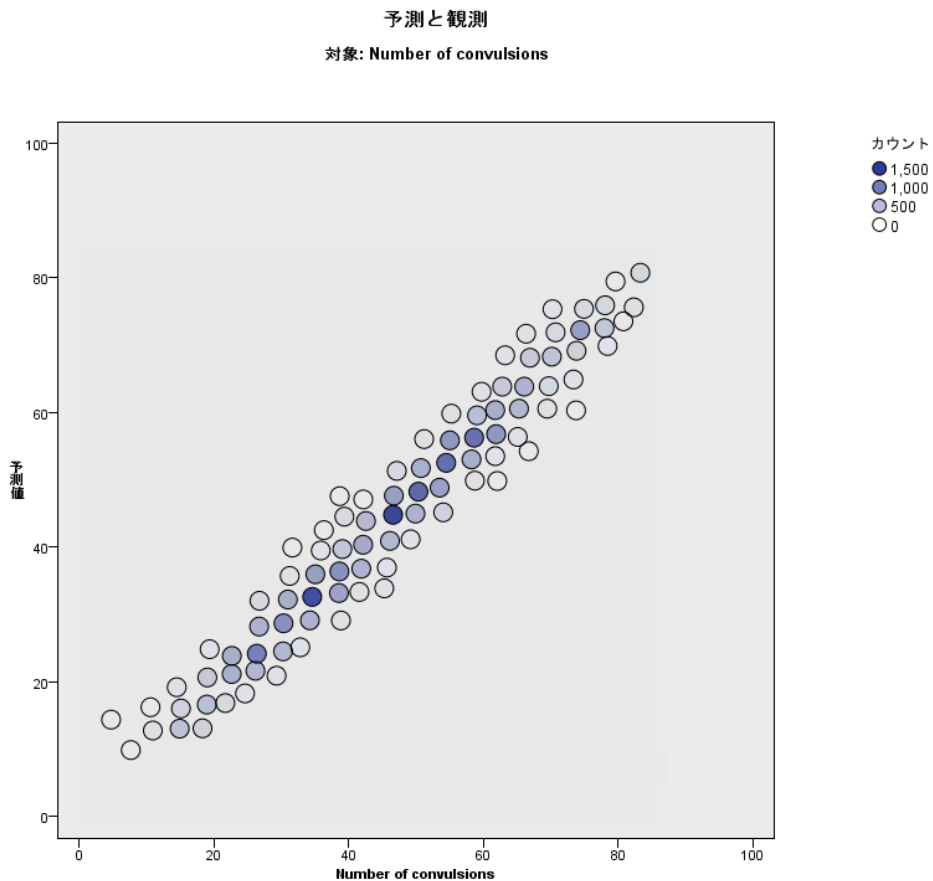
データ構造
目標: Number of convulsions

| | 被験者 | | 反復測定 | | 目標 |
|------------|-----------|------------------------|------------|------|-----------------------|
| | Center ID | Attending Physician ID | Patient ID | Week | Number of convulsions |
| 最初の被験者のデータ | 07057 | COMX | 1FSL | 0 | 2 |
| | 07057 | COMX | 1FSL | 1 | 6 |
| | 07057 | COMX | 1FSL | 2 | 4 |
| | 07057 | COMX | 1FSL | 3 | 4 |
| | 07057 | COMX | 1FSL | 4 | 6 |
| | 07057 | COMX | 1FSL | 5 | 3 |
| 水準数の合計 | 8 | 49 | 565 | 6 | |

このビューでは、指定されたデータ構造の概要を提供し、被験者と反復測定が正しく指定されていることを確認するのに役立ちます。最初の被験者の観測情報は、各被験者フィールドと反復測定のフィールド、および対象ごとに表示されます。さらに、それぞれの被験者フィールドと反復測定フィールドのレベルの数が表示されます。

予測対観測

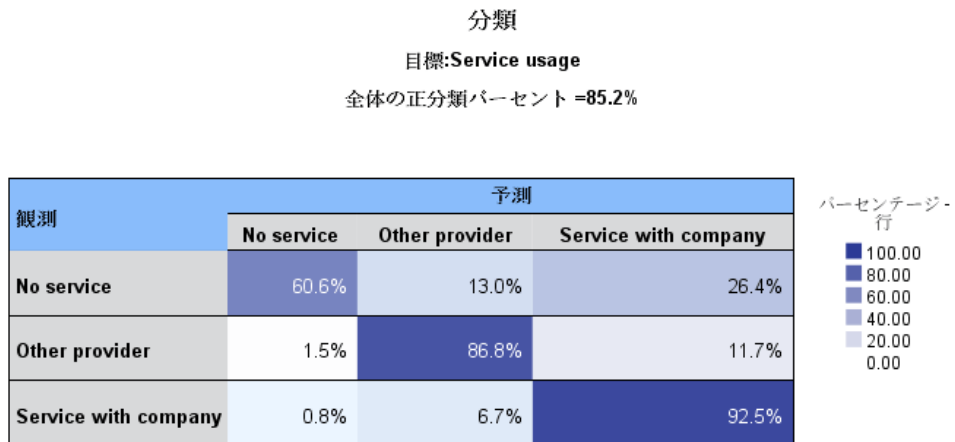
図 8-14
[予測対観測] ビュー



イベント/試行として指定された対象を含む連続型対象について、縦軸に予測値を、横軸に観測値を示す分割散布図を表示します。点は45度の線にあるのが理想です。このビューはレコードがモデルによって特に不正に予測されているかどうかを示します。

分類

図 8-15
分類ビュー



カテゴリ型対象の場合、観測値と予測値のクロス分類と、すべての正分類パーセントをヒートマップに表示します。

テーブルのスタイル:さまざまな表示スタイルがあり、[スタイル] ドロップダウン リストから指定できます。

- **行パーセント:**行パーセント（行合計のパーセントとして表示されるセル度数）がセルに表示されます。これはデフォルトです。
- **セル度数:**セル度数がセルに表示されます。ヒートマップの色の濃さは、行パーセントに基づいています。
- **ヒートマップ:**セルに値は表示されず、色の濃さだけで示します。
- **圧縮:**セルには行、列の見出し、値は表示されません。この方法は、対象にカテゴリ数が多い場合に役立ちます。

欠損値:レコードの対象フィールドに欠損値がある場合、すべての有効な行に [[欠損値]] と表示されます。欠損値のあるレコードは、すべての正分類パーセントには貢献しません。

複数の対象:複数のカテゴリ型対象がある場合、各対象は個別のテーブルに表示され、表示する対象を制御する [対象] ドロップダウン リストがあります。

大きいテーブル:表示される対象フィールドに 100 件を超えるカテゴリがある場合、テーブルは表示されません。

固定効果

図 8-16
[固定効果] ビュー、ダイアグラム スタイル

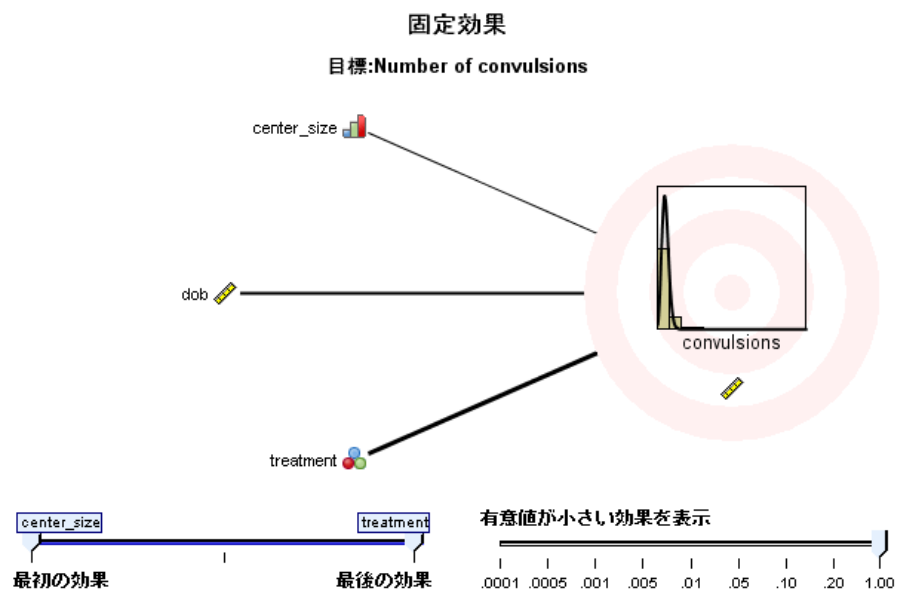
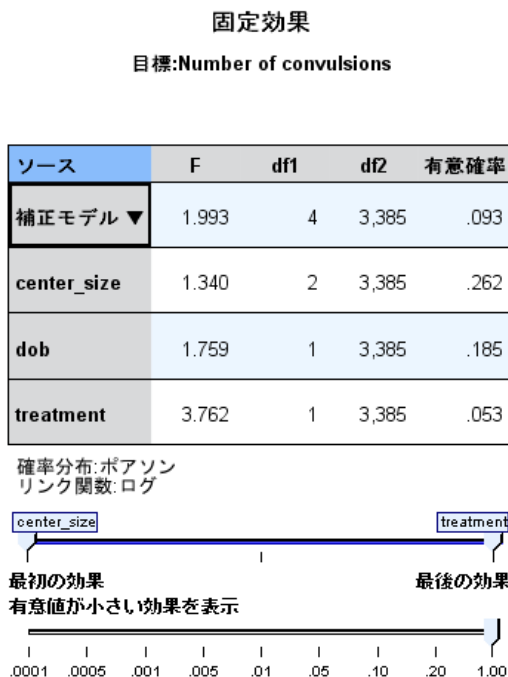


図 8-17
[固定効果] ビュー、テーブル スタイル



このビューには、モデルの各固定効果のサイズが表示されます。

スタイル: さまざまな表示スタイルがあり、[スタイル] ドロップダウン リストから指定できます。

- **ダイアグラム:** これは効果が固定効果の設定で指定された順序で上から下にソートされているチャートです。ダイアグラムで繋がった線は、効果の有意確率に基づいて重みづけられます。線の太いほど効果の有意確率は大きくなります (p 値は小さくなります)。これはデフォルトです。
- **テーブル:** モデル全体および個々のモデルの効果を示す ANOVA テーブルです。各効果は、固定効果の設定で指定された順序で上から下にソートされているチャートです。

有意確率: [重要度] スライダーで、ビュー内に表示される効果を制御します。有意確率の値がスライダーの値より大きい効果は表示されません。このスライダーを使用してもモデルは変更されませんが、最も重要な効果に焦点を当てることができます。デフォルトでは値が 1.00 になるため、有意確率に基づいてフィルタリング処理される効果はありません。

固定係数

図 8-18
[固定係数] ビュー、ダイアグラム スタイル

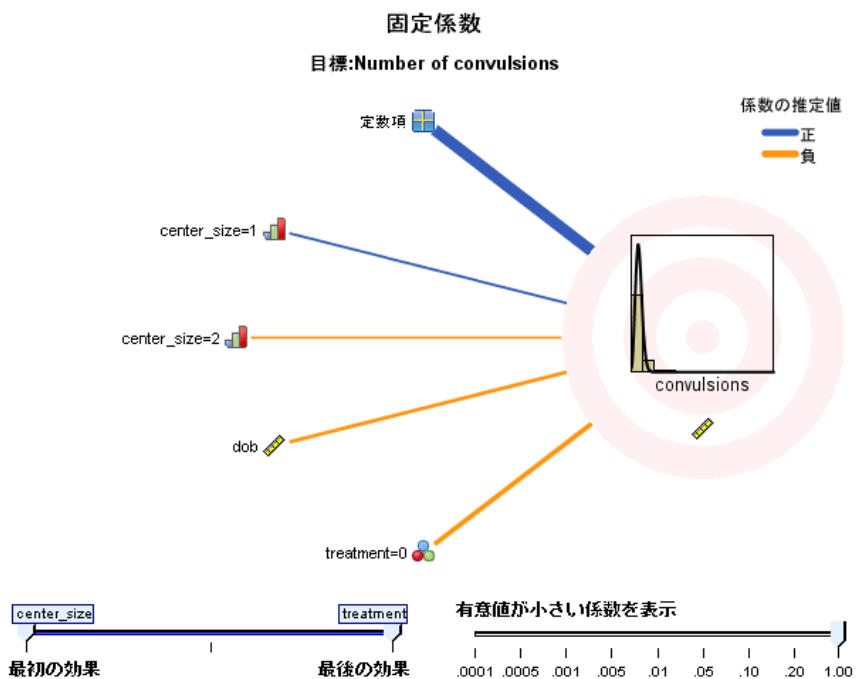


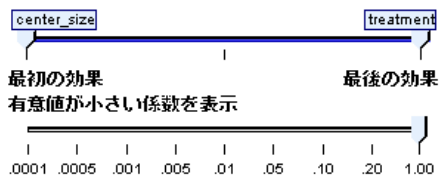
図 8-19
[固定係数] ビュー、テーブル スタイル

固定係数
目標: Number of convulsions

| モデルの項 | 係数 | 有意確率 | 指数(係数) |
|---------------|----------------|------|--------|
| 定数項 | 1.745 | .000 | 5.727 |
| center_size=1 | 0.078 | .610 | 1.081 |
| center_size=2 | -0.157 | .201 | 0.855 |
| center_size=3 | 0 ^a | | |
| dob | -0.000 | .185 | 1.000 |
| treatment=0 | -0.138 | .053 | 0.871 |
| treatment=1 | 0 ^a | | |

確率分布: ポアソン
リンク関数: ログ

^a この係数が冗長であるため、0 に設定されています。



このビューには、モデルの各固定係数の値が表示されます。因子（カテゴリ型予測フィールド）はモデル内で指標コード化されるため、因子を含む効果には通常複数の関連する係数があります。一方は冗長係数に対応するカテゴリを除いたものとなります。

スタイル: さまざまな表示スタイルがあり、[スタイル] ドロップダウン リストから指定できます。

- ダイアグラム:** これはまず定数項を表示し、固定効果の設定で指定された順序で上から下にソートされているチャートです。因子を含む効果内で、係数はデータ値が小さい順に並べ替えられます。ダイアグラムで繋がった線は色分けされ、係数の有意確率に基づいて重みがつけられま

す。線の太いほど係数の有意確率は大きくなります (p 値は小さくなります)。これはデフォルトのスタイルです。

- **テーブル:** 各モデル係数の値、有意性検定、信頼区間が表示されます。定数項の後、効果が固定効果の設定で指定された順序で上から下にソートされているチャートです。因子を含む効果内で、係数はデータ値が小さい順に並べ替えられます。

多項分布: 多項分布が有効になっている場合には、[多項] ドロップダウンリストは、表示する対象カテゴリを制御します。リスト内の値のソート順は、[作成オプション] の設定の仕様によって決定されます。

指数: 二項ロジスティック回帰 (二項分布とロジットリンク)、名義ロジスティック回帰 (多項分布とロジットリンク)、負の二項回帰 (負の二項分布と対数リンク)、および対数線型モデル (ポアソン分布と対数リンク) など、特定のモデルタイプの指数係数推定および信頼区間を表示します。

有意確率: [重要度] スライダーで、ビュー内に表示される係数を制御します。有意確率の値がスライダーの値より大きい係数は表示されません。このスライダーを使用してもモデルは変更されませんが、最も重要な係数に焦点を当てることができます。デフォルトでは値が 1.00 になるため、有意確率に基づいてフィルタリング処理される係数はありません。

変数効果共分散

変数効果共分散行列 (G) を表示します。

スタイル: さまざまな表示スタイルがあり、[スタイル] ドロップダウン リストから指定できます。

- **共分散値:** これは効果が固定効果の設定で指定された順序で上から下にソートされている共分散行列のヒート マップです。corrgram の色は、キーに示されているセルの値に対応します。これはデフォルトです。
- **Corrgram:** 共分散行列のヒート マップです。
- **圧縮:** 行および列の見出しのない共分散行列のヒート マップです。

ブロック: 複数の変数効果のブロックがある場合は、[ブロック] ドロップダウンリストで表示するブロックを選択します。

グループ: 変数効果のブロックにグループの指定がある場合は、[グループ] ドロップダウン リストで表示するグループ レベルを選択します。

多項分布: 多項分布が有効になっている場合には、[多項] ドロップダウンリストは、表示する対象カテゴリを制御します。リスト内の値のソート順は、[作成オプション] の設定の仕様によって決定されます。

共分散パラメータ

図 8-20
共分散パラメータビュー

共分散パラメータ
目標: Number of convulsions

| | | |
|----------|------|-----------------|
| 共分散パラメータ | 残差効果 | 2 |
| | 変量効果 | 2 |
| 計画行列の列 | 固定効果 | 7 |
| | 変量効果 | 10 ^a |
| 共通被験者 | | 8 |

共通被験者は、残差効果および変量効果の被験者指定に基づき、パフォーマンス効果のためにデータをひとまとめにするのに使用されます。

^aこれは共通被験者あたりの列の数です。

| 残差効果 | 推定値 | 標準誤差 | Z | 有意確率 | 95% の信頼区間 | |
|--------|-------|-------|--------|------|-----------|-------|
| | | | | | 下限 | 上限 |
| AR1 対角 | 5.379 | 0.243 | 22.151 | .000 | 4.923 | 5.876 |
| AR1 ロー | 0.805 | 0.010 | 82.823 | .000 | 0.785 | 0.824 |

共分散構造: 1 次自己回帰
被験者指定: center_id*attphys_id*patient_id

このビューには、残差とランダム効果の共分散パラメータの推定値と関連する統計情報が表示されます。これらは高度ですが、共分散構造画適しているかどうかに関する情報を提供する基本的な結果です。

集計表: 残差 (R) および変量効果 (G) 共分散行列のパラメータ数、固定効果 (X) および変量効果 (Z) デザイン行列の順位 (列の数)、データ構造を定義する被験者フィールドで定義されている被験者の数に関するクイック リファレンスです。

共分散パラメータ テーブル: 選択した効果について、推定値、標準誤差、信頼区間が共分散パラメータごとに表示されます。表示されるパラメータの数は、効果および変量効果ブロックの共分散構造、ブロックの効果の数に

よって異なります。非対角パラメータが重大ではないことが表示された場合、単純な共分散構造を使用することができる場合があります。

効果: 変量効果のブロックがある場合は、[効果] ドロップダウンリストで表示する残差または変量効果ブロックを選択します。残差効果は常に使用可能です。

グループ: 残差または変量効果のブロックにグループの指定がある場合は、[グループ] ドロップダウン リストで表示するグループ レベルを選択します。

多項分布: 多項分布が有効になっている場合には、[多項] ドロップダウンリストは、表示する対象カテゴリを制御します。リスト内の値のソート順は、[作成オプション] の設定の仕様によって決定されます。

推定平均:有意な効果

3 方向の相互作用から始まり、2 方向の相互作用、および最終的に主効果となる、10 個の「最も有意な」固定全因子効果について表示されるグラフです。グラフは横軸上の主効果（または相互作用で最初に表示されている効果）の各値について縦軸上の対象のモデル推定値を表示します。相互作用の 2 番目に記載されている効果の各値について別の線が生成されます、3 方向相互作用で 3 番目に表示される効果の値ごとにグラフが表示されます。他のすべての予測値は一定です。対象フィールドに対する各予測フィールドの係数の効果について、役立つ視覚化を提供します。予測値が重要でない場合、推定平均値は生成されません。

確信度: [作成オプション] の一部として指定された信頼レベルを使用し、周辺平均の信頼限界の上限および下限が表示されます。

推定平均:カスタム効果

ユーザが要求したすべての因子の効果を固定するためのテーブルおよびグラフです。

スタイル: さまざまな表示スタイルがあり、[スタイル] ドロップダウン リストから指定できます。

- **ダイアグラム:** このスタイルは、横軸上の主効果（または相互作用で最初に表示されている効果）の各値について縦軸上の対象のモデル推定値の線グラフを表示します。相互作用の 2 番目に記載されている効果の各値について別の線が生成されます、3 方向相互作用で 3 番目に表示される効果の値ごとにグラフが表示されます。他のすべての予測値は一定です。

対比が要求された場合、別のグラフが、対比のフィールドのレベルを比較するために表示されます。相互作用の場合、グラフが対比フィールド以外の効果の各レベルの組み合わせで表示されます。**ペアごとの対比**の場合、距離のネットワーク グラフです、つまり、ネットワーク内のノード間の距離は、サンプル間の差異に対応する比較のテーブルをグラフィカルに表現したものです。黄色の線は、統計的に有意な差に対応し、黒線が非有意差に対応しています。ネットワーク内の線の上にマウスを乗せると、線で接続されたノード間の相違の調整済み有意度とツールヒントが表示されます。

偏差の対比の場合、棒グラフに縦軸上のターゲットのモデル推定値と横軸に対比のフィールドの値が表示されます。相互作用の場合、グラフは対比フィールド以外の効果の各水準の組み合わせごとに表示されます。バーは、対比のフィールドのレベルと全体の平均値との差を示し、黒の水平線で表されます。

単純なの対比の場合、棒グラフに縦軸上のターゲットのモデル推定値と横軸に対比のフィールドの値が表示されます。相互作用の場合、グラフは対比フィールド以外の効果の各水準の組み合わせごとに表示されます。バーは、対比のフィールド（最後を除く）と最後のレベルとの間の差を示し、黒の水平線で表されます。

- **テーブル** : このスタイルは、モデルの推定対象の値、その標準誤差、および効果のフィールドの各レベルの組み合わせの信頼区間の表が表示されます。他のすべての予測変数が一定に保たれます。

対比が要求された場合、別のテーブルが推定、標準誤差、有意差検定、およびそれぞれのコントラストのための信頼区間とともに表示されます。相互作用の場合、対比フィールド以外の効果の各レベルの組み合わせの行が別途表示されます。さらに、全体的なテスト結果を含むテーブルが表示されます。相互作用の場合、対比のフィールド以外の効果の各レベルの組み合わせごとに全体的な検定があります。

確信度 : [作成オプション] の一部として指定された信頼レベルを使用して、周辺平均の信頼限界の上限および下限の表示を切り替えます。

レイアウト : ペアごとの対比ダイアグラムのレイアウトを切り替えます。サークル レイアウトは、ネットワーク レイアウトよりも対比は少ないですが、行の重複を回避できます。

モデル選択の対数線型分析

モデル選択の対数線型分析手続きでは、多重配列のクロス集計表（分割表）を分析します。反復比例当てはめアルゴリズムを用い、階層対数線型モデルを多次元のクロス集計表に当てはめます。この手続きを使用して、関連するカテゴリ変数を簡単に見つけることができます。モデルを構築する場合は、強制投入法と変数減少法が使用できます。飽和モデルの場合は、パラメータ推定値と偏関連の検定を要求できます。飽和モデルは、すべてのセルに 0.5 を加えます。

例。2 種類の洗濯用洗剤の一方に対するユーザーの好みについての調査で、研究者は各グループの人数を数え、軟水度（軟、中、硬）、2 つのブランドの使用経験の有無、洗濯温度（低温または高温）などのさまざまなカテゴリを組み合せました。これにより、洗濯温度は軟水度だけでなく、ブランドの好みとも関係があることがわかりました。

統計量。度数、残差、パラメータ推定値、標準誤差、信頼区間、および偏関連の検定。ユーザーの指定によるモデルの場合は、残差プロットと正規確率プロット。

データ。因子変数はカテゴリ型です。分析の対象となる変数はすべて数値型でなければなりません。文字型のカテゴリ変数は、モデル選択の分析を始める前に、数値型変数に値を再度割り当てることができます。

水準数の多い変数を多数指定することは避けてください。そのような指定をすると、複数のセルで、観測数が少数になり、カイ 2 乗値が役に立たなくなることがあります。

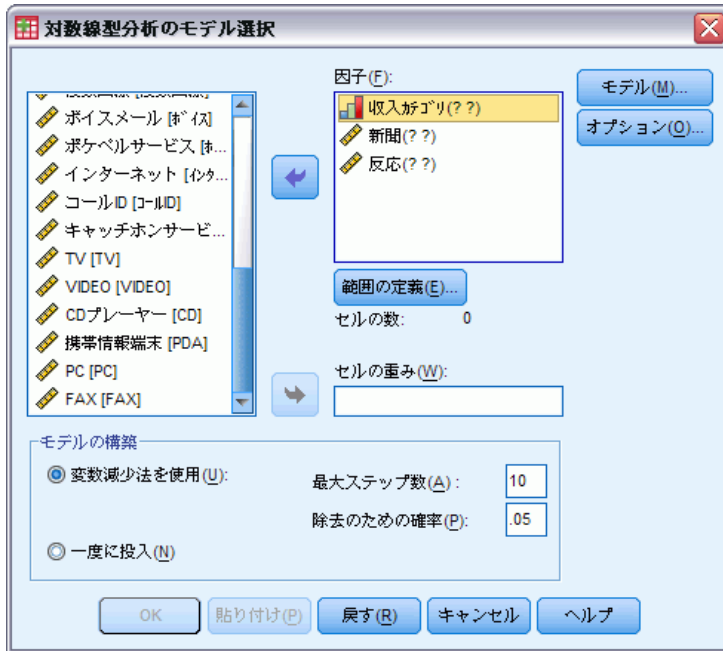
関連手続き。モデル選択手続きを用いると、モデルに必要な項を簡単に識別できます。その後、一般的な対数線型分析かロジット対数線型分析で、モデルの評価を継続できます。値の自動的な再割り当てを使用して、文字型変数の値を再度割り当てることができます。数値型変数のカテゴリが空白の場合は、値の再割り当てで連続した整数値を作成してください。

モデル選択の対数線型分析を行うには

メニューから次の項目を選択します。

分析 > 対数線型 > モデル選択...

図 9-1
[対数線型分析のモデル選択] ダイアログ ボックス

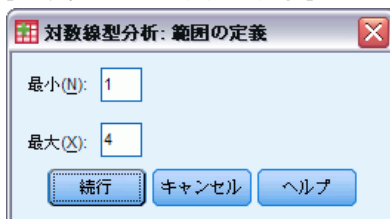


- ▶ 2 つ以上の数値型カテゴリ因子を選択します。
- ▶ [因子] リストで 1 つ以上の因子変数を選択し、[範囲の定義] をクリックします。
- ▶ 因子変数ごとに、値の範囲を定義します。
- ▶ [モデルの構築] でオプションを選択します。

オプションとして、セルの重み付け変数を選択して構造 0 を指定できます。

対数線型分析の範囲の定義

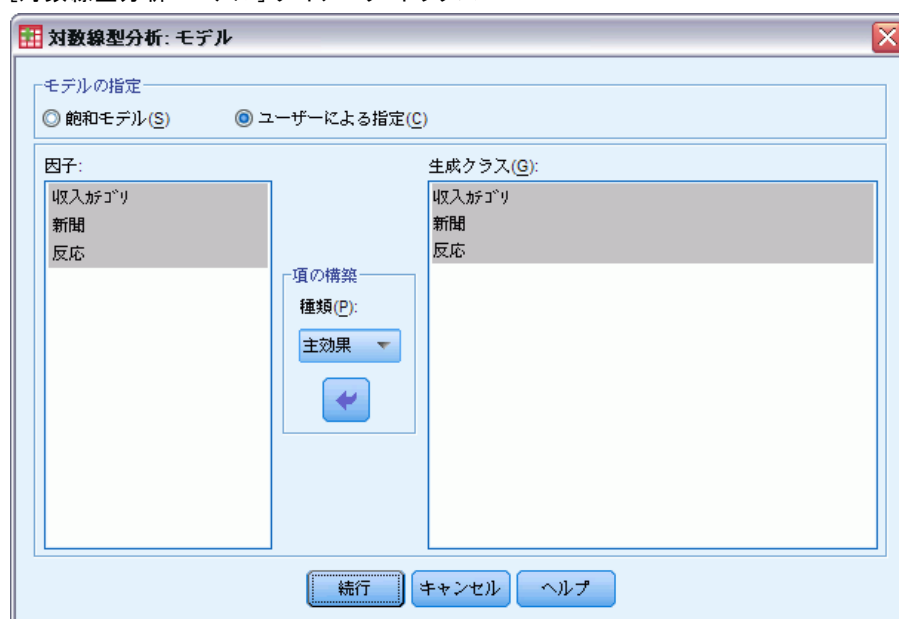
図 9-2
[対数線型分析: 範囲の定義] ダイアログ ボックス



それぞれの因子変数には、カテゴリの範囲を指定する必要があります。最小値と最大値は、因子変数の最低カテゴリと最高カテゴリに対応します。値は、両方とも整数で、最小値は最大値より小さい値でなければなりません。上下限を超える値を含むケースは除外されます。たとえば、最小値に 1、最大値に 3 を指定すると、1、2、3 以外は使用できません。各因子変数で、この手順を繰り返します。

対数線型分析のモデル

図 9-3
[対数線型分析: モデル] ダイアログ ボックス



モデルの指定。[飽和モデル] をクリックすると、因子の主効果と因子間の交互作用のすべてが含まれます。不飽和のモデルに対して生成するクラスを指定するには [ユーザー指定] をクリックします。

生成クラス。[生成クラス] ボックスの一覧には、因子が表示される最高次の項が表示されます。階層モデルには、生成クラスを定義する項とそれより低次の関係のすべてが含まれます。[因子] リストで変数 A、B、C を選択し、次に [項の構築] ドロップダウン リストから [交互作用] を選んだとします。構築されるモデルには、指定された 3 次の交互作用 $A*B*C$ 、2 次の交互作用 $A*B$ 、 $A*C$ 、および $B*C$ 、さらに A、B、C の主効果が含まれます。生成クラスには低次の関係を指定しないでください。

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。これはデフォルトです。

主効果。選択した変数のそれぞれに主効果の項を作成します。

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。

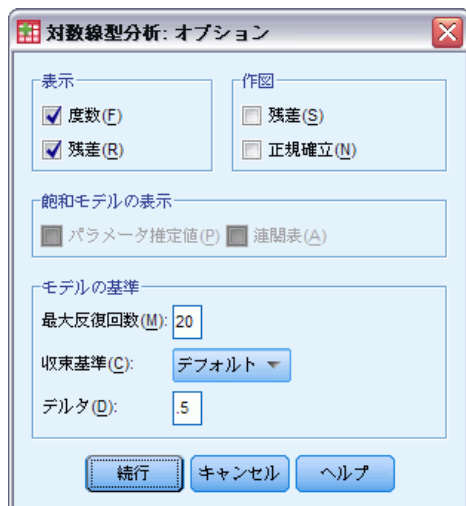
3 次まで。選択した変数に 3 次までの交互作用を作成します。

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。

モデル選択の対数線型分析のオプション

図 9-4
[対数線型分析: オプション] ダイアログ ボックス



表示。[度数] か [残差]、あるいはその両方を選択できます。飽和モデルでは、観測度数と期待度数は同じで、残差は 0 になります。

作図。ユーザーの指定によるモデルでは、[残差] または [正規確率]、あるいはその両方の作図を選択できます。モデルとデータとの適合度の確定に役立ちます。

飽和モデルの表示。飽和モデルの場合は、[パラメータ推定値] を選択できます。パラメータ推定値を使用すると、モデルから除外してもかまわない項の確定が容易になります。偏連関の検定を表示する [連関表] も選択できます。表に因子が多数含まれていると、このオプションでの計算は効率的ではありません。

モデルの基準。反復比例当てはめアルゴリズムによって、パラメータ推定値を出します。[最大反復回数]、[収束基準]、または [デルタ]（飽和モデルのセル度数のすべてに加える値）を指定すれば、1 つ以上の推定基準を変更できます。

HILOGLINEAR 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。

- 行列形式指定でのセルの重みの指定（CWEIGHT サブコマンドを使用）。
- 1 つのコマンドによる複数モデルの分析結果の生成（DESIGN サブコマンドを使用）。

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

一般的な対数線型分析

一般的な対数線型分析手続きでは、クロス集計表や分割表の各クロス分類カテゴリに分類される観測値の度数を分析します。表中の各クロス分類がセルを構成します。また、各カテゴリ変数を因子と呼びます。従属変数は、クロス集計表のセルに含まれるケースの数（度数）であり、説明変数は、因子と共変量です。この手続きでは、Newton-Raphson 法を使用して、階層対数線型モデルと非階層対数線型モデルの最尤パラメータの推定値を出します。ポアソン分布または多項分布を分析できます。

表中のセルを定義する場合、因子は 10 個まで選択できます。セル構造変数を使用して、不完全な表に構造 0 を定義する、モデルにオフセット項を含める、対数比モデルを当てはめる、あるいは周辺表の調整法を実施できます。対比変数を使用して、一般化対数オッズ比 (GLOR) を計算できます。

モデル情報および適合度統計量が自動的に表示されます。また、各種の統計やプロットを表示したり、アクティブ データセットに残差や予測値を保存したりすることもできます。

例。フロリダ州の自動車事故報告からのデータを使って、シートベルトの着用と怪我の程度（死亡か生存か）との関係を確定します。オッズ比によって、関係が有意であるという証拠が示されます。

統計量。観測度数と期待度数（元データ、調整済み、および最大対数尤度比残差）、計画行列、パラメータ推定値、オッズ比、対数オッズ比、GLOR、Wald 統計量、および信頼区間。作図（調整済み残差、最大対数尤度比残差、正規確率）。

データ。因子はカテゴリ型で、セル共変量は連続型です。モデル内に共変量があると、セルに含まれているケースの共変量の平均値をそのセルに適用します。対比変数は連続型です。対比変数は、一般化対数オッズ比を計算するときに使用します。対比変数の値は期待セル度数の対数を線型に組み合わせるための係数です。

セル構造変数は、重みを割り当てます。たとえば、セルの中に構造 0 のものがあると、セル構造変数の値は 0 か 1 のどちらかになります。集計済みデータに重みを付ける場合は、セル構造変数を使いません。代わりに、[データ] メニューの [ケースの重み付け] を使用します。

仮定。[一般的な対数線型分析] では、ポアソンと多項という 2 つの分布が利用できます。

ポアソン分布では、次のように仮定します。

- 研究に先立ってサンプル サイズは決まっています。または、分析はサンプル サイズに関して条件付きではありません。
- セルに存在する観測の事象は、統計上、他のセルのセル度数からは独立しています。

多項分布では、次のように仮定します。

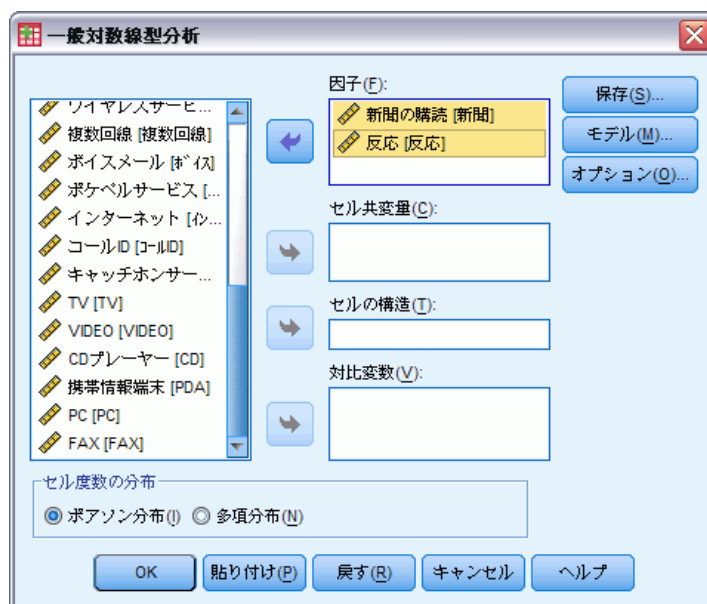
- サンプル サイズを決めます。つまり、分析はサンプル サイズによる条件付きです。
- 統計上、セル度数は独立していません。

関連手続き。クロス集計表を調べる場合は、クロス集計表手続きを使用します。1 つ以上のカテゴリ変数を回答変数と見なし、残りを説明変数と見なすのが自然な場合は、ロジット対数線型分析手続きを使用します。

一般的な対数線型分析を行うには

- ▶ メニューから次の項目を選択します。
分析 > 対数線型 > 一般的...

図 10-1
[一般的な対数線型分析] ダイアログ ボックス



- ▶ [一般的な対数線型分析] ダイアログ ボックスで、最高 10 個までの因子変数を選択します。

オプションとして、次の選択が可能です。

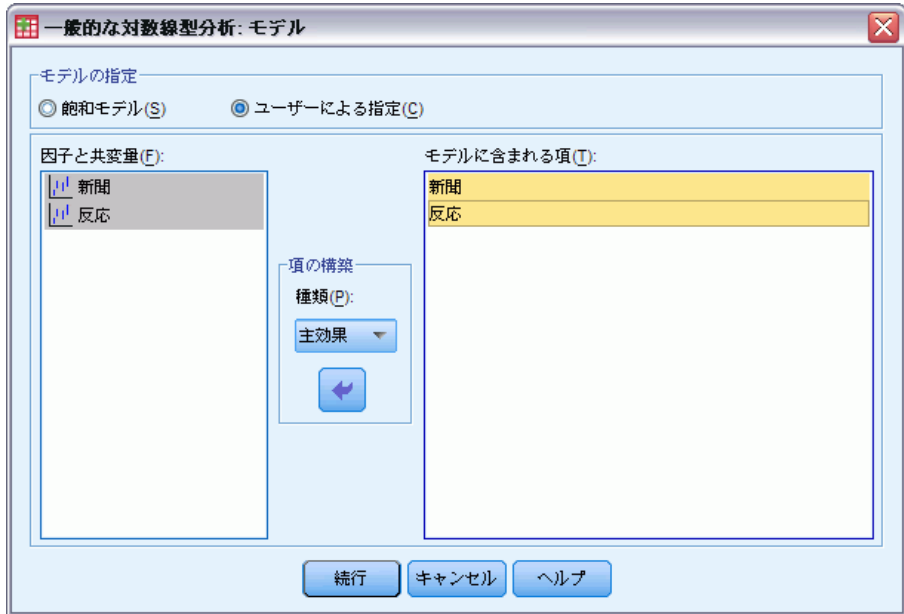
- セルの共変量を選択。

- [セルの構造] ボックスにセル構造変数の選択による構造 0 の定義、あるいはオフセット項を設定。
- [対比変数] ボックスに対比変数を選択。

一般的な対数線型分析のモデル

図 10-2

[一般的な対数線型分析: モデル] ダイアログ ボックス



モデルの指定。飽和モデルには、因子変数に関わるすべての主効果と交互作用が含まれます。共変量の項は含まれていません。交互作用の一部だけを指定する場合や、因子と共変量の交互作用を指定する場合は、[ユーザーの指定] をクリックします。

因子と共変量。要素および共変量はリストされます。

モデルに含まれる項。モデルは、使用するデータの性質によって異なります。[ユーザーの指定] をクリックすれば、分析対象の主効果と交互作用を選択できます。モデルに含める項目はすべて指示する必要があります。

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。

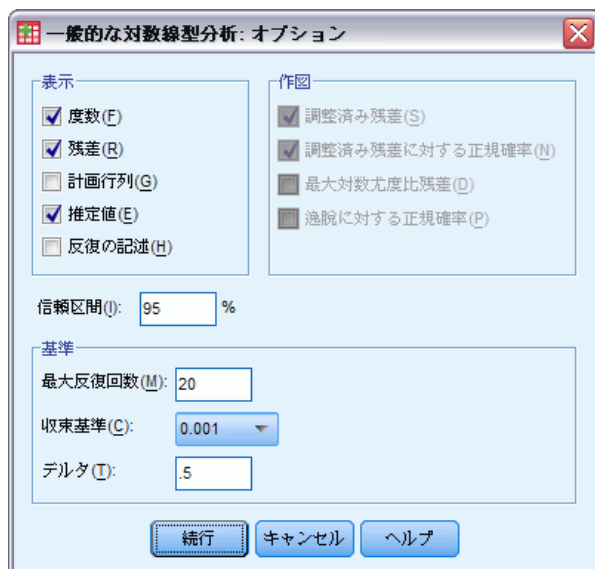
交互作用。選択した変数のすべてについて、最高水準の交互作用項を作成します。これはデフォルトです。

主効果。選択した変数のそれぞれに主効果の項を作成します。

- 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。

一般的な対数線型分析のオプション

図 10-3
[一般的な対数線型分析: オプション] ダイアログ ボックス



一般的な対数線型分析手続きでは、モデル情報と適合度統計量が表示されます。その他に、次の中から 1 つ以上を選択できます。

表示。 観測セル度数と期待セル度数（元データ、調整済み、および最大対数尤度比残差）、モデルの計画行列、そしてモデルのパラメータ推定値など、いくつかの統計量を表示できます。

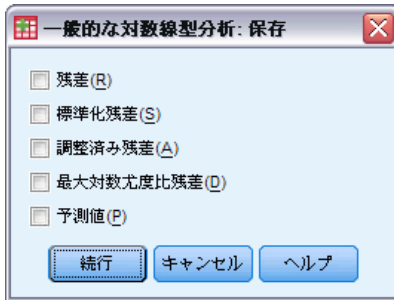
作図。 ユーザーの指定によるモデルに限り使用できる作図には、2 つの散布図行列（観測セル度数および予測セル度数に対する調整済み残差または最大対数尤度比残差）が含まれます。調整済み残差または最大対数尤度比残差の正規確率および傾向化除去正規プロットを表示することもできます。

信頼区間。 パラメータ推定値の信頼区間を調整できます。

基準。 Newton-Raphson 法を使用して、最尤パラメータ推定値を出します。[最大反復回数]、[収束基準]、[デルタ]（最初の近似のため、すべてのセルに加える定数）の各ボックスに新たな値を入力できます。飽和モデルのセルのデルタはそのままです。

一般的な対数線型分析の保存

図 10-4
[一般的な対数線型分析: 保存] ダイアログ ボックス



新変数としてアクティブ データセットに保存する値を選択します。新しい変数名の接尾辞 n を 1 ずつ増やしていくことで、保存変数に一意の名前を作成することができます。

集計データ（分割表のセル）がデータ エディタで個別の観測値ごとに記録されている場合でも、保存値はそのデータを参照します。未集計データの残差や予測値を保存すると、分割表のセルの保存値は、データ エディタ中のそのセルのケースごとに入力されます。保存値を理解するためには、データを集計してセル度数を出す必要があります。

保存できる残差は、[残差]、[標準化残差]、[調整済み残差]、[最大対数尤度比残差] の 4 種類です。[予測値] も保存できます。

- **残差.** 単純残差または生の残差とも呼ばれ、観測されるセルの度数とそれの期待度数の差を表します。
- **標準化残差.** 残差を標準誤差の推定値で割った値。標準化残差は、Pearson 残差としても知られています。
- **調整済み残差.** 調整済み残差は、標準化残差を推定された標準誤差で割り算したものです。モデルが正しいとき、調整済み残差は漸近的に標準正規分布になるので、正規性のチェックに使われます。
- **最大対数尤度比残差.** 尤度比カイ 2 乗統計量に対する個々の寄与率の符号付き（この符号は残差（観測度数から期待度数を引いた値）の符号です）平方根（G の 2 乗）。最大対数尤度比残差は、漸近標準正規分布になります。

GENLOG 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。

- 観測セル度数と期待セル度数の線型組み合わせの計算、およびその組み合わせの残差、標準化残差、調整済み残差の印刷（GERESID サブコマンドを使用）。

- デフォルトの冗長検査用しきい値の変更 (CRITERIA サブコマンドを使用)。
- 標準化残差の表示 (PRINT サブコマンドを使用)。

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

ロジット対数線型分析

ロジット対数線型分析手続きでは、従属（応答）変数と独立（説明）変数との間の関係を分析します。従属変数は常にカテゴリ変数ですが、独立変数はカテゴリ（因子）変数にすることもできます。その他の独立変数、セルの共変量は連続型にすることができますが、その場の状況に応じて適用することはできません。セルに対する共変量の重み付き平均値をそのセルに適用します。従属変数のオッズの対数は、パラメータを線型に組み合わせて表現します。多項分布が自動的に仮定されるので、多項ロジットモデルと呼ばれることもあります。この手続きでは、Newton-Raphson 法を採用してロジット対数線型モデルのパラメータを推定します。

従属変数と因子変数は 1 から 10 までの組み合わせを選択できます。セル構造変数を使用して、不完全な表に構造 0 を定義する、モデルにオフセット項を含める、対数比モデルを当てはめる、あるいは周辺表の調整法を実施できます。対比変数を使用して、一般化対数オッズ比 (GLOR) を計算できます。対比変数の値は期待セル度数の対数を線型に組み合わせるための係数です。

モデル情報および適合度統計量が自動的に表示されます。また、各種の統計やプロットを表示したり、アクティブ データセットに残差や予測値を保存したりすることもできます。

例。フロリダ州で行われたある研究に、219 匹のワニのデータがあります。その大きさや生息地である 4 か所の湖によって、ワニの食料の種類はどうか変化するのでしょうか。この研究からは、小型のワニは魚類より爬虫類を好むというオッズが大型のワニの場合の 0.7 倍であることと、魚類ではなくまず爬虫類を選択するというオッズが湖 3 で最高であることがわかります。

統計量。観測度数と期待度数；元データ、調整済み、および最大対数尤度比残差；計画行列；パラメータ推定値；一般化対数オッズ比；Wald 統計量；および信頼区間。作図：調整済み残差、最大対数尤度比残差、正規確率の各プロット。

データ。従属変数はカテゴリ型です。因子も同じくカテゴリ型です。セルの共変量は連続型にすることもできますが、共変量がモデルに含まれていると、セル中のケースの共変量の平均値をそのセルに適用します。対比変数は連続型です。対比変数は、一般化対数オッズ比 (GLOR) を計算するときに使用します。対比変数の値は期待セル度数の対数を線型に組み合わせるための係数です。

セル構造変数は、重みを割り当てます。たとえば、セルの中に構造 0 のものがあると、セル構造変数の値は 0 か 1 のどちらかになります。セル構造変数を使って集計データに重みをつけることはできません。その場合は、[データ] メニューの [ケースの重み付け] を使用します。

仮定。説明変数のカテゴリの各組み合わせ内の度数は、多項分布すると仮定します。多項分布では、次のように仮定します。

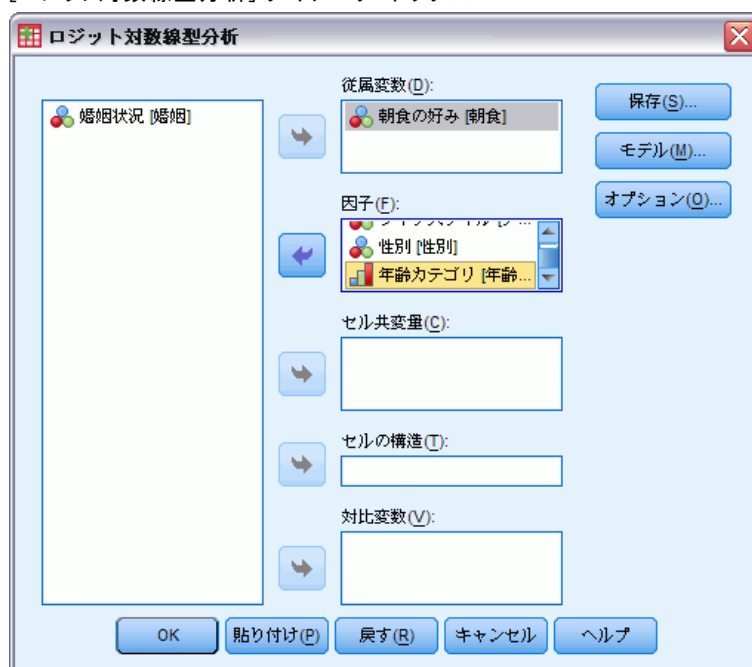
- サンプル サイズを決めます。つまり、分析はサンプル サイズによる条件付きです。
- 統計上、セル度数は独立していません。

関連手続き。分割表を表示するには、クロス集計表手続きを使用します。観測度数と一連の説明変数との間の関係を分析したい場合には、一般的な対数線型分析手続きを使用します。

ロジット対数線型分析を行うには

- ▶ メニューから次の項目を選択します。
分析 > 対数線型 > ロジット...

図 11-1
[ロジット対数線型分析] ダイアログ ボックス

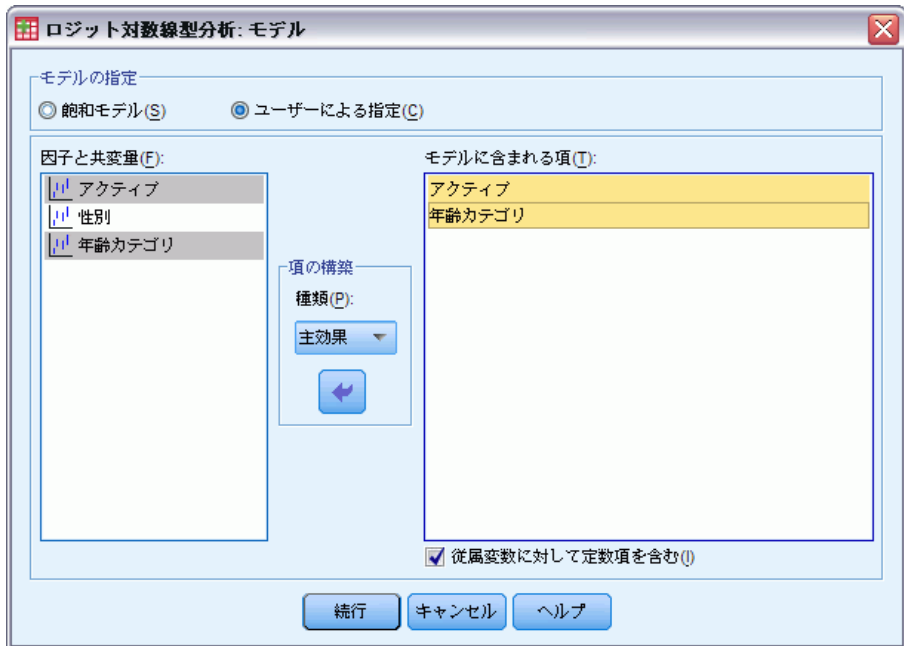


- ▶ [ロジット対数線型分析] ダイアログ ボックスで、1 つ以上の従属変数を選択します。
- ▶ [因子] ボックスに 1 つ以上の因子変数を選択します。
従属変数と因子変数の合計は 10 個以下でなければなりません。
オプションとして、次の選択が可能です。
 - セルの共変量を選択。

- [セルの構造] ボックスにセル構造変数の選択による構造 0 の定義、あるいはオフセット項を設定。
- 1 つ以上の対比変数を選択。

ロジット対数線型分析のモデル

図 11-2
[ロジット対数線型分析: モデル] ダイアログ ボックス



モデルの指定。飽和モデルには、因子変数に関わるすべての主効果と交互作用が含まれます。共変量の項は含まれていません。交互作用の一部だけを指定する場合や、因子と共変量の交互作用を指定する場合は、[ユーザーの指定] をクリックします。

因子と共変量。要素および共変量はリストされます。

モデルに含まれる項。モデルは、使用するデータの性質によって異なります。[ユーザーの指定] をクリックすれば、分析対象の主効果と交互作用を選択できます。モデルに含める項目はすべて指示する必要があります。

項は、従属項に可能なすべての組み合わせを使用し、モデルの一覧の各項にそれぞれの組み合わせを一致させて計画に追加します。[従属変数に対して定数項を含む] チェック ボックスをオンにすると、モデルの一覧には単位項 (1) も追加されます。

たとえば、変数 D1 と D2 が従属変数であると仮定します。ロジット対数線型分析手続きでは、従属項の一覧が作成されます (D1、D2、D1*D2)。[モデルに含まれる項] リストに M1 および M2 が含まれ、さらに定数が含

まれている場合、そのモデルの一覧には、1、M1、および M2 が含まれます。すると、計画には、次のように、各従属項と各モデル項の組み合わせが含まれています。

D1, D2, D1*D2

M1*D1, M1*D2, M1*D1*D2

M2*D1, M2*D2, M2*D1*D2

従属変数に対して定数項を含む。従属変数の定数をユーザー指定によるモデルに含めます。

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。これはデフォルトです。

主効果。選択した変数のそれぞれに主効果の項を作成します。

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。

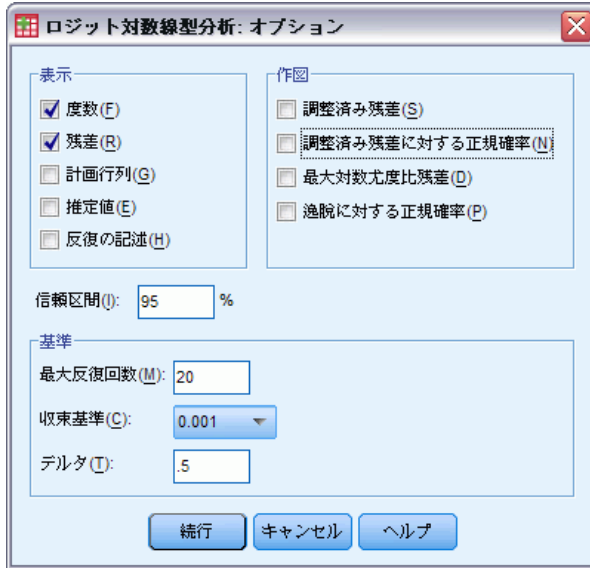
3 次まで。選択した変数に 3 次までの交互作用を作成します。

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。

ロジット対数線型分析のオプション

図 11-3
[ロジット対数線型分析: オプション] ダイアログ ボックス



ロジット対数線型分析手続きでは、モデル情報と適合度統計量を表示します。その他に、次のオプションから 1 つ以上を選択できます。

表示。 観測セル度数と期待セル度数（元データ、調整済み、および最大対数尤度比残差）、モデルの計画行列、およびモデルのパラメータ推定値など、いくつかの統計量を表示できます。

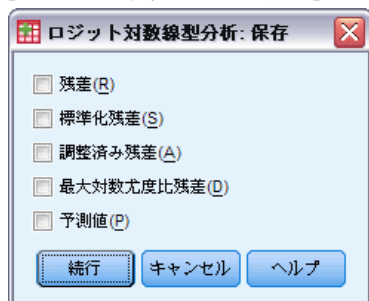
作図。 ユーザーの指定によるモデルで使用できる作図には、2 つの散布図行列（観測セル度数および予測セル度数に対する調整済み残差または最大対数尤度比残差）が含まれます。調整済み残差または最大対数尤度比残差の正規確率および傾向化除去正規プロットを表示することもできます。

信頼区間。 パラメータ推定値の信頼区間を調整できます。

基準。 Newton-Raphson 法を使用して、最尤パラメータ推定値を出します。[最大反復回数]、[収束基準]、[デルタ]（最初の近似のため、すべてのセルに加える定数）の各ボックスに新たな値を入力できます。飽和モデルのセルのデルタはそのままです。

ロジット対数線型分析の保存

図 11-4
[ロジット対数線型分析: 保存] ダイアログ ボックス



新変数としてアクティブ データセットに保存する値を選択します。新しい変数名の接尾辞 n を 1 ずつ増やしていくことで、保存変数に一意的な名前を作成することができます。

集計データ（分割表のセルのもの）がデータ エディタで個別の観測値ごとに記録されている場合でも、保存値はそのデータを参照します。未集計データの残差や予測値を保存すると、分割表のセルの保存値は、データ エディタ中のそのセルのケースごとに入力されます。保存値を理解するためには、データを集計してセル度数を出す必要があります。

保存できる残差は、[残差]、[標準化残差]、[調整済み残差]、[最大対数尤度比残差] の 4 種類です。[予測値] も保存できます。

- **残差.** 単純残差または生の残差とも呼ばれ、観測されるセルの度数とそれの期待度数の差を表します。
- **標準化残差.** 残差を標準誤差の推定値で割った値。標準化残差は、Pearson 残差としても知られています。
- **調整済み残差.** 調整済み残差は、標準化残差を推定された標準誤差で割り算したものです。モデルが正しいとき、調整済み残差は漸近的に標準正規分布になるので、正規性のチェックに使われます。
- **最大対数尤度比残差.** 尤度比カイ 2 乗統計量に対する個々の寄与率の符号付き（この符号は残差（観測度数から期待度数を引いた値）の符号です）平方根（ G の 2 乗）。最大対数尤度比残差は、漸近標準正規分布になります。

GENLOG 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。

- 観測セル度数と期待セル度数の線型組み合わせの計算、およびその組み合わせの残差、標準化残差、調整済み残差の印刷（GERESID サブコマンドを使用）。

- デフォルトの冗長検査用しきい値の変更 (CRITERIA サブコマンドを使用)。
- 標準化残差の表示 (PRINT サブコマンドを使用)。

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

生命表（生命表 オプション）

雇用期間（就職から退職までの期間）など、2つの事象の間の時間分布を調査したくなるさまざまな状況があります。しかし、この種のデータには2番目の事象が記録されないケースが数多くあります（たとえば、調査の終了時点でまだ調査対象者が会社に在籍している場合など）。このようなケースが発生する理由はいくつかあります。たとえば、調査が終了する前に2番目の事象が単に発生しなかっただけという場合、または、調査とは関連しない理由で、その事象が中断している場合（従業員が病気で休職している場合など）が考えられます。これらのケースをまとめて、**打ち切られたケース**と呼びます、この種の調査に、t検定や線型回帰分析など従来の手法を使用することは適切ではありません。

この種のデータに対しては、フォローアップ **生命表**という統計手法が有効です。この生命表の基本的な考えは、観測期間をより短い時間間隔に分割するということです。各時間間隔において、少なくともその間隔で観測されたすべての人を対象に、その時間間隔の中で発生する終結事象の確率を計算します。次に、各時間間隔から推定された確率を使って、さまざまな時点で発生する事象の確率全体を推定します。

例。新しいニコチンパッチ療法は、従来のパッチ療法に比べて禁煙効果が優れているのでしょうか。この場合、喫煙者を、従来どおりの治療を受けるグループと、試験的な治療を受けるグループの2つに分けて研究を実施できます。データから生命表を構築すれば、この2つのグループの全体的な禁煙率を比較し、試験的な治療が従来どおりの治療より優れているかどうかを判断できます。同時に、累積生存関数とハザード関数を作図して視覚的に比較し、さらに詳細な情報を得ることもできます。

統計量。各グループの各時間間隔に対する、参加数、離脱数、危険にさらされた数、終結事象の数、終結する比率、生存する比率、生存する累積比率（および標準誤差）、確率密度（および標準誤差）、ハザード比（および標準誤差）、およびグループ間の生存分布を比較するための Wilcoxon (Gehan) 検定。作図：累積生存関数、対数累積生存関数、密度関数、ハザード比関数、および 1 マイナス累積生存関数に対する作図。

データ。時間変数は、量的変数でなければなりません。状態変数は2分変数またはカテゴリ変数であり、整数としてコード化されます。この状態変数は、単一値または連続値の範囲としてコード化される事象を伴います。因子変数はカテゴリ変数でなければならず、整数としてコード化されます。

仮定。注目している事象の確率は、初期事象以降の時間にだけ依存し、絶対時間に対して一定であると仮定されます。つまり、異なる時間で調査に入るケース（たとえば、異なる時間に治療を始める患者など）は、同じように起

こる必要があります。また、調査済みケースとそうでないケースとの間にシステム上の違いがあってははいけません。たとえば、調査済みケースの多くがより深刻な状態の患者である場合、結果が偏向することがあります。

関連手続き。[生命表] 手続きは、この種の分析に対して保険統計の手法を使います（一般に生存分析と呼ばれます）。[Kaplan-Meier 生存分析] 手続きは、観測期間のより短い時間間隔への分割に依存しない、少し異なる生命表の計算方法を使います。この方法が推奨されるのは、観測数が少なく、各生存時間間隔における観測数がわずかしくないような場合です。生存時間との関連に疑いのある変数、または共変量に対して制御したい変数がある場合、[Cox 回帰] 手続きを使います。同一のケースに対して各時点で共変量の値が変わる可能性がある場合、[時間依存の Cox 回帰] を使います。

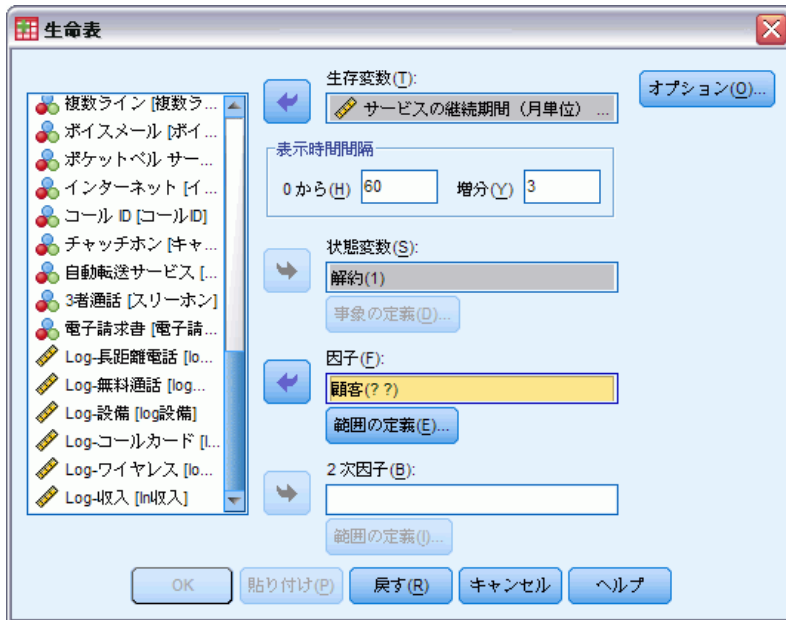
生命表を作成するには

- ▶ メニューから次の項目を選択します。

分析 > 生存分析 > 生命表...

図 12-1

[生命表] ダイアログ ボックス



- ▶ **数値型**の生存変数を 1 つ選択します。
- ▶ 調べる時間間隔を指定します。
- ▶ 状態変数を選択し、終結事象が発生したケースを定義します。
- ▶ [事象の定義] をクリックし、イベントが発生したことを示す状態変数の値を指定します。

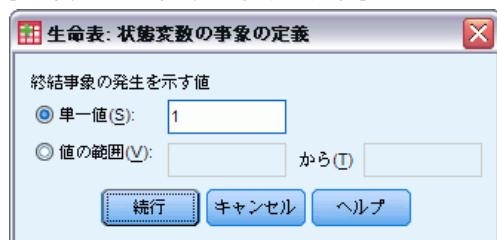
オプションとして、1 次因子変数を選択できます。生存変数に対する保険統計表が、因子変数の各カテゴリに対して生成されます。

また、2 次因子変数を選択することもできます。生存変数に対する保険統計表が、1 次および 2 次の因子変数のあらゆる組み合わせに対して生成されます。

生命表の状態変数の事象の定義

図 12-2

[生命表: 状態変数の事象の定義] ダイアログ ボックス

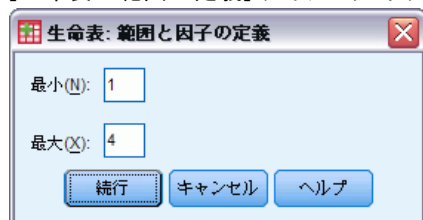


状態変数に対して選択された値が発生した場合、それらのケースに対して終結事象が発生したことを示します。その他のすべてのケースは、調査済みと見なされます。関連事象を識別する単一値または値の範囲を入力します。

生命表の範囲の定義

図 12-3

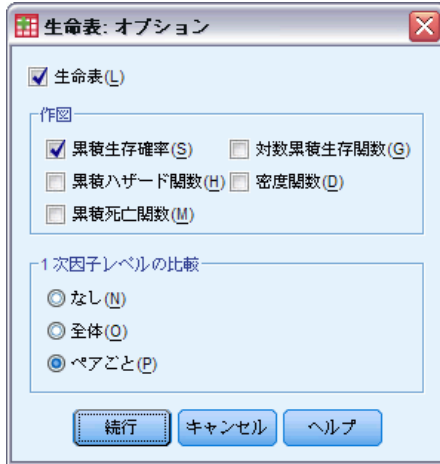
[生命表の範囲の定義] ダイアログ ボックス



指定範囲内の因子変数に対する値を持つケースは、分析の中に含まれます。また、各表 (要求があればプロットも) は範囲内の一意な値ごとに生成されます。

生命表のオプション

図 12-4
[生命表: オプション] ダイアログ ボックス



生命表分析は、さまざまな角度から管理できます。

生命表。出力中の生命表の表示を抑制するには、[生命表] を選択解除します。

作図。累積生存関数の作図を要求できます。因子変数を定義している場合、その因子変数によって定義されたサブグループごとに作図されます。利用できる作図は、累積生存関数、対数累積生存関数、ハザード関数、密度関数、および 1 マイナス累積生存関数です。

- **生存推定値。**累積生存関数を線型のスケールで表示します。
- **対数累積生存関数。**累積生存関数を対数のスケールで表示します。
- **累積ハザード関数。**累積ハザード関数を線型のスケールで表示します。
- **密度関数。**密度関数を表示します。
- **累積死亡関数。**1 次スケールで、1 マイナス累積生存関数を作図します。

1 次因子レベルの比較。1 次の制御変数がある場合、このグループ内の対象から 1 つを選択して Wilcoxon (Gehan) の検定を実行し、サブグループの累積生存確率を比較できます。検定は 1 次因子に対して実行されます。2 次因子を定義している場合、検定は 2 次変数のレベルごとに実行されます。

SURVIVAL 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。

- 複数の従属変数の指定。
- 不均一な間隔の指定。

- 複数の状態変数の指定。
- 必ずしもすべての因子および制御変数を含まない比較の指定。
- 厳密ではないおおよその比較の計算。

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

Kaplan-Meier 生存分析

雇用期間（就職から退職までの期間）など、2 つの事象の間の時間分布を調査したくなるさまざまな状況があります。ただし、この種のデータには、多少の調査済みケースも含まれているのが一般的です。調査済みケースとは、第 2 の事象がこれ以上記録されないケースのことです（たとえば、調査の終了時点でまだ調査対象者が会社に在籍している場合など）。Kaplan-Meier 手続きは、調査済みケースが存在する中で、事象までの時間のモデルを推定するための手法です。基本的に、Kaplan-Meier モデルは事象が発生する各時点で条件付き確率を推定し、これらの確率の積の限界を使って各時点における生存率を推定します。

例。AIDS の新しい治療法には延命に対する医療上の効果があるでしょうか。この場合、AIDS 患者を、従来の治療を受けるグループと、試験的な治療を受けるグループの 2 つに分けて研究を実施できます。データから Kaplan-Meier モデルを構築すれば、この 2 つのグループの全体的な生存率を比較し、試験的な治療が従来の治療より優れているかどうかを確定できます。同時に、累積生存関数とハザード関数を作図して視覚的に比較し、さらに詳細な情報を得ることもできます。

統計量。時間、状態、累積生存、および標準誤差を含む生命表、累積終結事象数、残存数；および標準誤差および 95% の信頼区間を伴う平均値および中央値の生存時間。作図：累積生存関数、ハザード関数、対数累積生存関数、および 1 マイナス累積生存関数。

データ。時間変数は量的変数でなければなりません、状態変数はカテゴリ変数でも量的変数でも構いません。因子変数とストラータ変数はカテゴリ変数でなければなりません。

仮定。注目している事象の確率は、初期事象以降の時間にだけ依存し、絶対時間に対して一定であると仮定されます。つまり、異なる時間で調査に入るケース（たとえば、異なる時間に治療を始める患者など）は、同じように起こる必要があります。また、調査済みケースとそうでないケースとの間にシステム上の違いがあってははいけません。たとえば、調査済みケースの多くがより深刻な状態の患者である場合、結果が偏向することがあります。

関連手続き。Kaplan-Meier 手続きでは、各事象の発生時間における生存確率関数またはハザード関数を推定する生命表計算の手法を使用します。[生命表] 手続きは、観測期間のより短い時間間隔への分割に依存する生存分析に対して、保険統計の手法を使います。大量のサンプルを扱う場合に有用な場合があります。生存時間との関連に疑いのある変数、または共変量に対して制御したい変数がある場合、[Cox 回帰] 手続きを使います。同一

のケースに対して各時点で共変量の値が変わる可能性がある場合、[時間依存の Cox 回帰] を使います。

Kaplan-Meier 生存分析を実行するには

- ▶ メニューから次の項目を選択します。
分析 > 生存分析 > Kaplan-Meier...

図 13-1
[Kaplan-Meier] ダイアログ ボックス



- ▶ 時間変数を選択します。
- ▶ 状態変数を選択し、終結事象が発生したケースを識別します。この変数は、数値型または **短い文字型** です。次に、[事象の定義] をクリックします。オプションとして、因子変数を選び、グループ間の差を調べることができます。また、ストラータ変数を選べば、変数のレベル別（階層）分析を生成できます。

Kaplan-Meier の状態変数の事象の定義

図 13-2

[Kaplan-Meier: 状態変数の事象の定義] ダイアログ ボックス

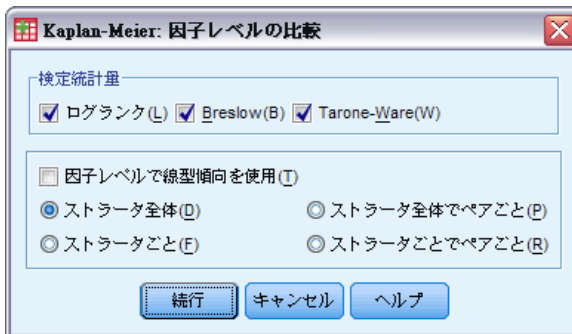


終結事象が発生したことを示す値（1 つ以上）を入力します。[単一値]、[値の範囲]、あるいは [値のリスト] のいずれか 1 つを選んで値を入力できます。[値の範囲] が利用できるのは、状態変数が数値型の場合に限られます。

Kaplan-Meier の因子レベルの比較

図 13-3

[Kaplan-Meier: 因子レベルの比較] ダイアログ ボックス



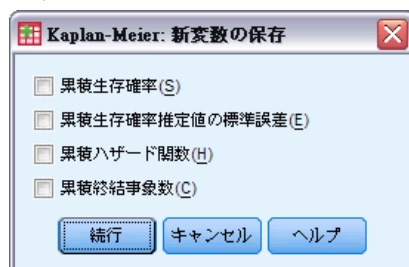
レベルが異なる因子について、累積生存確率分布の同等性を検定する統計を要求できます。利用できる統計は、[ログ ランク]、[Breslow]、[Tarone-Ware] です。これらの中から 1 つを選び、[ストラータ全体]、[ストラータごと]、[ストラータ全体でペアごと]、[ストラータごとでペアごと] から、実行する比較を指定します。

- **ログランク**. 生存分布の等質性を比較するための検定。この検定では、すべての時点に同じ重みが割り当てられます。
- **Breslow の検定**. 生存分布の等質性を比較するための検定。各時点での危険なケースの個数によって、時点に重みを付けます。
- **Tarone-Ware の検定**. 生存分布の等質性を比較するための検定。各時点における危険なケースの個数の平方根によって時点に重みを付けます。
- **ストラータ全体**. 生存曲線の相等性を検定するため、単一検定ですべての因子レベルを比較します。
- **ストラータ全体でペアごと**. 因子レベルの各個別ペアを比較します。ペアごとのトレンド検定は利用できません。
- **ストラータごと**. 各ストラータに対してすべての因子レベルの相等性を個別に検定します。ストラータ変数がない場合、検定は実行されません。
- **ストラータごとでペアごと**. 各ストラータに対して因子レベルの各個別ペアを比較します。ペアごとのトレンド検定は利用できません。ストラータ変数がない場合、検定は実行されません。

因子水準で線型傾向を使用. 因子のさまざまなレベルで線型傾向に対する検定を行うことができます。このオプションは、因子レベルを（ペアごとでなく）全体的に比較する場合に限り利用できます。

Kaplan-Meier の新変数の保存

図 13-4
[Kaplan-Meier: 新変数の保存] ダイアログ ボックス



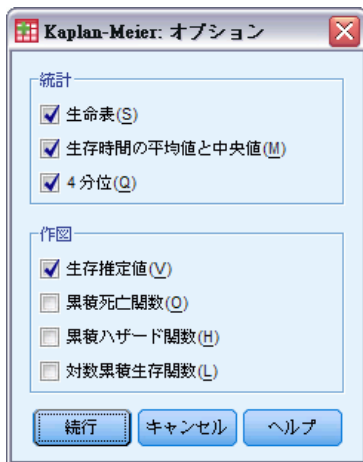
Kaplan-Meier 表の情報を新変数として保存できます。保存した新変数をその後の分析に使い、仮説を検定したり、仮定を確認したりできます。[累積生存確率] と、[累積生存確立推定値の標準誤差]、[累積ハザード関数]、[累積終結事象数] チェック ボックスをオンにし、新変数として保存できます。

- **生存推定値**. 累積生存確率推定値。デフォルトの変数名は、接頭辞が sur_ で、その後に連続番号が付加されます。たとえば、sur_1 がすでに存在している場合、Kaplan-Meier は sur_2 という変数名を付けます。

- **累積生存確率推定値の標準誤差.** 累積生存確率推定値の標準誤差。デフォルトの変数名は、接頭辞が `se_` で、その後に連続番号が付加されます。たとえば、`se_1` がすでに存在している場合、Kaplan-Meier は `se_2` という変数名を付けます。
- **累積ハザード関数.** 累積ハザード関数推定値。デフォルトの変数名は、接頭辞が `haz_` で、その後に連続番号が付加されます。たとえば、`haz_1` がすでに存在している場合、Kaplan-Meier は `haz_2` という変数名を付けます。
- **累積終結事象数.** ケースを生存時間およびステータス コードで並べ替えたときのイベントの累積度数。デフォルトの変数名は、接頭辞が `cum_` で、その後に連続番号が付加されます。たとえば、`cum_1` がすでに存在している場合、Kaplan-Meier は `cum_2` という変数名を付けます。

Kaplan-Meier のオプション

図 13-5
[Kaplan-Meier: オプション] ダイアログ ボックス



Kaplan-Meier 分析では、各種の出力を要求できます。

統計量. 生命表、平均値および中央値の生存時間、4 分位など、計算した累積生存確率関数について表示される統計量を選択できます。因子変数を含めている場合は、グループごとに統計量が生成されます。

作図. 作図することにより、累積生存関数、累積死亡関数、累積ハザード関数、および対数累積生存関数を視覚的に調べることができます。因子変数を含めている場合、関数はグループごとに作図されます。

- **生存推定値.** 累積生存関数を線型のスケールで表示します。
- **累積死亡関数.** 1 次スケールで、1 マイナス累積生存関数を作図します。

- **累積ハザード関数.** 累積ハザード関数を線型のスケールで表示します。
- **対数累積生存関数.** 累積生存関数を対数のスケールで表示します。

KM 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。

- フォローアップではなくなったケースを、調査済みケースとは別のカテゴリとして扱う度数分布表を取得する。
- 線型傾向の検定の際に不均等な間隔を指定する。
- 4 分位以外で累積生存時間変数のパーセンタイルを求める。

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

Cox 回帰分析

Cox 回帰分析では、事象までの時間のデータに関する予測モデルを作成します。このモデルは、予測変数の特定の値に対する特定の時間 t に関心のある事象が発生する確率を予測する生存関数を生成します。生存関数の形状および予測変数の回帰係数は、観測された被験者から推定されます。このモデルは、予測変数を測定した新しいケースに適用できます。打ち切られた被験者（つまり、観測時間中に関心のある事象を経験していない被験者）からの情報は、モデルの推定に大きく寄与することに注意してください。

例:喫煙による肺ガン発生の危険性は、男性と女性とで異なるでしょうか。Cox 回帰分析モデルを構築する際に喫煙量（1 日あたりに吸うタバコの本数）と性別を共変量として入力すれば、肺ガンが発病するまでの期間における、性別および喫煙量が与える影響についての仮説を検定できます。

統計量。各モデル（-2LL、尤度比統計量、および全体のカイ 2 乗）。モデルの変数（パラメータ推定、標準誤差、Wald 統計量）。モデルにない変数（スコア統計、残差カイ 2 乗）。

データ。時間変数は量的変数でなければなりません、状態変数はカテゴリ変数または量的変数のどちらでもかまいません。独立変数（共変量）は、量的変数またはカテゴリ変数のどちらでもかまいません。カテゴリ変数にする場合、独立変数はコード化されたダミーまたは指標でなければなりません（この手続き内には、カテゴリ変数を自動的に再コード化するオプションがあります）。ストラータ変数は、カテゴリ変数でなければなりません。これらは、整数または短い文字型としてコード化されます。

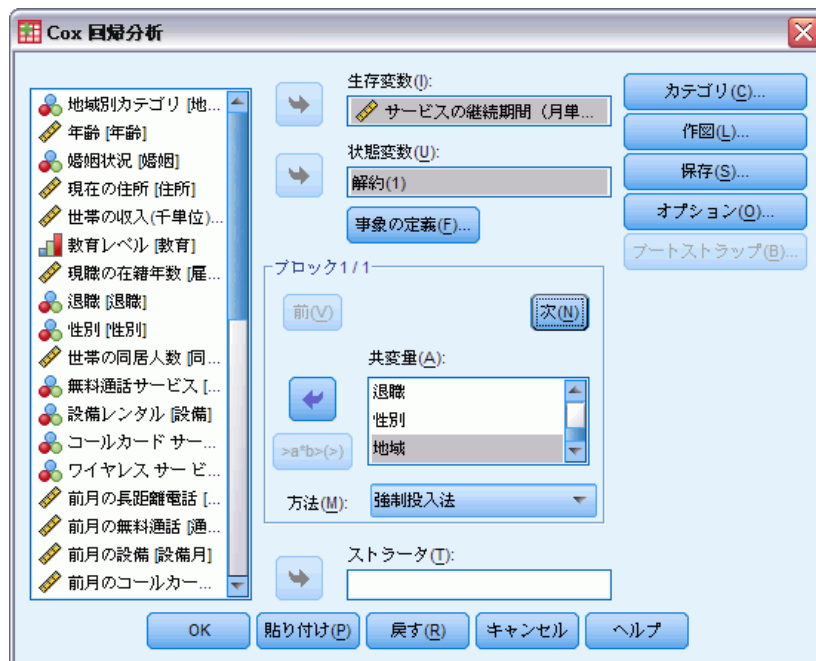
仮定。観測値は独立変数でなければなりません。またハザード比は時間に対して一定でなければなりません。つまり、あるケースと他のケースとのハザードの比例関係は時間に対して変化してはならないということです。後者の仮定は、**ハザードの比例仮定**と呼ばれます。

関連手続き。ハザードの比例仮定が一定でない場合（上記参照）、[時間依存の Cox 回帰] 手続きを使う必要がある場合があります。共変量がない場合、またはカテゴリ共変量が 1 つしかない場合、[生命表] 手続きまたは [Kaplan-Meier] 手続きを使って、サンプルに対する生存関数またはハザード関数を調べることができます。サンプル内に調査済みデータがない場合（つまり、すべてのケースで終結事象が実行されている場合）、[線型回帰] 手続きを使って予測と事象までの時間との間の関係をモデリングできます。

Cox 回帰分析の実行

- ▶ メニューから次の項目を選択します。
分析(A) > 生存分析 > Cox 回帰...

図 14-1
[Cox 回帰] ダイアログ ボックス



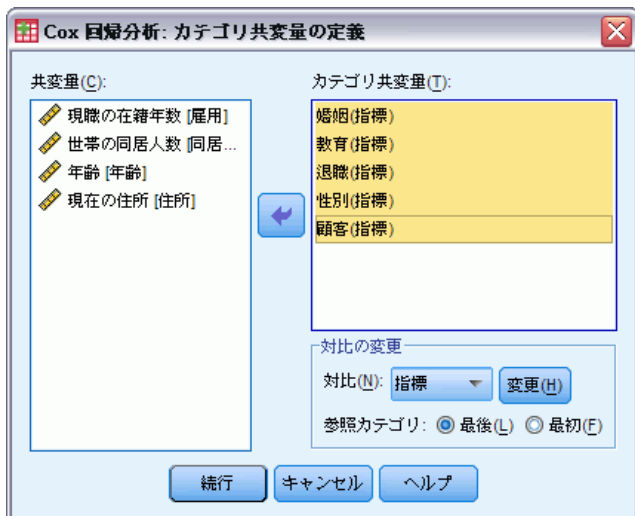
- ▶ 時間変数を選択します。時刻値が負になっているケースは分析されません。
- ▶ 状態変数を選択し、[事象の定義] をクリックします。
- ▶ 1 つ以上の共変量を選択します。交互作用項を含めるには、交互作用に含めるすべての変数を選択し、次に [>a*b>] をクリックします。

オプションとして、ストラータ変数を定義することにより、グループ別のモデルを計算できます。

Cox 回帰分析におけるカテゴリ変数の定義

図 14-2

[Cox 回帰分析: カテゴリ共変量の定義] ダイアログ ボックス



[Cox 回帰分析] 手続きでカテゴリ変数を処理する方法は、細かく指定できます。

共変量。層を問わず、単独で、または相互作用の一部として、メイン ダイアログ ボックスで指定したすべての共変量の一覧を表示します。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。各変数のかつこ内には、使用する対比コードが表示されています。文字型変数（変数名の後に記号く付き）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。

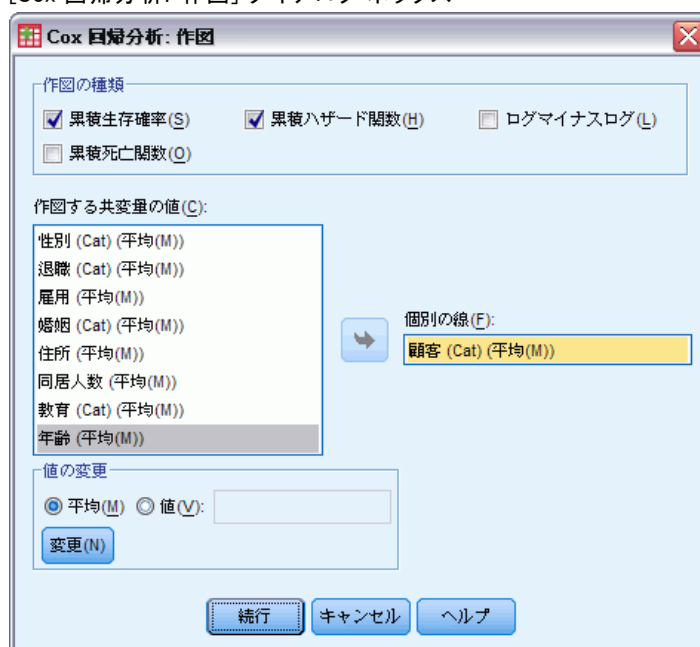
- **多項式**。直交多項式の対比。カテゴリは、一定の間隔を等しく空けて配置されていると仮定されます。多項式の対比は、数値型変数に限り利用できます。
- **偏差**。[参照カテゴリ] で指定したカテゴリを除く予測変数の各カテゴリを、全体的な結果と比べます。

[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。

Cox 回帰分析の作図

図 14-3
[Cox 回帰分析: 作図] ダイアログ ボックス



作図することにより、推定モデルを評価し、また結果を解釈できます。累積生存関数、ハザード関数、ログマイナスログ関数、および 1 マイナス累積生存関数を作図できます。

- **生存分析**。累積生存関数を線型のスケールで表示します。
- **累積ハザード関数(H)**。累積ハザード関数を線型のスケールで表示します。

- **ログマイナスログ**. $\ln(-\ln)$ 変換が推定値に適用された後の累積生存推定値を保存します。
- **累積死亡関数(M)**. 1 次スケールで、1 マイナス累積生存関数を作図します。

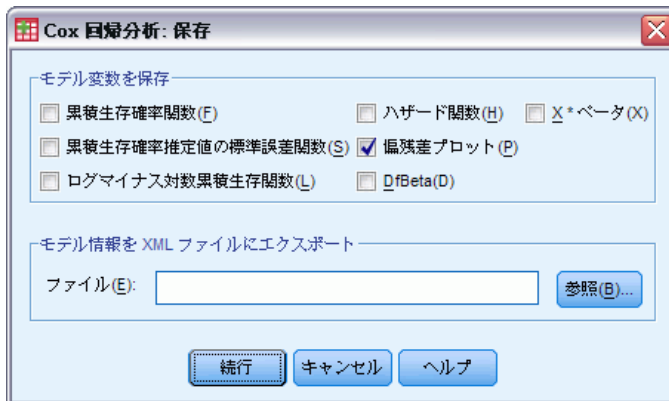
これらの関数は共変量の値に依存しているので、時間に対する関数を作図するには、共変量に対して定数値を使う必要があります。デフォルトは各共変量の平均を定数として用いますが、[値の変更]を使えば、作図に独自の値を入力できます。

共変量を [個別の線] テキスト ボックスに移動すれば、カテゴリ共変量の値ごとに個別の線を作図できます。このオプションは、カテゴリ共変量に限り選択できます。[作図する共変量の値] リストでは、カテゴリ共変量の名前の後に (カテゴリ) と表示されます。

Cox 回帰分析の新変数の保存

図 14-4

[Cox 回帰分析: 新変数の保存] ダイアログ ボックス



各種分析結果は、新変数として保存できます。変数を保存すれば、その後の分析に使用して仮説を検定したり、仮定を確認したりできます。

モデル変数の保存. 回帰に対する生存関数、標準誤差、ログマイナスログ推定値、ハザード関数、偏残差プロット、DfBeta、および線型予測 $X*Beta$ を新変数として保存できます。

- **生存推定値**. 与えられた時間 (時点) に対する累積生存関数の値。それは、その期間の生存確率に等しくなります。
- **ログマイナス対数累積生存関数(L)**. $\ln(-\ln)$ 変換が推定値に適用された後の累積生存推定値を保存します。
- **ハザード関数(H)**. 累積ハザード関数推定値 (Cox-Snell 残差ともいう) を保存します。

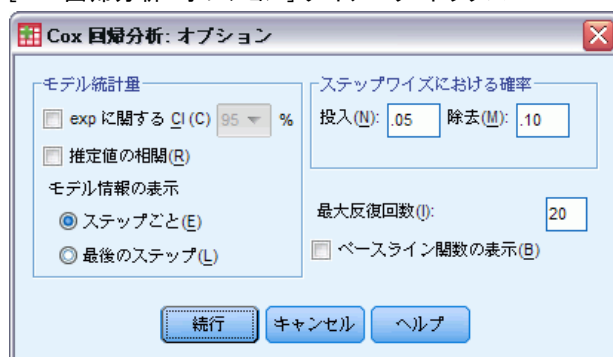
- **偏残差プロット(P)**. 偏残差を生存時間に対してプロットすると、比例ハザードの仮定を検定することができます。最後のモデルでは各共変量に対して 1 つの変数が保存されます。偏残差は、少なくとも 1 つの共変量を持つモデルに対してのみ使用することができます。
- **DfBeta**. ケースを除去した場合の係数の推定された変化。最後のモデルでは各共変量に対して 1 つの変数が保存されます。DfBetas は、少なくとも 1 つの共変量を持つモデルに対してのみ使用することができます。
- **X*ベータ**. 線型予測得点。個々のケースに対する、平均値中心化共変量の値とそれらに対応するパラメータ推定値の積和。

時間依存の共変量を使って Cox 回帰分析を実行する場合、保存できる変数は DfBeta および線型予測変数 X*Beta だけです。

モデル情報を XML ファイルにエクスポート. パラメータ推定は指定されたファイルに XML 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。

Cox 回帰分析のオプション

図 14-5
[Cox 回帰分析: オプション] ダイアログ ボックス



分析と出力は、いろいろな角度から管理できます。

モデル統計量. モデル パラメータには、[Exp(B) の信頼区間] や [推定値の相関] などの統計量を得ることができます。[各ステップごと]、または [最終ステップ] で、これらの統計量を要求できます。

ステップワイズにおける確率. ステップワイズの方法を選択すると、モデルへの投入または除去のどちらかで確率を指定できます。変数は、投入される F 値の有意水準が [投入] の値を下回ると投入され、有意水準が [除去] の値を上回ると除去されます。[投入] の値は [除去] の値未満である必要があります。

最大反復回数. モデルの最大反復回数を指定できます。これにより、手続きによる解の算出時間を制御できます。

ベースライン関数の表示。ベースラインのハザード関数と累積生存確率を共変量の平均値で表示できます。時間依存の共変量を指定していると、表示できません。

Cox 回帰分析の状態変数の事象の定義

終結事象が発生したことを示す値（1 つ以上）を入力します。[単一値]、[値の範囲]、あるいは [値のリスト] のいずれか 1 つを選んで値を入力できます。[値の範囲] が利用できるのは、状態変数が数値型の場合に限られます。

COXREG 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。

- フォローアップではなくなったケースを、調査済みケースとは別のカテゴリとして扱う度数分布表を取得する。
- 偏差、単純、指標の各対比のいずれかで、最初と最後以外のカテゴリを参照カテゴリとして選択。
- 直交多項式の対比で、カテゴリの不等間隔の指定。
- 反復回数の追加基準の指定。
- 欠損値の扱いの制御。
- 保存変数名の指定。
- 外部 IBM® SPSS® Statistics データ ファイルへの出力の書き出し。
- 処理の際、外部スクラッチ ファイルに分割ファイル グループのデータを保持。大量のデータセットを使って分析を実行する際に、メモリー リソースを効果的に節約できます。これは、時間依存の共変量には利用できません。

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。

時間依存の共変量の計算

Cox 回帰分析モデルを計算したいときにハザードの比例仮定が適用できない状況があります。具体的には、ハザード率が時間によって変化し、(1 以上の) 共変量の値が時点によって異なるような場合です。そのような場合には、拡張 Cox 回帰分析モデルを使う必要があります。それにより、**時間依存の共変量**を指定できます。

このようなモデルを分析するには、まず時間依存の共変量を定義しなければなりません。(コマンド シンタックスを使えば、複数の時間依存の共変量を指定できます)。この作業を容易にするために、時間を表すシステム変数を利用できます。この変数を `T_` といいます。この変数を使うと、次のように、一般的な 2 通りの方法で時間依存の共変量を定義できます。

- 特定の共変量について、ハザードの比例仮定を検定したい場合、あるいは比例しないハザードを可能にする拡張 Cox 回帰分析を推定したい場合は、時間変数 `T_` と対象の共変量の関数として時間依存の共変量を定義すれば、その作業を実行できます。時間変数と共変量の単純な結果などが一般的な例と言えますが、同様に、複雑な関数も指定できます。時間依存の共変量の係数の有意確率検定を行うことで、ハザードの比例仮定が理にかなうものであるかどうかわかります。
- 変数の中には、時間によって値が異なりますが、系統だった時間との関連性がないものもあります。このような場合、**セグメント化した時間依存の共変量**を定義する必要があります。これは、**論理式**を使って実行できます。この論理式では、真であれば 1、偽であれば 0 という値を取ります。一連の論理式を使うと、1 組の測定から時間依存の共変量を作成できます。たとえば、研究のため、週に一度の血圧測定を 4 週間続ける場合は、(BP1 から BP4 という名前で識別)、時間依存の共変量を $(T_ < 1) * BP1 + (T_ \geq 1 \ \& \ T_ < 2) * BP2 + (T_ \geq 2 \ \& \ T_ < 3) * BP3 + (T_ \geq 3 \ \& \ T_ < 4) * BP4$ と定義できます。どのケースでもかっこ内の条件の 1 つだけが 1 となり、他はすべて 0 となることに注意してください。つまり、この関数は、時間が 1 週間未満の場合は BP1 を使い、1 週間より多く 2 週間未満の場合には BP2 を使う、ということの意味をしています。

[時間依存の共変量の計算] ダイアログ ボックスで、関数作成用のコントロールを使うことにより、時間依存の共変量の式を作成できますが、[時間依存の共変量の式] テキスト領域に直接入力することもできます。文字型定数はダブル クォーテーションかアポストロフィで囲み、数値定数の場合は小数部を点で区切って米国方式で入力しなければならないで、注意してください。そのようにして得た変数は `T_COV_` と呼ばれ、Cox 回帰モデルに共変量として使用できます。

時間依存の共変量を計算するには

- ▶ メニューから次の項目を選択します。
分析 > 生存分析 > 時間依存の Cox 回帰...

図 15-1
[時間依存の共変量の計算] ダイアログ ボックス



- ▶ 時間依存の共変量の式を入力します。
- ▶ [モデル...] をクリックして、Cox 回帰分析を続行します。

注: Cox 回帰モデルの共変量として、新変数の T_COV_ を必ず指定してください。

詳細は、14 章 p.168 Cox 回帰分析 を参照してください。

時間依存の共変量を使用する Cox 回帰分析の追加機能

コマンド言語を使用すると、複数の時間依存の共変量を指定できます。Cox 回帰における時間依存の共変量の有無とは関係なく、コマンド シンタックスの追加機能を利用できます。

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。

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

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。

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

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。

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

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) |

$$\text{df}(k-1) \quad \left(\begin{array}{cccc} 0 & 0 & 0 & \dots \\ & & & 1 & -1 \end{array} \right)$$

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。

$$\left(\begin{array}{cccc} 1/4 & 1/4 & 1/4 & 1/4 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right)$$

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。

| | |
|---|-----------------------|
| $\left(\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array} \right)$ | 平均計算のための重み |
| $\left(\begin{array}{cccc} 3 & -1 & -1 & -1 \end{array} \right)$ | 1 番目と 2 ~ 4 番目の比較 |
| $\left(\begin{array}{cccc} 0 & 2 & -1 & -1 \end{array} \right)$ | 2 番目と 3 番目および 4 番目の比較 |
| $\left(\begin{array}{cccc} 0 & 0 & 1 & -1 \end{array} \right)$ | 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。

共分散構造

この項は、共分散構造に関する追加情報です。

前従属: 1 次。この共分散構造は、隣接する要素間で不均質の分散および不均質の相関を持ちます。隣接しない 2 つの要素の間の相関は、その間に存在する各要素間の相関の積を取ったものです。

$$\begin{bmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho_1 & \sigma_3\sigma_1\rho_1\rho_2 & \sigma_4\sigma_1\rho_1\rho_2\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_3\sigma_2\rho_2 & \sigma_4\sigma_2\rho_2\rho_3 \\ \sigma_3\sigma_1\rho_1\rho_2 & \sigma_3\sigma_2\rho_2 & \sigma_3^2 & \sigma_4\sigma_3\rho_3 \\ \sigma_4\sigma_1\rho_1\rho_2\rho_3 & \sigma_4\sigma_2\rho_2\rho_3 & \sigma_4\sigma_3\rho_3 & \sigma_4^2 \end{bmatrix}$$

AR(1)。等質の分散を持つ 1 次自己回帰構造です。2 つの要素間の相関は、両要素が隣接する場合は ρ に等しく、1 要素分隔たりがある場合は ρ^2 となり、以下同様に続きます。 ρ の値は、 $-1 < \rho < 1$ になるように制限されます。

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

AR(1): 不均質。不均質の分散を持つ 1 次自己回帰構造です。2 つの要素間の相関は、両要素が隣接する場合は ρ に等しく、1 要素分隔たりがある場合は ρ^2 となり、以下同様に続きます。 ρ の値は、 -1 と 1 の間になるように制限されます。

$$\begin{bmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho & \sigma_3\sigma_1\rho^2 & \sigma_4\sigma_1\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_3\sigma_2\rho & \sigma_4\sigma_2\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_4\sigma_3\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$$

ARMA(1,1)。1 次自己回帰の移動平均構造です。分散は等質です。2 つの要素間の相関は、両要素が隣接する場合は $\phi * \rho$ に等しく、1 要素分隔たりがある場合は $\phi * (\rho^2)$ となり、以下同様に続きます。 ρ と ϕ は、それぞれ自己回帰と平均パラメータであり、これらの値は常に $-1 \sim 1$ の間になります (-1 と 1 を含む)。

$$\sigma^2 \begin{bmatrix} 1 & \phi\rho & \phi\rho^2 & \phi\rho^3 \\ \phi\rho & 1 & \phi\rho & \phi\rho^2 \\ \phi\rho^2 & \phi\rho & 1 & \phi\rho \\ \phi\rho^3 & \phi\rho^2 & \phi\rho & 1 \end{bmatrix}$$

複合シムトリ。この構造は定数分散および定数共分散を持ちます。

$$\begin{bmatrix} \sigma^2 + \sigma_1^2 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1^2 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1^2 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1^2 \end{bmatrix}$$

複合シムトリ: 相関行列。この共分散構造は、要素間に等質の分散および等質の相関があります。

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

複合シムトリ: 不均質。この共分散構造は、要素間に不均質の分散および定数の相関があります。

$$\begin{bmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho & \sigma_3\sigma_1\rho & \sigma_4\sigma_1\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_3\sigma_2\rho & \sigma_4\sigma_2\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_4\sigma_3\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$$

対角。この共分散構造は、要素間に不均質の分散および 0 相関があります。

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

因子解析的: 1 次。この共分散構造は、要素間で不均質の項と等質の項で構成される不均質の分散を持ちます。2 つの要素間の共分散はすべて、不均質の分散項の積の平方根になります。

$$\begin{bmatrix} \lambda_1^2 + d & \lambda_2 \lambda_1 & \lambda_3 \lambda_1 & \lambda_4 \lambda_1 \\ \lambda_2 \lambda_1 & \lambda_2^2 + d & \lambda_3 \lambda_2 & \lambda_4 \lambda_2 \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & \lambda_3^2 + d & \lambda_4 \lambda_3 \\ \lambda_4 \lambda_1 & \lambda_4 \lambda_2 & \lambda_4 \lambda_3 & \lambda_4^2 + d \end{bmatrix}$$

因子解析的: 1次、不均質。この共分散構造は、要素間で不均質の2つの項で構成される不均質の分散を持ちます。2つの要素間の共分散は、不均質の分散項のうち、最初の項の積の平方根になります。

$$\begin{bmatrix} \lambda_1^2 + d_1 & \lambda_2 \lambda_1 & \lambda_3 \lambda_1 & \lambda_4 \lambda_1 \\ \lambda_2 \lambda_1 & \lambda_2^2 + d_2 & \lambda_3 \lambda_2 & \lambda_4 \lambda_2 \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & \lambda_3^2 + d_3 & \lambda_4 \lambda_3 \\ \lambda_4 \lambda_1 & \lambda_4 \lambda_2 & \lambda_4 \lambda_3 & \lambda_4^2 + d_4 \end{bmatrix}$$

Huynh-Feldt。これは「循環」行列であり、2つの要素間の共分散は、要素の分散の平均から定数を引いたものになります。分散および共分散は、定数ではありません。

$$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda & \frac{\sigma_1^2 + \sigma_3^2}{2} - \lambda & \frac{\sigma_1^2 + \sigma_4^2}{2} - \lambda \\ \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda & \sigma_2^2 & \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda & \frac{\sigma_2^2 + \sigma_4^2}{2} - \lambda \\ \frac{\sigma_1^2 + \sigma_3^2}{2} - \lambda & \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda & \sigma_3^2 & \frac{\sigma_3^2 + \sigma_4^2}{2} - \lambda \\ \frac{\sigma_1^2 + \sigma_4^2}{2} - \lambda & \frac{\sigma_2^2 + \sigma_4^2}{2} - \lambda & \frac{\sigma_3^2 + \sigma_4^2}{2} - \lambda & \sigma_4^2 \end{bmatrix}$$

計測された単位。この構造は定数分散を持ちます。どの要素間にも相関がないものと仮定されます。

$$\sigma^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Toeplitz。この共分散構造は、要素間に等質の分散および不均質の相関があります。隣接する要素間の相関は、どの対の間でも等質です。3番目の要素間の相関も等質です。

$$\sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}$$

Toeplitz: 不均質。この共分散構造は、要素間に不均質の分散および不均質の相関があります。隣接する要素間の相関は、どの対の間でも等質です。3番目の要素間の相関も等質です。

$$\begin{bmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho_1 & \sigma_3\sigma_1\rho_2 & \sigma_4\sigma_1\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_3\sigma_2\rho_1 & \sigma_4\sigma_2\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_4\sigma_3\rho_1 \\ \sigma_4\sigma_1\rho_3 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{bmatrix}$$

無構造。これは完全に一般的な分散共分散行列です。

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$$

無構造: 相関行列。この共分散構造は、不均質の分散および不均質の相関を持ちます。

$$\begin{bmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho_{21} & \sigma_3\sigma_1\rho_{31} & \sigma_4\sigma_1\rho_{41} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_3\sigma_2\rho_{32} & \sigma_4\sigma_2\rho_{42} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \sigma_4\sigma_3\rho_{43} \\ \sigma_4\sigma_1\rho_{41} & \sigma_4\sigma_2\rho_{42} & \sigma_4\sigma_3\rho_{43} & \sigma_4^2 \end{bmatrix}$$

分散成分。この構造は、指定した変量効果に、計測された単位 (ID) 構造を割り当てます。

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

- 最大対数尤度比残差
 - 一般化線型モデル, 75
- 一般化推定方程式, 79
 - 推定周辺平均, 99
 - 予測変数, 89
 - 推定基準, 93
 - 初期値, 95
 - 統計量, 97
 - 応答, 87
 - 2 値反応に関する参照カテゴリ, 88
 - アクティブなデータセットへの変数の保存, 101
 - カテゴリ因子に関するオプション, 90
 - モデル指定, 91
 - モデルの種類, 83
 - モデルをエクスポート, 103
- 分散共分散行列
 - 一般化推定方程式, 93, 98
 - 一般化線型モデル, 67, 71
- 多変量分散分析, 2
- 対数尤度収束
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67
- 対数線型分析
 - 一般化線型混合モデル, 106
- 対比係数行列
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 推定周辺平均
 - 一般化推定方程式, 99
 - GLM 多変量, 14
 - 一般化線型モデル, 72
- 共分散分析
 - GLM 多変量, 2
- 多変量回帰, 2
- 最小有意差
 - GLM 多変量, 10
- 観測平均値
 - GLM 多変量, 14
- 記述統計量
 - 一般化推定方程式, 98
 - GLM 多変量, 14
 - 一般化線型モデル, 71
- 交差相関
 - 仮説と誤差行列, 14
- 信頼区間
 - GLM 多変量, 14
- 分散分析
 - 一般化線型混合モデル, 106
- 反復回数
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67
- 多項分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- 尤度残差
 - 一般化線型モデル, 76
- 標準偏差
 - GLM 多変量, 14
- 標準誤差
 - GLM 多変量, 14
- 正規分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- 生存分析
 - Cox 回帰分析, 168
- 相関行列
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 共変量
 - Cox 回帰分析, 170
- 平方和
 - 仮説と誤差行列, 14
- 適合度
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 分離
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67
- 商標, 188
- 対比
 - Cox 回帰分析, 170
- 残差
 - 一般化推定方程式, 102
 - 一般化線型モデル, 75
- 段階 2 分
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67
- 多変量 GLM, 2
- 残差 SSCP
 - GLM 多変量, 14
- 分散分析(N)
 - GLM 多変量, 2
- 2 項分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- Bartlett の球面性の検定
 - GLM 多変量, 14
- Bonferroni (B)
 - GLM 多変量, 10
- Bonferroni の方法
 - GLM 反復測定, 28

- Box の M 検定
 - GLM 多変量, 14
- Breslow の検定
 - Kaplan-Meier, 164
- Cook の距離
 - 一般化線型モデル, 75
- Cook の距離
 - GLM, 12
 - GLM 反復測定, 30
- Cox 回帰分析, 168
 - 生存状態変数, 174
 - 反復回数, 173
 - 生存関数, 172
 - 共変量, 168
 - 統計量, 168, 173
 - 作図, 171
 - 対比, 170
 - 例, 168
 - DfBeta, 172
 - カテゴリ共変量, 170
 - コマンドの追加機能, 174
 - ステップワイズ法による投入と除去, 173
 - ストリング共変量, 170
 - 新変数の保存, 172
 - 事象の定義, 174
 - ハザード関数, 172
 - 偏残差プロット, 172
 - ベースライン関数, 173
- Cox 回帰分析
 - 時間依存の共変量, 175-176
- Duncan の多重範囲検定
 - GLM 多変量, 10
- Duncan の多重範囲検定
 - GLM 反復測定, 28
- Dunnett の C
 - GLM 多変量, 10
- Dunnett の t 検定
 - GLM 多変量, 10
- Dunnett の T3(3)
 - GLM 多変量, 10
- Dunnett の C
 - GLM 反復測定, 28
- Dunnett の T3
 - GLM 反復測定, 28
- Fisher の LSD
 - GLM 多変量, 10
- Fisher 得点
 - 線型混合モデル, 50
- Fisher の LSD
 - GLM 反復測定, 28
- Gabriel と Howell のペアごとの比較検定
 - GLM 多変量, 10
- Gabriel のペアごとの比較検定
 - GLM 多変量, 10
- Gabriel と Howell のペアごとの比較検定
 - GLM 反復測定, 28
- Gabriel のペアごとの比較検定
 - GLM 反復測定, 28
- Gehan の検定
 - 生命表, 160
- GLM
 - 行列の保存, 12
 - 変数の保存, 12
- GLM 多変量, 2, 15
 - 推定周辺平均, 14
 - 従属変数, 2
 - 共変量, 2
 - 因子, 2
 - 表示, 14
 - 診断, 14
 - オプション, 14
 - その後の検定, 10
 - プロファイル プロット, 8
- GLM 反復測定, 17
 - 因子の定義, 21
 - オプション, 32
 - コマンドの追加機能, 33
 - 診断, 32
 - 推定周辺平均, 32
 - その後の検定, 28
 - 表示, 32
 - プロファイル プロット, 27
 - 変数の保存, 30
 - モデル, 23
- GLOR
 - 一般的な対数線型分析, 144
- Hessian 収束 (H)
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67
- Hochberg の GT2(H)
 - GLM 多変量, 10
- Hochberg の GT2
 - GLM 反復測定, 28
- Kaplan-Meier, 162
 - 4 分位, 166
 - 因子の水準の比較, 164
 - 因子レベルの線型傾向, 164
 - コマンドの追加機能, 167
 - 作図, 166
 - 事象の定義, 164
 - 新変数の保存, 165
 - 生存状態変数, 164
 - 生命表, 166
 - 統計量, 162, 166
 - 平均値および中央値の生存関数時間, 166
 - 例, 162
- L 行列
 - 一般化推定方程式, 98

索引

- 一般化線型モデル, 71
- LaGrange 乗数検定
 - 一般化線型モデル, 72
- Levene の検定
 - GLM 多変量, 14
- Levene の検定
 - GLM 反復測定, 32
- Mauchly の球面性検定
 - GLM 反復測定, 32
- MINQUE (分散成分 オプション)
 - 分散成分, 38
- Newman-Keuls
 - GLM 多変量, 10
- Newman-Keuls
 - GLM 反復測定, 28
- Newton-Raphson 法
 - 一般的な対数線型分析, 144
 - ロジット対数線型分析, 150
- Pearson 残差
 - 一般化推定方程式, 102
 - 一般化線型モデル, 75
- R-E-G-W の F
 - GLM 多変量, 10
- R-E-G-W の Q
 - GLM 多変量, 10
- R-E-G-W の Q
 - GLM 反復測定, 28
- Ryan-Einot-Gabriel-Welsch 多重範囲
 - GLM 多変量, 10
- Ryan-Einot-Gabriel-Welsch 多重 F
 - GLM 多変量, 10
- Ryan-Einot-Gabriel-Welsch 多重 F
 - GLM 反復測定, 28
- Ryan-Einot-Gabriel-Welsch 多重範囲
 - GLM 反復測定, 28
- Scheffé の検定
 - GLM 多変量, 10
- Scheffé の検定
 - GLM 反復測定, 28
- Sidak の t 検定
 - GLM 多変量, 10
- Sidak の t 検定
 - GLM 反復測定, 28
- SSCP
 - GLM 多変量, 14
 - GLM 反復測定, 32
- Student-Newman-Keuls の検定
 - GLM 多変量, 10
- Student-Newman-Keuls の検定
 - GLM 反復測定, 28
- t 検定
 - GLM 多変量, 14
- Tamhane の T2(M)
 - GLM 多変量, 10
- amhane の T2
 - GLM 反復測定, 28
- Tarone-Ware の検定
 - Kaplan-Meier, 164
- Tukey の b 検定
 - GLM 多変量, 10
- Tukey の HSD 検定
 - GLM 多変量, 10
- Tukey の HSD 検定
 - GLM 反復測定, 28
- Tukey の b 検定
 - GLM 反復測定, 28
- Tweedie 分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- t 検定
 - GLM 反復測定, 32
- ald 統計量
 - 一般的な対数線型分析, 144
 - ロジット対数線型分析, 150
- Waller-Duncan の t 検定
 - GLM 多変量, 10
- Waller-Duncan の t 検定
 - GLM 反復測定, 28
- Wilcoxon の検定
 - 生命表, 160
- イータの 2 乗
 - GLM 反復測定, 32
- 一般化線型モデル
 - 2 値反応に関する参照カテゴリ, 62
 - カテゴリ因子に関するオプション, 64
 - 初期値, 69
- 一般化対数オッズ比
 - 一般的な対数線型分析, 144
- 一般的な対数線型分析
 - 因子, 144
 - 基準, 147
 - コマンドの追加機能, 148
 - 作図, 147
 - 残差, 148
 - 信頼区間, 147
 - セル共変量, 144
 - セル構造, 144
 - セル度数分布, 144
 - 対比, 144
 - 表示オプション, 147
 - 変数の保存, 148
 - モデル指定, 146
 - 予測値の保存, 148
- 入れ子項目
 - 線型混合モデル, 47

- 因子
 - GLM 反復測定, 21
- 因子レベル情報
 - 線型混合モデル, 51
- イータの 2 乗
 - GLM 多変量, 14
- 打ち切られたケース
 - Kaplan-Meier, 162
 - 生命表, 157
- オッズ比
 - 一般的な対数線型分析, 144
- オッズべき乗リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 重み付き予測値
 - GLM, 12
 - GLM 反復測定, 30
- 階層対数線型モデル, 139
- 階層的分解法, 6, 24
 - 分散成分, 39
- 逆ガウス分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- 参照カテゴリ
 - 一般化推定方程式, 88, 90
- 観測平均値
 - GLM 反復測定, 32
- ガンマ分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- 記述統計量
 - GLM 反復測定, 32
 - 線型混合モデル, 51
- 共分散構造, 183
 - 線型混合モデル, 183
- 共分散パラメータ検定
 - 線型混合モデル, 51
- クロス表
 - モデル選択の対数線型分析, 139
- ケース処理の要約
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 効果サイズの推定
 - GLM 反復測定, 32
- 交互作用項, 5, 24, 37, 142, 146, 153
 - 線型混合モデル, 46
- 項目の構築, 5, 24, 37, 142, 146, 153
- 固定効果
 - 線型混合モデル, 46
- 固定予測値
 - 線型混合モデル, 54
- 累積コーチットリンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 最小有意差
 - GLM 反復測定, 28
- 効果サイズの推定
 - GLM 多変量, 14
- 最尤推定法
 - 分散成分, 38
- 削除ケース残差
 - GLM, 12
 - GLM 反復測定, 30
- 作図
 - 一般的な対数線型分析, 147
 - ロジット対数線型分析, 154
- 残差
 - 一般的な対数線型分析, 148
 - 線型混合モデル, 54
 - モデル選択の対数線型分析, 142
 - ロジット対数線型分析, 155
- 残差 SSCP
 - GLM 反復測定, 32
- 残差プロット
 - GLM 反復測定, 32
- 残差分散共分散行列
 - 線型混合モデル, 51
- 参照カテゴリ
 - 一般化線型モデル, 62
- 信頼区間
 - GLM 反復測定, 32
 - 一般的な対数線型分析, 147
 - 線型混合モデル, 51
 - ロジット対数線型分析, 154
- 水準と広がり of 図
 - GLM 反復測定, 32
- 推定周辺平均
 - GLM 反復測定, 32
 - 線型混合モデル, 53
- スコアリング
 - 線型混合モデル, 50
- ストリング共変量
 - Cox 回帰分析, 170

索引

- すべての因子によるモデル
 - GLM 反復測定, 23
 - 分散成分, 37
- 正規確率プロット
 - モデル選択の対数線型分析, 142
- 制限された最尤法の推定
 - 分散成分, 38
- 生成クラス
 - モデル選択の対数線型分析, 141
- 生存分析
 - Kaplan-Meier, 162
 - 時間依存の Cox 回帰分析, 175
 - 生命表, 157
- 生命表 (生命表 オプション, 157
 - Wilcoxon (Gehan) の検定, 160
 - 因子の水準の比較, 160
 - 因子変数, 159
 - コマンドの追加機能, 160
 - 作図, 160
 - 生存状態変数, 159
 - テーブル表示の抑制, 160
 - 統計量, 157
 - ハザード比, 157
 - 累積生存関数, 157
 - 例, 157
- セグメント化した時間依存の共変量
 - Cox 回帰分析, 175
- 線型混合モデル, 42, 183
 - 共分散構造, 183
 - 交互作用項, 46
 - 項目の構築, 46-47
 - 固定効果, 46
 - コマンドの追加機能, 54
 - 推定基準, 50
 - 推定周辺平均, 53
 - 変数の保存, 54
 - 変量効果, 48
 - モデル, 51
- 相関行列
 - 線型混合モデル, 51
- 対数線型分析, 139
 - 一般的な対数線型分析, 144
 - ロジット対数線型分析, 150
- 対数尤度収束
 - 線型混合モデル, 50
- 対比
 - 一般的な対数線型分析, 144
 - ロジット対数線型分析, 150
- 段階 2 分
 - 線型混合モデル, 50
- 打ち切られたケース
 - Cox 回帰分析, 168
- てこ比の値
 - 一般化線型モデル, 75
- てこ比の値
 - GLM, 12
 - GLM 反復測定, 30
- 等分散性の検定
 - GLM 反復測定, 32
- 水準と広がり of 図
 - GLM 多変量, 14
- 特異性許容度
 - 線型混合モデル, 50
- 度数
 - モデル選択の対数線型分析, 142
- 法律に関する注意事項, 187
- 一般の推定可能関数
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 等分散性の検定
 - GLM 多変量, 14
- 反復の記述
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 負の 2 項分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- 負の 2 項リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- ハザード比
 - 生命表, 157
- ハザードの比例モデル
 - Cox 回帰分析, 168
- パラメータ分散共分散行列
 - 線型混合モデル, 51
- パラメータ収束基準
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67
- パラメータ推定値
 - 一般化推定方程式, 98
 - GLM 多変量, 14
 - 一般化線型モデル, 71
- 尺度パラメータ
 - 一般化推定方程式, 93
 - 一般化線型モデル, 67

- パラメータ収束基準
 - 線型混合モデル, 50
- パラメータ推定値
 - GLM 反復測定, 32
 - 一般的な対数線型分析, 144
 - 線型混合モデル, 51
 - モデル選択の対数線型分析, 142
 - ロジット対数線型分析, 150
- 反復回数
 - モデル選択の対数線型分析, 142
- 反復測定変数
 - 線型混合モデル, 44
- 反復の記述
 - 線型混合モデル, 50

- 被験者数
 - 線型混合モデル, 44
- 標準化されていない残差
 - GLM, 12
 - GLM 反復測定, 30
- 標準化残差
 - GLM, 12
 - GLM 反復測定, 30
- 標準誤差
 - GLM, 12
 - GLM 反復測定, 30, 32
- 標準偏差
 - GLM 反復測定, 32

- 複合モデル
 - 線型, 42
- 残差プロット
 - GLM 多変量, 14
- プロビット分析
 - 一般化線型混合モデル, 106
- プロビット リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 累積プロビット リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- プロファイル プロット
 - GLM 多変量, 8
- プロファイル プロット
 - GLM 反復測定, 27
- 分割表
 - 一般的な対数線型分析, 144
- 分散共分散行列
 - GLM, 12
 - 線型混合モデル, 51
- 分散成分, 35
 - オプション, 38
 - 結果の保存, 40
 - コマンドの追加機能, 40
 - モデル, 37
- 分散分析
 - 分散成分, 38
 - 分散分析 (分散成分 オプション)
 - GLM 反復測定, 17

- 平方和, 6, 24
 - 線型混合モデル, 47
 - 分散成分, 39
- べき乗推定値
 - GLM 多変量, 14
- べき乗推定値
 - GLM 反復測定, 32
- べき乗リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 変数減少法
 - モデル選択の対数線型分析, 139
- 変量効果
 - 線型混合モデル, 48
- 変量効果先行
 - 分散成分, 38
- 変量効果分散共分散行列
 - 線型混合モデル, 51

- ポアソン分布
 - 一般化推定方程式, 84
 - 一般化線型モデル, 58
- ポアソン回帰
 - 一般化線型混合モデル, 106
- ポアソン回帰
 - 一般的な対数線型分析, 144
- 飽和モデル
 - モデル選択の対数線型分析, 141

- マルチレベルモデル
 - 一般化線型混合モデル, 106

- 一般化線型混合モデル, 106
 - 変量効果共分散, 135
 - 推定周辺平均, 123
 - 予測対観測, 129
 - 推定平均値, 137
 - 回帰係数, 133
 - 固定効果, 114, 131
 - 変量効果, 117
 - 目標分布, 110
 - offset, 120
 - カスタム項, 115
 - データ構造, 128
 - 分類テーブル, 130

索引

- 分析の重み付け, 120
- 共分散パラメータ, 136
- フィールドの保存, 125
- 変量効果ブロック, 118
- モデル ビュー, 126
- モデルの要約, 127
- モデルのエクスポート, 125
- リンク関数, 110
- 一般化線型モデル, 56
 - 推定周辺平均, 72
 - 予測変数, 63
 - 推定基準, 67
 - 統計量, 70
 - 分布, 56
 - 応答, 61
 - アクティブなデータセットへの変数の保存, 74
 - 一般化線型混合モデル, 106
 - モデル指定, 65
 - モデルの種類, 56
 - モデルをエクスポート, 76
 - リンク関数, 56
- モデル情報
 - 一般化推定方程式, 98
 - 一般化線型モデル, 71
- 縦方向モデル
 - 一般化線型混合モデル, 106
- 混合モデル
 - 一般化線型混合モデル, 106
- 階層モデル
 - 一般化線型混合モデル, 106
- モデル ビュー
 - 一般化線型混合モデル, 126
- モデル選択の対数線型分析, 139
 - 因子範囲の定義, 140
 - オプション, 142
 - コマンドの追加機能, 143
 - モデル, 141
- ユーザーの指定によるモデル
 - GLM 反復測定, 23
 - 分散成分, 37
 - モデル選択の対数線型分析, 141
- 予測値
 - 一般的な対数線型分析, 148
 - 線型混合モデル, 54
 - ロジット対数線型分析, 155
- 対数補数リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- リンク関数
 - 一般化線型混合モデル, 110
- 同一リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 対数リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 累積生存関数
 - 生命表, 157
- 入れ子項目
 - 一般化推定方程式, 91
 - 一般化線型モデル, 65
- ログ ランク検定
 - Kaplan-Meier, 164
- 累積補ログ・マイナス・ログ関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 累積負ログ・マイナス・ログ関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 補ログ・マイナス・ログ関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- 負ログ・マイナス・ログ関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- ロジスティック回帰
 - 一般化線型混合モデル, 106
- 多項ロジスティック回帰
 - 一般化線型混合モデル, 106
- 累積ロジット リンク関数
 - 一般化推定方程式, 85
 - 一般化線型モデル, 59
- ロジット対数線型分析, 150
 - 因子, 150
 - 基準, 154
 - 作図, 154
 - 残差, 155
 - 信頼区間, 154
 - セル共変量, 150
 - セル構造, 150
 - セル度数分布, 150
 - 対比, 150
 - 表示オプション, 154
 - 変数の保存, 155
 - モデル指定, 152
 - 予測値, 155
- ロジット多項式モデル, 150

-
- ロジットリンク関数
一般化推定方程式, 85
一般化線型モデル, 59