バージョン 14

消費者調査

「真の発見の旅とは、新しい風景を探すことではなく、新たな視点を持つことである。」
マルセル・プルースト

JMP, A Business Unit of SAS
SAS Campus Drive
Cary, NC 27513

このマニュアルを引用する場合は、次の正式表記を使用してください: SAS Institute Inc. 2018. 『JMP® 14 消費者調査』Cary, NC: SAS Institute Inc.

JMP® 14 消費者調査

Copyright © 2018, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

印刷物の場合: この出版物のいかなる部分も、出版元であるSAS Institute Inc. の書面による許可なく、電子的、機械的、複写など、形式や方法を問わず、複製すること、検索システムへ格納すること、および転送することを禁止します。

Web からのダウンロードや電子本の場合: この出版物の使用については、入手した時点で、ベンダーが規定した条件が適用されます。

この出版物を、インターネットまたはその他のいかなる方法でも、出版元の許可なくスキャン、アップロード、および配布することは違法であり、法律によって罰せられます。正規の電子版のみを入手し、著作権を侵害する不正コピーに関与または加担しないでください。著作権の保護に関するご理解をお願いいたします。

米国 政府のライセンス権利、権利の制限: 本ソフトウェアとそのマニュアルは、私的な費用負担の下に開発された商業的コンピュータソフトウェアであり、米国政府に対して権利を制限した上で提供されます。米国政府による本ソフトウェアの使用、複製または開示は、該当する範囲でFAR 12.212, DFAR 227.7202-1(a)、DFAR 227.7202-3(a)、DFAR 227.7202-4に従った本合意書のライセンス条件に従うものとし、米国連邦法の下で求められる範囲において、FAR 52.227-19（2007年12月）で規定されている制限された最小限の権利に従うものとします。FAR 52.227-19が適用される場合、この条項は、その(c)項に基づく通告の役目を果たし、本ソフトウェアまたはマニュアルにその他の通告を添付する必要はありません。本ソフトウェアおよびマニュアルにおける政府の権利は、本合意書で規定されている権利に限られます。

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513-2414.

2018年3月

SAS® と、SAS Institute Inc. の他の製品名およびサービス名は、米国および他の国におけるSAS Institute Inc. の登録商標または商標です。®は、米国において登録されていることを示します。

他のブランド名および製品名は、それぞれの会社の商標です。

SASソフトウェアは、オープンソースのソフトウェアを含むがそれに限らない、特定のサードパーティ製ソフトウェアと共に提供される場合があります。かかるソフトウェアは、適用されるサードパーティソフトウェアライセンス契約に基づいてライセンスを得たものです。SASソフトウェアと共に配布されるサードパーティ製ソフトウェアに関する情報は、http://support.sas.com/thirdpartylicensesを参照してください。

テクノロジーライセンスに関する通知

- Scintilla - Copyright © 1998-2017 by Neil Hodgson <neilh@scintilla.org>.

  All Rights Reserved.

  何らかの目的でこのソフトウェアとそのマニュアルを手数料なしで使用、コピー、変更および配布することは、これをもって許可されます。ただし、すべてのコピーに上記の著作権に関する通知が記載されていること、および補助的なマニュアルに著作権に関する通知とこの許可に関する通知の両方が記載されていることを条件とします。

  NEIL HODGSONは、商業性および適合性の黙示的な保証を含め、このソフトウェアに関するすべての保証を放棄します。NEIL HODGSONは、いかなる場合においても、それが契約、過失、もしくは他の不法行為のどれであれ、このソフトウェアの使用もしくは性能から生じた、もしくはそれに関連して生じた使用、データ、もしくは利益の損失の結果として生じる特別損害、間接損害、もしくは付随的損害を始めとするいかなる損害に対しても責任を負いません。

- Telerik RadControls: Copyright © 2002-2012, Telerik. 含まれている Telerik RadControls を JMP 以外で使用することは許可されていません。
- ZLIB 圧縮ライブラリ - Copyright © 1995-2005, Jean-Loup Gailly and Mark Adler.
- Natural Earth を使用して作成。無料のベクトルおよびラスター地図データ @ naturalearthdata.com.
- パッケージ - Copyright © 2009-2010, Stéphane Sudre (s.sudre.free.fr). All rights reserved.

  ソースおよびバイナリの形で、そのまま、もしくは変更を加えて再配布および使用することは、次のような条件を満たす限り、許可されます。

  再配布するソースコードには、上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  バイナリ形式で再配布する場合は、共に提供されるマニュアルなどの資料に上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  事前に書面による許可を得ることなく、このソフトウェアから派生した製品の推奨または宣伝のために WhiteBox の名前やその貢献者の名前を使用することはできません。

  このソフトウェアは、著作権保有者および貢献者によって「現状のままで」提供され、商業性および特定の目的に対する適合性に関する黙示的な保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。いかなる場合においても、著作権保有者または貢献者は、損害の原因が何であれ、そして法的責任の根拠が何であれ、つまり、契約、厳格責任、不法行為（過失その他を含む）のどれであれ、かかる損害の発生する可能性を事前に知らされていたとしても、このソフトウェアをどのように使用して生じた損害であれ、いかなる直接損害、間接損害、付随的損害、特別損害、懲罰的損害、もしくは結果損害（代替品または代替サービスの調達、使用機会、データもしくは利益の損失、業務の中断を含むがそれに限らない）に対しても責任を負いません。

- iODBC ソフトウェア - Copyright © 1995-2006, OpenLink Software Inc and Ke Jin (www.iodbc.org). All rights reserved.

  ソースおよびバイナリの形で、そのまま、もしくは変更を加えて再配布および使用することは、次のような条件を満たす限り、許可されます。

  – 再配布するソースコードには、上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  – バイナリ形式で再配布する場合は、共に提供されるマニュアルなどの資料に上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  – 事前に書面による許可を得ることなく、このソフトウェアから派生した製品の推奨または宣伝のために OpenLink Software Inc. の名前やその貢献者の名前を使用することはできません。

  このソフトウェアは、著作権保有者および貢献者によって「現状のままで」提供され、商業性および特定の目的に対する適合性に関する黙示的な保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。いかなる場合においても、OPENLINK または貢献者は、損害の原因が何であれ、そして法的責任の根拠が何であれ、つまり、契約、厳格責任、不法行為（過失その他を含む）のどれであれ、かかる損害の発生する可能性を事前に知らされていたとしても、このソフトウェアをどのように使用して生じた損害であれ、いかなる直接損害、間接損害、付随的損害、特別損害、懲罰的損害、もしくは結果損害（代替品または代替サービスの調達、使用機会、データもしくは利益の損失、業務の中断を含むがそれに限らない）に対しても責任を負いません。

- bzip2、関連ライブラリの「libbzip2」、およびすべてのマニュアル : Copyright © 1996-2010, Julian R Seward. All rights reserved.

  ソースおよびバイナリの形で、そのまま、もしくは変更を加えて再配布および使用することは、次のような条件を満たす限り、許可されます。

  再配布するソースコードには、上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  このソフトウェアの供給源は正しく表記しなければならず、使用者が元のソフトウェアを記述したと主張することはできません。ある製品の中でこのソフトウェアを使用する場合は、その製品のマニュアルに謝辞を記載してもらえるとありがたいですが、必須ではありません。

  ソースに変更を加えたバージョンには、その旨を明記しなければならず、元のソフトウェアとは違うものであることを明確にしてください。

  事前に書面による許可を得ることなく、このソフトウェアから派生した製品の推奨または宣伝のために作成者の名前を使用することはできません。

  このソフトウェアは、作成者によって「現状のままで」提供され、商業性および特定の目的に対する適合性に関する黙示的な保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。いかなる場合においても、作成者は、損害の原因が何であれ、そして法的責任の根拠が何であ

れ、つまり、契約、厳格責任、不法行為（過失その他を含む）のどれであれ、かかる損害の発生する可能性を事前に知らされていたとしても、このソフトウェアをどのように使用して生じた損害であれ、いかなる直接損害、間接損害、付随的損害、特別損害、懲罰的損害、もしくは結果損害（代替品または代替サービスの調達、使用機会、データもしくは利益の損失、業務の中断を含むがそれに限らない）に対しても責任を負いません。

- R ソフトウェア : Copyright © 1999-2012, R Foundation for Statistical Computing.
- MATLAB ソフトウェア : Copyright © 1984-2012, The MathWorks, Inc. は米国特許法および国際特許法によって保護されています。www.mathworks.com/patents を参照してください。MATLAB および Simulink は、The MathWorks, Inc. の登録商標です。
  他の商標については、www.mathworks.com/trademarks を参照してください。他の製品名やブランド名は、それぞれの所有者の商標または登録商標である可能性があります。
- libopc: Copyright © 2011, Florian Reuter. All rights reserved.

  ソースおよびバイナリの形で、そのまま、もしくは変更を加えて再配布および使用することは、次のような条件を満たす限り、許可されます。

  – 再配布するソースコードには、上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。
  – バイナリ形式で再配布する場合は、共に提供されるマニュアルなどの資料に上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。
  – 事前に書面による許可を得ることなく、このソフトウェアから派生した製品の推奨または宣伝のために Florian Reuter の名前やその貢献者の名前を使用することはできません。

  このソフトウェアは、著作権保有者および貢献者によって「現状のままで」提供され、商業性および特定の目的に対する適合性に関する黙示的な保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。いかなる場合においても、著作権保有者または貢献者は、損害の原因が何であれ、そして法的責任の根拠が何であれ、つまり、契約、厳格責任、不法行為（過失その他を含む）のどれであれ、かかる損害の発生する可能性を事前に知らされていたとしても、このソフトウェアをどのように使用して生じた損害であれ、いかなる直接損害、間接損害、付随的損害、特別損害、懲罰的損害、もしくは結果損害（代替品または代替サービスの調達、使用機会、データもしくは利益の損失、業務の中断を含むがそれに限らない）に対しても責任を負いません。
- libxml2 - ソースコードに特に記載がある場合を除く（たとえば、使用しているライセンスは類似しているが、著作権の通知が異なる hash.c、list.c ファイルや trio ファイル）、すべてのファイル :

  Copyright © 1998 - 2003 Daniel Veillard. All Rights Reserved.

  これをもって、このソフトウェアのコピーと関連する文書ファイル（「本ソフトウェア」）を入手した人すべてに対し、無料で本ソフトウェアを使用、コピー、変更、マージ、パブリッシュ、配布、サブライセンスする、もしくはコピーを販売する権利を含むがそれに限定せず、本ソフトウェアを制限なく取り

扱う権利、および本ソフトウェアの供給相手に対してそうすることを許可する権利が付与されます。ただし、以下の条件を満たさなければなりません。

上記の著作権に関する通知とこの許可に関する通知が、本ソフトウェアのコピーのすべてまたは大部分に記載されていること。

このソフトウェアは、 現状のままで」提供され、商業性および特定の目的に対する適合性、および非侵害の保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。DANIEL VEILLARDは、いかなる場合においても、それが契約、過失、もしくは他の不法行為のどれであれ、本ソフトウェアから、もしくは本ソフトウェアに関連して、または本ソフトウェアの使用もしくは他の取り扱いに関連して生じた申し立て、損害賠償もしくは他の義務に対し、責任を負いません。

この通知に含まれているものを除き、Daniel Veillardから事前に書面による許可を得ることなく、本ソフトウェアの広告、またはその他の手段による本ソフトウェアの販売、使用もしくは他の取り扱いの宣伝にDaniel Veillardの名前を使用することはできません。

- UNIXファイルに使用された解凍アルゴリズムについて:

  Copyright © 1985, 1986, 1992, 1993

  カリフォルニア大学評議員。All rights reserved.

  このソフトウェアは、評議員および貢献者によって「現状のままで」提供され、商業性および特定の目的に対する適合性に関する黙示的な保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。いかなる場合においても、評議員または貢献者は、損害の原因が何であれ、そして法的責任の根拠が何であれ、つまり、契約、厳格責任、不法行為（過失その他を含む）のどれであれ、かかる損害の発生する可能性を事前に知らされていたとしても、このソフトウェアをどのように使用して生じた損害であれ、いかなる直接損害、間接損害、付随的損害、特別損害、懲罰的損害、もしくは結果損害（代替品または代替サービスの調達、使用機会、データもしくは利益の損失、業務の中断を含むがそれに限らない）に対しても責任を負いません。

  1. 再配布するソースコードには、上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  2. バイナリ形式で再配布する場合は、共に提供されるマニュアルなどの資料に上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

  3. 事前に書面による許可を得ることなく、このソフトウェアから派生した製品の推奨または宣伝のために大学の名前や貢献者の名前を使用することはできません。

- Snowball - Copyright © 2001, Dr Martin Porter, Copyright © 2002, Richard Boulton.

  All rights reserved.

  ソースおよびバイナリの形で、そのまま、もしくは変更を加えて再配布および使用することは、次のような条件を満たす限り、許可されます。

1. 再配布するソースコードには、上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

2. バイナリ形式で再配布する場合は、共に提供されるマニュアルなどの資料に上記の著作権に関する通知、この条件リスト、これに続く放棄声明が記載されていなければなりません。

3. 事前に書面による許可を得ることなく、このソフトウェアから派生した製品の推奨または宣伝のために著作権保有者の名前や貢献者の名前を使用することはできません。

このソフトウェアは、著作権保有者および貢献者によって「現状のままで」提供され、商業性および特定の目的に対する適合性に関する黙示的な保証を含むがそれに限らない、いかなる明示的もしくは黙示的な保証も行われません。いかなる場合においても、著作権保有者または貢献者は、損害の原因が何であれ、そして法的責任の根拠が何であれ、つまり、契約、厳格責任、不法行為（過失その他を含む）のどれであれ、かかる損害の発生する可能性を事前に知らされていたとしても、このソフトウェアをどのように使用して生じた損害であれ、いかなる直接損害、間接損害、付随的損害、特別損害、懲罰的損害、もしくは結果損害（代替品または代替サービスの調達、使用機会、データもしくは利益の損失、業務の中断を含むがそれに限らない）に対しても責任を負いません。

目次
消費者調査

1 JMP の概要 13

マニュアルとその他のリソース

表記規則 15
JMP のマニュアル 15
JMP ドキュメンテーションライブラリ 16
JMP ヘルプ 22
JMP を習得するためのその他のリソース 22
チュートリアル 22
サンプルデータテーブル 22
統計用語と JSL 用語の習得 23
JMP を使用するためのヒント 23
ツールヒント 23
JMP User Community 24
JMP 関連書籍 24
「JMP スターター」ウィンドウ 24
テクニカルサポート 24

2 消費者調査について 25

消費者調査データに対する分析の概要

3 カテゴリカルな応答の分析 27

アンケート調査データなどの度数表

「カテゴリカル」プラットフォームの使用例 29
「カテゴリカル」プラットフォームの起動 30
応答の役割 30
列の役割 32
起動ウィンドウのその他のオプション 33
「カテゴリカル」レポート 34
「カテゴリカル」プラットフォームのオプション 35
レポートに関するオプション 35
統計的検定に関するオプション 36
「カテゴリカル」プラットフォームのその他のオプション 38

クロス表の応答におけるオプション 41
比較を表すアルファベット文字 41
上位カテゴリ 43
［上位カテゴリ］のオプション 43
環境設定の変更 44
「カテゴリカル」プラットフォームの別例 45
応答の等質性に対する検定の例 45
多重応答（複数回答）に対する検定の例 46
セルのカイ2乗検定の例 48
標本の各ペア比較の例 49
セルの各ペア比較の例 50
分析者が興味のある比較を指定する例 52
条件付き関連と相対リスクの例 53
判定の一致性の例 56
反復測定の例 57
［多重応答］タブの例 58
スコア平均を比較する例 63
「表の構成」レポートの例 65
「カテゴリカル」プラットフォームの統計的詳細 66
Rao-Scott 修正 66

4 選択モデル 69

選択モデルのあてはめ

「選択モデル」プラットフォームの概要 71
「選択モデル」プラットフォームの例 72
「選択なし」がある1つのデータテーブル 73
複数のデータテーブル 75
「選択モデル」プラットフォームの起動 81
［1つのデータテーブル, 積み重ね］の起動ウィンドウ 82
［複数のデータテーブル, 相互参照］の起動ウィンドウ 84
「選択モデル」レポート 89
効果の要約 89
パラメータ推定値 90
尤度比検定 91
Bayes パラメータ推定値 91
「選択モデル」プラットフォームのオプション 93
支払意思額 96
「選択モデル」の別例 98
製品の仕様を決める分析例 98
セグメント化の例 107

「選択モデル」プラットフォームを使用したロジスティック回帰の例 111
対応のあるデータに対する条件付きロジスティック回帰の例 114
データを2つの分析テーブルに変換する例 116
データを1つの分析テーブルに変換する例 121
「選択モデル」プラットフォームの統計的詳細 124
分析対象のデータに対する特殊な規則 124
効用と確率 125
勾配 125

5 MaxDiff 127

MaxDiff モデルのあてはめ

「MaxDiff モデル」プラットフォームの概要 129
「MaxDiff」プラットフォームの例 129
1つのデータテーブル 130
複数のデータテーブル 133
「MaxDiff」プラットフォームの起動 136
［1つのデータテーブル, 積み重ね］の起動ウィンドウ 137
［複数のデータテーブル, 相互参照］の起動ウィンドウ 138
「MaxDiff モデル」レポート 143
効果の要約 143
MaxDiff 結果 144
パラメータ推定値 145
Bayes パラメータ推定値 147
尤度比検定 148
「MaxDiff」プラットフォームのオプション 148
「比較」レポート 150
Bayes チェーンの保存 151

6 アップリフトモデル 153

施策が消費者行動に及ぼす効果をモデル化

「アップリフト」プラットフォームの概要 155
「アップリフト」プラットフォームの使用例 156
「アップリフト」プラットフォームの起動 157
「アップリフトモデル」レポート 158
アップリフトモデルのグラフ 159
アップリフトレポートのオプション 161

7 多重因子分析 163
複数の回答者における回答の類似性を視覚化する
「多重因子分析」プラットフォームの概要 165
多重因子分析の例 166
「多重因子分析」プラットフォームの起動 167
データ形式 169
「多重因子分析」レポート 170
要約プロット 170
コンセンサスマップ 171
「多重因子分析」プラットフォームのオプション 171
「多重因子分析」プラットフォームの統計的詳細 174
A 参考文献 175
索引 177
消費者調査

第 1 章

JMP の概要

マニュアルとその他のリソース

この章には表記規則、各JMPドキュメンテーションの説明、ヘルプシステムなど、JMPマニュアルの詳細と、他のサポートの記載場所が掲載されています。

目次

表記規則 15
JMP のマニュアル 15
JMP ドキュメンテーションライブラリ 16
JMP ヘルプ 22
JMP を習得するためのその他のリソース 22
チュートリアル 22
サンプルデータテーブル 22
統計用語と JSL 用語の習得 23
JMP を使用するためのヒント 23
ツールヒント 23
JMP User Community 24
JMP 関連書籍 24
「JMP スターター」ウィンドウ 24
テクニカルサポート 24

表記規則

マニュアルの内容と画面に表示される情報を対応付けるために、次のような表記規則を使っています。

- サンプルデータ名、列名、パス名、ファイル名、ファイル拡張子、およびフォルダ名は「」で囲んで表記しています。
- スクリプトのコードは`Lucida Sans Typewriter`フォントで表記しています。
- スクリプトコードの結果（ログに表示されるもの）は`Lucida Sans Typewriter`フォントで表記し、先に示すコードよりインデントされています。
- クリックまたは選択する項目は［］で囲んで太字で表記しています。これには以下の項目があります。
  - ボタン
  - チェックボックス
  - コマンド
  - 選択可能なリスト項目
  - メニュー
  - オプション
  - タブ名
  - テキストボックス
- 次の項目の表記規則は下記のとおりです。
  - 重要な単語や句、JMPに固有の定義を持つ単語や句は太字または「」で囲んで表記
  - マニュアルのタイトルは『』で囲んで表記
  - 変数名は「」で囲んで太字で表記
- JMP Proのみの機能にはJMP Proアイコン JMP PRO がついています。JMP Proの機能の概要については https://www.jmp.com/ja_jp/software/predictive-analytics-software.html をご覧ください。

メモ：特別な情報および制限事項には、この文のように「メモ」という見出しがついています。

ヒント：役に立つ情報には「ヒント」という見出しがついています。

JMPのマニュアル

JMPでは、PDF形式のマニュアルが用意されています。

- PDF版は［**ヘルプ**］>［**ドキュメンテーション**］メニューまたはJMPオンラインヘルプのフッタから開くことができます。

- 検索しやすいようにすべてのドキュメンテーションが1つのPDFファイルにまとめられた『JMPドキュメンテーションライブラリ』と呼ばれるファイルがあります。『JMPドキュメンテーションライブラリ』のPDFファイルは［**ヘルプ**］>［**ドキュメンテーション**］メニューから開くことができます。

JMPドキュメンテーションライブラリ

以下の表は、JMPライブラリに含まれている各ドキュメンテーションの目的および内容をまとめたものです。

マニュアル	目的	内容
『はじめてのJMP』	JMPをあまりご存知ない方を対象とした入門ガイド	JMPの紹介と、データを作成および分析し始めるための情報や、結果の共有方法についても学びます。
『JMPの使用法』	JMPのデータテーブルと、基本操作を理解する	一般的なJMPの概念と、データの読み込み、列プロパティの変更、データの並べ替え、SASへの接続など、JMP全体にわたる機能の説明
『基本的な統計分析』	このマニュアルを見ながら、基本的な分析を行う	［分析］メニューからアクセスできる以下のプラットフォームの説明： • 一変量の分布 • 二変量の関係 • 表の作成 • テキストエクスプローラ ［分析］>［二変量の関係］で二変量分析、一元配置分散分析、分割表分析を実行する方法の説明。ブートストラップを使用した標本分布の近似やシミュレーションの機能を使用したパラメトリックな標本再抽出の説明も含まれています。

マニュアル	目的	内容
『グラフ機能』	データに合った理想的なグラフを見つける	［グラフ］メニューからアクセスできる以下のプラットフォームの説明： • グラフビルダー • 三次元散布図 • 等高線図 • バブルプロット • パラレルプロット • セルプロット • 散布図行列 • 三角図 • ツリーマップ • チャート • 重ね合わせプロット このマニュアルには背景マップやカスタムマップの作成方法も記載されています。
『プロファイル機能』	プロファイルの使い方を学ぶ。任意の応答曲面の断面を表示できるようになります。	［グラフ］メニューに表示されるすべてのプロファイルについて。誤差因子（ランダムな入力値）がある状況のシミュレーションについても説明されています。
『実験計画(DOE)』	実験計画と標本サイズ設計を学ぶ	［実験計画（DOE）］メニューと［分析］>［発展的なモデル］メニューの「発展的な実験計画モデル」に関するすべてのトピックについて。

マニュアル	目的	内容
『基本的な回帰モデル』	「モデルのあてはめ」プラットフォームとその多くの手法について学ぶ	[分析] メニューの「モデルのあてはめ」プラットフォームで使用できる、以下の手法の説明： • 標準最小2乗 • ステップワイズ • 一般化回帰 • 混合モデル • MANOVA • 対数線形-分散 • 名義ロジスティック • 順序ロジスティック • 一般化線形モデル

マニュアル	目的	内容
『予測モデルおよび発展的なモデル』	さらなるモデリング手法について学ぶ	［分析］>［予測モデル］メニューで使用できる以下のプラットフォームの説明： • モデル化ユーティリティ • ニューラル • パーティション • ブートストラップ森 • ブースティングツリー • K近傍法 • 単純Bayes • モデルの比較 • 計算式デポ ［分析］>［発展的なモデル］メニューで使用できる以下のプラットフォームの説明： • 曲線のあてはめ • 非線形回帰 • 関数データエクスプローラ • Gauss過程 • 時系列分析 • 対応のあるペア ［分析］>［スクリーニング］メニューで使用できる以下のプラットフォームの説明： • 応答のスクリーニング • 工程のスクリーニング • 説明変数のスクリーニング • アソシエーション分析 • プロセス履歴エクスプローラ ［分析］>［発展的なモデル］>［発展的な実験計画モデル］で使用できるプラットフォームについては、『実験計画(DOE)』に説明があります。

マニュアル	目的	内容
『多変量分析』	複数の変数を同時に分析するための手法について理解を深める	［分析］>［多変量］メニューで使用できる以下のプラットフォームの説明： • 多変量の相関 • 主成分分析 • 判別分析 • PLS • 多重対応分析 • 因子分析 • 多次元尺度構成 • 項目分析 ［分析］>［クラスター分析］メニューで使用できる以下のプラットフォームの説明： • 階層型クラスター分析 • K Means クラスター分析 • 正規混合 • 潜在クラス分析 • 変数のクラスタリング
『品質と工程』	工程を評価し、向上させるためのツールについて理解を深める	［分析］>［品質と工程］メニューで使用できる以下のプラットフォームの説明： • 管理図ビルダーと個々の管理図 • 測定システム分析 • 計量値/計数値ゲージチャート • 工程能力 • パレート図 • 特性要因図 • 仕様限界の管理

マニュアル	目的	内容
『信頼性/生存時間分析』	製品やシステムにおける信頼性を評価し、向上させる方法、および人や製品の生存時間データを分析する方法について学ぶ	［分析］>［信頼性/生存時間分析］メニューで使用できる以下のプラットフォームの説明： • 寿命の一変量 • 寿命の二変量 • 累積損傷 • 再生モデルによる分析 • 劣化分析と破壊劣化 • 信頼性予測 • 信頼性成長 • 信頼性ブロック図 • 修理可能システムのシミュレーション • 生存時間分析 • 生存時間(パラメトリック)のあてはめ • 比例ハザードのあてはめ
『消費者調査』	消費者選好を調査し、その洞察を使用してより良い製品やサービスを作成するための方法を学ぶ	［分析］>［消費者調査］メニューで使用できる以下のプラットフォームの説明： • カテゴリカル • 選択モデル • MaxDiff • アップリフト • 多重因子分析
『スクリプトガイド』	パワフルなJMPスクリプト言語（JSL）の活用方法について学ぶ	スクリプトの作成やデバッグ、データテーブルの操作、ディスプレイボックスの構築、JMPアプリケーションの作成など。
『スクリプト構文リファレンス』	JSL関数、その引数、およびオブジェクトやディスプレイボックスに送信するメッセージについて理解を深める	JSLコマンドの構文、例、および注意書き。

メモ：［ドキュメンテーション］メニューでは、印刷可能な2つのリファレンスカードも用意されています。『メニューカード』はJMPのメニューをまとめた表で、『クイックリファレンス』はJMPのショートカットキーをまとめた表です。

JMPヘルプ

JMPヘルプは、一連のマニュアルの簡易版です。JMPのヘルプは、次のいくつかの方法で開くことができます。

- Windowsでは、F1キーを押すとヘルプシステムウィンドウが開きます。
- データテーブルまたはレポートウィンドウの特定の部分のヘルプを表示します。［**ツール**］メニューからヘルプツール を選択した後、データテーブルやレポートウィンドウの任意の位置でクリックすると、その部分に関するヘルプが表示されます。
- JMPウィンドウ内で［**ヘルプ**］ボタンをクリックします。
- Windowsの場合、［ヘルプ］メニューの［**ヘルプの目次**］、［**ヘルプの検索**］、［**ヘルプの索引**］の各オプションを使用して、JMPヘルプ内を検索し、目的の内容を表示します。Macの場合、［**ヘルプ**］>［**JMPヘルプ**］を選択します。

JMPを習得するためのその他のリソース

JMPのマニュアルとJMPヘルプの他、次のリソースもJMPの学習に役立ちます。

- 「チュートリアル」
- 「サンプルデータテーブル」
- 「統計用語とJSL用語の習得」
- 「JMPを使用するためのヒント」
- 「ツールヒント」
- 「JMP User Community」
- 「JMP関連書籍」
- 「「JMPスターター」ウィンドウ」

チュートリアル

［**ヘルプ**］>［**チュートリアル**］を選択して、JMPのチュートリアルを表示できます。［**チュートリアル**］メニューの最初の項目は［**チュートリアルディレクトリ**］です。この項目を選択すると、すべてのチュートリアルをカテゴリ別に整理した新しいウィンドウが開きます。

JMPに慣れていない方は、まず［**初心者用チュートリアル**］を試してみてください。JMPのインターフェースおよび基本的な使用方法を学ぶことができます。

他のチュートリアルでは、実験の計画、標本平均と定数の比較など、JMPの具体的な活用法を学習できます。

サンプルデータテーブル

JMPのマニュアルで取り上げる例は、すべてサンプルデータを使用しています。サンプルデータディレクトリを開くには、［**ヘルプ**］>［**サンプルデータライブラリ**］を選択します。

サンプルデータテーブルを文字コード順に並べた一覧を表示する、またはカテゴリごとにサンプルデータを表示するには、［**ヘルプ**］>［**サンプルデータ**］を選択します。

サンプルデータテーブルは次のディレクトリにインストールされています。

Windowsの場合: C:¥Program Files¥SAS¥JMP¥14¥Samples¥Data

Macintoshの場合: \Library\Application Support\JMP\14\Samples\Data

JMP Proでは、サンプルデータが（JMPではなく）JMPPROディレクトリにインストールされています。シングルユーザーライセンス版のJMP（JMPシュリンクラップ）では、サンプルデータがJMPSWディレクトリにインストールされています。

サンプルデータの使用例を参照するには、［**ヘルプ**］>［**サンプルデータ**］を選択し、教育用セクションから検索してください。教育用リソースについては、http://jmp.com/toolsにも情報があります。

統計用語とJSL用語の習得

［**ヘルプ**］メニューには、次の索引が用意されています。

統計の索引　統計用語が説明されています。

スクリプトの索引　JSL関数、オブジェクト、ディスプレイボックスに関する情報を検索できます。スクリプトの索引からサンプルスクリプトを編集して実行することもできます。

JMPを使用するためのヒント

JMPを最初に起動すると、「使い方ヒント」ウィンドウが表示されます。このウィンドウには、JMPを使う上でのヒントが表示されます。

「使い方ヒント」ウィンドウを表示しないようにするには、［**起動時にヒントを表示する**］のチェックを外します。再表示するには、［**ヘルプ**］>［**使い方ヒント**］を選択します。または、「環境設定」ウィンドウで非表示に設定することもできます。

ツールヒント

次のような項目の上にカーソルを置くと、その項目を説明するツールヒントが表示されます。

- メニューまたはツールバーのオプション
- グラフ内のラベル
- レポートウィンドウ内の結果（テキスト）（カーソルで円を描くと表示される）
- 「ホームウィンドウ」内のファイル名またはウィンドウ名
- スクリプトエディタ内のコード

ヒント： Windowsでは、JMP環境設定でツールヒントを表示しないよう設定できます。［**ファイル**］>［**環境設定**］>［**一般**］を選択し、［**メニューのヒントを表示**］の選択を解除します。このオプションは、Macintoshでは使用できません。

JMP User Community

JMP User Communityでは、さまざまな方法でJMPをさらに習得したり、他のSASユーザとのコミュニケーションを図ったりできます。ラーニングライブラリには1ページガイド、チュートリアル、デモなどが用意されており、JMPを使い始める上でとても便利です。また、JMPのさまざまなトレーニングコースに登録して、自己教育を進めることも可能です。

その他のリソースとして、ディスカッションフォーラム、サンプルデータやスクリプトファイルの交換、Webcastセミナー、ソーシャルネットワークグループなども利用できます。

WebサイトのJMPリソースにアクセスするには、［**ヘルプ**］>［**JMP User Community**］を選択するか、https://community.jmp.com/をご覧ください。

JMP関連書籍

JMP関連書籍は、次のJMP Webページで紹介されています。

http://www.jmp.com/ja_jp/academic/books-for-jmp-users.html

「JMPスターター」ウィンドウ

JMPまたはデータ分析にあまり慣れていないユーザは、「JMPスターター」ウィンドウから開始するとよいでしょう。カテゴリ分けされた項目には説明がついており、ボタンをクリックするだけで該当の機能を起動できます。「JMPスターター」ウィンドウには、［分析］、［グラフ］、［テーブル］、および［ファイル］メニュー内の多くのオプションがあります。また、JMP Proの機能やプラットフォームのリストも含まれています。

- 「JMPスターター」ウィンドウを開くには、［**表示**］（Macintoshでは［**ウィンドウ**］）>［**JMPスターター**］を選択します。
- WindowsでJMPの起動時に自動的に「JMPスターター」を表示するには、［**ファイル**］>［**環境設定**］>［**一般**］を選び、「開始時のJMPウィンドウ」リストから［**JMPスターター**］を選択します。Macintoshでは、［**JMP**］>［**環境設定**］>［**起動時にJMPスターターウィンドウを表示する**］を選択します。

テクニカルサポート

JMPのテクニカルサポートは、JMPのエンジニアが担当し、その多くは、統計学などの技術的な分野の知識を有しています。

http://www.jmp.com/japan/support には、テクニカルサポートへの連絡方法などが記載されています。

第 2 章

消費者調査について

消費者調査データに対する分析の概要

JMP では、消費者調査のデータを分析する機能が備わっています。「どのように製品やサービスを消費者が利用しているか？」、「顧客の満足度はどのくらいか？」、また、「彼らはどんな新機能を望んでいるか？」などを、企業が調べるときがあります。そして、その結果にもとづき、製品やサービスの品質を良くし、顧客の満足度を向上させています。このような消費者調査を分析するプラットフォームが、［消費者調査］メニューに用意されています。以下のようなプラットフォームが用意されています。

- 「カテゴリカル」プラットフォームは、カテゴリカルデータに対して、度数表の作成、グラフの作成、度数の比較を行います。多重応答（multiple responses; 複数回答）も扱えます。アンケート調査の回答、工業製品の不適合原因や故障原因、調査参加者のデモグラフィック情報など、さまざまな分野のカテゴリカルデータを分析できます。「カテゴリカル」プラットフォームでは、さまざまな形式のデータを分析できます。詳細については、第 3 章「カテゴリカルな応答の分析」を参照してください。
- 「選択モデル」は、市場調査（マーケットリサーチ）の実験データを分析するためのプラットフォームです。消費者の選好構造を明らかにし、消費者に好まれる製品やサービスを企画することができます。詳細については、第 4 章「選択モデル」を参照してください。
- 「MaxDiff」プラットフォームは、各商品に対する好ましさに関する質問から、それらの商品の相対的な重要度を算出します。MaxDiff 法では、標準的な調査方法とは異なり、提示した選択肢を回答者に選んでもらうときに、最も好きな選択肢と最も嫌いな選択肢の 2 つを選んでもらいます。詳細については、第 5 章「MaxDiff」を参照してください。
- 「アップリフト」プラットフォームは、ダイレクトメールを送付したときのレスポンスが、送付しなかったときのレスポンスに比べて大きい集団を、探し出すのに役立ちます。この分析は、性別や年齢などの考慮すべき背景因子が多く、データ量が比較的大きい場合でも処理できます。マーケティング分野だけではなく、テイラーメイド治療や個別化医療に応用することも考えられます。詳細については、第 6 章「アップリフトモデル」を参照してください。
- 「多重因子分析」プラットフォームでは、官能検査データを分析し、検査者間の類似性を調べることができます。多重因子分析では、測定する項目が異なっていたり、測定装置・回答者・環境状況が異なっていたりしているデータを分析します。詳細については、第 7 章「多重因子分析」を参照してください。

第 3 章

カテゴリカルな応答の分析

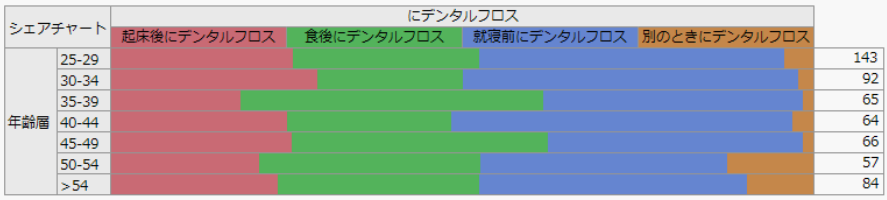
アンケート調査データなどの度数表

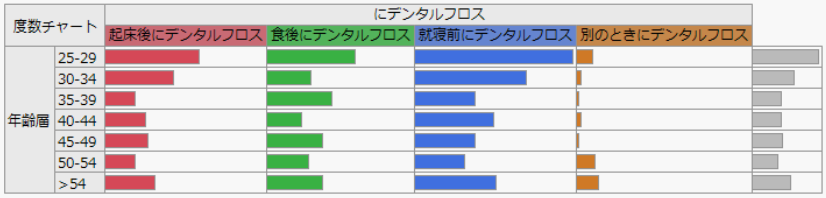
「カテゴリカル」プラットフォームは、カテゴリカルデータに対して、度数表の作成、グラフの作成、度数の比較を行います。多重応答（multiple responses; 複数回答）も扱えます。アンケート調査の回答、工業製品の不適合原因や故障原因、調査参加者のデモグラフィック情報など、さまざまな分野のカテゴリカルデータを分析できます。「カテゴリカル」プラットフォームでは、さまざまな形式のデータを分析でき、起動ウィンドウにおいてその形式を指定できます。

図 3.1 カテゴリカルなデータの分析例

指示変数(起床後にデンタルフロス, 食後にデンタルフロス, 就寝前にデンタルフロス, 別のときにデンタルフロス) By Age Group

度数		年齢層						
ケース		25-29	30-34	35-39	40-44	45-49	50-54	>54
にデンタルフロス	起床後にデンタルフロス	37	27	12	16	17	12	20
		32.7%	39.7%	24.5%	30.8%	32.7%	27.9%	28.2%
	食後にデンタルフロス	38	19	28	15	24	18	24
		33.6%	27.9%	57.1%	28.8%	46.2%	41.9%	33.8%
	就寝前にデンタルフロス	62	44	24	31	24	20	32
		54.9%	64.7%	49.0%	59.6%	46.2%	46.5%	45.1%
	別のときにデンタルフロス	6	2	1	2	1	7	8
		5.3%	2.9%	2.0%	3.8%	1.9%	16.3%	11.3%
	応答の合計数	143	92	65	64	66	57	84
	全ケース数	113	68	49	52	52	43	71
	応答している全ケース数	104	67	46	51	50	43	70





各応答に対するPoisson検定

年齢層, にデンタルフロス

にデンタルフロス	カイ2乗	p値(Prob>ChiSq)
起床後にデンタルフロス	2.6995	0.8455
食後にデンタルフロス	9.1746	0.1640
就寝前にデンタルフロス	4.0051	0.6760
別のときにデンタルフロス	13.2863	0.0387*

検定はPoisson分布を仮定しています。

目次

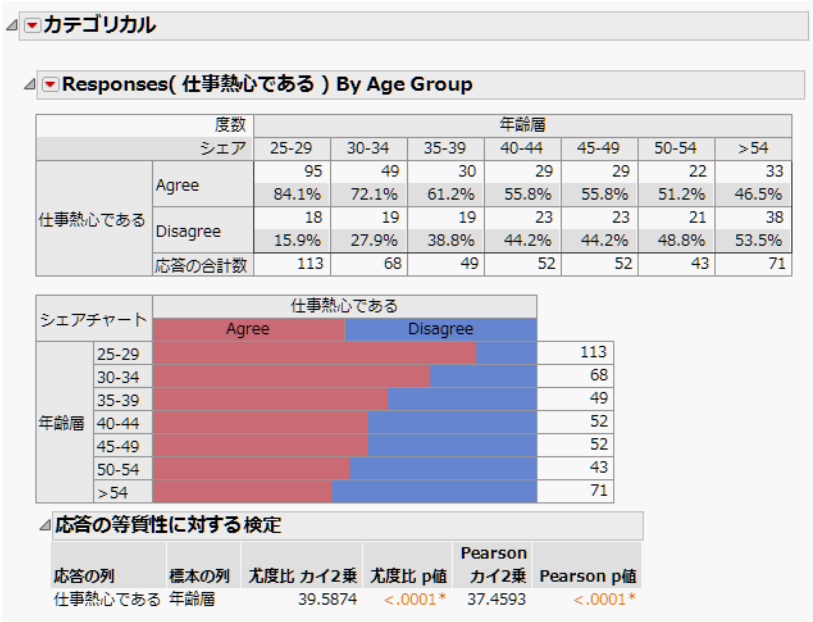
「カテゴリカル」プラットフォームの使用例 29
「カテゴリカル」プラットフォームの起動 30
応答の役割 30
列の役割 32
起動ウィンドウのその他のオプション 33
「カテゴリカル」レポート 34
「カテゴリカル」プラットフォームのオプション 35
レポートに関するオプション 35
統計的検定に関するオプション 36
「カテゴリカル」プラットフォームのその他のオプション 38
クロス表の応答におけるオプション 41
比較を表すアルファベット文字 41
上位カテゴリ 43
［上位カテゴリ］のオプション 43
環境設定の変更 44
「カテゴリカル」プラットフォームの別例 45
応答の等質性に対する検定の例 45
多重応答（複数回答）に対する検定の例 46
セルのカイ2乗検定の例 48
標本の各ペア比較の例 49
セルの各ペア比較の例 50
分析者が興味のある比較を指定する例 52
条件付き関連と相対リスクの例 53
判定の一致性の例 56
反復測定の例 57
［多重応答］タブの例 58
スコア平均を比較する例 63
「表の構成」レポートの例 65
「カテゴリカル」プラットフォームの統計的詳細 66
Rao-Scott 修正 66

「カテゴリカル」プラットフォームの使用例

この例では「Consumer Preferences.jmp」を使用します。このデータには、一般的な事柄に対する意見や態度、口腔衛生に関する質問への回答が含まれています。「カテゴリカル」プラットフォームを使うと、年齢層別に回答を比較することができます。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「仕事熱心である」**を選択して、［単純］タブの［**応答**］をクリックします。
4. **「年齢層」**を選択して［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「カテゴリカル」の赤い三角ボタンをクリックして、［**クロス表 転置**］を選択します。
7. 「カテゴリカル」の赤い三角ボタンをクリックして、［**応答の等質性に対する検定**］を選択します。

図3.2 質問に対する年齢層別の応答



カテゴリカル

Responses(仕事熱心である) By Age Group

度数		年齢層						
シェア		25-29	30-34	35-39	40-44	45-49	50-54	>54
仕事熱心である	Agree	95	49	30	29	29	22	33
		84.1%	72.1%	61.2%	55.8%	55.8%	51.2%	46.5%
	Disagree	18	19	19	23	23	21	38
		15.9%	27.9%	38.8%	44.2%	44.2%	48.8%	53.5%
	応答の合計数	113	68	49	52	52	43	71

応答の等質性に対する検定

応答の列	標本の列	尤度比 カイ2乗	尤度比 p値	Pearson カイ2乗	Pearson p値
仕事熱心である	年齢層	39.5874	<.0001*	37.4593	<.0001*

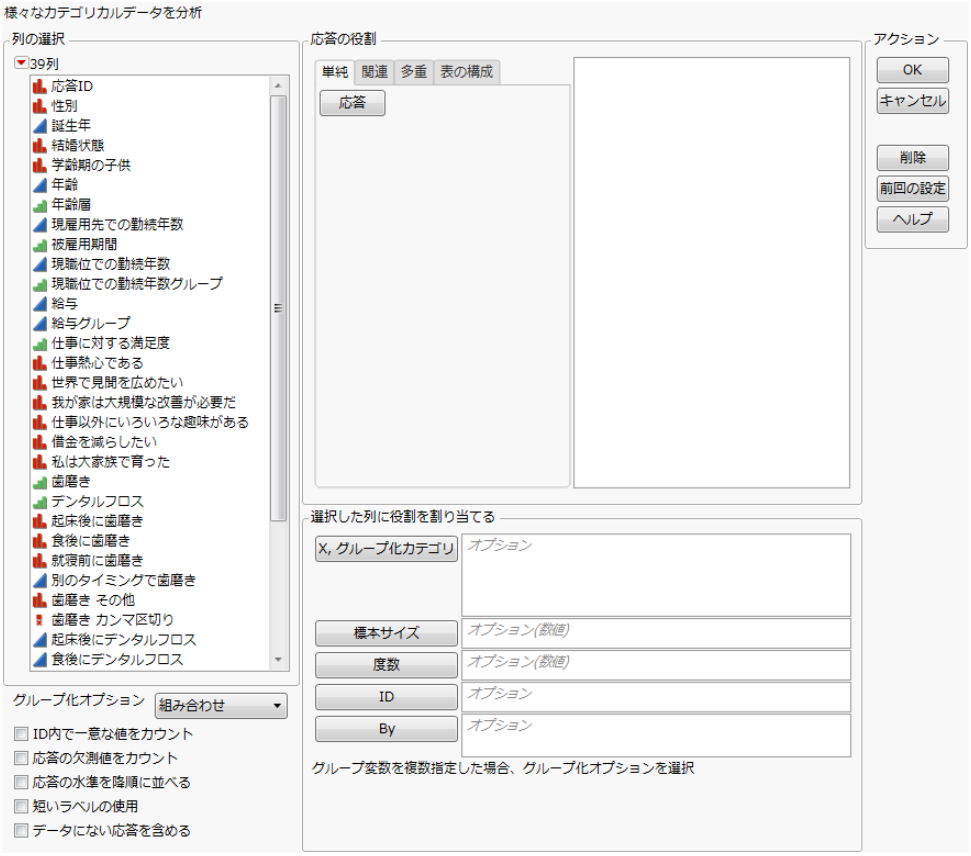
この質問に対する回答には、「Agree（同意）」と「Disagree（反対）」の2つの選択肢が用意されています。回答者は7つの年齢層に分けられ、各年齢層は43～113人です。最年少である年齢層（25～29才の年齢層）は、113名の回答者が属しており、この調査において最も回答者数が多い年齢層になっています。「Agree（同意）」の割合が最も高いのは、最年少の年齢層で、84.1%でした。シェアチャートには、同意／反対の割合が、年齢層ごとに描かれています。年齢が上がるにつれて、この質問に同意する回答者の割合は減少します。したがって、年長の回答者はより年少の回答者よりも仕事熱心ではない傾向にあると言えます。

ヒント: チャートの色は、「値の色」列プロパティで定義することができます。『JMP の使用法』の「列情報ウィンドウ」章を参照してください。

「カテゴリカル」プラットフォームの起動

「カテゴリカル」プラットフォームを起動するには、[**分析**] > [**消費者調査**] > [**カテゴリカル**] を選択します。

図3.3「カテゴリカル」プラットフォームの起動ウィンドウ



応答の役割

起動ウィンドウには、応答の役割について、3 種類のタブ（[単純]、[関連]、[多重]）があります。また、独自のクロス表を作成したいときに用いる［表の構成］タブがあります。これらの各タブは、分析対象とする応答の種類に対応しています。そして、それぞれのタブに用意されている各オプションは、データがどのような形式で保存されているかに対応しています。

［単純］タブ

［単純］タブは、1列ずつ個別にデータを集計したいときに用います。［単純］タブには、次の1つのオプションしかありません。

応答　1つまたは複数の列を分析に追加します。複数の列を選択した場合、「カテゴリカル」レポートには、個々の列に対するレポートが表示されます。

［関連］タブ

［関連］タブは、関係がある複数の列を集計したいときに用います。以下のオプションが用意されています。

共通の値をもつ応答　同じ水準を持つ複数列のデータを、1つの表に要約します。このオプションは、アンケート調査で、選択肢がまったく同じである質問を要約する場合に役立ちます。それらの質問を同時に要約・比較できます。

反復測定　異なる時点での応答が各列に含まれているデータを要約します。複数の時点で同一個体が応答している場合、標本が重複している（overlapping）と言います。標本に重複があるときは、検定を行うときにKish修正が使われます。Kish（1965; 12.4節）を参照してください。

判定の一致性　データが同一の対象物を判定した結果であり、各列が異なる回答者（判定者）に対応しているときに、そのデータを要約します。

［多重］タブ

［多重］タブには、1つまたは複数の列に記録された多重応答（mutiple responses; 複数回答）を分析するためのオプションが用意されています。「多重応答」は、アンケート調査において、複数の選択肢を選ぶことが許されている質問で見られます。ほかにも、「多重応答」は、製造業などで、1つの製品に複数の不適合が生じる場合にも見られます。［多重応答］タブの各オプションは、多重応答データがどのような形式で保存されているかに対応しています。

多重応答　多重応答データが複数列に含まれており、各列にそれぞれ1つの応答データが含まれている形式です。ここで選択された列の個数は、応答の最大数（選択肢の個数）に対応しています。実際のデータでは、セルの多くは空白になる可能性があります。

多重応答 ID別　回答者や部品のIDを含む列が別にあり、かつ、単一の列に積み重ねて多重応答データが保存されている形式です。

多重応答 区切り文字　多重応答データが、カンマ・セミコロン・タブのいずれかで区切られていて、単一の列に保存されている形式です。

指示変数　多重応答データが、指示変数で保存されている形式です。この形式は、各応答（質問の各選択肢）に対して1つの列があり、その列が0と1を含む指示変数となっているものです。

応答の度数　多重応答データが度数で集計されている形式です。この形式は、「指示変数」形式を集計したものです。

自由回答 テキストデータを要約します。［自由回答］オプションを使用すると、「カテゴリカル」レポートウィンドウの中にテキストエクスプローラが開きます。『基本的な統計分析』の「テキストエクスプローラ」章を参照してください。

［表の構成］タブ

［表の構成］タブは、分析者が望むクロス表を柔軟に構成したいときに用います。

- 横並び、交差、枝分かれのクロス表を対話的に指定できます。［表の構成］タブで作成されるクロス表では、クロス表の横側に位置する最も内側の項目が応答と見なされ、その他のすべての項目は因子と見なされます。
- ［表の構成］タブを使えば、1つの起動ウィンドウにて、複数の異なるクロス表を指定できます。
- 指定された列の尺度が［多重応答］に設定されている場合は、区切り文字で区切られた多重応答データとみなして分析を行います。列の尺度についての詳細は、『JMPの使用法』の「列情報ウィンドウ」章を参照してください。

詳細については、「「表の構成」レポートの例」（65ページ）を参照してください。

ヒント： 多重応答データで、応答の等質性に対する検定を行いたい場合は、［表の構成］タブを使用してください。

列の役割

次のような役割が用意されています。

X, グループ化カテゴリ グループを示す列を指定します。グループごとに度数がカウントされます。このオプションに複数の列が指定された場合に、デフォルト（［グループ化オプション］で［組み合わせ］が選択されている場合）では、枝分かれで集計されます。［グループ化オプション］によって、集計方法を変更できます。

標本サイズ 多重応答のデータを集計する場合に、属するグループの大きさ（含まれるユニットの数）を定義する列を指定します。たとえば、ある製品を100個検査したときに、50箇所で不適合がみつかった場合、「標本サイズ」（全ケース数）は100個で、「多重応答の度数」は50箇所となります。

度数 度数を含む列を指定します。このオプションは、分析対象のデータが要約されたものである場合に用いてください。

ID 回答者を識別する列を指定します。このオプションは、［多重応答ID別］が選択された場合にのみ必要で、他の応答の種類が指定された場合には使われません。

By By変数の水準ごとに個別のレポートが作成されます。複数のBy変数を指定した場合は、By変数の水準のすべての組み合わせごとにレポートが作成されます。

起動ウィンドウのその他のオプション

追加のオプションは、起動ウィンドウの左下にあります。また、これらのオプションは、起動ウィンドウで［OK］をクリックした後に、「カテゴリカル」の赤い三角ボタンのメニューから選択することもできます。

グループ化オプション　［X, グループ化カテゴリ］に複数の列が指定されている場合に、どのようにそれらの列を用いるかを定義します。

組み合わせ　指定されたグループ列に含まれるグループのすべての組み合わせに対して分析します。最初に指定された列が、クロス表で最も外側に位置するグループとなります。

それぞれを個別に　1つ1つのグループ列に対して、個別に分析します。

両方　グループ列を組み合わせた分析と、グループ列ごとの個別の分析が行われます。

ID内で一意な値をカウント　多重応答において、回答者内で重複している同一データを1つとしてカウントします。このオプションを用いるには、［ID］変数が指定されていなければなりません。

応答の欠測値をカウント　欠測値を1つのカテゴリとして扱います。欠測値は、空のセルまたは「欠測値のコード」列プロパティで定義された欠測コードのいずれかです。なお、欠測値しかない列の場合は、このオプションの状態とは無関係に欠測値がカウントされます。

メモ: このオプションが選択されていない場合、欠測値は分析から除かれます。

応答の水準を降順に並べる　応答を高い水準から低い水準に並べます。（デフォルトの順序は昇順です。）このオプションは、［X, グループ化カテゴリ］の列には適用されず、応答にのみ適用されます。

ヒント: ［値の順序］列プロパティを使うと、表示されるカテゴリの順序を変更できます。『JMPの使用法』の「列情報ウィンドウ」章を参照してください。

短いラベルの使用　すべてのデータ値（もしくは値ラベル）に共通する接頭部や接尾部を省略して、表示されるデータ値（もしくは値ラベル）を短くします。

メモ: このオプションは、データ値や値ラベルだけに適用されるものであり、列名には適用されません。

データにない応答を含める　データに存在していないカテゴリを、クロス表に含めます。そのカテゴリは、［値ラベル］列プロパティに指定しておく必要があります。このオプションは、応答にのみ適用されます。グループ化カテゴリには、データに存在しているカテゴリのみが含まれます。

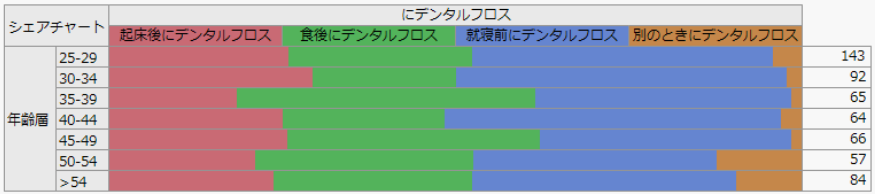
「カテゴリカル」レポート

最初の「カテゴリカル」レポートには、選択した応答列ごとに、二元クロス表（二元度数表）とシェアチャートが表示されます。[表の構成] タブを使用している場合、最初のレポートには、クロス表のみが表示されます。

図3.4 最初の「カテゴリカル」レポート

指示変数(起床後にデンタルフロス, 食後にデンタルフロス, 就寝前にデンタルフロス, 別のときにデンタルフロス) By Age Group

度数 / シェア / ケース		にデンタルフロス						
		起床後にデンタルフロス	食後にデンタルフロス	就寝前にデンタルフロス	別のときにデンタルフロス	応答の合計数	全ケース数	応答している全ケース数
年齢層	25-29	37	38	62	6	143	113	104
		25.9%	26.6%	43.4%	4.2%			
		32.7%	33.6%	54.9%	5.3%			
	30-34	27	19	44	2	92	68	67
		29.3%	20.7%	47.8%	2.2%			
		39.7%	27.9%	64.7%	2.9%			
	35-39	12	28	24	1	65	49	46
		18.5%	43.1%	36.9%	1.5%			
		24.5%	57.1%	49.0%	2.0%			
	40-44	16	15	31	2	64	52	51
		25.0%	23.4%	48.4%	3.1%			
		30.8%	28.8%	59.6%	3.8%			
	45-49	17	24	24	1	66	52	50
		25.8%	36.4%	36.4%	1.5%			
		32.7%	46.2%	46.2%	1.9%			
	50-54	12	18	20	7	57	43	43
		21.1%	31.6%	35.1%	12.3%			
		27.9%	41.9%	46.5%	16.3%			
	>54	20	24	32	8	84	71	70
		23.8%	28.6%	38.1%	9.5%			
		28.2%	33.8%	45.1%	11.3%			



クロス表の左上隅には、クロス表の各セルに含まれている統計量の名前（「度数」・「シェア」・「ケース」）が記載されます。これらの項目は、赤い三角ボタンのメニューにあるオプションを使って削除することができます。

- クロス表の各セルには、「度数」が表示されます。また、クロス表の右端には、それら度数の合計（「応答の合計数」）が表示されます。多重応答がある場合、クロス表の右端に、「全ケース数」（すなわち、行数）・「応答の合計数」・「応答している全ケース数」も表示されます。
- 各セルの「（応答の）シェア」は、応答の各度数を、その合計で割ったものです。これは、すべての応答に占める各応答の割合を示します。これは、クロス表における行パーセント（転置したクロス表では列パーセント）です。
- （欠測値を除外した）「ケース（あたりの比率）」は、応答の度数を応答している全ケース数で割ったものです。

図3.4には、年齢層ごとの応答数とケース数が表示されています。クロス表の最初の行には、25～29歳の年齢層の結果が記載されています。この行をもとに統計量を説明いたします。

- 最初のセルには、37の応答があります。これは、25 ～ 29歳において、「起床後にデンタルフロス」と回答している回答数が、37あることを示しています。
- 25 ～ 29 歳の年齢層における「応答の合計数」（回答総数）は 143 です。「起床後にデンタルフロス」の「シェア」は、143の応答のうちの25.9%（37/143）です。言い換えれば、25～29歳においては、デンタルフロスをするうちの25.9%が、起床後でのデンタルフロスであると言えます。
- 25～29歳の年齢層は全部で113人ですが、そのうち32.7%（37/113）が「起床後にデンタルフロス」と答えています。言い換えれば、25～29歳の32.7%が、起床後にデンタルフロスをしています。この結果は、1番左上のセルにおける「ケース」の行に表示されています。
- 「応答の合計数」（143）は、「応答しているケース数」（104）よりも多いので、25～29歳の年齢層においては、複数選択している回答者が存在していることが分かります。
- 「全ケース数」（113）は、「応答している全ケース数」（104）よりも多いです。この質問に回答していない25～29歳の回答者が9名いたということになります。

「カテゴリカル」プラットフォームのオプション

「カテゴリカル」の赤い三角ボタンのメニューには、必要に応じてレポートをカスタマイズできるオプションがあります。そこで使用できるオプションは、応答の役割、グループ列を指定したかどうか、起動ウィンドウで選択されたオプションによって異なります。

- 「レポートに関するオプション」
- 「統計的検定に関するオプション」
- 「「カテゴリカル」プラットフォームのその他のオプション」
- 「クロス表の応答におけるオプション」

レポートに関するオプション

度数　クロス表において、度数の表示／非表示を切り替えます。この「度数」は、各カテゴリにおける応答の頻度です。

応答のシェア（「シェア」）　クロス表において、「応答のシェア」の表示／非表示を切り替えます。「応答のシェア」は、全応答数に対して該当の応答数が占める割合です。

ケースあたりの比率（「ケース」）（多重応答の場合のみ使用可能。）クロス表において、「ケースあたりの比率」の表示／非表示を切り替えます。「ケースあたりの比率」は、全ケース数（応答しているかどうかに関わらない、すべてのケース数）で、該当の応答数を割ったものです。

非欠測ケースあたりの比率 （多重応答の場合のみ使用可能。）クロス表において、「非欠測ケースあたりの比率」の表示／非表示を切り替えます。「非欠測ケースあたりの比率」は、応答数を、何らかの応答をしているケースの総数で割ったものです。

シェアチャート 分割された棒グラフの表示／非表示を切り替えます。棒の長さは、各グループ内での応答の割合を表しています。棒グラフの右側には、各グループでの全応答数が表示されます。［X, グループ化カテゴリ］列が指定されていない場合、右側には総応答数が表示されます。

ヒント：「値の色」列プロパティを使用して、シェアチャートの色を変更することができます。『JMPの使用法』の「列情報ウィンドウ」章を参照してください。

度数チャート ［度数チャート］の表示／非表示を切り替えます。各棒は、応答の度数を示しています。スケールはチャート内において同じものが使われています。右端にあるグレーの棒グラフは、各グループでの総応答数を表します。

ヒント：「値の色」列プロパティを使用して、度数チャートの色を変更することができます。『JMPの使用法』の「列情報ウィンドウ」章を参照してください。

度数チャート 転置 転置した度数チャートの表示／非表示を切り替えます。各棒は、応答の度数を示しています。応答が行で、グループ変数が列に描かれます。各グループでの総応答数が、最後の行にグレーの棒グラフで描かれます。

クロス表 クロス表の表示／非表示を切り替えます。このクロス表では、列が応答、行が（指定されている場合は）グループとなっています。表の左上のセルに、表の各セルに表示されている統計量の名前（「度数」・「（応答の）シェア」・「ケース（あたりの比率）」）が記載されています。なお、［クロス表 転置］を選択すると、この元のクロス表は削除されます。

クロス表 転置 転置したクロス表の表示／非表示を切り替えます。この転置したクロス表では、行が応答、列が（指定されている場合は）グループとなっています。表の左上のセルに、表の各セルに表示されている統計量の名前（「度数」・「（応答の）シェア」・「ケース（あたりの比率）」）が記載されています。なお、［クロス表］オプションを選択すると、この転置したクロス表は削除されます。

統計的検定に関するオプション

統計的検定に関するオプションは、応答の役割、グループ変数があるかどうか、によって使用できるものが異なります。オプションには、応答の等質性に対する検定、関連性の指標、相対リスク、一致性の指標があります。

多重応答の検定 （1つまたは複数のグループ列が指定されており、かつ、多重応答の場合にのみ使用可能。）「多重応答（複数回答）に対する検定の例」（46ページ）を参照してください。各グループ間での応答カテゴリの差について、次の検定が用意されています。

Poisson分布の度数検定 Poisson回帰の枠組みで、ケースあたりの比率の差を検定します。ケースあたりの比率が、グループごとに異なるかどうかが検定されます。1ケースあたりで計算される比率がグループ列の水準間で異なるかが、尤度比カイ2乗検定によって計算されます。

二項分布の等質性検定　割合の差について尤度比カイ2乗検定を行います。検定においては、各カテゴリに対して、「選択されている」／「選択されていない」の二項分布が仮定されます。

応答の等質性に対する検定　（1つまたは複数のグループ列があり、かつ、単一選択の応答がある場合。もしくは、［表の構成］タブで、1つまたは複数のグループ列と多重応答がある場合。）応答の等質性に対する検定を行います。実行される検定は、次のように指定によって異なります。

– グループ間における応答の差に対する検定としては、尤度比検定とPearsonカイ2乗検定が実行されます。詳細については、「応答の等質性に対する検定の例」（45ページ）を参照してください。

– ［表の構成］タブで多重応答が指定された場合には、Rao-Scottのカイ2乗検定が実行されます。

セルのカイ2乗　クロス表の各セルにおける*p*値を求めます。これらのp値は、独立性に対するカイ2乗検定の枠組みで求められています。*p*値が小さい場合は、そのセルにおいて、「行と列とが独立である」という仮定のもとでの期待度数よりも、観測度数が大きいかまたは小さいことを示します。*p*値は、観測度数が期待度数よりも大きいかまたは小さいかに応じて、色分けされ、濃淡がつけられます。詳細については、「セルのカイ2乗検定の例」（48ページ）を参照してください。

標本の各ペア比較　（1つまたは複数のグループ列があり、かつ、単一選択の応答がある場合にのみ使用可能。）グループのペアごとに、応答の等質性に対する検定（行と列との独立性に対する検定）を行います。尤度比カイ2乗検定とPearsonのカイ2乗検定を行います。詳細については、「標本の各ペア比較の例」（49ページ）を参照してください。

セルの各ペア比較　（1つまたは複数のグループ列があり、かつ、単一選択の応答または多重応答がある場合にのみ使用可能。）グループのペアごとに、応答の特定のカテゴリを比較します。このとき、特定のカテゴリ以外のカテゴリが、すべて併合されて検定が計算されます。尤度比カイ2乗検定、Pearsonのカイ2乗検定、Fisherの正確検定を行います。詳細については、「セルの各ペア比較の例」（50ページ）を参照してください。

相対リスク　（グループ列が2水準であり、かつ、応答も2水準の場合。もしくは、グループ列が2水準であり、かつ、応答が多重応答であり、かつ、［ID内で一意な値をカウント］オプションが選択されている場合。）相対リスクを計算します。詳細については、「条件付き関連と相対リスクの例」（53ページ）を参照してください。

条件付き関連　（［ID内で一意な値をカウント］オプションが選択されている場合。）特定の応答水準で条件付けたもとで、ある応答水準が出現する条件付き確率を求めます。詳細については、「条件付き関連と相対リスクの例」（53ページ）を参照してください。

一致性の統計量　（［判定の一致性］応答の場合にのみ使用可能。）一致性のカッパ係数を計算し、また、対称性に対するBowkerの検定を行います。詳細については、「判定の一致性の例」（56ページ）を参照してください。

遷移レポート　（［反復測定］応答の場合にのみ使用可能。）時間経過に伴う応答の変化に関して、遷移度数と遷移率行列を求めます。詳細については、「反復測定の例」（57ページ）を参照してください。

検定オプション このメニューで使用できるオプションは、選択された分析によって異なります。

カイ2乗検定の選択 単一選択の応答に対するカイ2乗検定の種類を指定します。［尤度比とPearson］、［尤度比のみ］、［Pearsonのみ］から選択できます。

警告の表示 カイ2乗検定の標本サイズが小さい場合に警告を表示します。

p値の小さい順に並べ替え レポートを、統計的検定の有意確率が小さい順に並べ替えます。

有意でない結果を非表示にする 統計的に有意でないレポートを非表示にします。

「カテゴリカル」プラットフォームのその他のオプション

応答の合計数 クロス表とシェアチャートにおいて、総度数の表示／非表示を切り替えます。グループ列が使われている場合、この総度数はグループごとに集計されます。

応答水準の表示 クロス表とシェアチャートにおいて、応答列におけるカテゴリの表示／非表示を切り替えます。

上位カテゴリの表示 （1つまたは複数の「上位カテゴリ」列プロパティが定義されている場合のみ使用可能。）クロス表と度数チャートにおいて、「上位カテゴリ」を集計した列の表示／非表示を切り替えます。「上位カテゴリ」列プロパティの詳細については、「上位カテゴリ」（43ページ）を参照してください。

ヒント： このオプションは、上位カテゴリの表示／非表示を切り替えるものです。上位カテゴリ内にある下位カテゴリを非表示にするには、「上位カテゴリ」列プロパティで［非表示］オプションを用いてください。または、「カテゴリカル」レポートの赤い三角ボタンメニューより「応答水準の表示」オプションをオフにすると、上位カテゴリだけが表示され、下位にあるすべての応答水準が非表示になります。

全ケース数 （多重応答の場合のみ使用可能。）クロス表において、各グループにおけるケース数（回答者数）の表示／非表示を切り替えます。

応答している全ケース数 （多重応答列の場合のみ使用可能。）クロス表において、少なくとも1つは応答のあるケース（回答者）の度数の表示／非表示を切り替えます。全く応答しなかった人は、この度数には含まれません。「応答している全ケース数」は、必ず、「全ケース数」以下です。

スコアの平均 クロス表とシェアチャートにおいて、全体またはグループごとの平均の表示／非表示を切り替えます。この平均は、それぞれの応答のカテゴリに割り当てられた数値に基づいて計算されます。

- 数値の場合には、その数値がそのままスコアとして使われます。
- 数値以外の場合には、「値スコア」列プロパティで各カテゴリに割り当てられた数値が、スコアとして使われます。
- 「値スコア」列プロパティも指定されていな場合には、1 から順に振られた通し番号がスコアとして使われます。

詳細については、「スコア平均を比較する例」（63ページ）を参照してください。

スコア平均の比較　クロス表において、「平均の比較」列の表示／非表示を切り替えます。Satterthwaite の *t* 検定（プールしない誤差分散による t 検定）によって、グループ間のスコア平均を比較します。SAS Institute Inc.（2017）を参照してください。比較の結果は、アルファベット文字を使って表示されます。比較を表すアルファベット文字の詳細については、「比較を表すアルファベット文字」（41 ページ）を参照してください。比較グループを指定する方法については、「分析者が興味のある比較を指定する例」（52 ページ）を参照してください。

スコアの標準偏差　クロス表において、全体もしくはグループごとの標準偏差を含む列の表示／非表示を切り替えます。

スコア平均で並べ替え　（結果に複数の応答があり、かつ、グループ列がない場合にのみ使用可能。）複数の応答に対するレポートを、スコア平均順に並べます。

テーブルの保存　レポートの特定の部分を、新しいデータテーブルに保存します。各オプションを実行すると、該当の情報を含んだデータテーブルが作成されます。このメニューで使用できるオプションは、選択された分析によって異なります。

メモ： 上位カテゴリは、新しいテーブルに含まれません。

度数の保存　クロス表の「度数」を、新しいデータテーブルに保存します。

応答のシェアの保存　クロス表の「（応答の）シェア」を、新しいデータテーブルに保存します。

分割表の保存　クロス表全体を新しいデータテーブルに保存します。

ケースあたりの比率の保存　クロス表の「ケース（あたりの比率）」を、新しいデータテーブルに保存します。

転置した度数の保存　転置したクロス表の「度数」を、新しいデータテーブルに保存します。

転置した応答のシェアの保存　転置したクロス表の「（応答の）シェア」を、新しいデータテーブルに保存します。

転置したケースあたりの比率の保存　転置したクロス表の「ケース（あたりの比率）」を、新しいデータテーブルに保存します。

比率検定の保存　［多重応答の検定］オプションの結果を、新しいデータテーブルに保存します。

等質性検定の保存　［応答の等質性に対する検定］オプションの結果を、新しいデータテーブルに保存します。

スコア平均の保存　各標本グループのスコアの平均を、新しいデータテーブルに保存します。

t 検定と p 値の保存　「スコア平均の比較」レポートの *t* 検定統計量と *p* 値を、新しいデータテーブルに保存します。

Excel ファイルとして保存　Microsoft Excel スプレッドシートに、表を保存します。保存される表では、応答カテゴリを行とし、標本水準を列として、すべての表を 1 つのシートにまとめています。複数の表にまたがる標本の見出しは共有されます。表の各セルに複数の要素がある場合は、それを Microsoft Excel ファイルに複数のセルとして表示するか、単一のセルとして表示するかを選択できます。

フィルタ データをフィルタリングするためのローカルデータフィルタの表示／非表示を切り替えます。このとき、応答を1つも含まない標本の水準は、表示されなくなります。そのような表示されなくなった水準を表示するには、起動ウィンドウで［**データにない応答を含める**］を選択しておいてください。または、**「環境設定の変更」**の赤い三角メニューを選択し、［**データにない応答を含める**］を選択しておいてください。

要約の目次 「カテゴリカル」レポート最上部の「要約の目次」レポートの表示／非表示を切り替えます。「要約の目次」では、すべての検定とスコア平均を要約して、関連レポートへのリンクとともに表示しています。

レポートで使用された列を表示 「レポートで使用された列」という情報の表示／非表示を切り替えます。このオプションは、「SPSS名」や「SAS名」、または「SPSSラベル」や「SASラベル」の列プロパティを持つ列のみに影響します。SASやSPSSからデータセットを読み込む（インポートする）と、SASやSPSSのデータセットの変数名とラベル名が、自動的にJMPデータテーブルの列プロパティに追加されます。また、「その他」の列プロパティを使って、「SAS名」や「SPSS名」または「SAS ラベル」や「SPSSラベル」の列プロパティを追加することもできます。たとえば、「SAS名」や「SPSS名」の列プロパティを使って、短い列名を与えることができます。

要素の表示形式 度数・（応答の）シェア・ケース（あたりの比率）・ゼロ度数に対して、それらの表示形式を指定します。デフォルトでは、度数は固定小数点形式（総桁数7、小数桁数0）、「（応答の）シェア」と「ケース（あたりの比率）」はパーセント値形式（総桁数6、小数桁数1）です。

複数行に配置 レポートを、ページの下に移動させずに横方向に配置します。横方向に並べるレポートの数を入力してください。

環境設定の変更 次回以降に「カテゴリカル」プラットフォームを起動したときのために、現在や次回以降のJMPセッションでの環境設定を変更することができます。詳細については、「環境設定の変更」（44ページ）を参照してください。

カテゴリに関するオプション ［グループ化オプション］・［応答の欠測値をカウント］・［応答の水準を降順に並べる］・［短いラベルの使用］・［データにない応答を含める］のオプションがあります。これらのオプションは起動ウィンドウにも表示されます。これらのオプションがここで選択されると、その新しい設定でプラットフォームが更新されます。「カテゴリに関するオプション」の詳細については、「起動ウィンドウのその他のオプション」（33ページ）を参照してください。

クロス表を濃淡表示 環境設定で濃淡表示が指定されていなくても、クロス表を濃淡表示にします。このオプションが選択されていない場合、クロス表レポートの濃淡は、現在の環境設定における［テーブル行の濃淡表示］オプションに合わせて表示されます。

ダイアログの再起動 起動ウィンドウを呼び出します。起動ウィンドウに戻り、もう一度、分析をやり直すことができます。

以下のオプションについて詳しくは、『JMPの使用法』の「JMPのレポート」章を参照してください。

ローカルデータフィルタ 現在のレポートにおいて、データをフィルタリングするローカルデータフィルタの表示／非表示を切り替えます。

やり直し　分析を繰り返したり、やり直したりするオプションを含みます。また、［自動再計算］オプションに対応しているプラットフォームにおいては、［自動再計算］オプションを選択すると、データテーブルに加えた変更が、該当するレポートに即座に反映されるようになります。

スクリプトの保存　レポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。

By グループのスクリプトを保存　By 変数の全水準に対するレポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。起動ウィンドウで By 変数を指定した場合のみ使用可能です。

クロス表の応答におけるオプション

文字の表示　クロス表において、列名を識別するアルファベット文字の表示／非表示を切り替えます。これらのアルファベット文字は、統計的検定の結果を表すのに、どのグループとどのグループに差があるかを示すのに使われます。

比較の設定　検定において、どのグループとどのグループを比較するかを指定できます。スラッシュで区切られたグループが比較されます。各比較はカンマで区切ってください。グループは、アルファベットで指定してください。たとえば、A を E と、B を D と、C を F と、比較したい場合には、「A/E, B/D, C/F」のようにグループを指定します。［セルの各ペア比較］レポートに、指定された比較の結果が表示されます。詳細については、「分析者が興味のある比較を指定する例」（52 ページ）を参照してください。

削除　レポートウィンドウからレポートを削除します。

注意：［削除］オプションを元に戻すことはできません。

比較を表すアルファベット文字

［セルの各ペア比較］・［標本の各ペア比較］・［スコア平均の比較］オプションの結果では、それぞれの標本グループを A、B、C といったアルファベット文字で表しています。27 番目以降の標本グループには、「A1」、「B1」のように、アルファベット文字の後に数字が付けられます。比較に関するオプションが使われた場合、このアルファベット文字がクロス表の水準見出しに表示されます。

図3.5 比較の結果を表すアルファベット文字のあるクロス表

Responses(生産国) By type

度数			生産国				
シェア			米国	ヨーロッパ	日本	応答の合計数	比較
タイプ	ファミリー	A	74 47.7%	15 9.7%	66 42.6%	155	B
	スポーツ	B	23 23.0%	21 21.0%	56 56.0%	100	c
	ワーク	C	18 37.5%	4 8.3%	26 54.2%	48	*

デフォルトの比較: A/B/C

ペア比較で有意に異なっている水準が、後のほうの水準で表示されています
* 星1つで警告する度数の基準 100 大文字で示す有意水準 0.05
** 星2つで警告する度数の基準 30 小文字で示す有意水準 0.1

標本の各ペア比較

グループ間に有意差がある場合、シェアが小さいほうのグループのアルファベット文字が、もう一方のグループのセルに表示されます。アルファベットが大文字となっている場合、小文字よりも有意性が高いことを示します。デフォルトでは、p 値が0.05以下のときに大文字で、0.10以下のときに小文字で表示されます。図3.5における「B」は、生産国の割合における「スポーツ」と「ファミリー」との違いが、有意水準5%で有意になっていることを示しています。「ファミリー」の合計数（155）は、「スポーツ」の合計数（100）よりも多いため、「ファミリー」の行に「B」と表示されています。「スポーツ」行の「c」は、「スポーツ」と「ワーク」を比較した場合、生産国の割合には有意水準10%で有意差があることを示しています。「スポーツ」の応答数（100）は、「ワーク」の応答数（48）よりも多いため、この「c」は、「スポーツ」行に表示されます。

度数が小さい場合の警告も、比較のセルに表示されます。1つのアスタリスク（*）は、その水準に含まれる応答が100より少ないことを示し、2つのアスタリスク（**）は、30よりも少ないことを示します。図3.5では、「ワーク」行の合計数は48で、アスタリスクが1つ表示されています。有意水準や警告で使われる基準は、「カテゴリカル」プラットフォームの環境設定で変更できます。環境設定の変更についての詳細は、「環境設定の変更」（44ページ）を参照してください。

ヒント: 比較を表すアルファベット文字を、大文字と小文字の両方ではなく一方だけを使用する場合は、環境設定の「小文字で示す有意水準」を0に設定してください。

「標本の各ペア比較の例」（49ページ）、「セルの各ペア比較の例」（50ページ）、および「スコア平均を比較する例」（63ページ）を参照してください。

上位カテゴリ

「上位カテゴリ」とは、いくつかのカテゴリをまとめたものです。たとえば、5段階評価を使用している場合に、第1段階と第2段階を併合したものの割合を調べたいとします。そのようなときは、「上位カテゴリ」列プロパティによって、それら2つのカテゴリを併合した上位カテゴリを定義することができます。

上位カテゴリは、クロス表と度数チャートには表示されます。しかし、上位カテゴリはシェアチャートには表示されません。また、グループ列には適用されません。

上位カテゴリを作成するには、次の手順に従います。

1. データテーブルの中で、1つにまとめたいカテゴリを含んでいる列を選択します。
2. ［**列**］ > ［**列情報**］を選択します。
3. ［**列プロパティ**］をクリックし、［**上位カテゴリ**］を選択します。
4. （オプション）デフォルトで表示される上位カテゴリの名前は、変更が可能です。
5. 列の「カテゴリ」リストから1つまたは複数のカテゴリを選択します。
6. ［**追加**］をクリックします。
7. （オプション）上位カテゴリを選択し、「上位カテゴリ」の赤い三角ボタンのメニューをクリックして追加のオプションを表示します。

［上位カテゴリ］のオプション

列プロパティウィンドウにある「上位カテゴリ」の赤い三角ボタンのメニューには、次のオプションがあります。

非表示　クロス表と度数チャートにおいて、上位カテゴリ内のカテゴリを非表示にします。

ヒント: レポート内で下位カテゴリの表示／非表示を切り替えたい場合には、この［非表示］オプションを使わないでください。「カテゴリカル」の赤い三角ボタンのメニューにある［応答水準の表示］オプションを使ってください。

重複なし　（多重応答列の場合にのみ使用可能。）下位のカテゴリが、複数の上位カテゴリで表示される場合に、重複して数えられるのを防ぎます。

平均を追加　平均を含む列がクロス表に追加されます。

標準偏差を追加　標準偏差を含む列がクロス表に追加されます。

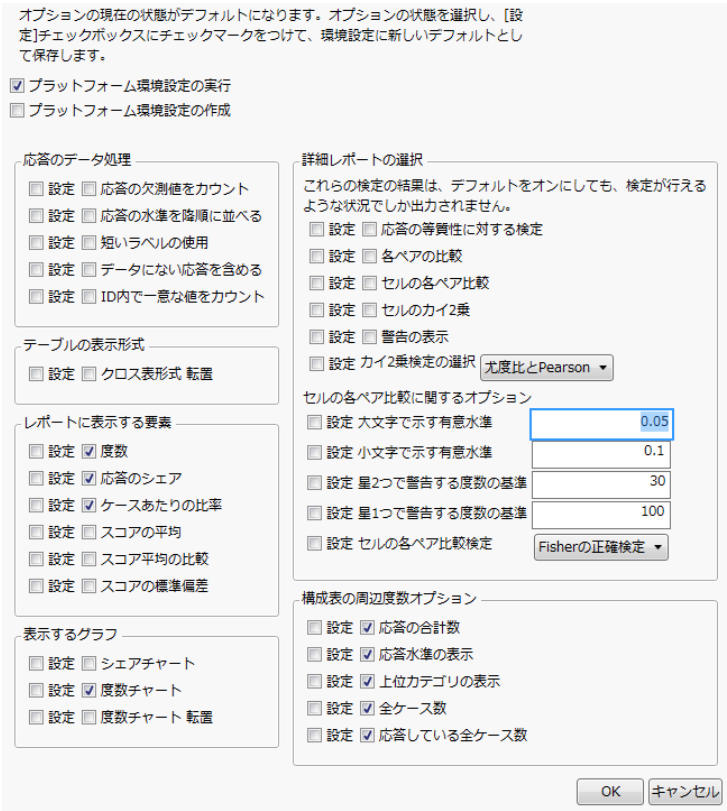
すべて追加　応答の合計数がクロス表に含まれます。なお、デフォルトでも、「応答の合計数」列は常に含まれています。

メモ: 「上位カテゴリ」の機能は、［反復測定］と［判定者の一致性］を除く、応答のすべての種類で使用できます。

環境設定の変更

「カテゴリカル」の赤い三角ボタンのメニューには、環境設定を指定できる［環境設定の変更］オプションがあります。

図3.6 環境設定の設定ウィンドウ



設定を変更したいオプションの［**設定**］チェックボックスを選択してください。このとき、デフォルトで実行したいオプションは、そのオプション名のすぐ左にあるチェックボックスを選択してください。逆に、デフォルトで実行したくないオプションは、そのチェックボックスの選択を解除してください。設定した変更を環境設定に適用したい場合には、［**プラットフォーム環境設定の実行**］ボックスを選択してください。また、環境設定を変更するスクリプトを生成したい場合には、［**プラットフォーム環境設定の作成**］ボックスを選択します。［プラットフォーム環境設定の実行］を選択すると、次回以降に、「カテゴリカル」プラットフォームを起動したときに、設定した環境設定が使われます。

メモ：［プラットフォーム環境設定の作成］で生成されたスクリプトを実行すると、そこでの設定がプラットフォーム環境設定に適用されます。このスクリプトを使えば、変更した環境設定を、複数のユーザーで共有したり、自分が使いたいときに利用したりできます。

「カテゴリカル」プラットフォームの別例

- 「応答の等質性に対する検定の例」
- 「多重応答（複数回答）に対する検定の例」
- 「セルのカイ2乗検定の例」
- 「標本の各ペア比較の例」
- 「セルの各ペア比較の例」
- 「分析者が興味のある比較を指定する例」
- 「条件付き関連と相対リスクの例」
- 「判定の一致性の例」
- 「反復測定の例」
- 「［多重応答］タブの例」
- 「スコア平均を比較する例」
- 「「表の構成」レポートの例」

応答の等質性に対する検定の例

ここでは、自家用車に関する調査データである「Car Poll.jmp」サンプルデータを使用します。このデータには、回答者の属性情報と所有する車についての情報が含まれています。既婚／未婚と車の生産国との間の関係を調べてみましょう。また、応答の等質性に対する検定も行ってみましょう。つまり、車の生産国の分布が、既婚と未婚の回答者で同じなのかどうかを調べます。

検定には、「Pearsonのカイ2乗検定」と「尤度比検定」の2つが用意されています。いずれの検定においても、カイ2乗統計量とp値が計算されます。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Car Poll.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「生産国」**を選択して、［単純］タブの［**応答**］をクリックします。
4. **「既婚/未婚」**を選択して、［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「カテゴリカル」の赤い三角ボタンをクリックして、［**応答の等質性に対する検定**］を選択します。

図3.7 応答の等質性に対する検定

Responses(生産国) By marital status

度数		生産国			
シェア		米国	ヨーロッパ	日本	応答の合計数
既婚/未婚	既婚	83 42.3%	26 13.3%	87 44.4%	196
	未婚	32 29.9%	14 13.1%	61 57.0%	107

応答の等質性に対する検定

応答の列	標本の列	尤度比 カイ2乗	尤度比 p値	Pearson カイ2乗	Pearson p値
生産国	既婚/未婚	5.14007	0.0765	5.08146	0.0788

「シェアチャート」を見ると、既婚グループが保有する車は、米国車と日本車の間で均等に分かれています。未婚グループでは、日本車が最も多く保有されています。応答の等質性に対する検定において、p 値は約 0.08 です。未婚と既婚との差は、有意水準を 0.05 としたとき、統計的に有意ではありません。

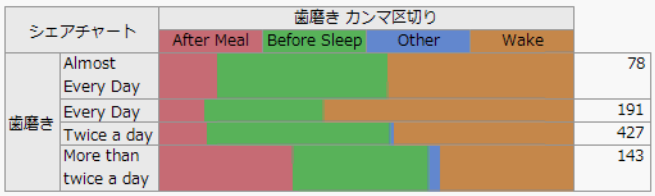
多重応答（複数回答）に対する検定の例

この例では、「Consumer Preferences.jmp」を使用します。このデータには、一般的な事柄に対する意見や態度が含まれています。それぞれの歯磨き時間帯（**「歯磨き カンマ区切り」**）に対する応答率が、**「歯磨き」**グループ間で同じであるかどうかを、［多重応答の検定］オプションを使って検定してみましょう。この「歯磨き」グループは、歯磨きを行う頻度で分類されているものです。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「歯磨き カンマ区切り」**を選択し、［多重］タブの［**多重応答 区切り文字**］をクリックします。
4. **「歯磨き」**を選択して［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「カテゴリカル」の赤い三角ボタンをクリックして、［**多重応答の検定**］>［**Poisson 分布の度数検定**］を選択します。

図3.8 多重応答の検定、Poisson分布

度数 シェア ケース		歯磨き カンマ区切り						
		After Meal	Before Sleep	Other	Wake	応答の合計数	全ケース数	応答している全ケース数
歯磨き	Almost Every Day	11	32	0	35	78	48	48
		14.1%	41.0%	0.0%	44.9%			
		22.9%	66.7%	0.0%	72.9%			
	Every Day	21	55	0	115	191	142	142
		11.0%	28.8%	0.0%	60.2%			
		14.8%	38.7%	0.0%	81.0%			
	Twice a day	49	189	4	185	427	209	209
		11.5%	44.3%	0.9%	43.3%			
		23.4%	90.4%	1.9%	88.5%			
	More than twice a day	46	47	4	46	143	49	49
		32.2%	32.9%	2.8%	32.2%			
		93.9%	95.9%	8.2%	93.9%			



各応答に対するPoisson検定

歯磨き, 歯磨き カンマ区切り

歯磨き カンマ区切り	カイ2乗	p値(Prob>ChiSq)
After Meal	59.5436	<.0001*
Before Sleep	39.1097	<.0001*
Other	12.7131	0.0053*
Wake	1.8793	0.5978

検定はPoisson分布を仮定しています。

pを見ると、「**After Meal**（食後）」、「**Before Sleep**（就寝前）」、「**Other**（その他）」の応答率において、「歯磨き」グループ間で有意差があります。「**Wake**（起床後）」においては、有意差は見られません。クロス表を見ると、歯磨きの頻度とは関係なく、大半の人が起床後には歯磨きをしています。

7. 「カテゴリカル」の赤い三角ボタンをクリックして、［**多重応答の検定**］>［**二項分布の等質性検定**］を選択します。

図3.9 多重応答の検定、二項分布

各応答に対する二項検定

歯磨き, 歯磨き カンマ区切り

歯磨き カンマ区切り	カイ2乗	p値(Prob>ChiSq)
After Meal	113.315	<.0001*
Before Sleep	131.169	<.0001*
Other	12.9822	0.0047*
Wake	12.2461	0.0066*

各検定は、該当カテゴリを含んでいるか、含んでいないかを応答としたものです。

［二項分布の等質性検定］オプションでは、［Poisson 分布の度数検定］よりも、必ず検定統計量が大きくなります（つまり、p 値が小さくなります）。［二項分布の等質性検定］は、応答が生じた回数（起床後に歯磨きをすると答えた人の数）だけを考慮するのではなく、応答が生じなかった回数（起床後に歯磨きをすると答えなかった人の数）も考慮します。

この例では、各応答（「After Meal（食後）」、「Before Sleep（就寝前）」、「Wake（起床後）」、「Other（その他）」）の割合は、年齢層間で異なっています。それらの p 値は、0.05 よりも小さくなっています。

メモ： JMP は、「多重応答」の尺度または「多重応答」の列プロパティを持つ列を多重応答列として認識します。

セルのカイ２乗検定の例

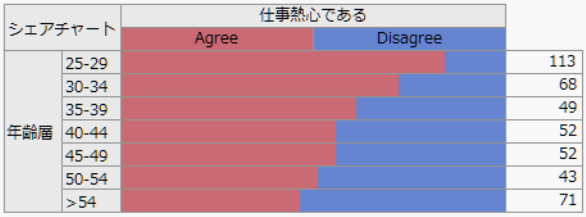
この例では、「Consumer Preferences.jmp」を使用します。このデータには、一般的な事柄に対する意見や態度が含まれています。年齢層によって「仕事熱心である」かどうかの分布が異なるかどうかを調べてみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「仕事熱心である」**を選択して、［単純］タブの［**応答**］をクリックします。
4. **「年齢層」**を選択して［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「カテゴリカル」の赤い三角ボタンをクリックして、［**クロス表 転置**］を選択します。
7. 「カテゴリカル」の赤い三角ボタンをクリックして、［**セルのカイ２乗**］を選択します。

図3.10 セルのカイ２乗

Responses(仕事熱心である) By Age Group

度数 シェア セルのカイ2乗p値		年齢層						
		25-29	30-34	35-39	40-44	45-49	50-54	>54
仕事熱心である	Agree	95	49	30	29	29	22	33
		84.1%	72.1%	61.2%	55.8%	55.8%	51.2%	46.5%
		0.00788	0.41003	0.80398	0.45495	0.45495	0.29058	0.06415
	Disagree	18	19	19	23	23	21	38
		15.9%	27.9%	38.8%	44.2%	44.2%	48.8%	53.5%
		0.00039	0.27136	0.74035	0.31848	0.31848	0.15823	0.01345
	応答の合計数	113	68	49	52	52	43	71



各セルにおいて、*p* 値が小さい場合は、そのセルの観測度数と期待度数との間に有意差があることを示しています。また、各セルの色は、*p* 値を基準にして、期待値よりも有意に度数が大きいほど濃い赤に、期待値よりも有意に度数が小さいほど濃い青に色分けされます。なお、各セルの期待度数は、観測された行和と列和から算出されます。

たとえば、25～29 歳で「Agree（同意する）」と回答する人の期待度数は、(287×113)/448 = 72.4 人で、実際の観測度数は 95 人です。その *p* 値は 0.00788 ですので、観測度数は期待度数よりも有意に大きいと言えます。「仕事熱心である」に同意する 25～29 歳の人数は、独立であるとの仮定から算出された期待度数よりも有意に多い、ということが分かります。

標本の各ペア比較の例

この例では、「Consumer Preferences.jmp」を使用します。このデータには、一般的な事柄に対する意見や態度が含まれています。年齢層によって「仕事熱心である」という質問に対する回答の分布が異なるかどうかを調べてみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「仕事熱心である」**を選択して、［単純］タブの［**応答**］をクリックします。
4. **「年齢層」**を選択して［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「カテゴリカル」の赤い三角ボタンをクリックして、［**標本の各ペア比較**］を選択します。

図 3.11 標本の各ペア比較

Responses(仕事熱心である) By Age Group

度数			仕事熱心である			
シェア			Agree	Disagree	応答の合計数	比較
年齢層	25-29	A	95 84.1%	18 15.9%	113	b,C,D,E,F,G
	30-34	B	49 72.1%	19 27.9%	68	d,e,F*
	35-39	C	30 61.2%	19 38.8%	49	*
	40-44	D	29 55.8%	23 44.2%	52	*
	45-49	E	29 55.8%	23 44.2%	52	*
	50-54	F	22 51.2%	21 48.8%	43	*
	>54	G	33 46.5%	38 53.5%	71	B*

デフォルトの比較: A/B/C/D/E/F/G

ペア比較で有意に異なっている水準が、後のほうの水準で表示されています
* 星1つで警告する度数の基準　100　大文字で示す有意水準　0.05
** 星2つで警告する度数の基準　30　小文字で示す有意水準　0.1

標本の各ペア比較

英文字は、Pearson カイ2乗の結果を表しています。 尤度比 p値　　Pearson p値
年齢層, 仕事熱心である

ペアに対する尤度比カイ2乗のp値

	A	B	C	D	E	F	G
A	1.0000	0.0552	0.0020	0.0001	0.0001	<.0001	<.0001
B	0.0552	1.0000	0.2183	0.0641	0.0641	0.0261	0.0020
C	0.0020	0.2183	1.0000	0.5781	0.5781	0.3312	0.1108
D	0.0001	0.0641	0.5781	1.0000	1.0000	0.6540	0.3083
E	0.0001	0.0641	0.5781	1.0000	1.0000	0.6540	0.3083
F	<.0001	0.0261	0.3312	0.6540	0.6540	1.0000	0.6276
G	<.0001	0.0020	0.1108	0.3083	0.3083	0.6276	1.0000

ペアに対するPearsonカイ2乗のp値

	A	B	C	D	E	F	G
A	1.0000	0.0523	0.0015	<.0001	<.0001	<.0001	<.0001
B	0.0523	1.0000	0.2170	0.0638	0.0638	0.0255	0.0022
C	0.0015	0.2170	1.0000	0.5783	0.5783	0.3314	0.1119
D	<.0001	0.0638	0.5783	1.0000	1.0000	0.6540	0.3087
E	<.0001	0.0638	0.5783	1.0000	1.0000	0.6540	0.3087
F	<.0001	0.0255	0.3314	0.6540	0.6540	1.0000	0.6276
G	<.0001	0.0022	0.1119	0.3087	0.3087	0.6276	1.0000

クロス表には、各年齢層における「仕事熱心である」の回答が要約されています。クロス表の各セルには、年齢層ごとに、「Agree（同意する）」と「Disagree（反対する）」の度数とシェア（割合）が表示されています。また、比較の検定結果を示すアルファベット文字も表示されています。アルファベット文字がどのグループを示しているかは、グループ名のすぐ右側の列に表示されています。また、一番右側にある比較列では、それらのアルファベット文字によって、検定結果が表示されています。

［標本の各ペア比較］アウトラインには、ペアごとの比較に対して、尤度比検定とPearson カイ2 乗検定のp値が表示されています。これらのp値は、アルファベット文字のラベルを付加された行列の形式で表示されています。

この例の場合、次のようなことがわかります。

- 25～29歳の比較列には、すべてのアルファベット文字（b～g）が含まれています。つまり、25～29歳層は、他の年齢層と比較して、「仕事熱心である」という質問に対する回答には有意差があります。アルファベット文字のb は小文字なので、25～29歳と30～34歳との違いにおける有意水準は、0.10です。他のアルファベット文字は大文字なので、有意水準0.05で差があります。
- 54歳以上の年齢層（G）は、30～34歳（B）と有意差があります。グループGの応答数は、グループBの応答数よりも多いので（71対68）、比較を表すアルファベット文字（B）が、Gの年齢層のほうに表示されています。
- 比較セルにある1つのアスタリスクは、標本サイズが小さいことを表す警告です。1つのアスタリスクがある場合、そのグループの度数が、30以上で、100未満であることを示します。
- この例では見られませんが、2つのアスタリスクがある場合、グループサイズが30よりも少ないことを示します。

セルの各ペア比較の例

この例では、「Consumer Preferences.jmp」を使用します。このデータには、一般的な事柄に対する意見や態度が含まれています。「被雇用期間」グループによって、仕事に対する満足度の分布が異なるかどうかを調べてみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「仕事に対する満足度」**を選択して、［単純］タブの**「応答」**をクリックします。
4. **「被雇用期間」**を選択して、［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「カテゴリカル」の赤い三角ボタンのメニューをクリックして、［**セルの各ペア比較**］を選択します。

図3.12 セルの各ペア比較

Responses(仕事に対する満足度) By Employee Tenure

			度数 シェア 比較	仕事に対する満足度			
				Not at all satisfied	Somewhat satisfied	Extremely satisfied	応答の合計数
被雇用期間	less than 5 years	A		11 6.9% 	97 60.6% d	52 32.5% 	160
	5 to 10 years	B		13 9.4% 	72 51.8% 	54 38.8% 	139
	10 to 20 years	C		4 4.7% *	43 50.0% *	39 45.3% a*	86
	more than 20 years	D		4 6.3% *	30 47.6% *	29 46.0% a*	63

デフォルトの比較: A/B/C/D

ペア比較で有意に異なっている水準が、後のほうの水準で表示されています
* 星1つで警告する度数の基準　100　大文字で示す有意水準　0.05
** 星2つで警告する度数の基準　30　小文字で示す有意水準　0.1

シェアチャート		仕事に対する満足度			
		Not at all satisfied	Somewhat satisfied	Extremely satisfied	
被雇用期間	less than 5 years				160
	5 to 10 years				139
	10 to 20 years				86
	more than 20 years				63

セルの各ペア比較 - 詳細

英文字は、Fisherの正確検定の結果を表しています。
被雇用期間, 仕事に対する満足度

尤度比 ペア比較

		AB	BB	AC	BC	CC	AD	BD	CD	DD
Not at all satisfied	1.0000	0.4323	1.0000	0.4781	0.1806	1.0000	0.8872	0.4658	0.6515	1.0000
Somewhat satisfied	1.0000	0.1246	1.0000	0.1092	0.7931	1.0000	0.0783	0.5820	0.7739	1.0000
Extremely satisfied	1.0000	0.2526	1.0000	0.0476	0.3366	1.0000	0.0607	0.3377	0.9341	1.0000

Pearson ペア比較

		AB	BB	AC	BC	CC	AD	BD	CD	DD
Not at all satisfied	1.0000	0.4316	1.0000	0.4870	0.1948	1.0000	0.8878	0.4763	0.6496	1.0000
Somewhat satisfied	1.0000	0.1246	1.0000	0.1086	0.7931	1.0000	0.0774	0.5820	0.7740	1.0000
Extremely satisfied	1.0000	0.2524	1.0000	0.0465	0.3360	1.0000	0.0585	0.3364	0.9341	1.0000

Fisher正確 ペア比較

		AB	BB	AC	BC	CC	AD	BD	CD	DD
Not at all satisfied	1.0000	0.5235	1.0000	0.5853	0.2990	1.0000	1.0000	0.5912	0.7221	1.0000
Somewhat satisfied	1.0000	0.1302	1.0000	0.1373	0.8909	1.0000	0.0983	0.6493	0.8685	1.0000
Extremely satisfied	1.0000	0.2764	1.0000	0.0531	0.4034	1.0000	0.0649	0.3574	1.0000	1.0000

ペア比較の *p* 値が、表形式で表示されます。これらのp値は、尤度比検定、Pearsonのカイ2乗検定、およびFisherの正確検定のものです。p値の表には、アルファベット文字によって、どのグループとどのグループが比較されているかが示されています。クロス表のグループ名のすぐ右側に、そのグループを示すアルファベット文字が表示されています。また、クロス表のセルには、検定結果を示すアルファベット文字が表示されています。

勤続年数5年未満の回答者は、勤続年数20年の回答者よりも、「Somewhat satisfied（やや満足）」の割合が高いです。このことは、クロス表の最初の行における「Somewhat satisfied」のセルに、「d」と示されていることから分かります。また、クロス表の最終行にある「Extremely satisfied（かなり満足）」のセルには、「a」と示されています。Dグループは、Aグループよりも、「Extremely satisified」の割合が低いです。なお、検定結果を示すアルファベット文字は、応答のシェア（割合）が高いほうのセルに配置されます。

分析者が興味のある比較を指定する例

この例では、「Consumer Preferences.jmp」を使用します。このデータには、一般的な事柄に対する意見や態度が含まれています。「被雇用期間」グループによって、「仕事熱心である」かどうかの分布が異なるかどうかを調べてみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「仕事熱心である」**を選択して、［単純］タブの**「応答」**をクリックします。
4. **「被雇用期間」**を選択して、［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。
6. 「Responses （応答）」の赤い三角ボタンのメニューをクリックし、［**文字の表示**］を選択します。
7. 「Responses （応答）」の赤い三角ボタンのメニューをクリックし、［**比較の設定**］を選択します。
8. 「A/B, B/C, C/D」と入力します。
9. 「カテゴリカル」の赤い三角ボタンのメニューをクリックして、［**応答の等質性に対する検定**］を選択します。

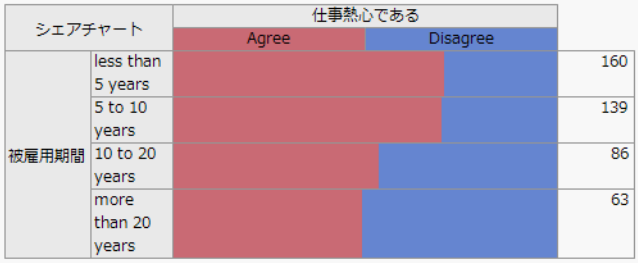
図3.13 比較の設定の例

Responses(仕事熱心である) By Employee Tenure

度数 / シェア / 比較			仕事熱心である		
			Agree	Disagree	応答の合計数
被雇用期間	less than 5 years	A	113 70.6%	47 29.4%	160
	5 to 10 years	B	97 69.8% C	42 30.2%	139
	10 to 20 years	C	46 53.5% *	40 46.5% B*	86
	more than 20 years	D	31 49.2% *	32 50.8% *	63

ユーザが指定した比較: A/B, B/C, C/D

ペア比較で有意に異なっている水準が、後のほうの水準で表示されています
* 星1つで警告する度数の基準 100 大文字で示す有意水準 0.05
** 星2つで警告する度数の基準 30 小文字で示す有意水準 0.1



応答の等質性に対する検定

応答の列	標本の列	尤度比 カイ2乗	尤度比 p値	Pearson カイ2乗	Pearson p値
仕事熱心である	AB, 被雇用期間	0.02515	0.8740	0.02515	0.8740
仕事熱心である	BC, 被雇用期間	6.04171	0.0140*	6.09112	0.0136*
仕事熱心である	CD, 被雇用期間	0.267	0.6054	0.26699	0.6054

応答の等質性に対する検定では、AとB、BとC、そしてCとDを比較します。グループB（5～10年）は、グループC（10～20年）よりも「仕事熱心である」という質問に同意する割合が高いです。Pearsonのp値が0.0136であるため、これらの割合には有意な差があると言えます。

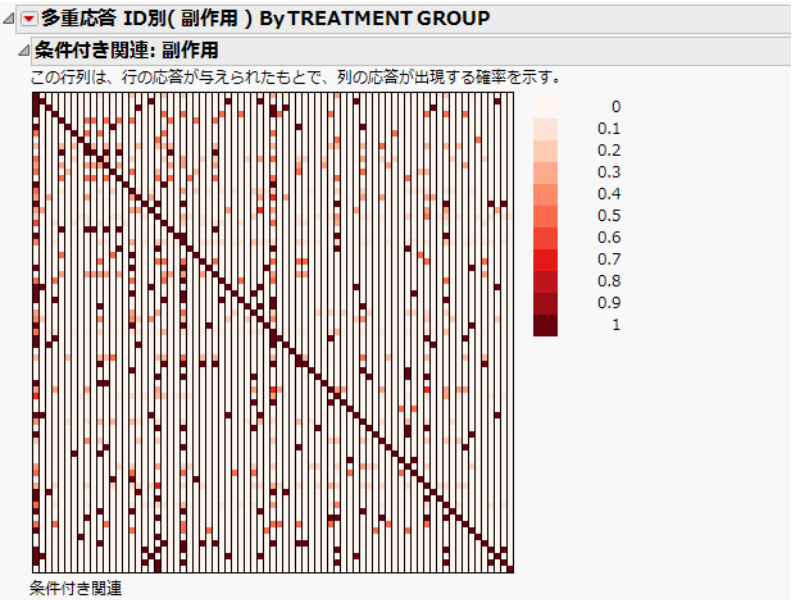
条件付き関連と相対リスクの例

この例では「AdverseR.jmp」を使用します。このデータには、臨床試験での有害事象（副作用）についての情報が含まれています。このデータを使用して、有害事象の条件付き関連を調べてみましょう。また、対照群に対する処置群の相対リスクを求めてみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「AdverseR.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. 「**副作用**」を選択して、［多重］タブの［**多重応答 ID 別**］をクリックします。

4. 「**処置グループ**」を選択して［**X, グループ化カテゴリ**］をクリックします。
5. 「**患者の ID**」を選択し、［**ID**］をクリックします。
6. ［**ID 内で一意な値をカウント**］を選択して、［**OK**］をクリックします。
7. 「カテゴリカル」の赤い三角ボタンのメニューをクリックして、［**条件付き関連**］を選択します。

図 3.14「条件付き関連」レポート（部分レポート）



条件付き関連

	ABDOMINAL PAIN	ABNORMAL VISION	ALOPECIA	AMBLYOPIA	ANEMIA
ABDOMINAL PAIN	1.0000	0.0323	0.0323	0.0645	0.0000
ABNORMAL VISION	1.0000	1.0000	0.0000	0.0000	0.0000
ALOPECIA	1.0000	0.0000	1.0000	0.0000	0.0000
AMBLYOPIA	1.0000	0.0000	0.0000	1.0000	0.0000
ANEMIA	0.0000	0.0000	0.0000	0.0000	1.0000
ANGINA PECTORIS	1.0000	0.0000	0.0000	0.0000	0.0000
ANOREXIA	0.5000	0.0000	0.0000	0.0000	0.0000
ANXIETY	0.0000	0.0000	0.0000	0.0000	0.0000
ASTHENIA	0.2500	0.0000	0.0000	0.0000	0.2500
ASTHMA	0.0000	0.0000	0.0000	0.0000	1.0000
BACK PAIN	0.5000	0.0000	0.0000	0.0000	0.0000
BRONCHITIS	0.3333	0.0000	0.0000	0.0000	0.3333
CHEST PAIN	0.5000	0.0000	0.0000	0.0000	0.0000

条件付き関連行列は、ある副作用が発生したときの、別の副作用が発生する条件付き確率を示します。これらの条件付き確率は、全グループにおけるものです。腹痛（ABDOMINAL PAIN）が報告された患者においては、0.0323 の確率で視覚異常（ABNORMAL VISION）が報告されています。

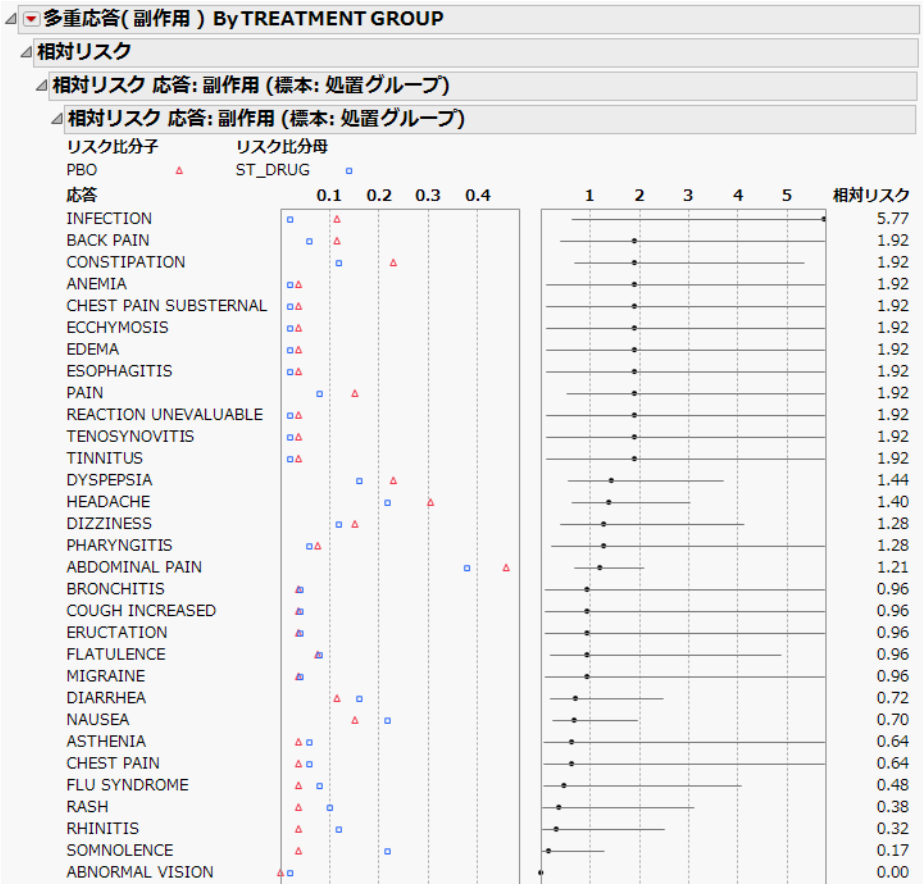
ヒント：ヒートマップにポインタを合わせると、条件付き確率が表示されます。

8. 「カテゴリカル」の赤い三角ボタンのメニューをクリックして、［**相対リスク**］を選択します。
9. 表示されるウィンドウで［**PBO**］を選択して、［**OK**］をクリックします。

10. ウィンドウ内の「相対リスク」レポートを右クリックして［**列の値で並べ替え**］を選択します。

11.「**相対リスク**」を選択して、［**OK**］をクリックします。

図3.15「相対リスク」レポート（一部）



［相対リスク］オプションによって、各応答の相対リスクが計算されます。デフォルトの「相対リスク」レポートには、応答名、各グループのリスク（割合）、相対リスク、相対リスクの95%信頼区間が表示されます。この例では、副作用の相対リスクが計算されています。感染症（INFECTION）が生じる、「ST_DRUG」に対する「PBO」の相対リスクは、5.7倍です。ただし、信頼区間が非常に広く、その信頼区間には1.0の相対リスクが含まれています。相対リスクが1.0となるのは、各グループのリスクが等しい場合です。

なお、［**列**］>［**下側95%信頼限界**］および［**列**］>［**上側95%信頼限界**］を右クリックして選択すると、相対リスク推定値の95%信頼区間がレポートに追加されます。

判定の一致性の例

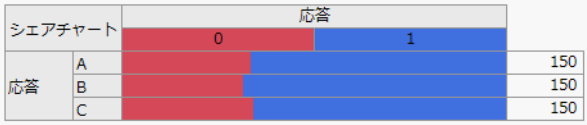
この例では「Attribute Gauge.jmp」を使用します。このデータは、3名の判定者が、50個の部品を3回、0／1で判定したものです。

1. ［ヘルプ］>［サンプルデータライブラリ］を選択し、「Attribute Gauge.jmp」を開きます。
2. ［分析］>［消費者調査］>［カテゴリカル］を選択します。
3. **A**、**B**、および**C**を選択します。
4. ［関連］タブで、［判定の一致性］をクリックします。
5. ［OK］をクリックします。

図3.16「判定の一致性」レポート

判定の一致性(A, B, C)

度数		応答		
シェア		0	1	応答の合計数
応答	A	50	100	150
		33.3%	66.7%	
	B	47	103	150
		31.3%	68.7%	
	C	51	99	150
		34.0%	66.0%	



判定の一致性統計量

一致性の度合い

一致性の度合い

応答1	応答2	カッパ	標準誤差	Bowkerの対称性	Bowkerのp値
A	B	0.862944	0.044198	1	0.3173
A	C	0.776119	0.054671	0.066667	0.7963
B	C	0.788007	0.053702	1.142857	0.2850

2x2表の場合、Bowker検定はMcNemar検定と等価です。

詳細

判定者の行	判定者の列	水準	0	1
B	A	0	44	3
B	A	1	6	97
C	A	0	43	8
C	A	1	7	92
C	B	0	42	9
C	B	1	5	94

カッパ値を見ると、3者の判定はかなり一致していることが分かります。カッパ値は、判定が独立のときに0となり、判定が完全に一致しているときに1.0となります。「詳細」アウトラインには、判定者の各ペアに対して2×2の分割表が表示されます。Bowkerの対称性検定は、「度数表が対角線に関して対称である」（つまり、「すべてのi、jに対して$p_{ij} = p_{ji}$である」）という帰無仮説に対する検定を行います。ここでのBowker検定のp値はすべて0.05よりも大きくなっており、対称性の帰無仮説は棄却されません。

反復測定の例

この例では 「Presidential Elections.jmp」を使用します。このデータには、1980年～2012年の米国の大統領選挙の各州における結果が含まれています。このデータを反復測定データとみなして、分析してみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Presidential Elections.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. **「1980年の勝者」**から**「2012年の勝者」**までを選択します。
4. ［関連］タブで、［**反復測定**］をクリックします。
5. ［**OK**］をクリックします。
6. レポートウィンドウの下のほうにあるグレーの「遷移レポート」の開閉アイコンをクリックして、遷移レポートを開きます。

図3.17 反復測定の「遷移レポート」

遷移レポート

すべて From	To	遷移度数			遷移率			
すべて 1980年の勝者	1984年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**1984年の勝者**
		Democrat	1	5	Democrat	1.0000	0.1020	
		Republican	0	44	Republican	0.0000	0.8980	
すべて 1984年の勝者	1988年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**1988年の勝者**
		Democrat	1	0	Democrat	0.1000	0.0000	
		Republican	9	40	Republican	0.9000	1.0000	
すべて 1988年の勝者	1992年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**1992年の勝者**
		Democrat	10	0	Democrat	0.3125	0.0000	
		Republican	22	18	Republican	0.6875	1.0000	
すべて 1992年の勝者	1996年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**1996年の勝者**
		Democrat	29	3	Democrat	0.9355	0.1579	
		Republican	2	16	Republican	0.0645	0.8421	
すべて 1996年の勝者	2000年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**2000年の勝者**
		Democrat	20	11	Democrat	1.0000	0.3667	
		Republican	0	19	Republican	0.0000	0.6333	
すべて 2000年の勝者	2004年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**2004年の勝者**
		Democrat	18	2	Democrat	0.9474	0.0645	
		Republican	1	29	Republican	0.0526	0.9355	
すべて 2004年の勝者	2008年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**2008年の勝者**
		Democrat	19	0	Democrat	0.6786	0.0000	
		Republican	9	22	Republican	0.3214	1.0000	
すべて 2008年の勝者	2012年の勝者		**Democrat**	**Republican**		**Democrat**	**Republican**	**2012年の勝者**
		Democrat	26	2	Democrat	1.0000	0.0833	
		Republican	0	22	Republican	0.0000	0.9167	

1980年～1984年において、「民主党」から「共和党」に遷移した州が5つありました。それらの5州は、1980年には「民主党」が選ばれていましたが、1984年には「共和党」が選ばれています。2008年～2012年において、「民主党」から「共和党」に遷移した州が2つありました。他のすべての州は、2008年と2012年の両方の選挙で同じ党が選ばれています。

［多重応答］タブの例

次に示すいくつかの例では、同じ情報を異なる形式で保持している、いくつかのサンプルデータを分析します。それらのサンプルデータは、異なる2つの条件において、3日間にわたり、生産ラインを検査したものです。各条件において、50ユニットが検査されました。7種類の不適合について、検査されています。1つのユニットにおいて、不適合が1つもない場合もあれば、複数の不適合が見つかる場合もあります。また、1つのユニットで、同じ種類の不適合が複数、生じる場合もあります。

多重応答

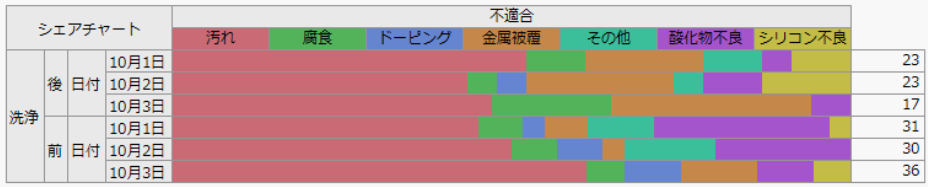
「Failure3MultipleField.jmp」のデータには、1ユニットが1行に保存されており、また、不適合データを含む列が複数あります。この例では、それらの列は3列になっています。これは、どのユニットも、多くて3つまでの不適合しか生じていないことを意味しています。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Quality Control」フォルダ内の「Failure3MultipleField.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. 「**不適合1**」、「**不適合2**」、および「**不適合3**」を選択します。
4. ［多重］タブの［**多重応答**］をクリックします。
5. 「**洗浄**」と「**日付**」を選択し、［**X, グループ化カテゴリ**］をクリックします。
6. ［**OK**］をクリックします。

図3.18「多重応答」レポート

多重応答(不適合1, 不適合2, 不適合3) By clean*date

			度数 / シェア / ケース	不適合									
				汚れ	腐食	ドーピング	金属被覆	その他	酸化物不良	シリコン不良	応答の合計数	全ケース数	応答している全ケース数
洗浄	後	日付	10月1日	12	2	0	4	2	1	2	23	50	23
				52.2%	8.7%	0.0%	17.4%	8.7%	4.3%	8.7%			
				24.0%	4.0%	0.0%	8.0%	4.0%	2.0%	4.0%			
			10月2日	10	1	1	5	1	2	3	23	50	23
				43.5%	4.3%	4.3%	21.7%	4.3%	8.7%	13.0%			
				20.0%	2.0%	2.0%	10.0%	2.0%	4.0%	6.0%			
			10月3日	8	3	0	5	0	1	0	17	50	17
				47.1%	17.6%	0.0%	29.4%	0.0%	5.9%	0.0%			
				16.0%	6.0%	0.0%	10.0%	0.0%	2.0%	0.0%			
	前	日付	10月1日	14	2	1	2	3	8	1	31	50	28
				45.2%	6.5%	3.2%	6.5%	9.7%	25.8%	3.2%			
				28.0%	4.0%	2.0%	4.0%	6.0%	16.0%	2.0%			
			10月2日	15	2	2	1	4	6	0	30	50	30
				50.0%	6.7%	6.7%	3.3%	13.3%	20.0%	0.0%			
				30.0%	4.0%	4.0%	2.0%	8.0%	12.0%	0.0%			
			10月3日	22	2	3	4	0	3	2	36	50	36
				61.1%	5.6%	8.3%	11.1%	0.0%	8.3%	5.6%			
				44.0%	4.0%	6.0%	8.0%	0.0%	6.0%	4.0%			



クロス表は、バッチごとに１つの行、不適合の種類ごとに１つの列で構成されています。各セルには、度数、応答のシェア、１ユニットあたりの比率（１ケースあたりの比率）が表示されます。たとえば、「10月１日」の「洗浄後」におけるバッチでは、「汚れ」における問題が12個ありました。そのシェアは、このバッチ全体の12/23、つまり52.2%になっています。また、50個のユニットにおいて、「汚れ」の不適合は12だけ生じているので、１ユニットあたりの比率（「ケース」）は24%（=12/50）となります。

多重応答 ID別

「Failure3ID.jmp」のデータは、バッチにおける１つの不適合ごとに１つの行となっており、各不適合の種類を示す列と、バッチのIDを示す列で構成されています。

図3.19「Failure3ID」データテーブル（一部）

	不適合	度数	洗浄	日付	標本サイズ	ID
1	汚れ	14	前	10月1日	50	OCT 1 before
2	腐食	2	前	10月1日	50	OCT 1 before
3	ドーピング	1	前	10月1日	50	OCT 1 before
4	金属被覆	2	前	10月1日	50	OCT 1 before
5	その他	3	前	10月1日	50	OCT 1 before
6	酸化物不良	8	前	10月1日	50	OCT 1 before
7	シリコン不良	1	前	10月1日	50	OCT 1 before
8	ドーピング	0	後	10月1日	50	OCT 1 after
9	腐食	2	後	10月1日	50	OCT 1 after
10	金属被覆	4	後	10月1日	50	OCT 1 after

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Quality Control」フォルダ内の「Failure3ID.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. 「**不適合**」を選択し、［多重］タブの［多重応答 ID 別］をクリックします。
4. 「**洗浄**」と「**日付**」を選択し、［**X, グループ化カテゴリ**］をクリックします。
5. 「**標本サイズ**」を選択し、［**標本サイズ**］をクリックします。
6. 「**度数**」を選択し、［**度数**］をクリックします。
7. 「**ID**」を選択し、［**ID**］をクリックします。
8. ［**OK**］をクリックします。

結果のレポートは、図3.18で表示されるレポートと同じです。

多重応答 区切り文字

「Failures3Delimited.jmp」のデータは、ユニットごとに1つの行があり、不適合の列には、不適合の種類がカンマで区切られて記録されています。図3.20のデータテーブルでわかるように、ユニットの一部は「**不適合**」列が空になっています。これは、不適合がまったく観測されなかったことを示します。

図3.20「Failure3Delimited.jmp」データテーブル（一部）

	不適合	洗浄	日付	ID	IDラベル
1		前	10月1日	1	10月1日 前
2	酸化物不良	前	10月1日	1	10月1日 前
3	汚れ,酸化物不良	前	10月1日	1	10月1日 前
4		前	10月1日	1	10月1日 前
5	汚れ	前	10月1日	1	10月1日 前
6	酸化物不良	前	10月1日	1	10月1日 前
7	汚れ	前	10月1日	1	10月1日 前
8		前	10月1日	1	10月1日 前
9		前	10月1日	1	10月1日 前
10	金属被覆,汚れ	前	10月1日	1	10月1日 前

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Quality Control」フォルダ内の「Failures3Delimited.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. 「**不適合**」を選択し、［多重］タブの［**多重応答 区切り文字**］をクリックします。
4. 「**洗浄**」と「**日付**」を選択し、［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。

［**OK**］をクリックすると、図3.18と同じレポートが表示されます。

メモ：区切り文字を使った列を複数指定した場合、列ごとに分析が行われます。

指示変数

「Failures3Indicators.jmp」のデータは、ユニットごとに1つの行と、各不適合の種類に対する指示変数の列で構成されています。各不適合列のデータは、該当のユニットに不適合が観測されなかった場合には「0」、不適合が観察された場合には「1」が入力されます。

図3.21「Faliure3Indicators.jmp」データテーブル（一部）

	洗浄	日付	ID	IDラベル	汚れ	腐食	ドーピング	金属被覆	その他	酸化物不良	シリコン不良
1	前	10月1日	1	10月1日 前	0	0	0	0	0	0	0
2	前	10月1日	1	10月1日 前	0	0	0	0	0	1	0
3	前	10月1日	1	10月1日 前	1	0	0	0	0	1	0
4	前	10月1日	1	10月1日 前	0	0	0	0	0	0	0
5	前	10月1日	1	10月1日 前	1	0	0	0	0	0	0
6	前	10月1日	1	10月1日 前	0	0	0	0	0	1	0

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Quality Control」フォルダ内の「Failures3Indicators.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. 「**汚れ**」、「**腐食**」、「**ドーピング**」、「**金属被覆**」、「**その他**」、「**酸化物不良**」、「**シリコン不良**」を選択し、［多重］タブの［**指示変数**］をクリックします。
4. 「**洗浄**」と「**日付**」を選択し、［**X, グループ化カテゴリ**］をクリックします。
5. ［**OK**］をクリックします。

［**OK**］をクリックすると、図3.18と同じレポートが表示されます。

応答の度数

「Failure3Freq.jmp」のデータは、バッチごとに1つの行、各不適合の種類ごとに複数の列、バッチサイズについての1つの列で構成されています。各不適合列には、バッチにおける不適合の発生度数が入力されています。

図3.22「Failure3Freq.jmp」データテーブル

	洗浄	日付	汚れ	腐食	ドーピング	金属被覆	その他	酸化物不良	シリコン不良	標本サイズ
1	後	10月1日	12	2	0	4	2	1	2	50
2	後	10月2日	10	1	1	5	1	2	3	50
3	後	10月3日	8	3	0	5	0	1	0	50
4	前	10月1日	14	2	1	2	3	8	1	50
5	前	10月2日	15	2	2	1	4	6	0	50
6	前	10月3日	22	2	3	4	0	3	2	50

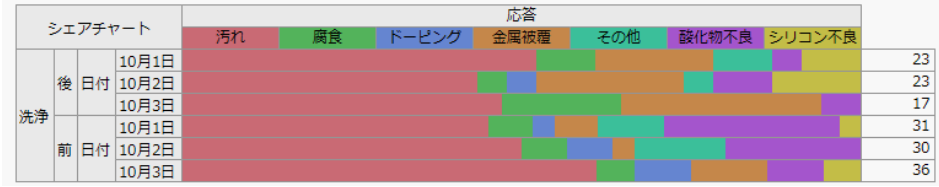
1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Quality Control」フォルダ内の「Failure3Freq.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. 度数変数（「**汚れ**」、「**腐食**」、「**ドーピング**」、「**金属被覆**」、「**その他**」、「**酸化物不良**」、「**シリコン不良**」）を選択します。
4. ［多重］タブの［**応答の度数**］をクリックします。
5. 「**洗浄**」と「**日付**」を選択し、［**X, グループ化カテゴリ**］をクリックします。
6. 「**標本サイズ**」を選択し、［**標本サイズ**］をクリックします。
7. ［**OK**］をクリックします。

図3.23 不適合数の表

標本サイズ: 標本サイズ

応答の度数(汚れ, 腐食, ドーピング, 金属被覆, その他, 酸化物不良, シリコン不良) By clean*date

			度数 / シェア / ケース	応答									
				汚れ	腐食	ドーピング	金属被覆	その他	酸化物不良	シリコン不良	応答の合計数	全ケース数	応答している全ケース数
洗浄	後	日付	10月1日	12	2	0	4	2	1	2	23	50	50
				52.2%	8.7%	0.0%	17.4%	8.7%	4.3%	8.7%			
				24.0%	4.0%	0.0%	8.0%	4.0%	2.0%	4.0%			
			10月2日	10	1	1	5	1	2	3	23	50	50
				43.5%	4.3%	4.3%	21.7%	4.3%	8.7%	13.0%			
				20.0%	2.0%	2.0%	10.0%	2.0%	4.0%	6.0%			
			10月3日	8	3	0	5	0	1	0	17	50	50
				47.1%	17.6%	0.0%	29.4%	0.0%	5.9%	0.0%			
				16.0%	6.0%	0.0%	10.0%	0.0%	2.0%	0.0%			
	前	日付	10月1日	14	2	1	2	3	8	1	31	50	50
				45.2%	6.5%	3.2%	6.5%	9.7%	25.8%	3.2%			
				28.0%	4.0%	2.0%	4.0%	6.0%	16.0%	2.0%			
			10月2日	15	2	2	1	4	6	0	30	50	50
				50.0%	6.7%	6.7%	3.3%	13.3%	20.0%	0.0%			
				30.0%	4.0%	4.0%	2.0%	8.0%	12.0%	0.0%			
			10月3日	22	2	3	4	0	3	2	36	50	50
				61.1%	5.6%	8.3%	11.1%	0.0%	8.3%	5.6%			
				44.0%	4.0%	6.0%	8.0%	0.0%	6.0%	4.0%			



クロス表の［応答している全ケース数］列を除き、図3.18と同じ結果です。この例では、分析対象の不適合データがすでに要約されていました。この要約されたデータには、不適合が0であるユニット数の記録はありません。そのため、この例の［応答している全ケース数］には、バッチサイズである50が表示されています。

スコア平均を比較する例

この例では「Consumer Preferences.jmp」を使用します。このデータで、被雇用期間と、学齢期の子供がいることの関係を調べます。**「被雇用期間」**列には、便宜的な通し番号として、1,2,3,4の数値が含まれています。そして、これらの値に対して、「値ラベル」列プロパティによって、ラベルが割り当てられています。「カテゴリカル」プラットフォームの［スコア平均］オプションで、被雇用期間の平均を求めるには、この列の値（1,2,3,4）に、被雇用期間を「値スコア」として割り当てる必要があります。列プロパティについての詳細は、『JMPの使用法』の「列情報ウィンドウ」章を参照してください。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. データテーブルで、**「被雇用期間」**列見出しを右クリックし、［**列プロパティ**］>［**値スコア**］を選択します。
3. **「値」**に1、**「スコア」**に3を入力してから、［**追加**］をクリックします。
4. **「値」**に2、**「スコア」**に7.5を入力してから、［**追加**］をクリックします。
5. **「値」**に3、**「スコア」**に15を入力してから、［**追加**］をクリックします。

6. 「**値**」に4、「**スコア**」に25を入力してから、［**追加**］をクリックします。
7. ［**OK**］をクリックします。
8. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
9. 「**被雇用期間**」を選択して、［単純］タブの［**応答**］をクリックします。
10. 「**学齢期の子供**」を選択して［**X, グループ化カテゴリ**］をクリックします。
11. ［**OK**］をクリックします。
12. 「カテゴリカル」の赤い三角ボタンをクリックして、［**スコアの平均**］を選択します。
13. 「カテゴリカル」の赤い三角ボタンをクリックして、［**スコア平均の比較**］を選択します。

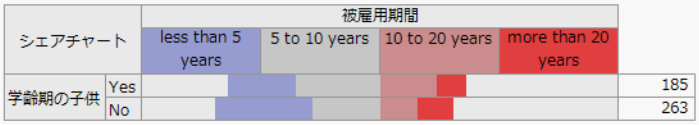
図3.24 スコアの平均が表示された「カテゴリカル」レポート

Responses(被雇用期間) By School Age Children

度数 / シェア			被雇用期間 less than 5 years	5 to 10 years	10 to 20 years	more than 20 years	応答の合計数	平均	平均の比較
学齢期の子供	Yes	A	53 28.6%	65 35.1%	44 23.8%	23 12.4%	185	10.170	
	No	B	107 40.7%	74 28.1%	42 16.0%	40 15.2%	263	9.529	

デフォルトの比較: A/B

ペア比較で有意に異なっている水準が、後のほうの水準で表示されています
* 星1つで警告する度数の基準 100 大文字で示す有意水準 0.05
** 星2つで警告する度数の基準 30 小文字で示す有意水準 0.1



学齢期の子供を持つ人の被雇用期間の平均は10.17、学齢期の子供がいない人の平均は9.53です。クロス表における［平均の比較］列が空なので、統計的に有意差があるとは言えません。もしいずれかのグループ間に有意差があるならば、その有意差があるグループにはアルファベット文字が表示されます。

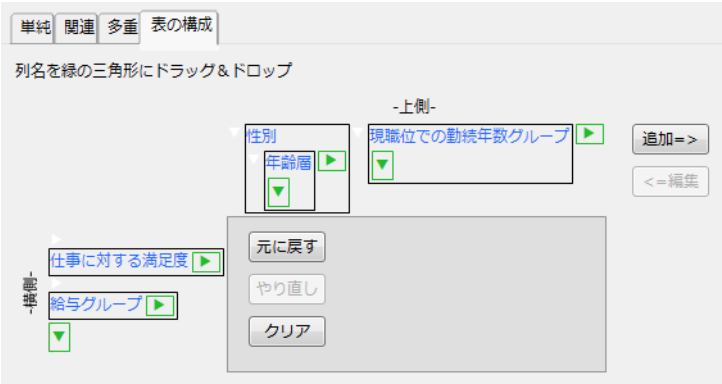
ヒント：スコア平均オプションを使用している場合は、データの記録方法に注意してください。データが値ラベルでコード化された数値で保存されている場合、平均値は、その数値データに基づいて計算されます。その数値データに意味がない場合、「**値スコア**」を使って、各水準に意味のあるスコア値を割り当ててください。

「表の構成」レポートの例

この例では「Consumer Preferences.jmp」を使用します。このデータで、仕事に対する満足度を、［表の構成］タブを使って、性別×年齢層、および現職での勤続年数で比較して調べます。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Consumer Preferences.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**カテゴリカル**］を選択します。
3. ［表の構成］タブを選択します。
4. **「性別」**を表の**「上側」**にある緑のドロップゾーンにドラッグします。
5. **「年齢層」**を**「性別」**の下にある緑のドロップゾーンにドラッグします。
6. **「現職位での勤続年数グループ」**を**「上側」**の**「性別」**の隣にある緑のドロップゾーンにドラッグします。
7. **「仕事に対する満足度」**を**「横側」**にある緑のドロップゾーンにドラッグします。
8. **「給与グループ」**を**「横側」**の**「仕事に対する満足度」**の下にある緑のドロップゾーンにドラッグします。

図3.25［表の構成］タブのレポートの設定



9. ［**追加 =>**］をクリックします。
10. ［**OK**］をクリックします。
11. 「カテゴリカル」の赤い三角ボタンをクリックして、［**応答の等質性に対する検定**］を選択します。

図3.26［表の構成］タブのレポートの例

仕事に対する満足度+給与グループ By 性別*年齢層+現職位での勤続年数グループ

度数		性別																
シェア		M							F									
		年齢層							年齢層							現職位での勤続年数グループ			
		25-29	30-34	35-39	40-44	45-49	50-54	>54	25-29	30-34	35-39	40-44	45-49	50-54	>54	less than 5 years	5 to 10 years	10 to 20 years	more than 20 years
仕事に対する満足度	Not at all satisfied	3	3	1	3	3	2	2	3	1	0	5	2	1	3	15	11	6	0
		5.5%	10.3%	3.3%	8.8%	7.9%	8.3%	4.1%	5.2%	2.6%	0.0%	27.8%	14.3%	5.3%	13.6%	7.2%	8.2%	6.8%	0.0%
	Somewhat satisfied	32	14	15	19	19	10	26	34	30	12	6	5	11	9	122	70	40	10
		58.2%	48.3%	50.0%	55.9%	50.0%	41.7%	53.1%	58.6%	76.9%	63.2%	33.3%	35.7%	57.9%	40.9%	58.7%	52.2%	45.5%	55.6%
	Extremely satisfied	20	12	14	12	16	12	21	21	8	7	7	7	7	10	71	53	42	8
		36.4%	41.4%	46.7%	35.3%	42.1%	50.0%	42.9%	36.2%	20.5%	36.8%	38.9%	50.0%	36.8%	45.5%	34.1%	39.6%	47.7%	44.4%
	応答の合計数	55	29	30	34	38	24	49	58	39	19	18	14	19	22	208	134	88	18
給与グループ	less than 40000	20	6	4	8	6	5	10	35	19	6	6	2	7	4	80	40	15	3
		36.4%	20.7%	13.3%	23.5%	15.8%	20.8%	20.4%	60.3%	48.7%	31.6%	33.3%	14.3%	36.8%	18.2%	38.5%	29.9%	17.0%	16.7%
	40000 to 60000	16	14	11	12	14	7	13	15	10	8	6	7	6	12	67	38	37	9
		29.1%	48.3%	36.7%	35.3%	36.8%	29.2%	26.5%	25.9%	25.6%	42.1%	33.3%	50.0%	31.6%	54.5%	32.2%	28.4%	42.0%	50.0%
	60000 to 80000	8	5	7	7	8	8	10	5	5	3	3	3	4	5	32	25	19	5
		14.5%	17.2%	23.3%	20.6%	21.1%	33.3%	20.4%	8.6%	12.8%	15.8%	16.7%	21.4%	21.1%	22.7%	15.4%	18.7%	21.6%	27.8%
	80000 - 120000	9	3	6	4	7	3	7	1	2	1	2	1	1	0	19	17	10	1
		16.4%	10.3%	20.0%	11.8%	18.4%	12.5%	14.3%	1.7%	5.1%	5.3%	11.1%	7.1%	5.3%	0.0%	9.1%	12.7%	11.4%	5.6%
	greater than 120000	2	1	2	3	3	1	9	2	3	1	1	1	1	1	10	14	7	0
		3.6%	3.4%	6.7%	8.8%	7.9%	4.2%	18.4%	3.4%	7.7%	5.3%	5.6%	7.1%	5.3%	4.5%	4.8%	10.4%	8.0%	0.0%
	応答の合計数	55	29	30	34	38	24	49	58	39	19	18	14	19	22	208	134	88	18

応答の等質性に対する検定

応答の列	標本の列	尤度比 カイ2乗	尤度比 p値	Pearson カイ2乗	Pearson p値
仕事に対する満足度	性別 = M, 年齢層	4.65556	0.9685	4.59049	0.9703
仕事に対する満足度	性別 = F, 年齢層	23.8406	0.0214*	24.9917	0.0149*
給与グループ	性別 = M, 年齢層	22.0707	0.5750	23.8074	0.4727
給与グループ	性別 = F, 年齢層	27.6209	0.2764	26.3992	0.3332
仕事に対する満足度	現職位での勤続年数グループ	8.00461	0.2378	6.76025	0.3436
給与グループ	現職位での勤続年数グループ	25.8336	0.0113*	24.0752	0.0199*

［表の構成］タブのレポートには、［表の構成］タブで指定した表が含まれます。応答の等質性に対する検定は、グループ変数のそれぞれの組み合わせに対して行われます。男性の場合、年齢層間には仕事に対する満足度に関して統計的有意差がありません（Pearson の p 値 = 0.9703）。女性の場合、年齢層間には仕事に対する満足度に関して統計的有意差があります（Pearson の p 値 = 0.0149）。中年女性は、仕事に対する満足度が最も低い傾向があります。これらの結果を視覚的に見るために、シェアチャートや度数チャートをレポートに追加してもよいでしょう。

「カテゴリカル」プラットフォームの統計的詳細

Rao-Scott 修正

多重応答データ（複数回答データ）における［応答の等質性に対する検定］では、Rao-Scott 修正が適用されます。Rao-Scott 修正については、Lavassani et al.（2009）を参照してください。なお、多重応答データにおいては、［応答の等質性に対する検定］オプションは、［表の構成］タブでのみ使うことができます。

多重応答データでは、データの重複が許されています。つまり、1 人の回答者から複数の応答を得ることが許されています。Pearson のカイ 2 乗検定は、多重応答データに適していません。なぜなら、多重応答データは、Pearson のカイ 2 乗検定の前提となっている、データの独立性を満たさないためです。また、周辺和は多重応答が許されない場合よりも大きくなるため、周辺和から計算される期待値も影響を受けます。

Rao-Scottのカイ２乗統計量は、次のように定義されます。

$$\chi_C^{\ 2} = \frac{\chi^2}{\bar{\delta}}$$

この式で、

χ^2 は、通常のPearsonカイ２乗検定統計量です。$\bar{\delta}$ は、次式で定義される補正係数です。

$$\bar{\delta} = 1 - \frac{m_{++}}{n_+ C}$$

この式で、

m_{++} は、多重応答の総度数です。

n_+ は、回答者（ケース）の総数です。

C は、応答変数の水準数（クロス表における列数）です。

自由度は、（R-1）C（つまり、行数から１を引いたものに、列数を掛けたもの）です。

第 4 章

選択モデル

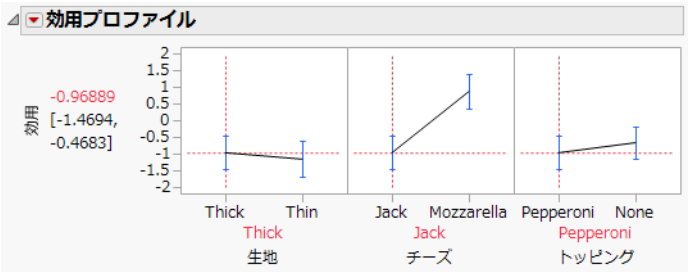
選択モデルのあてはめ

「選択モデル」プラットフォームは、マーケットリサーチ（市場調査）のひとつである選択実験（choise experiment）で得られたデータを分析します。選択実験は、製品やサービスのどのような属性を消費者が好むかを明らかにするための実験です。分析結果から、消費者に好まれる製品やサービスを設計することができます。

「選択モデル」プラットフォームでは次のようなことができます。

- 製品の属性だけでなく、消費者（被験者）の属性（たとえば性別や年齢）に関する情報も利用する。
- 「どの製品も選ばない」という選択肢が用意されている実験を分析する。
- 分析対象のデータにおいて、いくつかの形式がサポートされている。
- プロファイルを使って消費者の効用を視覚化・最適化する。
- 消費者ごとにスコアを求め、求められたスコアを用いて消費者をセグメント（クラスター）に分ける。
- JMP PRO Bayes（ベイズ）流の方法で、消費者ごとの係数値を求める。
- バイアス修正を伴う最尤推定（Firth 1993）を行う。

図 4.1「選択モデル」プラットフォームの効用プロファイル



目次

「選択モデル」プラットフォームの概要 71
「選択モデル」プラットフォームの例 72
「選択なし」がある1つのデータテーブル 73
複数のデータテーブル 75
「選択モデル」プラットフォームの起動 81
［1つのデータテーブル, 積み重ね］の起動ウィンドウ 82
［複数のデータテーブル, 相互参照］の起動ウィンドウ 84
「選択モデル」レポート 89
効果の要約 89
パラメータ推定値 90
尤度比検定 91
Bayesパラメータ推定値 91
「選択モデル」プラットフォームのオプション 93
支払意思額 96
「選択モデル」の別例 98
製品の仕様を決める分析例 98
セグメント化の例 107
「選択モデル」プラットフォームを使用したロジスティック回帰の例 111
対応のあるデータに対する条件付きロジスティック回帰の例 114
データを2つの分析テーブルに変換する例 116
データを1つの分析テーブルに変換する例 121
「選択モデル」プラットフォームの統計的詳細 124
分析対象のデータに対する特殊な規則 124
効用と確率 125
勾配 125

「選択モデル」プラットフォームの概要

McFadden（1974）が先駆者となって考案された選択モデルは、提示された選択肢の中から個人がどのような選択を行うかを推定する、強力な分析方法です。選択モデルは、「コンジョイントモデル」、「離散選択モデル」、「条件付きロジスティック回帰」とも呼ばれています。

選択実験では、製品やサービスの属性を組み合わせたもの（プロファイル）を回答者に提示して、どの組み合わせを回答者が好むかを尋ねます。選択実験では、いくつかのプロファイルを回答者に提示します （回答者に提示する、いくつかのプロファイルを集めたものを、「選択肢集合」と言います）。そして、回答者は、選択肢集合のなかから、最も好きなプロファイルを選択します。各回答者には、選択肢集合を提示して選択を行ってもらうことを、複数回、行ってもらうのが普通です。「選択モデル」プラットフォームは、このような選択実験のデータを分析します。

メモ: 選択実験は、「選択モデル計画」プラットフォームで作成できます。『実験計画 (DOE)』の「選択モデル計画」章を参照してください。

どの属性を重視するかは消費者によって異なるため、マーケットリサーチ関連の選択実験を分析する場合、消費者をセグメント（クラスター）に分けることが重要でしょう。セグメント化を行わないで製品やサービスを設計すると、実在しない「平均的な」消費者には好まれても、実在する消費者の嗜好を無視したものになるかもしれません。

選択モデルの背景については、Louviere et al.（2015）、Train（2009）、Rossi et al.（2005）を参照してください。

「選択モデル」プラットフォーム

「選択モデル」プラットフォームは、条件付きロジスティック回帰によって、ある製品が好まれる確率を推定します。「選択モデル」は、通常のロジスティック回帰とは異なり、消費者の属性（たとえば性別や年齢）だけでなく、製品の属性も考慮したモデルになっています。たとえば自動車を例にすると、価格・乗車人数・カップホルダーの数・色・カーナビ・燃費・盗難防止システム・取り外し可能なシート・安全機能の数・保険料といった、自動車の属性を選択モデルは考慮します。

「選択モデル」プラットフォームでは、選択肢の集合のいずれも回答者が選択しない場合も扱うことができます。この方法では、回答者がどれも選択しない状況を、単一の属性（「選択なし」という属性）をもつ1つの製品を選んだとみなしてモデル化されます。この「選択なし」に対するパラメータ推定値は、モデルの仮定に応じてさまざまな解釈があります。また、「選択モデル」プラットフォームでは、各回答者に関する情報を求めることもできます。これは回答者の嗜好パターンをセグメント化するのに役立ちます。

「選択モデル」プラットフォームでは、モデルを推定する方法として、Firth（1993）によるバイアス修正を伴う最尤推定も行えます。この推定法は、通常の最尤法に比べ、推定や検定がより良い性質をもちます。また、バイアス修正を伴う最尤推定法を用いることにより、ロジスティックモデルなどで生じる分離（separation）の問題が改善できます。ロジスティック回帰における分離問題については、Heinze and Schemper（2002）を参照してください。

メモ: 「選択モデル」プラットフォームは、順位やスコアに対するモデルや、枝分かれ階層の選択肢を含むモデルはサポートしていません。そういった分析には、SAS/ETSのPROC MDCを使用してください。

選択実験による製品・サービスの開発

満足度調査では、製品やサービスの評判が良いのか悪いのかを知ることはできますが、製品の各属性に対する人々の好みは明らかになりません。製品開発において、数百、数千という細かい仕様をエンジニアは決めていきます。その際、選択実験に協力してくれる消費者を確保できるのであれば、選択実験が製品開発の手助けとなるでしょう。

調査やプロトタイプ作成が安価に行えるのであれば、製品開発の段階で、多数の属性や代替案を消費者に評価してもらえるでしょう。選択実験をシックスシグマプログラムで実施することにより、多くの消費者に求められるように製品を改善できるでしょう。選択実験でデータを収集し、そのデータを分析することにより消費者の選好を明らかにできるでしょう。

セグメント化

マーケットリサーチ（市場調査）では、選好のパターンによって、消費者をグループに分けたい場合があります。しかし、通常のモデルで被験者ごとにパラメータ推定値を求めるには、被験者1人あたりに多くのデータが必要です。被験者1人あたりに多くのデータがあるなら、被験者IDを「応答データ」セクションで［By］変数に指定するか、被験者IDをモデル項として含めることができます。ただし、後者の方法は、被験者数が多いと計算が難しくなります。

［By］変数が指定できるほどのデータがない場合は、［被験者ごとの勾配を保存］オプションを使って被験者をクラスターに分け、セグメント化することができます。このオプションは、ヘッセ行列で尺度化した各パラメータの傾きを被験者ごとに平均し、新しいデータテーブルに表示します。例として、「セグメント化の例」（107ページ）を参照してください。傾きの値についての詳細は、「勾配」（125ページ）を参照してください。

JMP PRO また、JMP Proでは、階層型Bayesモデルもサポートされています。この階層型Bayesモデルでは、選択モデルにおける製品のパラメータ（このパラメータを、「部分効用 part worth」と呼ぶこともあります）の推定値が被験者ごとに計算されます。その結果をクラスター分析などを行えば、いくつかのセグメントに消費者を分類できます。

「選択モデル」プラットフォームの例

どのようなピザが好まれるかを調べた調査で、各回答者に2つのピザのどちらが好きかを、4回尋ねました。「選択モデル」プラットフォームでは、複数のデータテーブルに分かれたデータでも、1つのデータテーブルにまとめたデータでも分析できます。複数のデータテーブルを使う場合は、応答、選択肢集合、被験者に関する情報をそれぞれ個別のデータテーブルに保存しておきます。1つのデータテーブルを使う場合は、すべての情報を1つのデータテーブルにまとめます。

- 「「選択なし」がある1つのデータテーブル」（73ページ）に、1つのデータテーブルにまとめたデータに基づいて、選択モデル分析を行う操作方法を説明しています。
- 「複数のデータテーブル」（75ページ）に、複数のデータテーブルに記録されたデータに基づいて、選択モデル分析を行う操作方法を説明しています。

「選択なし」がある 1 つのデータテーブル

この例では、回答にどちらのピザも選ばなかった場合もあります。つまり、「選択なし」という回答です。どちらのピザも選ばなかった場合は、指示変数の値が欠測値になっています。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Pizza Combined No Choice.jmp」を開きます。

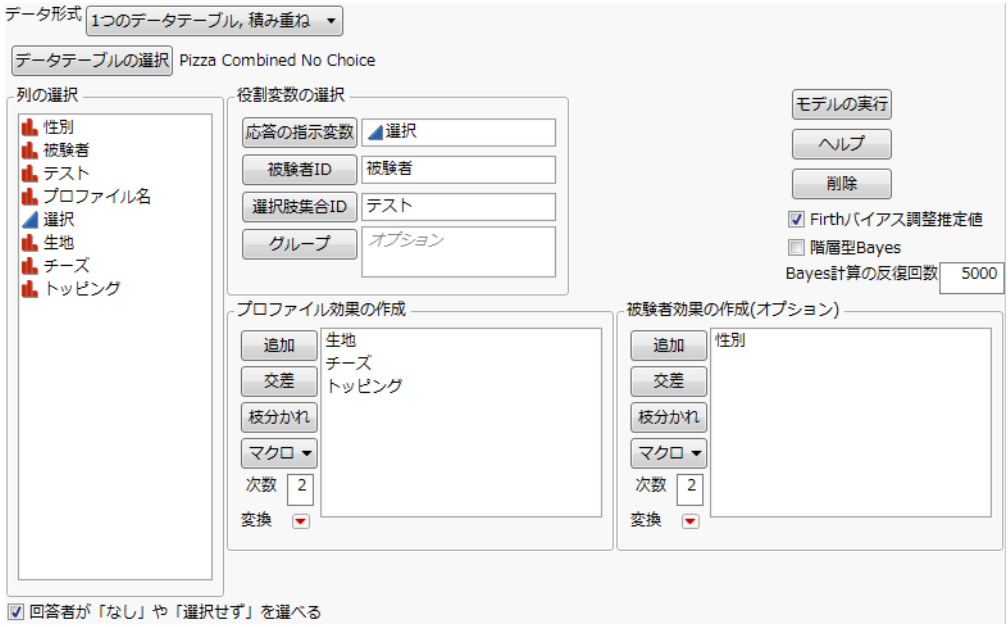
   選択肢集合は、「**被験者**」と「**テスト**」の組み合わせで定義します。一部の選択肢集合で、「**選択**」列が欠測値になっていることに注目してください。

2. ［**分析**］>［**消費者調査**］>［**選択モデル**］を選択します。

   ［1 つのデータテーブル , 積み重ね］がデフォルトです。

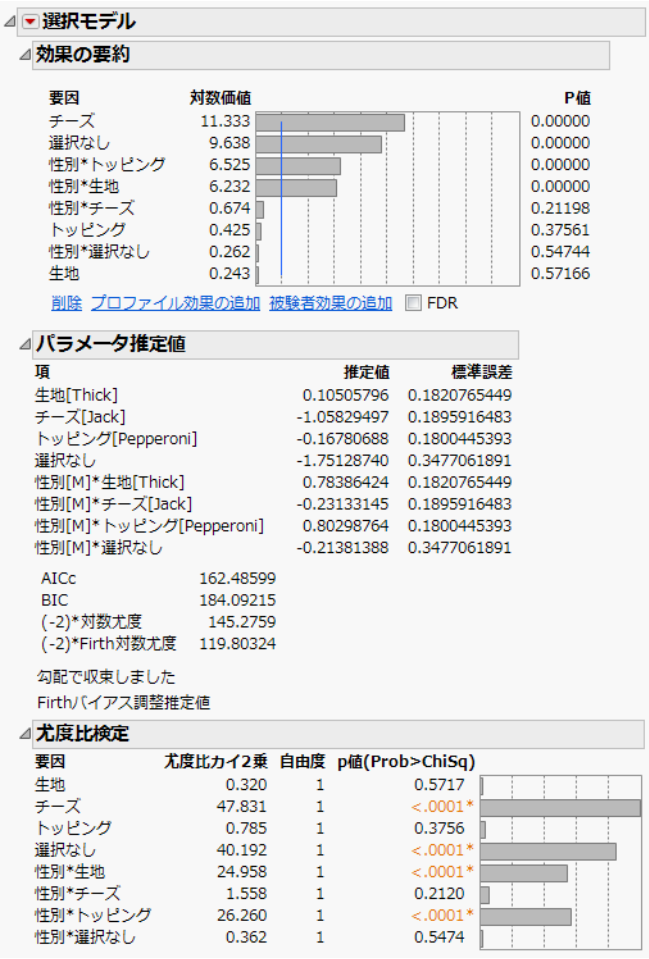
3. ［**データテーブルの選択**］をクリックします。
4. 「**Pizza Combined No Choice**」を選択し、［**OK**］をクリックします。
5. 起動ウィンドウで次のように入力を完了します。
   - 「**選択**」を選択し、［**応答の指示変数**］をクリックします。
   - 「**被験者**」を選択し、［**被験者 ID**］をクリックします。
   - 「**テスト**」を選択し、［**選択肢集合 ID**］をクリックします。
   - 「**生地**」、「**チーズ**」、「**トッピング**」を選択し、「プロファイル効果の作成」パネルで［**追加**］をクリックします。
   - 「**性別**」を選択し、「被験者効果の作成 (オプション)」パネルで［**追加**］をクリックします。

図 4.2 入力が完了した起動ウィンドウ



6. ［**回答者が「なし」や「選択せず」を選べる**］にチェックマークをつけます。
7. ［**モデルの実行**］をクリックします。

図4.3「選択なし」を効果として表示したレポート



選択モデル

効果の要約

要因	対数価値	P値
チーズ	11.333	0.00000
選択なし	9.638	0.00000
性別*トッピング	6.525	0.00000
性別*生地	6.232	0.00000
性別*チーズ	0.674	0.21198
トッピング	0.425	0.37561
性別*選択なし	0.262	0.54744
生地	0.243	0.57166

削除 プロファイル効果の追加 被験者効果の追加 FDR

パラメータ推定値

項	推定値	標準誤差
生地[Thick]	0.10505796	0.1820765449
チーズ[Jack]	-1.05829497	0.1895916483
トッピング[Pepperoni]	-0.16780688	0.1800445393
選択なし	-1.75128740	0.3477061891
性別[M]*生地[Thick]	0.78386424	0.1820765449
性別[M]*チーズ[Jack]	-0.23133145	0.1895916483
性別[M]*トッピング[Pepperoni]	0.80298764	0.1800445393
性別[M]*選択なし	-0.21381388	0.3477061891

AICc	162.48599
BIC	184.09215
(-2)*対数尤度	145.2759
(-2)*Firth対数尤度	119.80324

勾配で収束しました

Firthバイアス調整推定値

尤度比検定

要因	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
生地	0.320	1	0.5717
チーズ	47.831	1	<.0001*
トッピング	0.785	1	0.3756
選択なし	40.192	1	<.0001*
性別*生地	24.958	1	<.0001*
性別*チーズ	1.558	1	0.2120
性別*トッピング	26.260	1	<.0001*
性別*選択なし	0.362	1	0.5474

「効果の要約」レポートは、効果を有意性の順に表示しています。**「チーズ」**が最も有意な効果で、次に、モデル効果として扱われている「選択なし」が続きます。被験者効果の交互作用である**「性別*トッピング」**、**「性別*生地」**も有意なことから、**「トッピング」**と**「生地」**に対する好みが**「性別」**というマーケットセグメントに依存すると考えられます。

どんなときにいずれのピザも選択されなかったかを調べるために、元データを確認してみましょう。

8. データテーブルの**「選択」**列で、値が欠測値となっているセルを右クリックし、［**一致するセルを選択**］を選択します。
9. 「行」パネルで**「選択されている行」**を右クリックし、［**データビュー**］を選択します。

図4.4「選択なし」の応答を含む選択肢集合

	性別	被験者	テスト	プロファイル名	選択	生地	チーズ	トッピング
1	F	2	4	TrimPepperjack	•	Thin	Jack	Pepperoni
2	F	2	4	TrimOni	•	Thin	Mozzarella	Pepperoni
3	M	7	2	Trimella	•	Thin	Mozzarella	None
4	M	7	2	TrimJack	•	Thin	Jack	None
5	M	7	3	Trimella	•	Thin	Mozzarella	None
6	M	7	3	TrimJack	•	Thin	Jack	None
7	F	8	2	ThickElla	•	Thick	Mozzarella	None
8	F	8	2	ThickJack	•	Thick	Jack	None
9	M	11	4	ThickOni	•	Thick	Mozzarella	Pepperoni
10	M	11	4	ThickJackoni	•	Thick	Jack	Pepperoni
11	F	14	2	TrimOni	•	Thin	Mozzarella	Pepperoni
12	F	14	2	TrimPepperjack	•	Thin	Jack	Pepperoni
13	F	18	3	ThickJack	•	Thick	Jack	None
14	F	18	3	ThickElla	•	Thick	Mozzarella	None
15	F	24	3	ThickJack	•	Thick	Jack	None
16	F	24	3	TrimOni	•	Thin	Mozzarella	Pepperoni
17	M	29	1	TrimPepperjack	•	Thin	Jack	Pepperoni
18	M	29	1	ThickJackoni	•	Thick	Jack	Pepperoni

図4.4のデータテーブルにおいて、**「被験者」**と**「テスト」**の組み合わせで定義された最初の7つの選択肢集合（行1～14）を見てください。各選択肢集合において、2つのピザは、**「チーズ」**だけが異なっています。これらの回答者は、チーズの違いを見極められなかった可能性があります。しかし、分析結果においては、「選択なし」が考慮されるため、**「チーズ」**の効果が統計的に有意になっています。

このようなデータをさらに分析する方法については、「最適なプロファイルを見つける」（79ページ）を参照してください。

複数のデータテーブル

この節でも、3つの属性をもつピザを調べた実験を例にします。回答者は、好きなピザを必ず選択しなければならないものとします。ここでは、「Pizza Profiles.jmp」、「Pizza Responses.jmp」、「Pizza Subjects.jmp」の3つのデータテーブルを使用します。選択実験のデータは、常に1つのデータテーブルにまとめることもできますが、特に分析に含めたいプロファイル変数や被験者変数が他にもある場合は、複数のデータテーブルを使用する方が便利です。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Pizza Profiles.jmp」、「Pizza Responses.jmp」、「Pizza Subjects.jmp」を開きます。
   - プロファイルのデータテーブルである「Pizza Profiles.jmp」には、被験者に提示した選択肢がどのような属性のピザであるかの情報が含まれています。それぞれの選択肢に、**「ID」**を割り当てています。
   - 応答のデータテーブルである「Pizza Responses.jmp」には、提示した2つの選択肢と、どちらの選択肢を選択したかの結果が含まれています。実験では、各被験者に4つの選択肢集合が提示されます。各集合は、選択肢である2つのプロファイル（**「選択肢1」**と**「選択肢2」**）から成ります。被験者は、各選択肢集合の中から好きな方（**「選択」**）を選びます。選択モデル計画の作成方法については、『実験計

画 (DOE)』の「選択モデル計画」章を参照してください。応答のデータテーブルにおける「**選択**」列の値は、プロファイルのデータテーブルにおける「**ID**」列の値と対応しています。

– 被験者のデータテーブルである「Pizza Subjects.jmp」には、被験者IDを値とする「**被験者**」列と、被験者の特性である「**性別**」が保存されています。「Pizza Subjects.jmp」の「**被験者**」列の値は、「Pizza Responses.jmp」の「**被験者**」列の値に対応しています。

2. ［**分析**］>［**消費者調査**］>［**選択モデル**］を選択し、起動ウィンドウを開きます。

メモ: これは、開いている3つのデータテーブルのどれからでも実行できます。

3. 「データ形式」メニューから［**複数のデータテーブル, 相互参照**］を選択します。

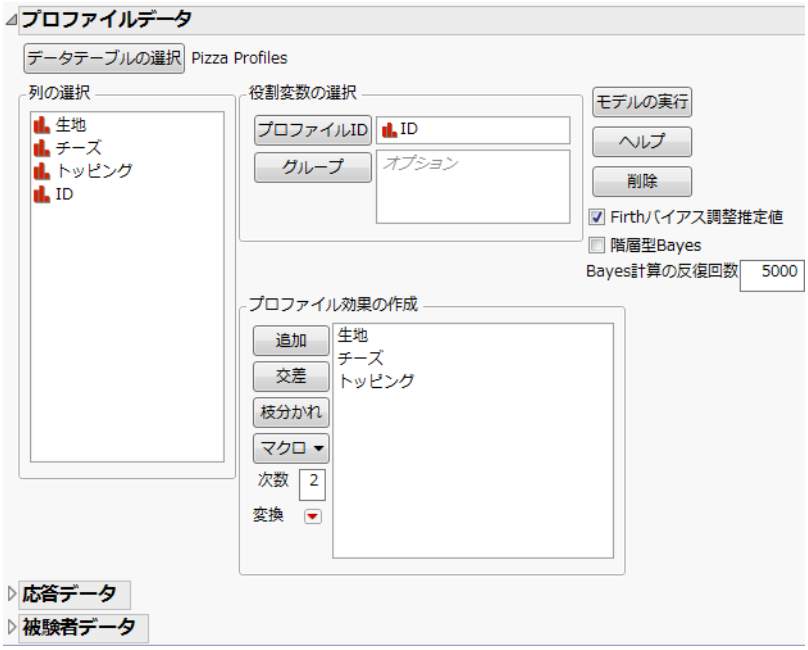
各データソースに対応する3つのアウトラインがあります。

4. 「プロファイルデータ」アウトラインにある［**データテーブルの選択**］をクリックします。

「プロファイルデータ」ウィンドウが開き、ここでプロファイルデータとして使うデータテーブルを指定します。

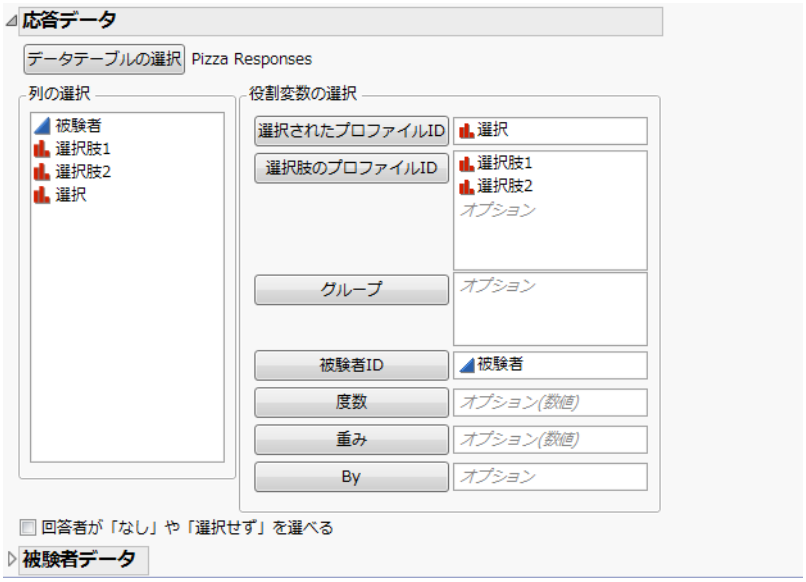
5. 「Pizza Profiles.jmp」を選択し、［**OK**］をクリックします。
6. 「**ID**」を選択し、［**プロファイルID**］をクリックします。
7. 「**生地**」、「**チーズ**」、「**トッピング**」を選択し、［**追加**］をクリックします。

図4.5 プロファイルデータ



8. 「応答データ」の開閉アイコンをクリックしてアウトラインを開き、［**データテーブルの選択**］をクリックします。
9. 「Pizza Responses.jmp」を選択し、［**OK**］をクリックします。
10. 次の手順を行います。
   – **「選択」**を選択し、［**選択されたプロファイル ID**］をクリックします。
   – **「選択肢 1」**と**「選択肢 2」**を選択し、［**選択肢のプロファイル ID**］をクリックします。
   – **「被験者」**を選択し、［**被験者 ID**］をクリックします。

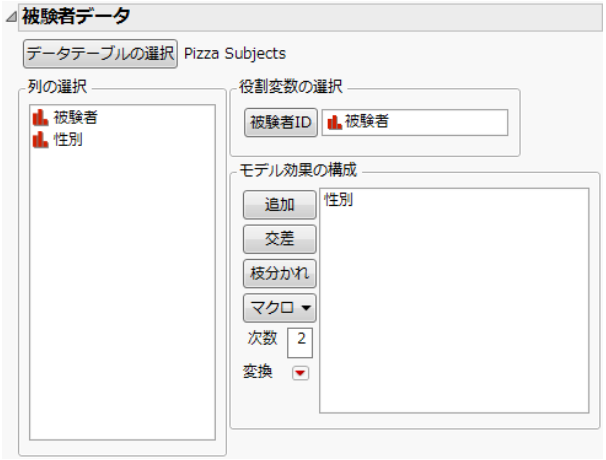
図 4.6「応答データ」パネル



「選択肢 1」と**「選択肢 2」**という列は、各被験者に 4 回ずつ提示される 2 つの選択肢を表しています。また、**「選択」**という列には、**「選択肢 1」**と**「選択肢 2」**のどちらが選択されたかが入力されています。

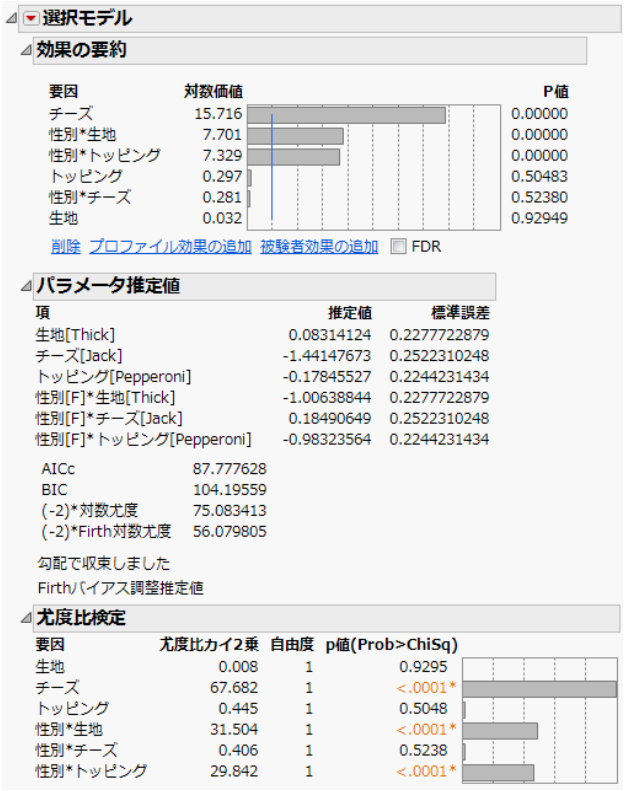
11. 「被験者データ」の開閉アイコンをクリックしてアウトラインを開き、［**データテーブルの選択**］をクリックします。
12. 「Pizza Subjects.jmp」を選択し、［**OK**］をクリックします。
13. **「被験者」**を選択し、［**被験者 ID**］をクリックします。
14. **「性別」**を選択し、［**追加**］をクリックします。

図4.7「被験者データ」パネル



15. ［**モデルの実行**］をクリックします。

図4.8「選択モデル」の結果



モデルに6つの効果が含まれています。**「生地」**、**「チーズ」**、**「トッピング」**は商品属性です。**「性別*生地」**、**「性別*チーズ」**、**「性別*トッピング」**は、被験者効果と商品属性の交互作用です。これらの交互作用を検討することにより、各マーケットセグメントに好まれる商品を探すことができます。

メモ：「選択モデル」の場合、被験者効果を主効果として含めることはできません。被験者効果は、交互作用項としてのみ表示されます。

「効果の要約」と「尤度比検定」の各レポートから、**「性別」**と**「生地」**、**「性別」**と［**トッピング**］の間に強い交互作用があることがわかります。**「生地」**と**「トッピング」**の主効果は、有意ではありません。この例で被験者効果を含めなければ、マーケットのセグメント化に役立つ重要な情報を見逃したかもしれません。

最適なプロファイルを見つける

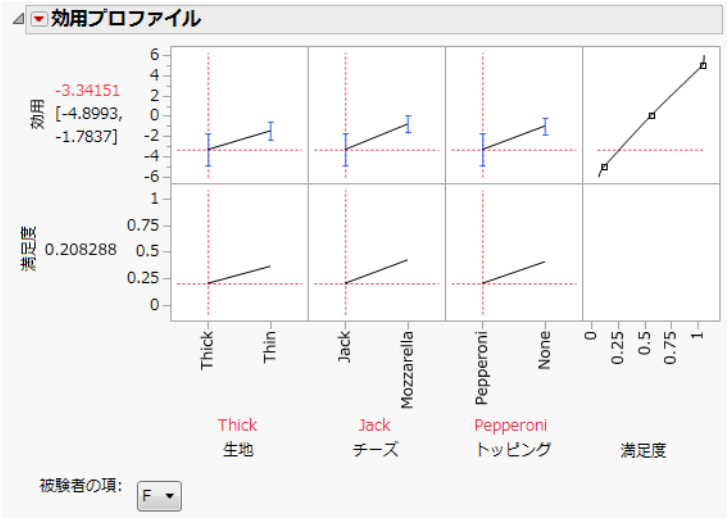
次に、「効用プロファイル」を使って結果を分析し、属性の最適設定を見つけましょう。

1. 「選択モデル」の赤い三角ボタンをクリックし、［**効用プロファイル**］を選択します。

   プロファイルの下にある「被験者の項」は、女性の結果が表示されていることを示します。

2. 「効用プロファイル」の赤い三角ボタンをクリックし、［**最適化と満足度**］>［**満足度関数**］を選択します。

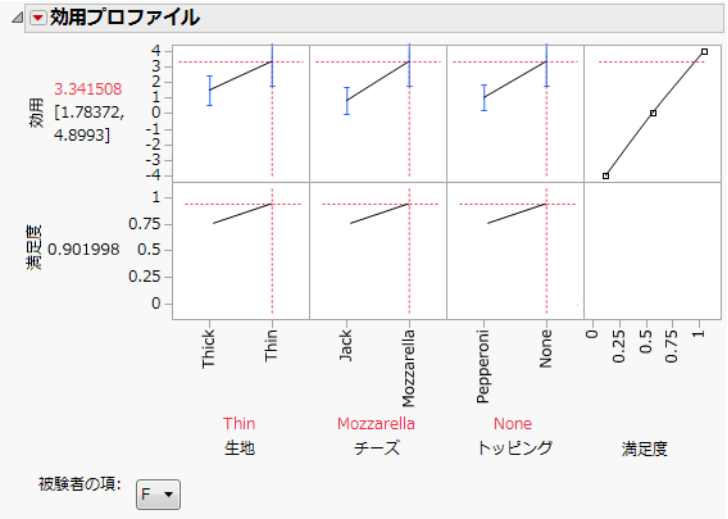
図4.9 満足度関数を表示した効用プロファイル



プロファイルに、効用を最大化する満足度関数が加わりました。『プロファイル機能』の「プロファイル」章を参照してください。

3. 「効用プロファイル」の赤い三角ボタンをクリックし、［**最適化と満足度**］>［**満足度の最大化**］を選択します。

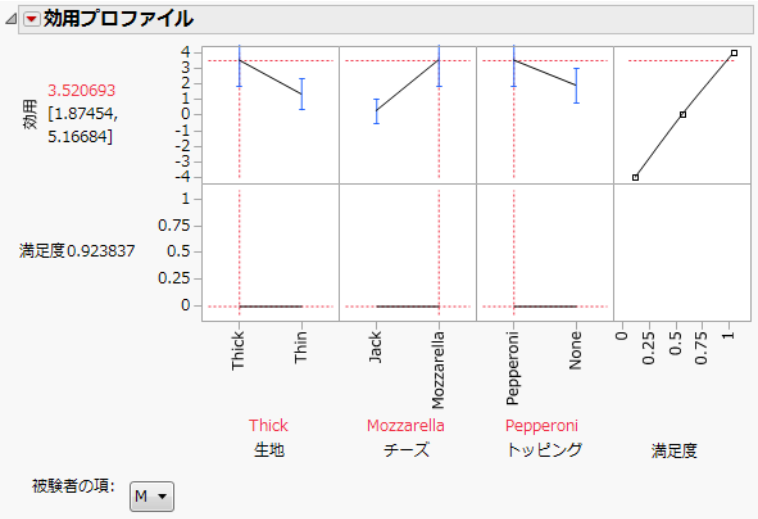
図4.10 女性に最適な設定の効用プロファイル



女性向けの最適な設定は、薄い生地、モッツァレラチーズ、トッピングなしです。

4. 「被験者の項」メニューから［**M**］を選択します。

図4.11「性別」の水準を「M」（男性）に固定したときのユーティリティプロファイル



男性向けの最適な設定は、厚い生地、モッツァレラチーズ、ペパロニのトッピングです。

この例では、マーケットを性別で2つのセグメントに分け、その嗜好を理解して、それぞれに最適なピザを特定しました。

「選択モデル」プラットフォームの起動

「選択モデル」プラットフォームを起動するには、［**分析**］>［**消費者調査**］>［**選択モデル**］を選択します。

「選択モデル」プラットフォームで使用するデータは、1つのデータテーブルにまとめたものでも、2、3個のデータテーブルに分かれているものでもかまいません。「選択モデル」起動ウィンドウの冒頭で「データ形式」を指定します。

1つのデータテーブル, 積み重ね

この形式では、データを1つのデータテーブルにまとめてください。被験者に提示されるプロファイルごとに1つの行を作成し、また、特定のプロファイルが選択されたかどうかを示す指示変数を作成してください。

データが1つのデータテーブルにまとめられている形式の例は、「「選択なし」がある1つのデータテーブル」（73ページ）を参照してください。詳細については、「［1つのデータテーブル, 積み重ね］の起動ウィンドウ」（82ページ）を参照してください。

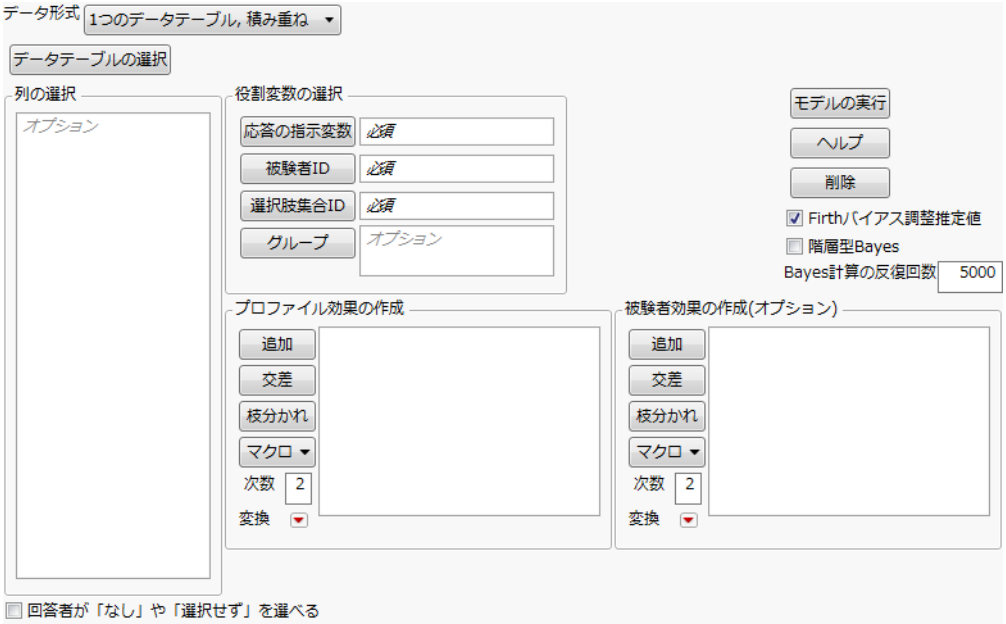
複数のデータテーブル, 相互参照

この形式では、データを2つまたは3つのデータテーブルに分けてください。プロファイルデータと応答データは必須で、被験者データはオプションで使うことができます。「選択モデル」起動ウィンドウは、各データテーブルに対応する3つのセクションで構成されています。各セクションは、必要に応じて開いたり閉じたりできます。

データが複数のデータテーブルに分かれている形式については、「複数のデータテーブル」（75ページ）を参照してください。詳細については、「［複数のデータテーブル, 相互参照］の起動ウィンドウ」（84ページ）を参照してください。

［1つのデータテーブル，積み重ね］の起動ウィンドウ

図4.12［1つのデータテーブル，積み重ね］データ形式の起動ウィンドウ



データテーブルの選択　このボタンをクリックした後、データが保存されているデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他...］を選択してください。

応答の指示変数　その選択肢が選択されたかどうかを示す値を含んだ列です。「1」は選択されたプロファイル、「0」は選択されなかったプロファイルを示します。回答者が何も選択しないことを許している場合は、いずれの選択肢も選択されなかった集合には欠測値を入力しておいてください。詳細については、「回答者が「なし」や「選択せず」を選べる」（83ページ）を参照してください。

被験者ID　調査に参加した人のID。

選択肢集合ID　被験者に提示された選択肢集合のID。被験者は、この選択肢集合のなかから、最も好きなものを選択します。

グループ　「選択肢集合ID」列と共に使用したときに、各選択肢集合を一意に示すことができる列。たとえば、**「選択肢集合ID」**の値が1である選択肢集合が2つあったとします。データテーブルにおいて、一方の選択肢集合に対しては**「調査」**列の値がAで、他方の選択肢集合に対しては**「調査」**列がBであるなら、その**「調査」**列を［グループ］列に指定します。

プロファイル効果の作成　プロファイルの属性に関する効果を追加します。

「プロファイル効果の作成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

被験者効果の作成(オプション) 被験者に関連する効果を追加します。

「被験者効果の作成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

Firthバイアス調整推定値 Firth法は、バイアス修正を伴う最尤推定であり、通常の最尤推定に比べ、推定や検定がより良い性質をもちます。さらに、ロジスティックモデルなどで生じる分離（separation）の問題も改善できます。ロジスティック回帰における分離問題については、Heinze and Schemper（2002）を参照してください。

JMP PRO **階層型Bayes** Bayes流の方法で被験者ごとの係数値を求めます。詳細については、「Bayesパラメータ推定値」（91ページ）を参照してください。

JMP PRO **Bayes計算の反復回数** （[階層型Bayes] を選択した場合のみ有効。）被験者に関する係数の推定値を計算する際に使う適応型Bayesアルゴリズムの反復回数。ここで指定した反復回数には、推定からは破棄されるバーンインの反復回数も含まれています。バーンインの反復回数は、起動ウィンドウで指定した「Bayes計算の反復回数」の半分です。

回答者が「なし」や「選択せず」を選べる 応答が欠測値である行を「選択なし」としてモデルに含めます。[1つのデータテーブル, 積み重ね] 形式の場合、回答者がどれも選択しなかったときは、データテーブルにおける「応答の指示変数」列を（数値の）欠測値にしてください。このオプションは、起動ウィンドウの一番下に表示されます。

［複数のデータテーブル，相互参照］の起動ウィンドウ

図4.13［複数のデータテーブル，相互参照］データ形式の起動ウィンドウ

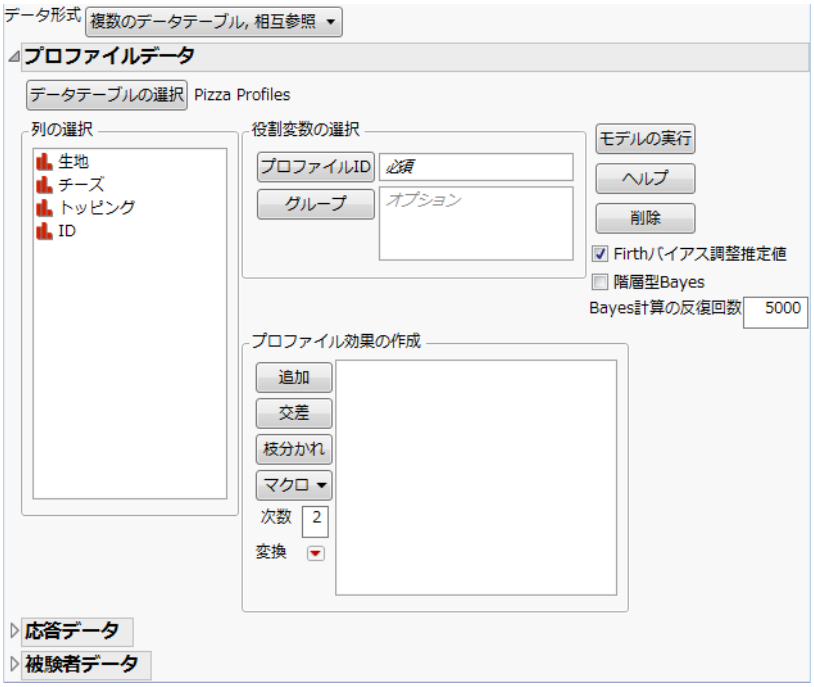


図4.13は、複数のデータテーブル用の起動ウィンドウで、「Pizza Profiles.jmp」をプロファイルのデータテーブルとして使用しています。

［複数のデータテーブル，相互参照］の場合、起動ウィンドウは3つのセクションに分かれています。

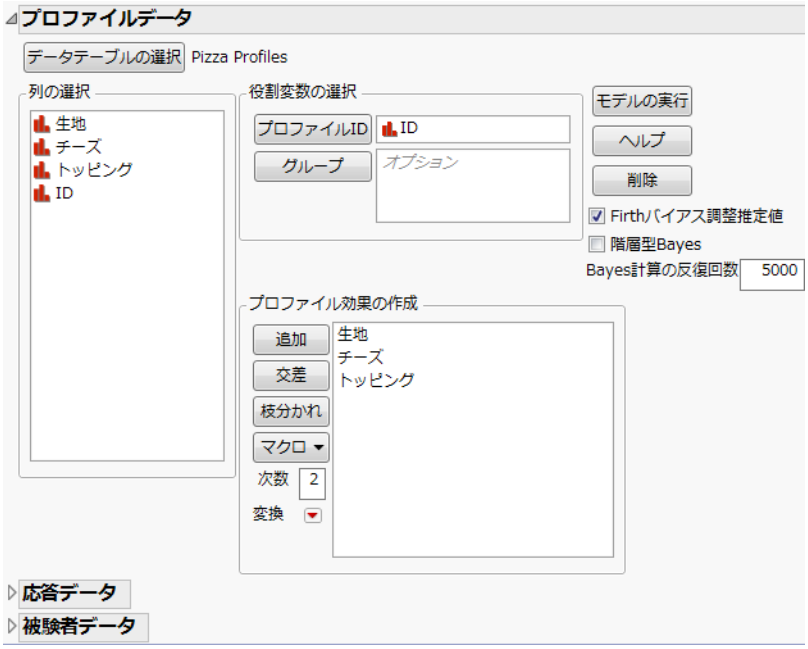
- 「プロファイルデータ」（84ページ）
- 「応答データ」（86ページ）
- 「被験者データ」（88ページ）

プロファイルデータ

プロファイルのデータテーブルには、各選択肢の属性を示すデータを保存してください。データテーブルの各列が1つの属性に対応するように、また、各行が1つのプロファイルに対応するように、データを作成してください。さらに、各プロファイルのID を含んだ列を設けてください。図4.14は、「Pizza Profiles.jmp」データテーブルと、列を指定した後の「プロファイルデータ」パネルです。

図4.14 プロファイルデータテーブルと設定後の「プロファイルデータ」アウトライン

	生地	チーズ	トッピング	ID
1	Thick	Mozzar…	Pepperoni	ThickOni
2	Thick	Mozzar…	None	ThickEllа
3	Thick	Jack	Pepperoni	ThickJackoni
4	Thick	Jack	None	ThickJack
5	Thin	Mozzar…	Pepperoni	TrimOni
6	Thin	Mozzar…	None	Trimella
7	Thin	Jack	Pepperoni	TrimPepperj…
8	Thin	Jack	None	TrimJack



データテーブルの選択　このボタンをクリックした後、プロファイルのデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他 ...］を選択してください。

プロファイル ID　属性の組み合わせ（プロファイル）を識別するための ID。［**プロファイル ID**］によってプロファイルのデータテーブルにおける各行を一意に識別できない場合は、［**グループ**］変数も指定する必要があります。その場合、［**グループ**］列と［**プロファイル ID**］列の組み合わせによって各行が一意に識別できるように、［**グループ**］列を追加してください。

グループ　「プロファイル ID」列とともに使用したときに、各選択肢集合を一意に示すことができる列。たとえば、［**プロファイル ID**］の値が「1」である行が2つあったとします。データテーブルにおいて、一方の選択肢集合に対しては「**調査**」列が「A」で、他方の選択肢集合に対しては「**調査**」列が「B」であるなら、その「**調査**」列を［**グループ**］列に指定します。

プロファイル効果の作成　プロファイルの属性で構成された効果を追加します。

「プロファイル効果の作成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

Firth バイアス調整推定値　Firth 法は、バイアス修正を伴う最尤推定であり、通常の最尤推定に比べ、推定や検定がより良い性質をもちます。さらに、ロジスティックモデルなどで生じる分離（separation）の問題も改善できます。ロジスティック回帰における分離問題については、Heinze and Schemper（2002）を参照してください。

JMP PRO **階層型 Bayes**　Bayes 流の方法で被験者ごとの係数値を求めます。詳細については、「Bayes パラメータ推定値」（91 ページ）を参照してください。

JMP PRO **Bayes 計算の反復回数**　（［階層型 Bayes］を選択した場合のみ有効。）被験者に関する係数の推定値を計算する際に使う適応型 Bayes アルゴリズムの反復回数。ここで指定した反復回数には、推定からは破棄されるバーンインの反復回数も含まれています。バーンインの反復回数は、起動ウィンドウで指定した「Bayes 計算の反復回数」の半分です。

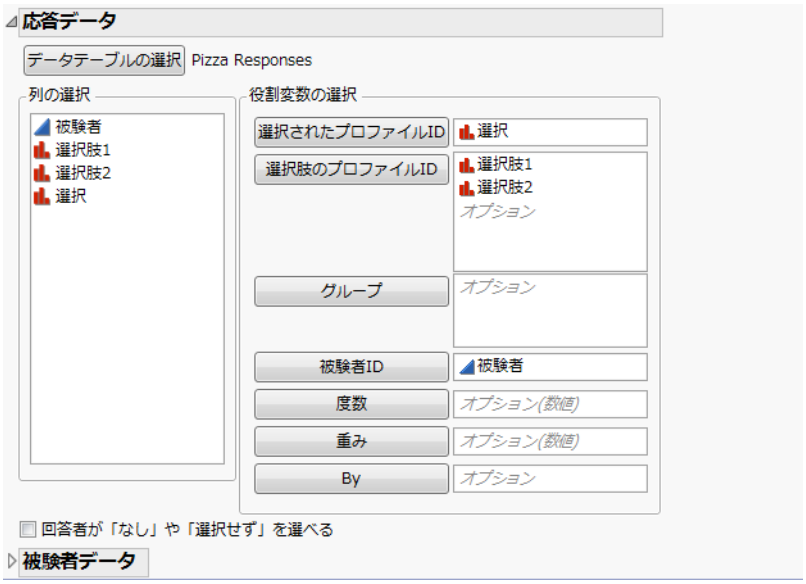
応答データ

応答のデータテーブルには、被験者 ID・選択肢のプロファイル ID・選択されたプロファイル ID の列を設けてください。そして、各被験者と選択肢集合ごとに 1 行ずつ保存してください。選択肢集合のグループが複数ある場合は、グループ変数を使って選択肢集合を区別してください。図 4.15 は、「Pizza Responses.jmp」データテーブルと、列を指定した後の「応答データ」パネルです。

応答データとプロファイルデータにおける選択肢 ID を対応させるのにグループ化が必要な場合には、その変数をグループ変数に指定してください。

図4.15 応答データテーブルと設定後の「応答データ」アウトライン

	被験者	選択肢1	選択肢2	選択
1	1	ThickJack	TrimPepperjack	TrimPepperjack
2	1	TrimPepperj…	ThickEllа	ThickEllа
3	1	TrimOni	Trimella	TrimOni
4	1	ThickEllа	ThickJack	ThickEllа
5	2	Trimella	ThickJackoni	Trimella
6	2	TrimJack	ThickEllа	ThickEllа
7	2	Trimella	TrimPepperjack	Trimella
8	2	TrimPepperj…	TrimOni	TrimOni
9	3	TrimOni	ThickJackoni	TrimOni
10	3	TrimPepperj…	ThickEllа	ThickEllа
11	3	ThickJackoni	TrimPepperjack	ThickJackoni
12	3	ThickOni	Trimella	ThickOni
13	4	ThickEllа	ThickOni	ThickEllа
14	4	TrimPepperj…	ThickJack	ThickJack



データテーブルの選択　このボタンをクリックした後、応答のデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他 ...］を選択してください。

選択されたプロファイル ID　被験者が選択したプロファイルを示す ID。この ID は、プロファイルのデータテーブルにおけるプロファイル ID に対応したものです。

グループ　**「選択されたプロファイル ID」**列と共に使用したときに、各選択肢集合を一意に示すことができる列。

選択肢のプロファイル ID　被験者に提示された複数のプロファイルを示す ID。

被験者 ID　調査に参加した人の ID。

度数　度数を含んだ列。度数が n である行は、データにn回登場しているものとして計算に使用されます。度数が1未満である行や、欠測値である行は、分析に使用されません。

重み　データテーブルの各行に対する重みを含んだ列。0より大きい重みだけが分析に使われます。

By　この列の水準に従ってデータがグループ化され、それぞれ個別に分析されます。指定した列の水準ごとに、個別に分析されます。分析結果は、水準ごとに個別のレポートにまとめられます。複数のBy変数を指定した場合は、By変数の水準のすべての組み合わせごとに分析が行われます。

回答者が「なし」や「選択せず」を選べる　応答が欠測値である行を「選択なし」としてモデルに含めます。［複数のデータテーブル, 相互参照］データ形式の場合、回答者がどれも選択しなかったときは、「選択されたプロファイルID」列のデータ値を欠測値にしてください。このオプションは、「応答データ」パネルの最後に表示されます。

被験者データ

被験者のデータテーブルは、必須ではなく、モデルに被験者効果を含めたいときに使用します。このデータテーブルの列には、被験者ID（応答データテーブルでも使用されているもの）、および、被験者の属性を含めてください。被験者に関するデータは応答のデータテーブルに含めてもかまいませんが、その場合も、「被験者データ」アウトラインで被験者効果を指定する必要があります。図4.16は、「Pizza Subjects.jmp」データテーブルと、列を指定した後の「プロファイルデータ」パネルです。

図4.16 被験者データテーブルと設定後の「被験者データ」アウトライン

データテーブルの選択　このボタンをクリックした後、被験者のデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他 ...］を選択してください。

被験者 ID　各被験者を一意に識別する ID。

モデル効果の構成　被験者に関する効果を追加します。

「モデル効果の構成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

「選択モデル」レポート

- 「効果の要約」
- 「パラメータ推定値」
- 「尤度比検定」
- 「Bayes パラメータ推定値」

効果の要約

モデルに 1 つ以上の効果があり、計算にあまり時間がかからない場合には、「効果の要約」レポートが表示されます。（表示されない場合は、赤い三角ボタンのメニューから［尤度比検定］を選択すると、「尤度比検定」と「効果の要約」の両方のレポートが表示されます。）モデルにおいて推定された効果と、それらの効果の対数価値（または FDR 対数価値）のプロットで構成されます。また、このレポートでは、モデルに効果を追加したり削除したりできます。このとき、モデルをあてはめた結果は、「効果の要約」レポートでの変更に合わせて自動的に更新されます。詳細については、『基本的な回帰モデル』の「標準最小 2 乗のレポートとオプション」章にある「効果の要約レポート」の節を参照してください。

起動ウィンドウで［Bayes 流の被験者効果］チェックボックスをオンにすると、「効果の要約」レポートは表示されません。Bayes 推定においては、尤度比検定が行われないためです。

「効果の要約」表の列

「効果の要約」表には次の列があります。

要因　モデル内の効果。*p* 値の小さい順に並べられます。

対数価値　各効果の対数価値（LogWorth）。$-\log_{10}(p\text{ 値})$ という式で計算されています。*p* 値を対数価値に変換すると、解釈がしやすくなります。対数価値が 2 以上あるものは、有意水準 0.01 で有意です（$-\log_{10}(0.01) = 2$）。

FDR 対数価値　各効果のFDR（False Discovery Rate; 偽発見率）の対数価値で、-$\log_{10}$(FDR調整p値)という式で計算されています。対数価値やFDR対数価値は、検定の有意性をグラフに表すのに適している統計量です。「対数価値」を「**FDR 対数価値**」に変更するには、［**FDR**］チェックボックスをオンにします。

棒グラフ　対数価値（またはFDR 対数価値）の棒グラフ。このグラフには整数値に縦の点線が、2の位置に青の参照線が引かれています。

p値　各モデル効果のp値。このp値は、「尤度比検定」レポートにある有意性検定のp値です。

FDR p値　Benjamini-Hochberg法で計算された、各効果のFDR調整p値。FDRは、検定の多重性を考慮して、生のp値を調整したものです。「**P値**」列を「**FDR p値**」に変更するには、［**FDR**］チェックボックスをオンにします。

FDRについては、Benjamini and Hochberg（1995）を参照してください。偽発見率の詳細については、『予測モデルおよび発展的なモデル』の「応答のスクリーニング」章、またはWestfall et al.（2011）を参照してください。

「効果の要約」表のオプション

要約表の下にあるオプションによって、効果を追加や削除できます。

削除　モデルから選択した効果を削除します。1つまたは複数の効果を削除するには、削除する効果に対応する列を選択し、［削除］ボタンをクリックします。

プロファイル効果の追加　パネルが開き、［1つのデータテーブル, 積み重ね］データ形式ならデータテーブルのすべての列、［複数のデータテーブル, 相互参照］データ形式ならプロファイルのデータテーブルにある列がリストされます。モデルに追加する列を選択し、リストの下にある［追加］ボタンをクリックすると、その効果がモデルに追加されます。［閉じる］をクリックすると、パネルが閉じます。

被験者効果の追加　パネルが開き、［1つのデータテーブル, 積み重ね］データ形式ならデータテーブルのすべての列、［複数のデータテーブル, 相互参照］データ形式なら被験者のデータテーブルにある列がリストされます。モデルに追加する列を選択し、リストの下にある［追加］ボタンをクリックすると、その効果がモデルに追加されます。［閉じる］をクリックすると、パネルが閉じます。

パラメータ推定値

「パラメータ推定値」レポートには、「項」としてリストされている効果ごとに、効用関数に関する係数の推定値と標準誤差が表示されます。製品の属性に対するパラメータは、「部分効用（part-worth）」と呼ぶこともあります。起動ウィンドウで［Firthバイアス調整推定値］オプションをオンにした場合は、Firthのバイアス調整を伴う最尤推定が行われます。この推定値は、通常の最尤推定値に比べ、より良い性質をもちます。効用関数の詳細については、「効用と確率」（125ページ）を参照してください。

比較規準

適合度統計量として、AICc（修正された赤池情報量規準）、BIC（ベイズの情報量規準）（– 2）× 対数尤度、（– 2）×Firth 対数尤度が表示されます。これらは、モデルを比較するための統計量です。詳細と計算式については、『基本的な回帰モデル』の付録「統計的詳細」の「尤度・AICc・BIC」の節を参照してください。

起動ウィンドウで［Firth バイアス調整推定値］オプションをオンにすると、レポートに（– 2）×Firth 対数尤度が表示されます。このオプションは、デフォルトではオンになっています。なお、Firth のバイアス調整を伴う最尤推定を行った場合も、行わなかった場合も、AICc や（–2）× 対数尤度の計算には、計算式自体は同じものが使われます。

メモ： 適合度統計量は、いずれも値が小さいほどよくあてはまっていることを示します。

尤度比検定

「尤度比検定」レポートは、モデルが5秒以内にあてはめられるならデフォルトで表示されます。表示されない場合は、「選択モデル」の赤い三角ボタンのメニューから「尤度比検定」オプションを選択すると表示されます。レポートには次のものが表示されます。

要因　モデルに含まれる効果。

尤度比カイ2乗　該当する効果に対する検定の尤度比カイ2乗の値。

自由度　カイ2乗検定の自由度。

p 値 (Prob>ChiSq)　カイ2乗検定に対する *p* 値。

棒グラフ　カイ2乗検定統計量の棒グラフ。

JMP PRO Bayes パラメータ推定値

（この結果は、起動ウィンドウで［階層型 Bayes］を選択した場合のみ表示されます。）「Bayes パラメータ推定値」レポートには、階層型 Bayes モデルで求められた、モデル効果に関する結果が表示されます。階層型 Bayes モデルでは、被験者による違いを尤度関数に組み込み、製品属性のパラメータにおける被験者の違いを推定します。Bayes 推定のアルゴリズムには、適応型 Metropolis-Hastings 法（Train 2001）の一種が使われています。各モデル効果に対し、事後分布の平均と分散が計算されます。また、各モデル効果に関して、被験者ごとの係数値も出力できます。詳細については、「被験者推定値の保存」（95ページ）を参照してください。

この Bayes 流の推定では、被験者ごとの効果が確率変数（変量効果）とみなされます。製品の属性に関するパラメータが、被験者ごとに異なっており、特定の平均ベクトルと共分散行列を持つ多変量正規分布に従うと仮定されます。そして、ある被験者が複数の選択肢から選択することを表す尤度関数には、通常の最尤推定と同じように、多項ロジットモデルが仮定されます。（被験者ごとの効果が従う）多変量正規分布の平均ベクトルに対する事前分布には、平均ベクトルが 0 で、共分散行列が対角要素がすべて等しい対角行列である多変量正規分布が仮定されます。また、（被験者ごとの効果が従う）多変量正規分布の共分散行列が従う事前分布には、対角要素がすべて等しい対角行列を尺度行列とする逆 Wishart 分布が仮定されます。

反復計算では、最初にバーンイン期間が設けられており、最初のほうの反復の結果は破棄されます。デフォルトでは、そのバーンインの反復回数は、起動ウィンドウで指定した「Bayes計算の反復回数」の半分です。

図4.17 Bayesパラメータ推定値レポート

選択モデル: 選択

Bayesパラメータ推定値

項	事後 平均	事後 標準偏差	被験者 標準偏差
生地[Thick]	0.27529781	0.724891965	2.815958945
チーズ[Jack]	-7.01110292	4.697169974	2.593184268
トッピング[Pepperoni]	-1.06702410	0.941602303	2.260232580

全反復回数	5000
バーンイン反復回数	2500
回答者数	32
バーンイン後の平均対数尤度	-14.19342

項 モデルの項。

事後 平均 該当の項に対する事後分布を、被験者に関しても平均したもの。各反復において、各被験者に対する係数の値が生成されます。「事後 平均」は、その係数値を各反復で被験者に関して平均し、バーンイン反復後のそれらの全被験者平均をさらに平均したものです。

ヒント: 赤い三角ボタンのメニューから［Bayesチェーンの保存］を選択すると、反復ごとに生成された係数値を被験者に関して平均した値を確認できます。

事後 標準偏差 各反復において、生成された被験者ごとの係数値から平均が求められます。「事後 標準偏差」は、バーンイン反復後のそれらの全被験者平均から計算される標準偏差です。

被験者 標準偏差 まず、被験者ごとに、生成された係数値の平均をバーンイン後のものから計算します。「被験者 標準偏差」は、その標準偏差です。

ヒント: 赤い三角ボタンのメニューから［被験者推定値の保存］を選択すると、生成された係数値の、バーンイン後における、被験者ごとの平均を確認できます。

全反復回数 バーンイン反復を含め、実行された反復の合計数。

バーンイン反復回数 バーンイン反復の回数。バーンインの反復回数は、起動ウィンドウで指定したBayes反復回数の半分に相当します。

回答者数 被験者の人数。

バーンイン後の平均対数尤度 バーイン反復後の生成された値から求めた対数尤度関数の平均。

「選択モデル」プラットフォームのオプション

「選択モデル」プラットフォームには、数多くのオプションがあります。これらのオプションを表示するには、「選択モデル」の赤い三角ボタンをクリックしてください。

メモ：［階層型Bayes］を指定した場合、被験者ごとの係数値は、モンテカルロ標本を使って計算されます。そのため、以下のオプションの計算結果は、実行するたびに異なります。

尤度比検定　詳細については、「尤度比検定」（91ページ）を参照してください。

JMP PRO **最尤推定値の表示**　（起動ウィンドウで［階層型Bayes］を選択した場合のみ表示。）モデル項の係数を、通常の最尤法で推定した結果（パラメータ推定値と標準誤差）。これらの値は、階層型Bayesアルゴリズムの開始値として使用されます。

複合因子検定　（起動ウィンドウで［階層型Bayes］を選択しなかった場合のみ表示。）モデル内の各因子を検定する目的で、その因子が関連しているすべての効果の尤度比検定を行います。複合因子検定について詳しくは、『基本的な回帰モデル』の「標準最小2乗のレポートとオプション」章を参照してください。

信頼区間　（［Firthバイアス調整推定値］を選択した場合のみ）「パラメータ推定値」レポートにパラメータの信頼区間を表示します。

信頼限界　（［階層型Bayes］を選択した場合のみ）「Bayesパラメータ推定値」レポートにパラメータの信頼限界を表示します。この信頼限界は、事後分布の2.5と97.5の各分位点に基づいて計算されます。

推定値の相関　［階層型Bayes］が選択されていないときには、最尤法に基づいて、パラメータ推定値間の相関行列を表示します。

［階層型Bayes］が選択されている場合は、係数値の事後平均における相関行列を表示します。この相関行列は、バーンイン後の反復から求めたものです。バーンイン後、反復ごとに全被験者における事後平均が計算されます。これらの事後平均から求めた相関行列が、「推定値の相関」表に表示されます。

全水準の推定値　モデルに含まれている各効果の周辺確率と限界効用を表示します。周辺確率とは、他の属性をすべて平均またはデフォルトの水準に固定したときに、人が属性因子Bではなく属性因子Aを選択する確率を指します。

図4.18において、（生地とトッピングは同じという条件で）被験者がモンテレージャックのピザではなく、モッツァレラのピザを選ぶ確率は、0.9470です。

図4.18「全水準の推定値」の例

効用プロファイル　異なる因子設定における効用の予測値を示します。効用とは、モデルにおいて線形式で表されている部分の値です。「最適なプロファイルを見つける」（79ページ）に効用プロファイルの例があります。効用関数の詳細については、「効用と確率」（125ページ）を参照してください。［効用プロファイル］のオプションについては、『プロファイル』の「プロファイル」章にある「予測プロファイルのオプション」の節を参照してください。

確率プロファイル　2つの製品が提示されたときに一方の製品を選択する確率を求めることができます。この確率は、次式で計算されます。

$$(\exp(U))/(\exp(U)+\exp(U_b))$$

Uは該当する製品の効用、U_bは基準となる製品の効用です。なお、この式に基づくと、2つの製品がまったく同じ場合、各製品が選択される確率は50%ずつです。詳細については、「効用と確率」（125ページ）を参照してください。

確率プロファイルを使用する例は、「基準との比較」（103ページ）で紹介しています。［確率プロファイル］のオプションについては、『プロファイル』の「プロファイル」章にある「予測プロファイルのオプション」の節を参照してください。

多選択プロファイル　指定した数の確率プロファイルが表示されます。複数の製品が提示されたときに、各製品が選択される確率が求められます。各製品のプロファイルは自由に設定できます。多選択プロファイルを使用する例は、「多選択プロファイルの比較」（105ページ）で紹介しています。［多選択プロファイル］のオプションについては、『プロファイル』の「プロファイル」章にある「予測プロファイルのオプション」の節を参照してください。

比較　設定された選択肢の比較を行います。比較したい因子と値を選択することができます。なお、［**すべて**］チェックボックスをチェックした場合、すべての水準が比較されます。被験者効果がある場合は、製品の比較を行いたい、被験者の水準を指定できます。［すべて］チェックボックスを複数の因子でチェックした場合、それらすべての因子の水準での比較ではなく、該当する1因子のすべての水準が、その他の因子を左側で選択されている水準に固定した上で、比較されます。

図4.19「効用の比較」ウィンドウ

支払意思額 このコマンドは、データテーブルに連続尺度の価格の列がある場合だけ使えます。製品の機能を新しいものに変更したときに、これまでの製品と同じ効用を与えるのに、価格をどれぐらい変更できるかを計算します。結果は、「基準」に指定された値を基準にして、背景因子の水準ごとに計算されます。

効用計算式の保存 複数のデータテーブルに対して分析を実行している場合は、効用の計算式列を含んだ、新しいデータテーブルを作成します。作成されたデータテーブルでは、各行がプロファイルと被験者の組み合わせに対応しており、プロファイルと被験者を示す列があります。1つのデータテーブルに対して分析を実行している場合は、「効用計算式」列がそのデータテーブルに追加されます。

被験者ごとの勾配を保存 各パラメータの変化量を被験者ごとに含んだデータテーブルを作成します。パラメータの変化量は、ヘッセ行列で尺度化した各パラメータの傾きを、被験者ごとに平均して求められています。これは、該当する被験者と他の被験者との差異に対するラグランジュ乗数検定に対応します。作成されたデータテーブルを、付随しているスクリプトでクラスター分析すれば、マーケットセグメント（同じ選好構造をもつ消費者集団）を特定することができます。詳細については、「勾配」（125ページ）を参照してください。例として、「セグメント化の例」（107ページ）を参照してください。

JMP PRO **被験者推定値の保存** （起動ウィンドウで［階層型Bayes］を選択した場合のみ表示。）各効果における被験者ごとのパラメータ推定値を含んだデータテーブルを新たに作成します。この値は、この被験者ごとに係数値を平均したものであり、「Bayesパラメータ推定値」レポートにある「事後 平均」です。「被験者受容率」は、Metropolis-Hastings法のステップにおいて、現在のステップとは異なる係数値に移動した割合（受容率）を示します。一般に、0.20ぐらいが良好な受容率とみなされています。詳細については、「Bayesパラメータ推定値」（91ページ）を参照してください。

JMP PRO **Bayesチェーンの保存** （起動ウィンドウで［階層型Bayes］を選択した場合のみ表示。）Bayes推定において使用されたチェーンの情報を含んだデータテーブルを、新たに作成します。詳細については、「Bayesチェーンの保存」（98ページ）を参照してください。

モデルダイアログ 「選択モデル」起動ウィンドウが開き、モデルを修正したり、別のモデルをあてはめたりすることができます。データテーブルや**ID**、モデル効果を新しく指定できます。

以下のオプションについて詳しくは、『JMPの使用法』の「JMPのレポート」章を参照してください。

やり直し 分析を繰り返したり、やり直したりするオプションを含みます。また、［自動再計算］オプションに対応しているプラットフォームにおいては、［自動再計算］オプションを選択すると、データテーブルに加えた変更が、該当するレポートに即座に反映されるようになります。

スクリプトの保存 レポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。

Byグループのスクリプトを保存 By 変数の全水準に対するレポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。起動ウィンドウでBy変数を指定した場合のみ使用可能です。

支払意思額

「支払意思額」（willingness to pay）とは、製品の新しい機能に対して顧客が支払うであろう価格を指します。「支払意思額」は、基準機能の効用と等しくなる、新機能の価格です。たとえば、40 GBのハードディスクを搭載したパソコンが、1000ドルだったとします。80 GBのハードディスクに対する支払意思額は、ハードディスクを80 GBにした場合、1000ドルである40 GBのパソコンと同じ効用をもたらすのに、どれぐらいの価格にできるかを示します。

「支払意思額」起動ウィンドウのオプション

［支払意思額］オプションを選択すると、「支払意思額」起動ウィンドウが開きます。図4.20は、「Laptop Profile.jmp」データテーブルにある**「選択モデル」**スクリプトを実行した後、赤い三角ボタンメニューから［支払意思額］オプションを選択したときに表示されるウィンドウです。

因子　分析に含まれる変数。製品の機能や被験者の属性などです。

基準　各因子の基準設定。カテゴリカルな因子に対しては、リストから基準値を選択してください。連続尺度の因子に対しては、基準値を入力してください。

役割　因子の種類。次の中から選択できます。

機能因子　支払意思額を求めたい製品やサービスの機能。

価格因子　選択実験で提示された価格の因子。価格因子は連続尺度でなければならず、1回の「支払意思額」分析において1つの価格因子しか指定できません。

背景因子 定数　「支払意思額」の計算で固定にする因子。通常は、被験者に関する因子を指定します。

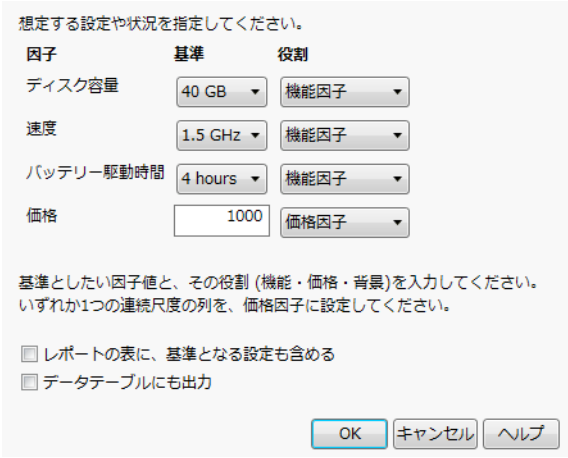
背景因子 変数　「支払意思額」の計算で水準ごとに固定する因子。通常は、被験者に関する因子を指定します。被験者に関する因子を［背景因子 定数］ではなく［背景因子 変数］に指定すると、その因子の全水準に対して支払意思額が計算されます。

レポートの表に、基準となる設定も含める　「支払意思額」レポートに、価格変更を加える前の基準設定が表示されます。

ヒント：このチェックボックスをオンにしてデータテーブルに出力した場合、そのデータテーブルには基準となる設定も出力されます。

データテーブルにも出力　「支払意思額」レポートを含んだデータテーブルが作成されます。

図4.20「支払意思額」起動ウィンドウ

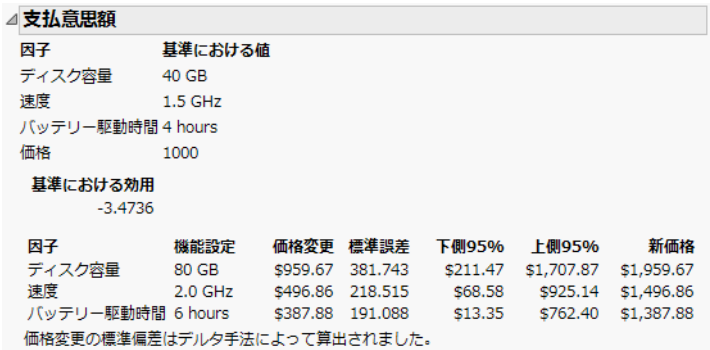


1回目の「支払意思額」の計算が完了すると、選択した基準値と役割が記憶されます。そのため、基準情報を一度入力するだけで、「支払意思額」の比較を何度も行うことができます。「価格」という名前の因子はないが、分析に使われている連続尺度の因子が1つしかない場合は、「支払意思額」ウィンドウにて、その因子に自動的に「価格因子」の役割が割り当てられます。実際の金銭的な価格以外にも、旅行時間や距離などの金銭的ではないコストも価格因子と考えられます。

「支払意思額」レポート

「支払意思額」レポートには、各因子の基準値と、基準における効用が表示されます。そして、各因子について、機能設定、価格変更（価格がどれだけ変化するか）、新価格がリストされます。交互作用や2次の効果がない場合は、標準誤差と信頼区間も表示されます。標準誤差や信頼区間は、デルタ法で求められます。

図4.21「支払意思額」レポート



支払意思額

因子	基準における値
ディスク容量	40 GB
速度	1.5 GHz
バッテリー駆動時間	4 hours
価格	1000

基準における効用
-3.4736

因子	機能設定	価格変更	標準誤差	下側95%	上側95%	新価格
ディスク容量	80 GB	$959.67	381.743	$211.47	$1,707.87	$1,959.67
速度	2.0 GHz	$496.86	218.515	$68.58	$925.14	$1,496.86
バッテリー駆動時間	6 hours	$387.88	191.088	$13.35	$762.40	$1,387.88

価格変更の標準偏差はデルタ手法によって算出されました。

JMP PRO Bayesチェーンの保存

Bayesチェーンのデータテーブルを使うと、MCMCアルゴリズムで生成される係数値の安定性を調べることができます。作成されるデータテーブルには、(起動ウィンドウで指定した)「Bayes計算の反復回数」に1を足した数の行が含まれます。最初の行(反復1)は、開始値を示します。2行目以降には、生成された乱数が反復の順に保存されます。各列には、次のような情報が出力されます。

反復 反復の番号。最初の行は、開始値です。

対数尤度 その反復におけるモデルの対数尤度。各反復に対する対数尤度をプロットすれば、バーンイン期間(調整段階)での振る舞いを確認できます。

<モデル効果> 適応型シグマ 逆Wishart分布の尺度行列における対角要素の平方根。

<モデル効果> 受容率 MCMCアルゴリズムでの受容率。

<モデル効果> 平均 被験者ごとに生成された係数値、の平均。

<モデル効果> 分散 被験者ごとに生成された係数値、の分散。

「選択モデル」の別例

- 「製品の仕様を決める分析例」
- 「セグメント化の例」
- 「「選択モデル」プラットフォームを使用したロジスティック回帰の例」
- 「対応のあるデータに対する条件付きロジスティック回帰の例」
- 「データを2つの分析テーブルに変換する例」
- 「データを1つの分析テーブルに変換する例」

製品の仕様を決める分析例

「選択モデル」プラットフォームを利用すると、製品がもつ属性の相対的な重要性を知ることができます。製品について、重要な属性がどれであるかは明らかになっていても、それらの属性におけるトレードオフの関係はわかっていない場合があります。選択モデルで分析することにより、最適なトレードオフを明らかにし、その情報を製品の企画・設計に反映させることができます。このようなプロセスで製品を設計することの利点を、例を挙げて解説していきます。

ノートパソコンの設計においては、ディスク容量・プロセッサの速度・バッテリーの寿命・販売価格の4つの属性が重要であることがわかっています。この調査のために収集したデータを使い、4つの属性(**「ディスク容量」**、**「速度」**、**「バッテリー駆動時間」**、**「価格」**)の中でどれが最も重要かを調べます。また、これらの属性に対する嗜好に**「性別」**や**「職種」**による違いがあるかどうかも調べます。

この例は、次のような節に分かれています。

- 「起動ウィンドウへの入力」（99 ページ）
- 「モデルの分析」（102 ページ）
- 「基準との比較」（103 ページ）
- 「多選択プロファイルの比較」（105 ページ）

起動ウィンドウへの入力

1. ［**ヘルプ**］ > ［**サンプルデータライブラリ**］を選択し、「Laptop Runs.jmp.」を開きます。

メモ : ここで説明している手順をスキップしたいときは、**「性別を含む選択モデル」**スクリプトの隣にある緑色の三角ボタンをクリックしてモデルを実行し、「モデルの分析」（102 ページ）に進んでください。

2. **「プロファイルと被験者のデータを開く」**スクリプトの隣にある緑色の三角ボタンをクリックします。

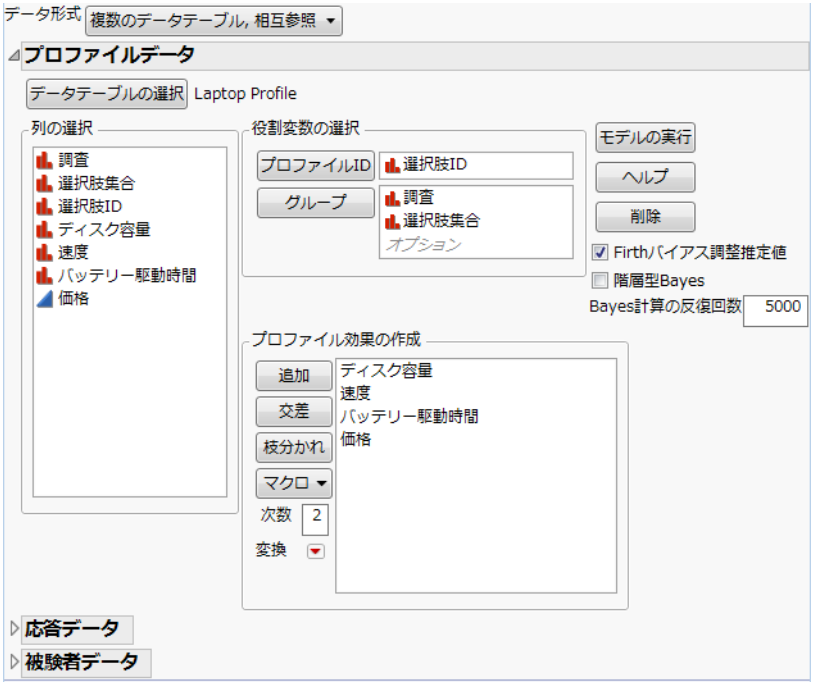
   「Laptop Profile.jmp」と「Laptop Subjects.jmp」のデータテーブルが開きます。

3. ［**分析**］ > ［**消費者調査**］ > ［**選択モデル**］を選択します。

   メモ : これは、開いている 3 つのデータテーブルのどれからでも実行できます。

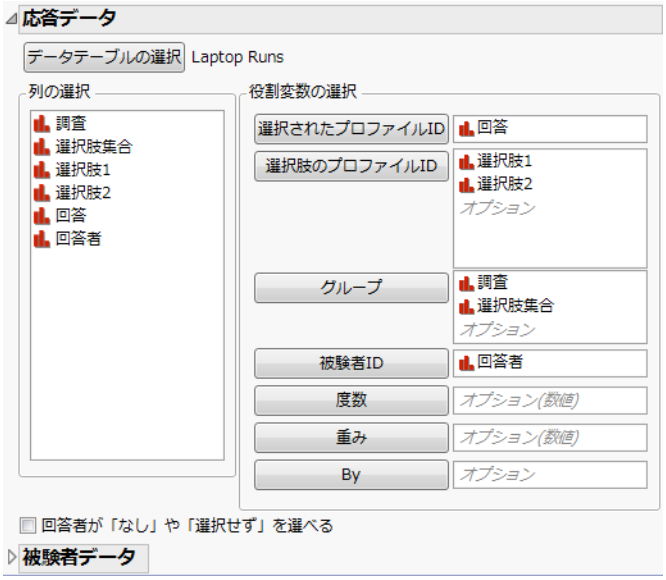
4. 「データ形式」メニューから［**複数のデータテーブル , 相互参照**］を選択します。
5. 「プロファイルデータ」の［**データテーブルの選択**］をクリックし、「Laptop Profile.jmp」を選択します。**「選択肢 ID」**を選択し、［**プロファイル ID**］をクリックします。
6. **「ディスク容量」**、**「速度」**、**「バッテリー駆動時間」**、**「価格」**を選択し、［**追加**］をクリックします。
7. **「調査」**と**「選択肢集合」**を選択し、［**グループ**］をクリックします。

図4.22 ノートパソコン調査の「プロファイルデータ」パネル



8. 「**応答データ**」アウトラインを開きます。
9. ［**データテーブルの選択**］をクリックし、リストから「Laptop Runs.jmp」を選択します。
10. 「応答データ」パネルに次のように入力します。
   - 「**回答**」を選択し、［**選択されたプロファイルID**］をクリックします。
   - 「**選択肢1**」と「**選択肢2**」を選択し、［**選択肢のプロファイルID**］をクリックします。
   - 「**調査**」と「**選択肢集合**」を選択し、［**グループ**］をクリックします。
   - 「**回答者**」を選択し、［**被験者ID**］をクリックします。「応答データ」ウィンドウは、図4.23のようになります。

図 4.23 ノートパソコン調査の「応答データ」パネル



11.「**被験者データ**」アウトラインを開きます。

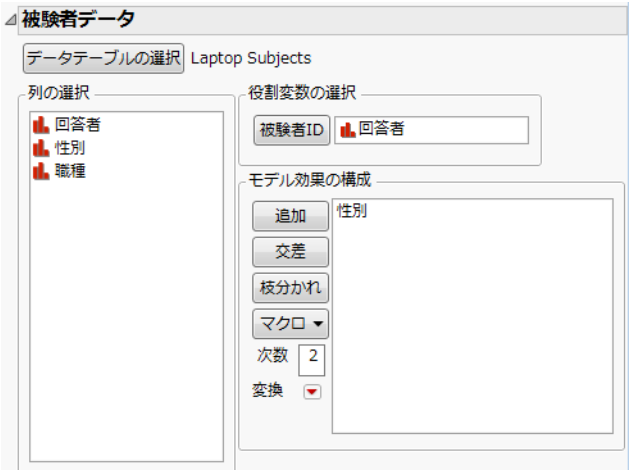
12.［**データテーブルの選択**］をクリックし、リストから「Laptop Subjects.jmp」を選択します。

13.「**回答者**」を選択し、［**被験者 ID**］をクリックします。

14.「**性別**」を選択し、［**追加**］をクリックします。

「被験者データ」ウィンドウは、図 4.24 のようになります。

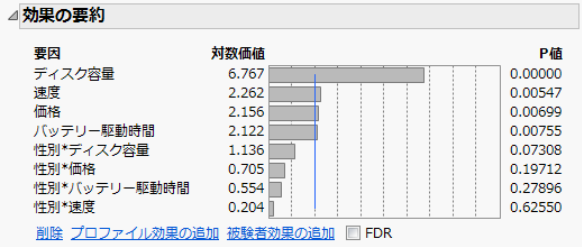
図 4.24 ノートパソコン調査の「被験者データ」パネル



モデルの分析

1. ［**モデルの実行**］をクリックします。

図4.25 ノートパソコン調査の効果の要約



効果の要約

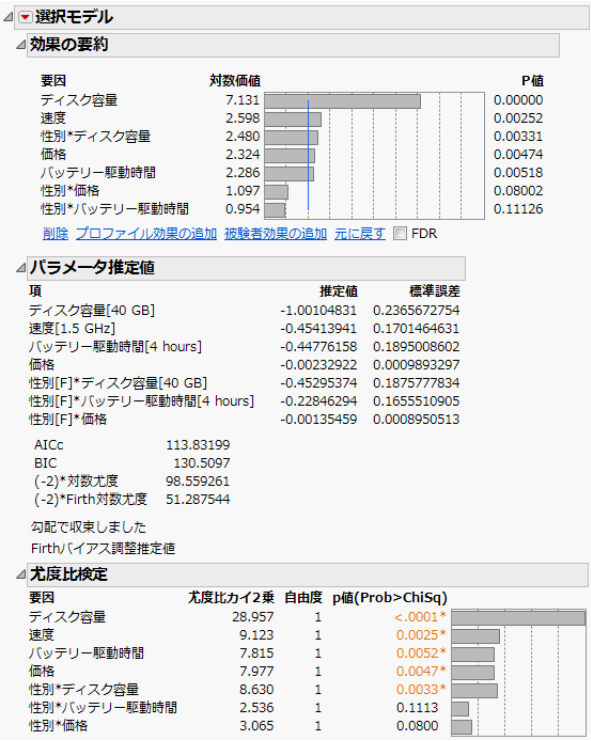
要因	対数価値	P値
ディスク容量	6.767	0.00000
速度	2.262	0.00547
価格	2.156	0.00699
バッテリー駆動時間	2.122	0.00755
性別*ディスク容量	1.136	0.07308
性別*価格	0.705	0.19712
性別*バッテリー駆動時間	0.554	0.27896
性別*速度	0.204	0.62550

削除 プロファイル効果の追加 被験者効果の追加 FDR

「効果の要約」レポートを見ると、「**ディスク容量**」が最も有意な効果であることがわかります。ここでは、p値が0.15を超えている効果を削除して、モデルをより単純なものにします。このとき、効果は1つずつ削除するほうがよいでしょう。この例では、「**性別*速度**」の効果が最も有意性が低く、p値が0.625となっています。

2. 「効果の要約」レポートで、「**性別*速度**」を選択し、［**削除**］をクリックします。

図4.26 ノートパソコン調査の結果



選択モデル

効果の要約

要因	対数価値	P値
ディスク容量	7.131	0.00000
速度	2.598	0.00252
性別*ディスク容量	2.480	0.00331
価格	2.324	0.00474
バッテリー駆動時間	2.286	0.00518
性別*価格	1.097	0.08002
性別*バッテリー駆動時間	0.954	0.11126

削除 プロファイル効果の追加 被験者効果の追加 元に戻す FDR

パラメータ推定値

項	推定値	標準誤差
ディスク容量[40 GB]	-1.00104831	0.2365672754
速度[1.5 GHz]	-0.45413941	0.1701464631
バッテリー駆動時間[4 hours]	-0.44776158	0.1895008602
価格	-0.00232922	0.0009893297
性別[F]*ディスク容量[40 GB]	-0.45295374	0.1875777834
性別[F]*バッテリー駆動時間[4 hours]	-0.22846294	0.1655510905
性別[F]*価格	-0.00135459	0.0008950513

AICc	113.83199
BIC	130.5097
(-2)*対数尤度	98.559261
(-2)*Firth対数尤度	51.287544

勾配で収束しました
Firthバイアス調整推定値

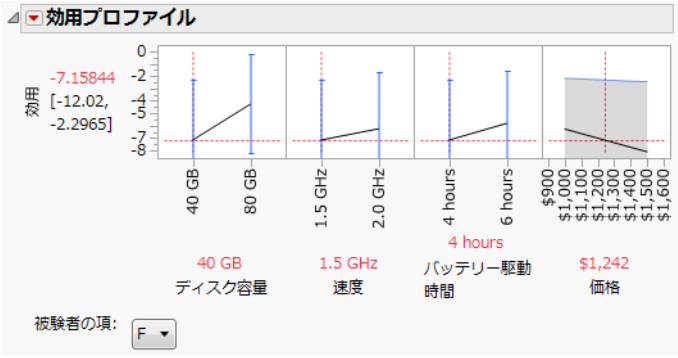
尤度比検定

要因	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
ディスク容量	28.957	1	<.0001*
速度	9.123	1	0.0025*
バッテリー駆動時間	7.815	1	0.0052*
価格	7.977	1	0.0047*
性別*ディスク容量	8.630	1	0.0033*
性別*バッテリー駆動時間	2.536	1	0.1113
性別*価格	3.065	1	0.0800

モデルから「**性別*速度**」を削除すると、すべての効果の p 値が0.15以下になります。そのため、これを最終モデルとして使用してみましょう。

3. 「選択モデル」の赤い三角ボタンをクリックし、［**効用プロファイル**］を選択します。

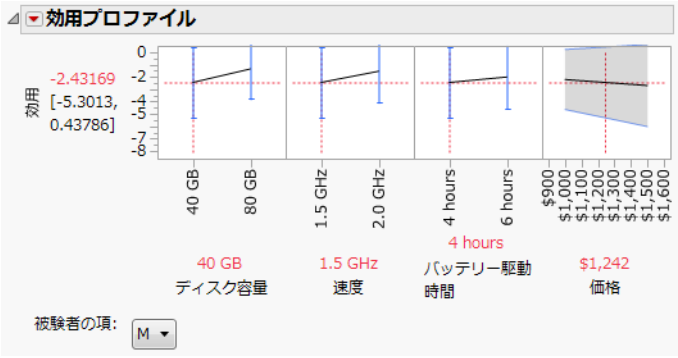
図4.27 ノートパソコン調査の「F」（女性）のプロファイル



ヒント：効用プロファイルが図4.27のようにならなかった場合は、効用プロファイルの赤い三角ボタンをクリックし、［**表示方法**］＞［**Y軸を自動調整**］を選択します。

4. 「被験者の項」メニューから［**M**］を選択します。

図4.28 ノートパソコン調査の「M」（男性）のプロファイル



「**性別**」と「**ディスク容量**」との交互作用は、p 値が0.0033となっており、高度に有意です。詳細については、図4.26（102ページ）を参照してください。「効用プロファイル」で「**性別**」の両水準の「**ディスク容量**」を確認すると、男性よりも女性のほうが傾きが大きいことがわかります。

基準との比較

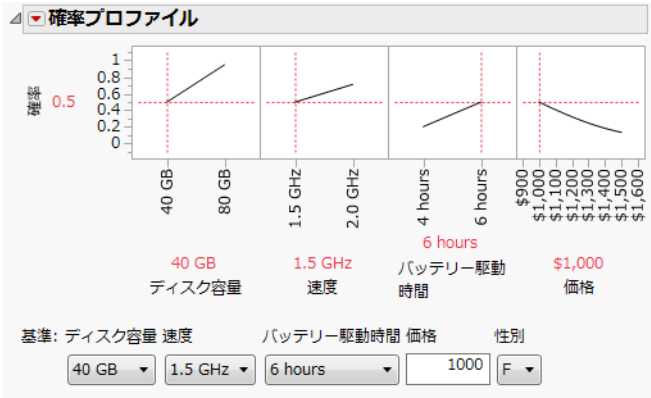
新しい製品を開発しているとしましょう。消費者が旧製品や競合製品ではなく新製品を選択するであろう確率を求めましょう。その場合、「確率プロファイル」を使ってプロファイルを基準プロファイルと比較します。

この例では、40 GBのハードディスク、1.5 GHzのCPU、6時間のバッテリー駆動時間であるノートパソコンを1,000ドルで現在、販売しているとします。そして、現在の製品を基準として、機能のいくつかを変更して、製品の魅力を高めるとします。なお、以下で行われる分析において、もし比較する2つの製品がまったく同じであった場合は、各製品が選ばれる確率は50%ずつと計算されます。基準となる製品を変更することにより、消費者に選択される確率がどれぐらい大きくなるかを調べます。

1. 次のいずれかを実行します。
   - 「起動ウィンドウへの入力」（99 ページ）の手順に従います。次に、「モデルの分析」（102 ページ）のステップ1とステップ2を実行します。
   - 「Laptop Runs.jmp」データテーブルで、**「縮小モデルの選択モデル」**スクリプトの緑色の三角ボタンをクリックします。
2. 「選択モデル」の赤い三角ボタンのメニューから［**確率プロファイル**］を選択します。

   この確率プロファイルでは、**「性別」**が「F」に設定されています。これは、後で変更することができます。
3. プロファイルの下にある「基準」エリアのメニューやテキストボックスを使用し、基準構成を40 GB、1.5 GHz、6 hours、1000に設定します。
4. これらの値を「確率プロファイル」の値としても指定します。価格を$1000に設定するには、右端のプロファイルの下にある「価格」のすぐ上の「$1,242」という値をクリックし、「1000」と入力します。次に、テキストボックスの外をクリックします。

図4.29 確率プロファイルにある「価格」のテキスト入力エリア

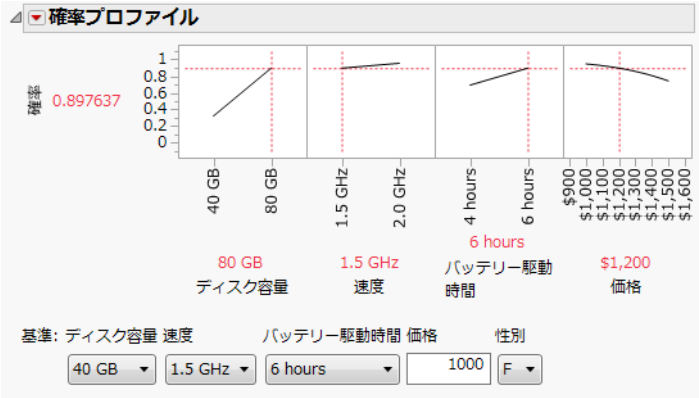


この構成の確率は0.5です。

5. 確率プロファイルで、**「ディスク容量」**のスライダを80 GBまで動かします。

   変更された設定において、たとえ**「価格」**が増加しても、選択される確率はさほど変化しないことがわかります。
6. プロファイルの**「価格」**の上にある「1000」という赤色の数字をクリックし、「**1200**」と入力してテキストボックスの外をクリックします。

図4.30 基準的なノートパソコンに対する確率プロファイル



「ディスク容量」を40 GBから80 GBへと変更し、「価格」を$1,200にすると、基準の製品とこの新製品の2つが選択肢にある場合、女性が新製品のほうを選択する確率は0.90です。**「性別」**を［**M**］（男性）に変更してみましょう。男性におけるこの確率を見ると0.71になっています。

多選択プロファイルの比較

「多選択プロファイル」を使って製品プロファイルを比較します。

- 現在の自社製品は、ディスク容量が小さく、CPU 速度が遅く、バッテリー駆動時間が短いですが、価格が$1,000と最も安価になっています。
- 会社Aの製品は、CPUが高速で、バッテリー駆動時間が長く、価格は$1,200と手頃です。
- 会社Bの製品は、ディスク容量が最大で、CPU速度も最速ですが、価格が$1,500と高く、バッテリー駆動時間が短くなっています。

分析者は、性能や価格のいずれか1つを改善し、市場シェアを高めたいと考えています。

1. 次のいずれかを実行します。
   - 「起動ウィンドウへの入力」（99 ページ）の手順に従います。次に、「モデルの分析」（102 ページ）のステップ1とステップ2を実行します。
   - 「Laptop Runs.jmp」データテーブルで、**「縮小モデルの選択モデル」**スクリプトの横にある緑色の三角ボタンをクリックします。
2. 「選択モデル」の赤い三角ボタンのメニューから［**多選択プロファイル**］を選択します。

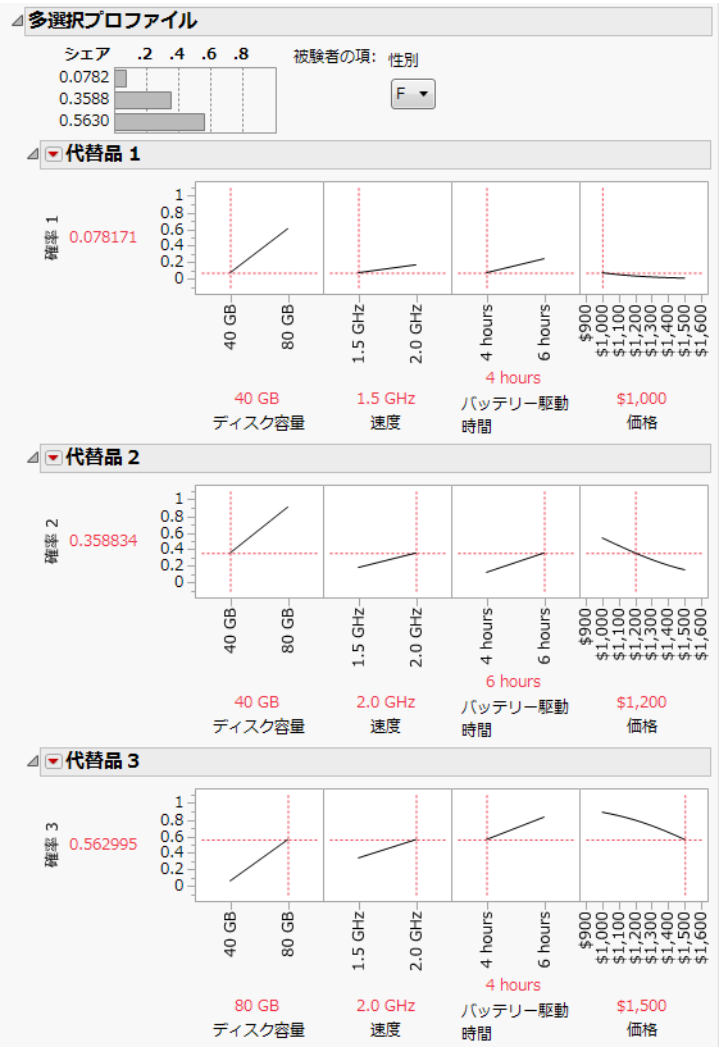
   ウィンドウが開いたら、プロファイルに含める選択肢の数を入力します。ここでは、デフォルトの「3」をそのまま使います。
3. ［**OK**］をクリックします。

   3つの「代替品」のプロファイルが表示されます。このプロファイルでは、性別が「F」に設定されています。

   各プロファイルで、すべての因子がデフォルトの値に設定されています。「代替品1」は、自社がこれから開発する製品です。「代替品2」は、会社Aの製品です。「代替品3」は、会社Bの製品です。

4. 「代替品 1」の「**ディスク容量**」を40 GB、「**速度**」を1.5 GHz、「**バッテリー駆動時間**」を4 hours、「**価格**」を$1,000に設定します。
5. 「代替品 2」の「**ディスク容量**」を40 GB、「**速度**」を2.0 GHz、「**バッテリー駆動時間**」を6 hours、「**価格**」を$1,200に設定します。
6. 「代替品 3」の「**ディスク容量**」を80 GB、「**速度**」を2.0 GHz、「**バッテリー駆動時間**」を4 hours、「**価格**」を$1,500に設定します。

図4.31「F」（女性）の多選択プロファイル

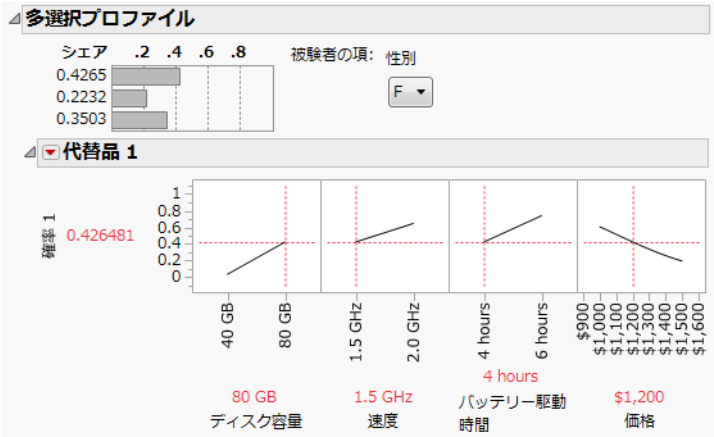


会社Bのシェアが最大で、0.5630です。今のままでは、自社製品を買ってくれる女性が非常に少ないことがわかります。

価格を少し高めに変更しながらも、性能のいずれか1つを改善することにより、自社のシェアを増やすことを考えてみましょう。「代替品1」を見ると、「ディスク容量」の直線の傾きが大きいので、ディスク容量を増やせば最も効果的にシェアを高められるようです。

7. そこで、「代替品1」の**「ディスク容量」**を**80 GB**、**「価格」**を**$1,200**にしてみましょう。

図4.32 改善したノートパソコンの多選択プロファイル



ディスク容量を増やしたことで、ノートパソコンの価格を上げても女性客においておよそ43%のシェアが見込めます。この数字は、会社Bの高性能ノートパソコンをも上回り、図4.31にある現在の設定よりずっと高シェアです。

なお、女性ではなく男性におけるシェアの変化を調べてみると、「ディスク容量」と「速度」の両方を高めた場合には、価格を$1,200にしても44%のシェアが見込めます。

セグメント化の例

この例では、ピザの選択実験データをもとに、消費者のセグメント化を行います。

起動ウィンドウの入力方法については、「複数のデータテーブル」（75ページ）のステップ1～ステップ15を参照してください。または、以下の指示に従います。

クラスターの定義

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Pizza Responses.jmp」を開きます。
2. **「選択モデル」**スクリプトの緑色の三角ボタンをクリックします。
3. 「選択モデル」の赤い三角ボタンをクリックし、［**被験者ごとの勾配を保存**］を選択します。

   データテーブルが開き、主効果と被験者の各交互作用の勾配が保存されます。

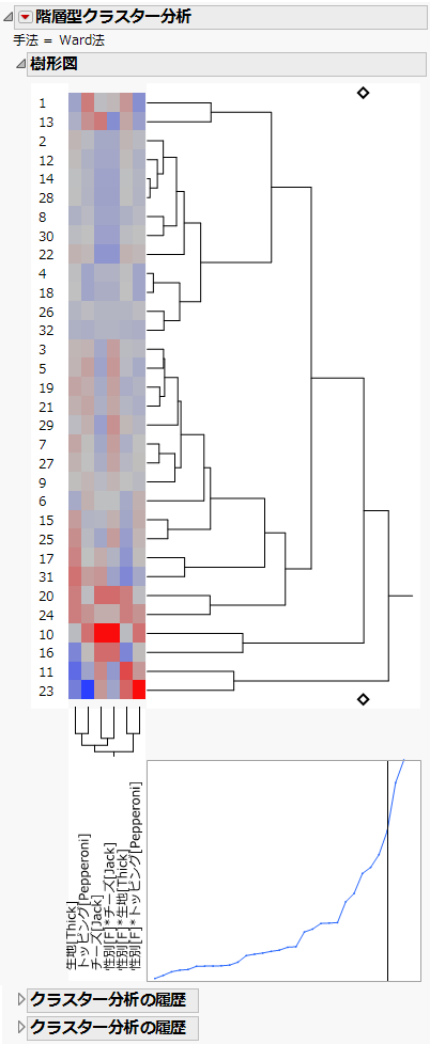
図 4.33 ピザのデータから計算された被験者ごとの勾配（一部）

無題
階層型クラスター分析
元データに結果をマージ

	被験者	生地[Thick]	チーズ[Jack]	トッピング[Pepperoni]	性別[F]*生地[Thick]	性別[F]*チーズ[Jack]	性別[F]*トッピング[Pepperoni]
1	1	-0.00959	-0.00168	0.014876	0.009585	0.001685	-0.01488
2	2	0.002373	-0.00758	-0.00239	0.002373	-0.00758	-0.00239
3	3	0.002129	-0.0079	0.003031	-0.00213	0.007899	-0.00303
4	4	-0.00106	-0.00485	-0.00901	-0.00106	-0.00485	-0.00901
5	5	0.002828	-0.00945	0.00725	-0.00283	0.009453	-0.00725
6	6	-0.0073	-0.00089	0.003761	-0.0073	-0.00089	0.003761

4. 「**階層型クラスター分析**」スクリプトの緑色の三角ボタンをクリックします。

図 4.34 ピザのデータで作成された被験者クラスターの樹形図



このスクリプトは、勾配のテーブルにある**「被験者」**列を除くすべての列を対象に、階層型クラスター分析を行います。いずれかのひし形をクリックすると、行が3つのクラスターに分かれていることがわかります。

5. 「階層型クラスター分析」の赤い三角ボタンをクリックし、［**クラスターの保存**］を選択します。

   勾配を保存したデータテーブルに**「クラスター」**という列が追加されます。同じぐらいの勾配を持つ被験者が、同じクラスターに分けられます。階層型クラスター分析のその他のオプションについては、『多変量分析』の「階層型クラスター分析」章を参照してください。

   勾配の列は、クラスター分けのために作成したものなので、削除してかまいません。

6. **「被験者」**と**「クラスター」**を除くすべての列を選択します。選択した列を右クリックし、［**列の削除**］を選択します。

7. **「元データに結果をマージ」**スクリプトの横にある緑色の三角ボタンをクリックします（図4.33）。

   クラスター情報が「被験者」データテーブルにマージされます。これで、「被験者」データテーブルの列は**「被験者」**、**「性別」**、**「クラスター」**の3つになりました（**図4.35**）。

図4.35「被験者」データテーブルに加わった「クラスター」列

	被験者	性別	クラスター
1	1	M	1
2	2	F	1
3	3	M	1
4	4	F	1
5	5	M	1
6	6	F	1

このデータテーブルを元に、分析を続けることができます。

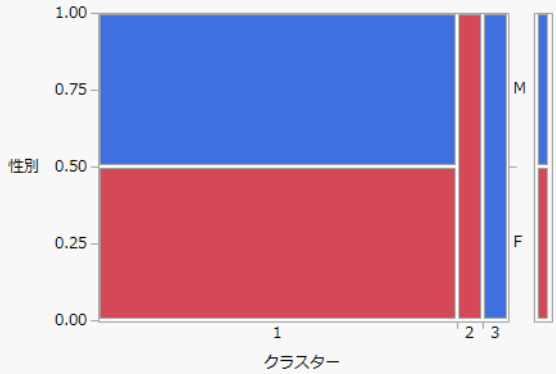
クラスターの検討

1. 「列」パネルで、**「クラスター」**変数の左にあるアイコンをクリックし、［**名義尺度**］を選択します。
2. ［**分析**］>［**二変量の関係**］を選択します。
3. **「性別」**を選択し、［**Y, 目的変数**］をクリックします。
4. **「クラスター」**を選択し、［**X, 説明変数**］をクリックします。
5. ［**OK**］をクリックします。

図4.36 クラスターと性別の分割表に対する分析

クラスターと性別の分割表に対する分析

モザイク図



分割表

性別

クラスター	度数 全体% 列% 行%	F	M	合計
1		14	14	28
		43.75	43.75	87.50
		87.50	87.50	
		50.00	50.00	
2		2	0	2
		6.25	0.00	6.25
		12.50	0.00	
		100.00	0.00	
3		0	2	2
		0.00	6.25	6.25
		0.00	12.50	
		0.00	100.00	
合計		16	16	32
		50.00	50.00	

検定

N	自由度	(-1)*対数尤度	R2乗(U)
32	2	2.7725887	0.1250

検定	カイ2乗	p値(Prob>ChiSq)
尤度比	5.545	0.0625
Pearson	4.000	0.1353

警告: セルのうち20%の期待度数が5未満です。カイ2乗に問題がある可能性があります。

次のようなことがわかります。

– クラスター1は男性と女性の間で均等に分かれています。

– クラスター2には女性しか含まれていません。

– クラスター3には男性しか含まれていません。

場合によっては、「**クラスター**」変数を含めた選択モデルをあてはめてみてもよいでしょう。

「選択モデル」プラットフォームを使用したロジスティック回帰の例

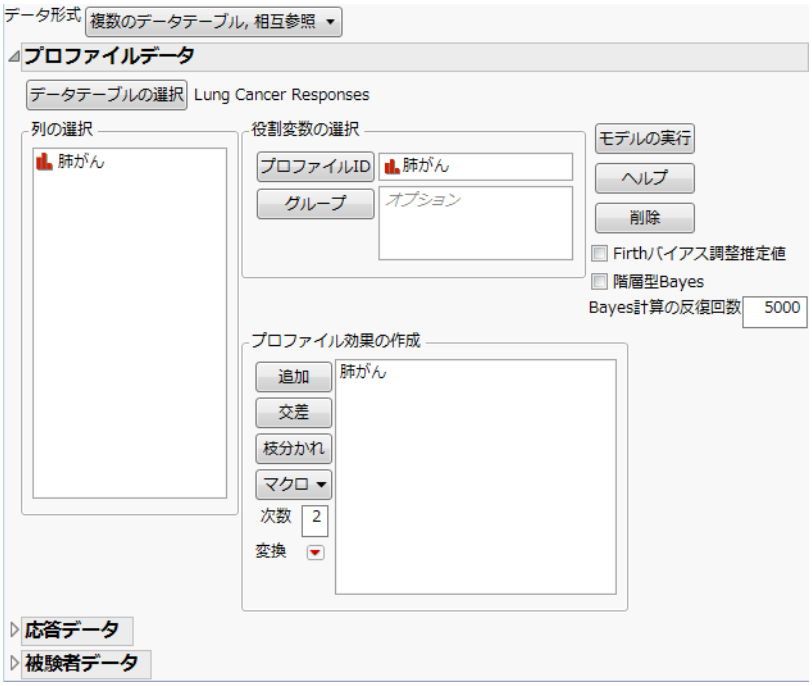
「選択モデル」プラットフォームによるロジスティック回帰

1. ［ヘルプ］>［サンプルデータライブラリ］を選択し、「Lung Cancer Responses.jmp」と「Lung Cancer Choice.jmp」を開きます。

   「Lung Cancer Responses.jmp」は、1列（「**肺がん**」）、2行（「**Cancer**」、「**NoCancer**」）しかありません。

2. ［分析］>［消費者調査］>［選択モデル］を選択します。
3. 「データ形式」メニューから［**複数のデータテーブル , 相互参照**］を選択します。
4. ［**データテーブルの選択**］をクリックし、「Lung Cancer Responses.jmp」を選択して［**OK**］をクリックします。
5. 「**肺がん**」を選択し、［**プロファイル ID**］をクリックします。
6. 「**肺がん**」を選択し、［**追加**］をクリックします。
7. ［Firthバイアス調整推定値］のチェックマークを外します。

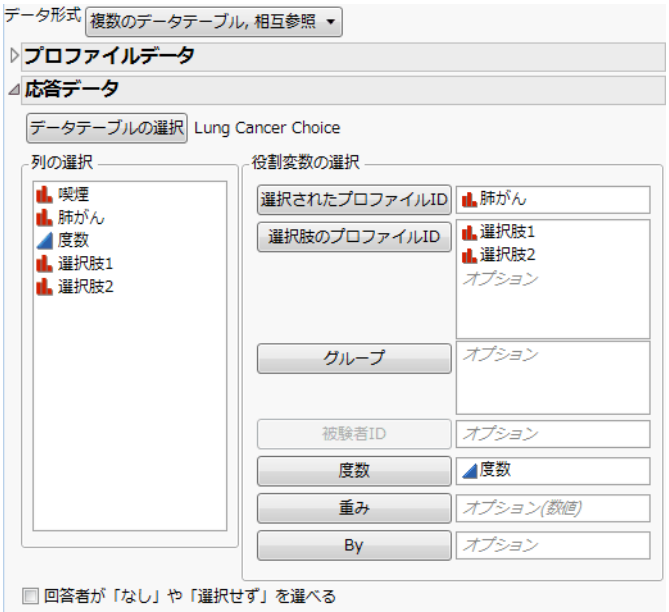
図4.37 入力後の「プロファイルデータ」パネル



8. 「**応答データ**」アウトラインを開きます。
9. ［**データテーブルの選択**］をクリックし、「Lung Cancer Choice.jmp」を選択して［**OK**］をクリックします。

10. 次の手順を行います。
   - 「**肺がん**」を選択し、［**選択されたプロファイル ID**］をクリックします。
   - 「**選択肢 1**」と「**選択肢 2**」を選択し、［**選択肢のプロファイル ID**］をクリックします。
   - 「**度数**」を選択し、［**度数**］をクリックします。

図 4.38 入力後の「応答データ」パネル

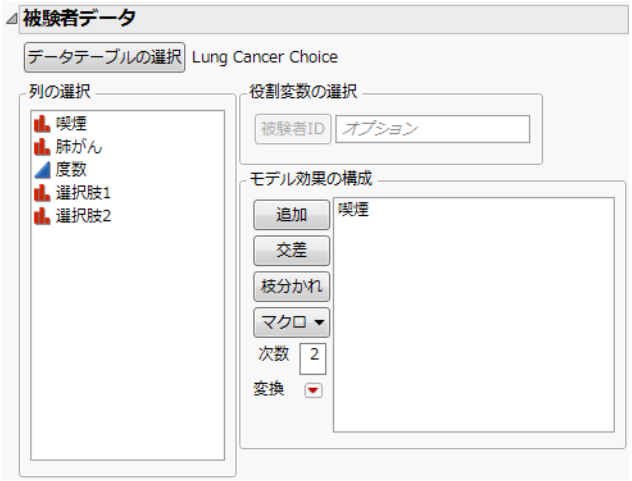


11.「**被験者データ**」アウトラインを開きます。

12.［**データテーブルの選択**］をクリックし、「Lung Cancer Choice.jmp」を選択して［**OK**］をクリックします。

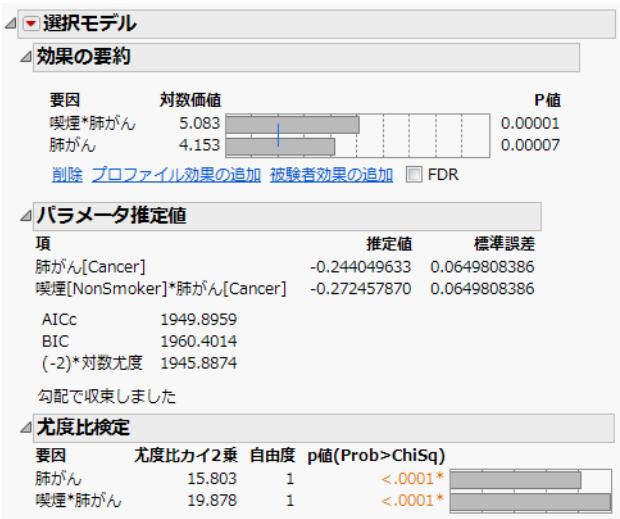
13.「**喫煙**」を選択し、［**追加**］をクリックします。

図 4.39 入力後の「被験者データ」パネル



14. ［**モデルの実行**］をクリックします。

図 4.40「選択モデル」を使ったロジスティック回帰



「モデルのあてはめ」プラットフォームによるロジスティック回帰

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Lung Cancer.jmp」を開きます。
2. ［**分析**］>［**モデルのあてはめ**］を選択します。

   データテーブルに「モデル」スクリプトが含まれているため、「モデルの指定」ウィンドウへの入力は自動的に行われます。［**名義ロジスティック**］の手法が指定されています。

3. ［実行］をクリックします。

図4.41「モデルのあてはめ」を使ったロジスティック回帰

名義ロジスティックのあてはめ 肺がん

効果の要約

勾配で収束しました, 4回の反復
度数: 度数

反復履歴

モデル全体の検定

モデル	(-1)*対数尤度	自由度	カイ2乗	p値(Prob>ChiSq)
差	9.93901	1	19.87802	<.0001*
完全	972.94369			
縮小	982.88270			

R2乗(U)	0.0101
AICc	1949.9
BIC	1960.4
オブザベーション(または重みの合計)	1418

あてはめの詳細

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片	-0.4880993	0.1299617	14.11	0.0002*
喫煙[NonSmoker]	-0.5449157	0.1299617	17.58	<.0001*

推定値は次の対数オッズに対するものです： Cancer/NoCancer

推定値の共分散

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
喫煙	1	1	19.878016	<.0001*

選択モデルにおける**「喫煙*肺がん」**の尤度比カイ2乗検定が、ロジスティックモデルにおける「喫煙」の尤度比カイ2乗検定と一致しています。図4.40と図4.41のレポートから、喫煙が肺がんの発病に対して高度に有意であることが分かります。詳細については、『基本的な回帰モデル』の「ロジスティック回帰」章を参照してください。

対応のあるデータに対する条件付きロジスティック回帰の例

ここでは、「選択モデル」プラットフォームを使い、「ケース（症例）」と「コントロール（対照）」とでペアごとに対応している子宮内膜がんの調査結果に対して、条件付きロジスティック回帰分析を行う例を紹介します。使用するデータは、Breslow and Day（1980）で取り上げられている「Los Angeles Study of the Endometrial Cancer Data」からの引用です。このケースコントロール分析の目標は、高血圧による影響を考慮しながら、胆嚢疾患の相対リスクを推定することです。「アウトカム」の1の値は、子宮内膜がんの発症（ケース群）を示し、0はコントロール群を示します。胆嚢疾患と高血圧も、同様に1と0で示されています。

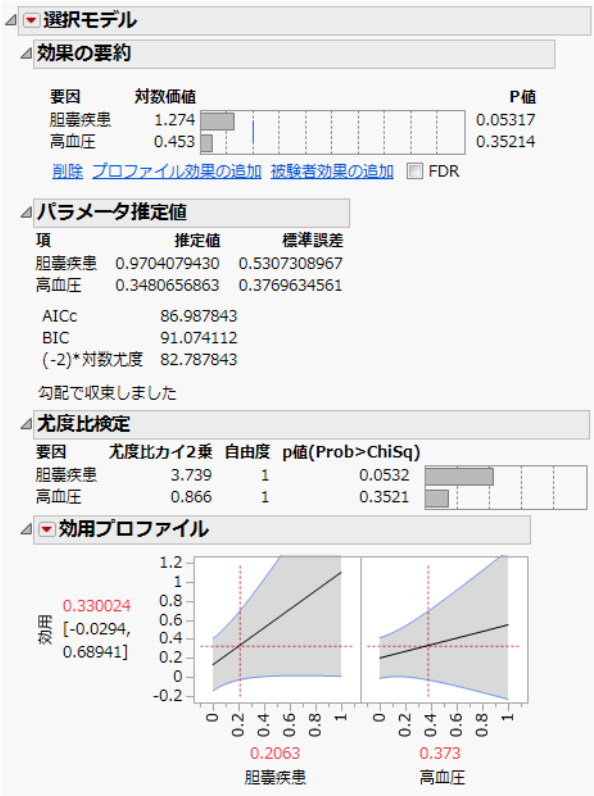
「選択モデル」プラットフォームによって、通常のロジスティック回帰を行う手順については、「ロジスティック回帰」（124ページ）を参照してください。

1. ［ヘルプ］>［サンプルデータライブラリ］を選択し、「Endometrial Cancer.jmp」を開きます。
2. ［分析］>［消費者調査］>［選択モデル］を選択します。

3. 「データ形式」が［**1 つのデータテーブル , 積み重ね**］になっていることを確認します。
4. ［**データテーブルの選択**］ボタンをクリックします。
5. プロファイルのデータテーブルとして「Endometrial Cancer」を選択します。［**OK**］をクリックします。
6. 「**アウトカム**」を選択し、［**応答の指示変数**］をクリックします。
7. 「**ペア**」を選択し、［**グループ**］をクリックします。
8. 「**胆嚢疾患**」と「**高血圧**」を選択し、「プロファイル効果の作成」パネルで［**追加**］をクリックします。
9. ［**Firth バイアス調整推定値**］のチェックマークを外します。
10. ［**モデルの実行**］をクリックします。
11. 「選択モデル」の赤い三角ボタンをクリックし、［**効用プロファイル**］を選択します。

    図 4.42 のようなレポートが作成されます。

図 4.42 子宮内膜がんデータのロジスティック回帰



要因ごとに尤比度検定が実行されています。「**胆嚢疾患**」が、α 水準を 0.05 としたとき、統計的にほぼ有意であることがわかります（p 値は 0.0532）。また、効用プロファイルによって、応答に対する要因の影響を視覚的に確認できます。

データを2つの分析テーブルに変換する例

「Daganzo Trip.jmp」に保存されているDaganzoのデータを見てみましょう。このデータテーブルには、3通りの交通手段のそれぞれにかかる時間と、各被験者が選択した交通手段が記録されています。

「選択した交通手段」と「被験者」の追加

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Daganzo Trip.jmp」データテーブルを開きます。

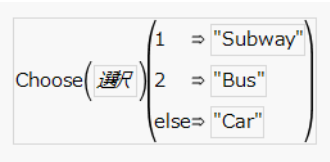
図4.43「Daganzo Trip.jmp」データテーブル（一部）

	Subway	Bus	Car	選択
1	16.481	16.196	23.89	2
2	15.123	11.373	14.182	2
3	19.469	8.822	20.819	2
4	18.847	15.649	21.28	2
5	12.578	10.671	18.335	2

まず、「**選択**」の番号を交通手段の名称に変換する必要があります。これは、計算式エディタにある**Choose**関数を使えば、以下の手順で簡単に行えます。

2. ［**列**］>［**列の新規作成**］を選択します。
3. 「列名」に「**選択した交通手段**」と入力し、「尺度」を［**名義尺度**］に設定します。
4. ［**列プロパティ**］をクリックし、［**計算式**］を選択します。
5. ［関数(グループ別)］の中から［**条件付き**］を選択し、［**Choose**］をクリックします。次に、カンマキーを2回押して関数に引数を追加します。
6. 「式」の枠をクリックした後、「**選択**」列をクリックしてChooseの計算式（「式」）に挿入します。また、「節」入力ボックスをダブルクリックして「**"Subway"**」、「**"Bus"**」、「**"Car"**」と（引用符も含めて）入力します（図4.44）。

図4.44 Daganzoのデータの「選択した交通手段」列にChoose関数を適用



7. 「計算式エディタ」ウィンドウで［**OK**］をクリックします。
8. 「列の新規作成」ウィンドウで［**OK**］をクリックします。

データテーブルに新しい「**選択した交通手段**」列が表示されます。各行には、各被験者によって選択された選択肢が含まれています。次に、各被験者の識別番号が含まれている列を作成します。

9. ［**列**］>［**列の新規作成**］を選択します。

10.「列名」に「**被験者**」と入力します。

11.「**データの初期化**」のドロップダウンメニューを開き、［**シーケンスデータ**］を選択します。

12.［**OK**］をクリックします。

図4.45「Daganzo Trip.jmp」データテーブルに「選択した交通手段」と「被験者」の列を追加

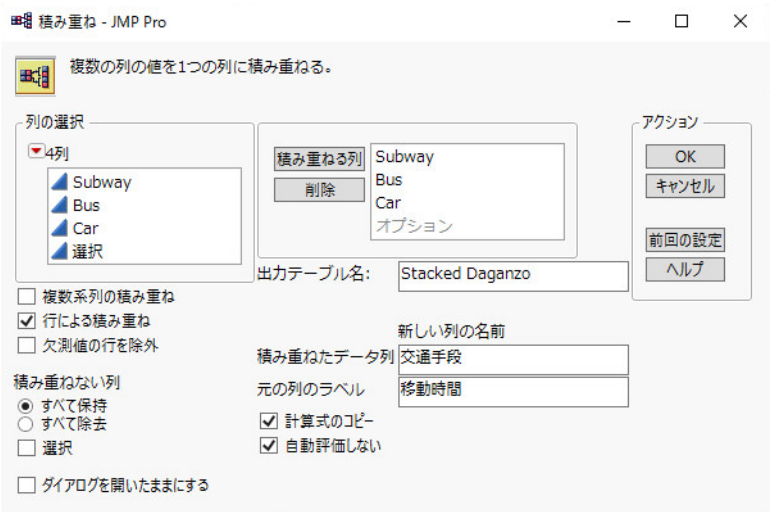
	Subway	Bus	Car	選択	選択した交通手段	被験者
1	16.481	16.196	23.89	2	Bus	1
2	15.123	11.373	14.182	2	Bus	2
3	19.469	8.822	20.819	2	Bus	3
4	18.847	15.649	21.28	2	Bus	4
5	12.578	10.671	18.335	2	Bus	5
6	11.513	20.582	27.838	1	Subway	6

データの積み重ね

プロファイルデータには、選択肢をそれぞれ1行ずつに収める必要があります。

1. ［**テーブル**］>［**列の積み重ね**］を選択します。
2. 「**Subway**」、「**Bus**」、「**Car**」を選択し、［**積み重ねる列**］をクリックします。
3. 「出力テーブル名」に「**Stacked Daganzo**」と入力します。「積み重ねたデータ列」に「**移動時間**」、「元の列のラベル」に「**交通手段**」と入力します。

図4.46 Daganzoのデータで行う積み重ねのダイアログボックス



4.［**OK**］をクリックします。

図4.47「Stacked Daganzo」データテーブル（一部）

	選択	選択した交通手段	被験者	交通手段	移動時間
1	2	Bus	1	Subway	16.481
2	2	Bus	1	Bus	16.196
3	2	Bus	1	Car	23.89
4	2	Bus	2	Subway	15.123
5	2	Bus	2	Bus	11.373
6	2	Bus	2	Car	14.182
7	2	Bus	3	Subway	19.469

プロファイルデータの作成

プロファイルのデータテーブルには、**「被験者」**、**「交通手段」**、**「移動時間」**の各列が必要です。

1. **「被験者」**、**「交通手段」**、**「移動時間」**の各列を選択し、［**テーブル**］＞［**サブセット**］を選択します。
2. ［**すべての行**］と［**選択されている列**］を選択し、［**OK**］をクリックします。

   図4.48は、データの一部です。デフォルトのテーブル名は、「Stacked Daganzoのサブセット」です。

図4.48 Stacked Daganzoのデータのサブセット（一部）

	被験者	交通手段	移動時間
1	1	Subway	16.481
2	1	Bus	16.196
3	1	Car	23.89
4	2	Subway	15.123
5	2	Bus	11.373
6	2	Car	14.182
7	3	Subway	19.469

応答データの作成

応答のデータテーブルには、**「被験者」**、**「選択した交通手段」**の各列と、各選択肢のための列が必要です。

3. 「Daganzo Trip.jmp」から、**「被験者」**と**「選択した交通手段」**の各列を選択します。
4. ［**テーブル**］＞［**サブセット**］を選択します。
5. ［**すべての行**］と［**選択されている列**］を選択し、［**OK**］をクリックします。

   デフォルトのテーブル名は、「Daganzo Tripのサブセット」です。
6. ［**列**］＞［**列の新規作成**］を選択します。
7. 「列名」に**「選択肢」**と入力します。
8. ［**データタイプ**］＞［**文字**］を選択します。
9. 「追加する列の数」に「3」と入力します。
10. ［**OK**］をクリックします。

    「選択肢 1」、**「選択肢 2」**、**「選択肢 3」**という列が追加されています。

11. **「選択肢 1」**の最初の行に「Bus」と入力します。セルを右クリックし、**［初期値を埋める / 挿入］** > **［テーブルの末尾まで入力］**を選択します。
12. **「選択肢　2」**の最初の行に「Subway」と入力します。セルを右クリックし、**［初期値を埋める / 挿入］** > **［テーブルの末尾まで入力］**を選択します。
13. **「選択肢 3」**の最初の行に「Car」と入力します。セルを右クリックし、**［初期値を埋める / 挿入］** > **［テーブルの末尾まで入力］**を選択します。

    図4.49のようなテーブルになります。

図4.49 Daganzoのデータのサブセットに選択肢集合を追加

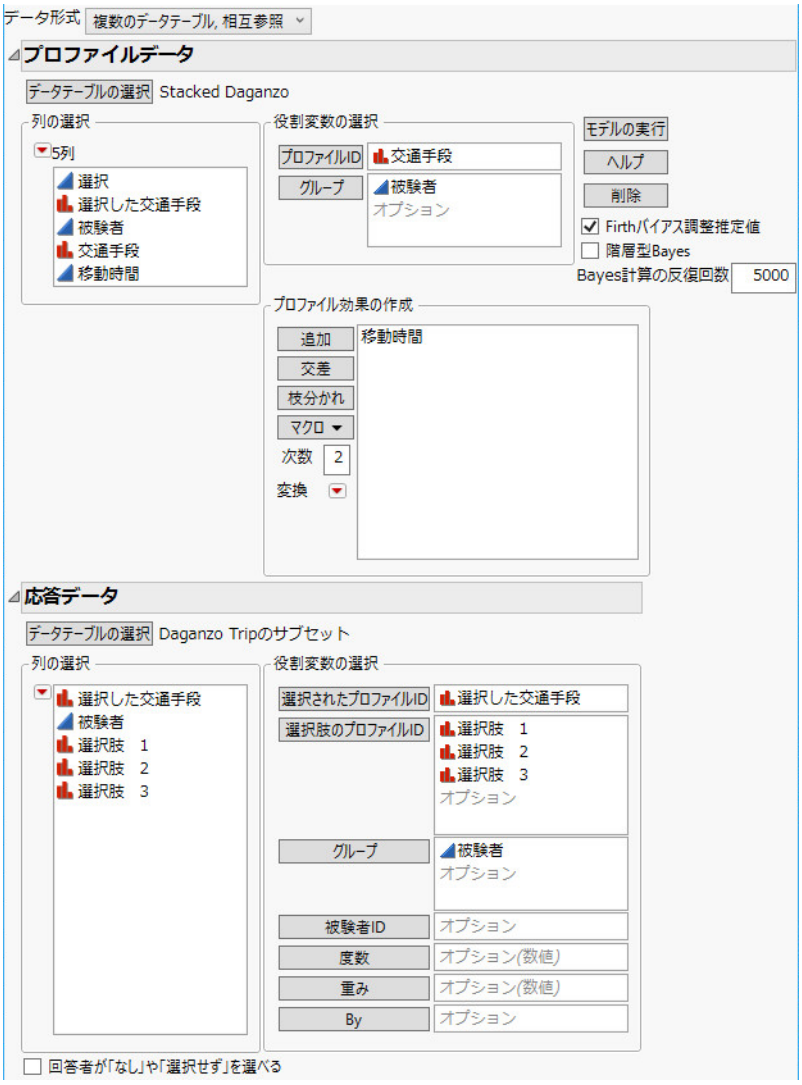
	選択した交通手段	被験者	選択肢 1	選択肢 2	選択肢 3
1	Bus	1	Bus	Subway	Car
2	Bus	2	Bus	Subway	Car
3	Bus	3	Bus	Subway	Car
4	Bus	4	Bus	Subway	Car
5	Bus	5	Bus	Subway	Car
6	Subway	6	Bus	Subway	Car
7	Subway	7	Bus	Subway	Car

モデルをあてはめる

元の「Daganzo Trip.jmp」テーブルを2つのテーブルに分けたところで、「選択モデル」プラットフォームを実行します。

1. **［分析］** > **［消費者調査］** > **［選択モデル］**を選択します。
2. 「データ形式」メニューから**［複数のデータテーブル , 相互参照］**を選択します。
3. 図4.50のとおりにモデルを指定します。

図4.50 Daganzoのデータのサブセットを使ったときの「選択モデル」ダイアログボックス



4. ［**モデルの実行**］をクリックします。

計算されたパラメータ推定値は、効用関数における**「移動時間」**の係数を表します（図4.51）。

図4.51 Daganzoのデータの「移動時間」に対するパラメータ推定値

選択モデル

パラメータ推定値

項	推定値	標準誤差
移動時間	-0.341768586	0.0745222259

AICc	68.766653
BIC	70.595342
(-2)*対数尤度	66.683319
(-2)*Firth対数尤度	61.490004

勾配で収束しました
Firthバイアス調整推定値

尤度比検定

要因	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
移動時間	43.178	1	<.0001*

係数が負の値になっているのは、移動時間の増加が消費者の効用（満足度）に対して負の効果を持つことを表します。尤度比検定において、「移動時間」の効果は有意になっています。

データを1つの分析テーブルに変換する例

2～3つのデータテーブルを作成するより、1つのデータテーブルにまとめた方が効率的な場合があります。1つのデータテーブルにする場合、被験者の列は先ほどの例と同じ手順で追加します。また、選択肢の集合として3つの列（**選択肢1**、**選択肢2**、**選択肢3**）を使用する代わりに、応答の指示変数（応答のインジケータの列）を作成します。1つのテーブルに変換する場合は、以下の手順で操作します。

1. 「データを2つの分析テーブルに変換する例」（116ページ）の「データの積み重ね」手順に従って「Stacked Daganzo.jmp」を作成するか、すでに作成したものを開きます。
2. ［**列**］>［**列の新規作成**］を選択します。
3. 「列名」に「**応答**」と入力します。
4. ［**列プロパティ**］をクリックし、［**計算式**］を選択します。
5. 関数リストから［**条件付き**］を選択し、［**If**］を選択します。
6. 「**選択した交通手段**」の列を選択し、計算式（「式」）に挿入します。
7. 「=」を入力して「**交通手段**」を選択します。
8. ［**then節**］に「1」を入力し、［**else節**］に「0」を入力します。
9. 「計算式エディタ」ウィンドウで［**OK**］をクリックします。「列の新規作成」ウィンドウで［**OK**］をクリックします。

   入力が完了した計算式は、図4.52のようになります。

図4.52 指示変数を作成する計算式

```
If( 選択した交通手段 == 交通手段 ⇒ 1
    else                         ⇒ 0 )
```

10.「**被験者**」、「**移動時間**」、「**応答**」の各列を選択し、［**テーブル**］>［**サブセット**］を選択します。

11.［**すべての行**］と［**選択されている列**］を選択し、［**OK**］をクリックします。

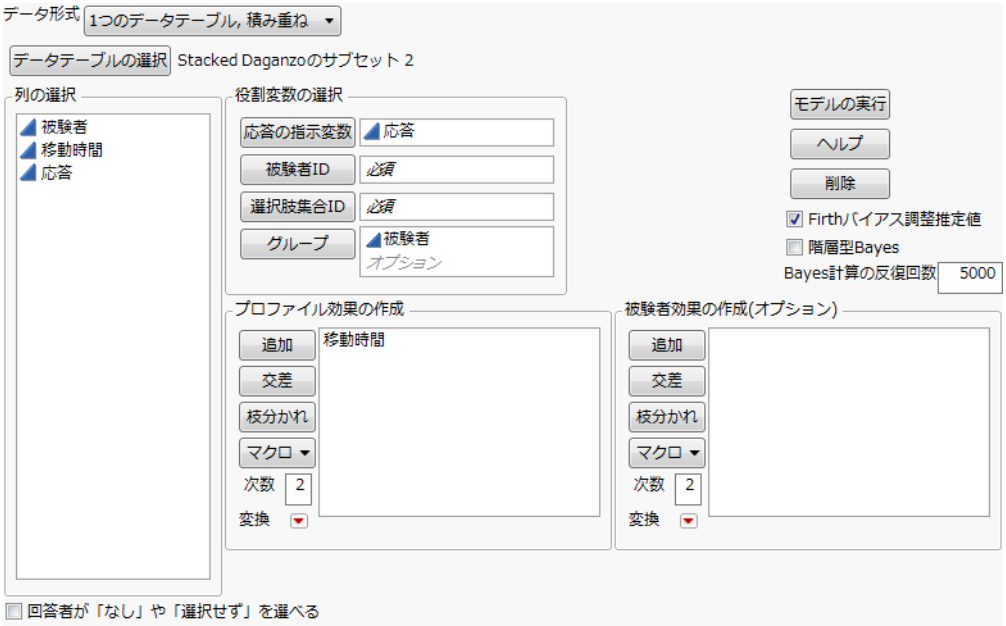
新しいデータテーブルは、図4.53のようになります。

図4.53「Stacked Daganzo」データテーブルのサブセット（一部）

	被験者	移動時間	応答
1	1	16.481	0
2	1	16.196	1
3	1	23.89	0
4	2	15.123	0
5	2	11.373	1
6	2	14.182	0
7	3	19.469	0

12.［**分析**］>［**消費者調査**］>［**選択モデル**］の順にクリックして起動ウィンドウを開き、図4.54のようにモデルを指定します。

図4.54「Stacked Daganzo」データのサブセットを使い、テーブル1つで分析するときの「選択モデル」ダイアログボックス



13. ［**モデルの実行**］をクリックします。

図4.55 1つのテーブルで分析したときの「移動時間」のパラメータ推定値

選択モデル

パラメータ推定値

項	推定値	標準誤差
移動時間	-0.341768586	0.0745222259

AICc	68.766653
BIC	70.595342
(-2)*対数尤度	66.683319
(-2)*Firth対数尤度	61.490004

勾配で収束しました
Firthバイアス調整推定値

尤度比検定

要因	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
移動時間	43.178	1	<.0001*

結果は、テーブルを2つ使って分析した場合（図4.51）と同じです。

この章では、単純な例を使って「選択モデル」プラットフォームの機能を紹介しましたが、もっと複雑な変換や交互作用項を含む、より高度なモデルも実行できます。

「選択モデル」プラットフォームの統計的詳細

- 「分析対象のデータに対する特殊な規則」
- 「効用と確率」
- 「勾配」

分析対象のデータに対する特殊な規則

デフォルトの選択肢集合

プロファイルのデータテーブルに保存されているプロファイルが、すべての質問で共通の選択肢になっている場合には、「選択モデル」起動ウィンドウの「応答データ」パネルにある**「役割変数の選択」**での［**選択肢のプロファイル ID**］を指定しなくてもかまいません。省略した場合は、プロファイルのデータテーブルに保存されているプロファイルが、すべての質問において共通して提示されているとみなされます。

被験者情報も含む応答データテーブル

応答のデータテーブルに被験者の情報も一緒に保存している場合は、別々に保存している場合と同じように、「被験者データ」セクションの［**データテーブルの選択**］でそのデータテーブルを指定してください。この場合、［**被験者 ID**］列は指定しなくてもかまいません。たとえ被験者 ID を指定しても、分析の計算には使われません。一般的には、被験者の属性データは、同一被験者であれば質問が変わっても変化しないでしょう。

ロジスティック回帰

「選択モデル」プラットフォームでは、通常のロジスティック回帰も実行できます。

メモ： ロジスティック回帰モデルをあてはめるには、「二変量の関係」または「モデルのあてはめ」プラットフォームの方が便利です。ただし、必要であれば「選択モデル」プラットフォームでロジスティック回帰を実行することもできるので、以下にその方法を紹介します。

「選択モデル」によってロジスティック回帰分析を実行するには、まず、データが選択モデル用の形式になっている必要があります。データの準備として、次の 3 つの作業が必要となります。

- ロジスティック回帰における応答変数の各水準を、1 行ずつに収めたプロファイルのデータテーブルを作成します。このプロファイルのデータテーブルには、応答変数の水準のみを含めます。
- 説明変数を含んだデータテーブルを保存し、それを応答のデータテーブルに指定します。
- 応答のデータテーブルを再び、今度は被験者のデータテーブルとして指定します。

「選択モデル」プラットフォームを使ってロジスティック回帰を実施する例については、「「選択モデル」プラットフォームを使用したロジスティック回帰の例」（111 ページ）および「対応のあるデータに対する条件付きロジスティック回帰の例」（114 ページ）を参照してください。

効用と確率

選択モデルでは、消費者の効用（utility）がパラメータ推定値から算出されます。「効用」とは、特定の属性をもつ製品から消費者が得る満足度のことで、モデルのパラメータ推定値から算出されます。なお、効用が線形関数で表されている場合、パラメータ推定値は限界効用（marginal utility）を示します。

選択モデルは、次のように表わされます。

$X[k]$を、切片も含んだ、被験者属性を示す行ベクトルとし、

$Z[j]$を、切片のない、選択属性を示す行ベクトルとします。

この場合、第k被験者が、m個の選択肢のなかから、j番目のものを選択する確率は、次式で表されます。

$$P_i[jk] = \frac{\exp(\beta'(X[k] \otimes Z[j]))}{\sum_{l=1}^{m} \exp(\beta'(X[k] \otimes Z[l]))}$$

この式で、

- $\otimes$は、Kronecker積です。
- 分子は、実際に選択されたj番目の選択肢に対する値です。
- 分母は、その質問で被験者に提示されたm個の選択肢すべての合計を示します。

勾配

［被験者ごとの勾配を保存］オプションで保存される勾配は、Newton-Raphson法のステップを被験者ごとに平均したものです。対数尤度の最大化が行われた段階での勾配の合計は0ベクトルになっています。また、この勾配の合計にヘッセ行列の逆行列を掛けたもの（Newton-Raphson法のステップ幅）も0ベクトルです。$\Delta = H^{-1}g = 0$。ここで、gは最尤法の対数尤度から計算された勾配の合計、H^{-1}はヘッセ行列の逆行列です。ヘッセ行列とは、対数尤度の二階偏導関数（対数尤度を2回編微分したもの）を要素とする行列のことです。

ここで、Δを分解すると次のように表わされます。

$$\Delta = \Sigma_{ij}\Delta_{ij} = \Sigma H^{-1}g_{ij} = 0$$

上式において、iは被験者を示す番号、jは各被験者における選択を示す番号、ΔijはNewton-Raphson法におけるステップ、g_{ij}は各選択における対数尤度の勾配を示します。

そこで、各被験者の平均ステップは、次のように計算できます。

$$\bar{\Delta}_i = \Sigma_j \frac{\Delta_{ij}}{n_i},$$

この式で、n_{i}はi番目の被験者が行った選択の回数を示します。$\overline{\Delta}_i$は、被験者iがパラメータに与える影響力に関係しています。あるセグメントに属する被験者の選好構造が、他の被験者と異なる場合には、そのセグメントの被験者がパラメータに与える影響力は強くなります。よって、この平均ステップをクラスター分析に利用できると考えられます。$\overline{\Delta}_i$はそのような影響力を示します。この値は、データテーブルに保存することができますので、保存したあとに「クラスター分析」プラットフォームなどで分析できます。

第 5 章

MaxDiff

MaxDiff モデルのあてはめ

MaxDiff 法（maximum difference scaling）は、消費者（回答者、被験者）の選択データをもとに、製品の各属性がもつ相対的な重要度を求める方法のひとつです。MaxDiff 法の実験では、回答者は最も好きな選択肢（最高）と、最も好きではない選択肢（最悪）を選びます。最高と最悪を選ぶので、最高だけを選ぶ実験よりも、順位付けに関する情報がデータに含まれています。

「MaxDiff」プラットフォームでは、次のようなことが実行できます。

- 製品の属性だけでなく、消費者の属性（たとえば性別や年齢）に関する情報も利用する。
- 分析対象のデータにおいて、いくつかの形式がサポートされている。
- 消費者ごとにスコアを求め、求められたスコアを用いて消費者をセグメント（クラスター）に分ける。
- JMP PRO Bayes（ベイズ）流の方法で、消費者ごとの係数値を求める。
- バイアス修正を伴う最尤推定（Firth 1993）を行う。

図 5.1 MaxDiff 法による「全水準の比較レポート」

全水準の比較レポート

差 (行 − 列) 差の標準誤差 Wald p値	Gyro	Sour Cream and Onion	Truffle Fries	Biscuits and Gravy	Reuben	Ketchup	All Dressed	Dill Pickle	Barbecue	Southern Barbecue
Gyro	0	-1.1747	-1.3082	-1.2688	-0.6992	-0.8229	-0.9467	-1.027	-2.4417	-1.9535
	0	0.33327	0.35477	0.34456	0.34758	0.32802	0.33571	0.35251	0.41499	0.38939
		0.00062	0.00035	0.00036	0.04668	0.01356	0.00569	0.00432	4.32e-8	2.01e-6
Sour Cream and Onion	1.17469	0	-0.1335	-0.0941	0.47548	0.35175	0.22804	0.14766	-1.267	-0.7788
	0.33327	0	0.31379	0.30716	0.30932	0.31915	0.28158	0.30707	0.37654	0.34896
	0.00062		0.67137	0.75983	0.12709	0.27279	0.41975	0.63156	0.00105	0.02764
Truffle Fries	1.30817	0.13348	0	0.03935	0.60896	0.48523	0.36152	0.28114	-1.1335	-0.6453
	0.35477	0.31379	0	0.31345	0.31817	0.31157	0.31136	0.30871	0.39213	0.34904
	0.00035	0.67137		0.90032	0.05821	0.12223	0.24809	0.36444	0.00462	0.06715
Biscuits and Gravy	1.26882	0.09413	-0.0393	0	0.56961	0.44588	0.32217	0.24179	-1.1729	-0.6847
	0.34456	0.30716	0.31345	0	0.30639	0.29752	0.29974	0.28249	0.36267	0.34038
	0.00036	0.75983	0.90032		0.06567	0.1368	0.28477	0.39388	0.00161	0.0467
Reuben	0.69921	-0.4755	-0.609	-0.5696	0	-0.1237	-0.2474	-0.3278	-1.7425	-1.2543
	0.34758	0.30932	0.31817	0.30639	0	0.30893	0.30148	0.30328	0.37117	0.35324
	0.04668	0.12709	0.05821	0.06567		0.68956	0.41355	0.28208	7.69e-6	0.00056
Ketchup	0.82294	-0.3517	-0.4852	-0.4459	0.12373	0	-0.1237	-0.2041	-1.6188	-1.1305
	0.32802	0.31915	0.31157	0.29752	0.30893	0	0.29728	0.27938	0.36911	0.34808
	0.01356	0.27279	0.12223	0.1368	0.68956		0.67812	0.46661	2.64e-5	0.00154
All Dressed	0.94665	-0.228	-0.3615	-0.3222	0.24744	0.12371	0	-0.0804	-1.4951	-1.0068
	0.33571	0.28158	0.31136	0.29974	0.30148	0.29728	0	0.29447	0.36838	0.3418
	0.00569	0.41975	0.24809	0.28477	0.41355	0.67812		0.78538	9.22e-5	0.00393
Dill Pickle	1.02703	-0.1477	-0.2811	-0.2418	0.32782	0.20409	0.08038	0	-1.4147	-0.9264
	0.35251	0.30707	0.30871	0.28249	0.30328	0.27938	0.29447	0	0.3539	0.35532
	0.00432	0.63156	0.36444	0.39388	0.28208	0.46661	0.78538		0.00012	0.01038
Barbecue	2.44172	1.26703	1.13355	1.1729	1.7425	1.61877	1.49507	1.41468	0	0.48824
	0.41499	0.37654	0.39213	0.36267	0.37117	0.36911	0.36838	0.3539	0	0.39164
	4.32e-8	0.00105	0.00462	0.00161	7.69e-6	2.64e-5	9.22e-5	0.00012		0.21514
Southern Barbecue	1.95348	0.77879	0.6453	0.68465	1.25426	1.13053	1.00683	0.92644	-0.4882	0
	0.38939	0.34896	0.34904	0.34038	0.35324	0.34808	0.3418	0.35532	0.39164	0
	2.01e-6	0.02764	0.06715	0.0467	0.00056	0.00154	0.00393	0.01038	0.21514	

目次

「MaxDiff モデル」プラットフォームの概要 129
「MaxDiff」プラットフォームの例 129
1 つのデータテーブル 130
複数のデータテーブル 133
「MaxDiff」プラットフォームの起動 136
［1 つのデータテーブル , 積み重ね］の起動ウィンドウ 137
［複数のデータテーブル , 相互参照］の起動ウィンドウ 138
「MaxDiff モデル」レポート 143
効果の要約 143
MaxDiff 結果 144
パラメータ推定値 145
Bayes パラメータ推定値 147
尤度比検定 148
「MaxDiff」プラットフォームのオプション 148
「比較」レポート 150
Bayes チェーンの保存 151

「MaxDiff モデル」プラットフォームの概要

MaxDiff 法は、選択実験データに対する統計手法です。MaxDiff 法は、「最良 - 最悪スケーリング」Best-Worst Scaling (BWS) とも呼ばれています。MaxDiff 法の実験では、複数の選択肢から最良のものと最悪のものを回答者に選んでもらいます。MaxDiff 法では、好きなものだけを選んでもらった実験に比べ、選好についてより詳しい情報が得られます。MaxDiff 法については、Louviere et al.（2015）を参照してください。また、選択モデルについては、Louviere et al.（2015）、Train（2009）、Rossi et al.（2005）を参照してください。

MaxDiff モデルは、「各自の効用に基づいて、いずれかの選択肢を確率的に人々は選択している」と仮定しています。一つひとつの選択肢に、その人にとっての価値（効用）があると仮定しています。MaxDiff 法は、その効用を推定します。推定された効用に基づいて、ある選択肢が他の選択肢より好まれる確率も推定できます。モデルとしては、条件付きロジスティック回帰が仮定されます。McFadden（1974）を参照してください。

メモ： 1 因子の MaxDiff 調査は、「MaxDiff 計画」プラットフォームで計画することができます。『実験計画 (DOE)』の「MaxDiff 計画」章を参照してください。

セグメント化や被験者効果の Bayes 推定

マーケットリサーチ（市場調査）では、選好のパターンによって、消費者（被験者）をグループに分けたい場合があります。しかし、通常のモデルで被験者ごとのパラメータ推定値を求めるには、被験者 1 人あたりに多くのデータが必要です。被験者 1 人あたりに多くのデータがあるなら、被験者 ID を「応答データ」セクションで［By］変数に指定するか、被験者 ID をモデル項として含めることができます。ただし、後者の方法は、被験者数が多いと計算が難しくなります。これに代わる統計手法として、統計学の文献では、Bayes モデルや混合モデルなどが提案されています。

［By］変数が指定できるほどのデータがない場合は、［**被験者ごとの勾配を保存**］オプションを使って被験者をクラスターに分け、セグメント化することができます。このオプションは、ヘッセ行列で尺度化した各パラメータの傾きを被験者ごとに平均し、新しいデータテーブルに表示します。例として、「選択モデル」章の「セグメント化の例」（107 ページ）を参照してください。傾きの値についての詳細は、「選択モデル」章の「勾配」（125 ページ）を参照してください。

JMP PRO MaxDiff 法では、階層型 Bayes モデルを使って被験者効果を推定することもできます。この Bayes 法による統計分析も、マーケットのセグメント化に役立ちます。

「MaxDiff」プラットフォームの例

30 人の回答者を対象に、MaxDiff 法を使い、7 種類の風味のポテトチップスを比較する調査を行いました。風味が異なるポテトチップスを 3 つずつ各回答者に提示しました。それら 3 つの選択肢のなかで回答者が最も好きなものを「1」、最も好きでないものを「-1」として記録しました。それ以外の選択肢は、「0」として記録しました。

「MaxDiff」プラットフォームでは、1つのデータテーブルにまとめたデータでも、複数のデータテーブルに分かれたデータでも分析できます。複数のデータテーブルを使う場合は、応答、選択肢集合、被験者に関する情報をそれぞれ個別のデータテーブルに保存しておきます。1つのデータテーブルを使う場合は、すべての情報を1つのデータテーブルにまとめます。

- 「1つのデータテーブル」(130ページ)に、1つのデータテーブルにまとめたデータに基づいて、MaxDiff分析を行う操作方法を説明しています。この例の分析を発展させて、プロファイルや被験者の情報を加えることもできます。
- 「複数のデータテーブル」(133ページ)に、複数のデータテーブルに記録されたデータに基づいて、MaxDiff分析を行う操作方法を説明しています。

1つのデータテーブル

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Potato Chip Combined.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**MaxDiff**］を選択します。

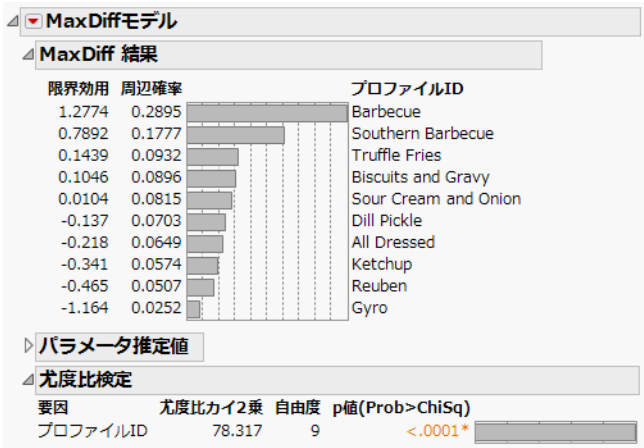
   「データ形式」は、デフォルトでは［1つのデータテーブル, 積み重ね］に設定されています。
3. ［**データテーブルの選択**］をクリックします。
4. 「Potato Chip Combined.jmp」を選択し、［**OK**］をクリックします。
5. 次の手順に従って、列に役割を割り当てます。入力後の起動ウィンドウは、図5.2のようになります。
   - 「**選択**」を選択し、［**応答の指示変数**］をクリックします。
   - 「**回答者**」を選択し、［**被験者ID**］をクリックします。
   - 「**選択肢集合ID**」を選択し、［**選択肢集合ID**］をクリックします。
   - 「**プロファイルID**」を選択し、「プロファイル効果の作成」パネルで［**追加**］をクリックします。

図5.2 入力後の「MaxDiff」起動ウィンドウ

「最悪」の値は、応答の列を［応答の指示変数］に指定した時点で自動的に「-1」に変わります。ここでは、「最良」を示す値を「1」、「最悪」を示す値を「-1」にしたいので、自動的に設定された値をそのまま使います。

6. ［**モデルの実行**］をクリックします。

図5.3「Potato Chip Combined.jmp」のMaxDiffレポート



MaxDiffモデル

MaxDiff 結果

限界効用	周辺確率	プロファイルID
1.2774	0.2895	Barbecue
0.7892	0.1777	Southern Barbecue
0.1439	0.0932	Truffle Fries
0.1046	0.0896	Biscuits and Gravy
0.0104	0.0815	Sour Cream and Onion
-0.137	0.0703	Dill Pickle
-0.218	0.0649	All Dressed
-0.341	0.0574	Ketchup
-0.465	0.0507	Reuben
-1.164	0.0252	Gyro

パラメータ推定値

尤度比検定

要因	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
プロファイルID	78.317	9	<.0001*

レポートは、「**プロファイルID**」が有意であること、つまり異なる風味のポテトチップスに対する好みが有意に異なることを示しています。限界効用が最も高いのは、「Barbecue」風味です。他の風味と一緒に提示されたときに「Barbecue」が選択される確率の推定値は、0.2895です。

7. 「MaxDiffモデル」の赤い三角ボタンをクリックし、［**全水準の比較レポート**］を選択します。

図5.4 全水準の比較レポート

全水準の比較レポート

差 (行 − 列) 差の標準誤差 Wald p値	Gyro	Sour Cream and Onion	Truffle Fries	Biscuits and Gravy	Reuben	Ketchup	All Dressed	Dill Pickle	Barbecue	Southern Barbecue
Gyro	0	-1.1747	-1.3082	-1.2688	-0.6992	-0.8229	-0.9467	-1.027	-2.4417	-1.9535
	0	0.33327	0.35477	0.34456	0.34758	0.32802	0.33571	0.35251	0.41499	0.38939
		0.00062	0.00035	0.00036	0.04668	0.01356	0.00569	0.00432	4.32e-8	2.01e-6
Sour Cream and Onion	1.17469	0	-0.1335	-0.0941	0.47548	0.35175	0.22804	0.14766	-1.267	-0.7788
	0.33327	0	0.31379	0.30716	0.30932	0.31915	0.28158	0.30707	0.37654	0.34896
	0.00062		0.67137	0.75983	0.12709	0.27279	0.41975	0.63156	0.00105	0.02764
Truffle Fries	1.30817	0.13348	0	0.03935	0.60896	0.48523	0.36152	0.28114	-1.1335	-0.6453
	0.35477	0.31379	0	0.31345	0.31817	0.31157	0.31136	0.30871	0.39213	0.34904
	0.00035	0.67137		0.90032	0.05821	0.12223	0.24809	0.36444	0.00462	0.06715
Biscuits and Gravy	1.26882	0.09413	-0.0393	0	0.56961	0.44588	0.32217	0.24179	-1.1729	-0.6847
	0.34456	0.30716	0.31345	0	0.30639	0.29752	0.29974	0.28249	0.36267	0.34038
	0.00036	0.75983	0.90032		0.06567	0.1368	0.28477	0.39388	0.00161	0.0467
Reuben	0.69921	-0.4755	-0.609	-0.5696	0	-0.1237	-0.2474	-0.3278	-1.7425	-1.2543
	0.34758	0.30932	0.31817	0.30639	0	0.30893	0.30148	0.30328	0.37117	0.35324
	0.04668	0.12709	0.05821	0.06567		0.68956	0.41355	0.28208	7.69e-6	0.00056
Ketchup	0.82294	-0.3517	-0.4852	-0.4459	0.12373	0	-0.1237	-0.2041	-1.6188	-1.1305
	0.32802	0.31915	0.31157	0.29752	0.30893	0	0.29728	0.27938	0.36911	0.34808
	0.01356	0.27279	0.12223	0.1368	0.68956		0.67812	0.46661	2.64e-5	0.00154
All Dressed	0.94665	-0.228	-0.3615	-0.3222	0.24744	0.12371	0	-0.0804	-1.4951	-1.0068
	0.33571	0.28158	0.31136	0.29974	0.30148	0.29728	0	0.29447	0.36838	0.3418
	0.00569	0.41975	0.24809	0.28477	0.41355	0.67812		0.78538	9.22e-5	0.00393
Dill Pickle	1.02703	-0.1477	-0.2811	-0.2418	0.32782	0.20409	0.08038	0	-1.4147	-0.9264
	0.35251	0.30707	0.30871	0.28249	0.30328	0.27938	0.29447	0	0.3539	0.35532
	0.00432	0.63156	0.36444	0.39388	0.28208	0.46661	0.78538		0.00012	0.01038
Barbecue	2.44172	1.26703	1.13355	1.1729	1.7425	1.61877	1.49507	1.41468	0	0.48824
	0.41499	0.37654	0.39213	0.36267	0.37117	0.36911	0.36838	0.3539	0	0.39164
	4.32e-8	0.00105	0.00462	0.00161	7.69e-6	2.64e-5	9.22e-5	0.00012		0.21514
Southern Barbecue	1.95348	0.77879	0.6453	0.68465	1.25426	1.13053	1.00683	0.92644	-0.4882	0
	0.38939	0.34896	0.34904	0.34038	0.35324	0.34808	0.3418	0.35532	0.39164	0
	2.01e-6	0.02764	0.06715	0.0467	0.00056	0.00154	0.00393	0.01038	0.21514	

この表は、行に記されている風味と列に記されている風味を比較し、効用の推定値の差を求めたものです。*p*値が小さいものは、濃い青色（マイナス）または赤色（プラス）で表示されています。たとえば、「Gyro」の行が全体に青色であることから、「Gyro」風味のポテトチップスの効用が他のすべての風味より有意に低いとわかります。「Barbecue」風味は、他のすべての味より効用が高いものの、「Southern Barbecue」風味との差は有意ではありません。

メモ：「全水準の比較レポート」の*p*値は、多重比較に対する調整がまったく行われていません。これらのp値の結果は、大まかな目安として使用してください。

複数のデータテーブル

ここでは、「Potato Chip Profiles.jmp」、「Potato Chip Responses.jmp」、「Potato Chip Subjects.jmp」の3つのデータテーブルを使用します。MaxDiff実験のデータは、常に1つのデータテーブルにまとめることもできますが、特に分析に含めたいプロファイル変数や被験者変数が他にもある場合は、複数のデータテーブルを使用する方が便利です。

起動ウィンドウへの入力

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Potato Chip Responses.jmp」データテーブルを開きます。

メモ: 起動ウィンドウへの入力をスキップしたいときは、**「風味による MaxDiff」**スクリプトの緑色の三角ボタンをクリックします。その後、「モデルを調べる」（135ページ）に進んでください。

---

2. **「プロファイルと被験者のデータを開く」**スクリプトの隣にある緑色の三角ボタンをクリックします。
   - プロファイルのデータテーブルである「Potato Chip Profiles.jmp」には、調査の対象となったポテトチップスの種類（**「風味」**）と原産国（**「原産国」**）の情報がリストされています。各選択肢に**「プロファイル ID」**がついています。
   - 被験者のデータテーブルである「Potato Chip Subjects.jmp」には、回答者がリストされています。回答者の**「市民権」**と**「性別」**も含まれています。
   - 応答のデータテーブルである「Potato Chip Responses.jmp」には、回答者がリストされています。そして、回答者ごとに、各プロファイルセットの**「調査 ID」**と**「選択肢集合 ID」**、各選択肢集合の**「プロファイル ID」**がまとめてあります。このデータテーブルには、応答値を含む**「最良のプロファイル」**と**「最悪のプロファイル」**の列もあります。
3. 3つのデータテーブルのいずれかで、［**分析**］>［**消費者調査**］>［**MaxDiff**］を選択します。
4. 「データ形式」メニューから［複数のデータテーブル, 相互参照］を選択します。

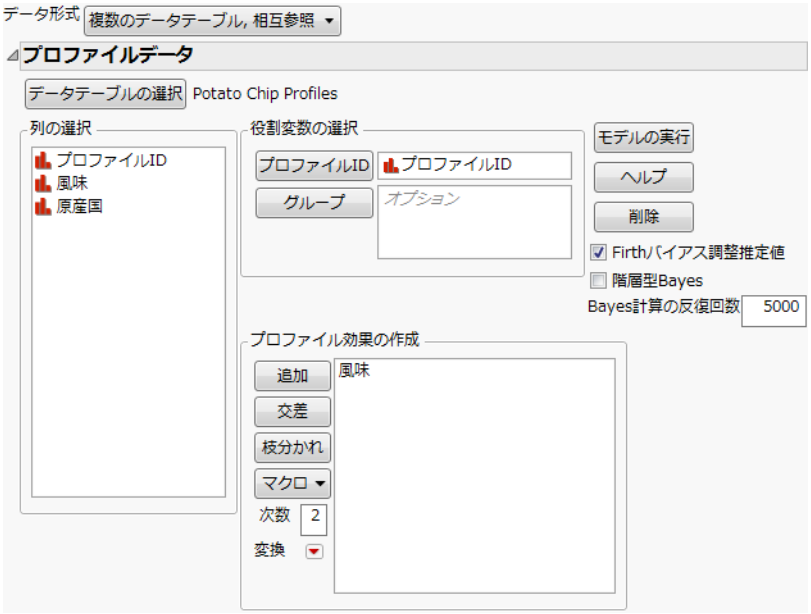
   各データソースに対応する3つのアウトラインがあります。
5. 「プロファイルデータ」アウトラインにある［**データテーブルの選択**］をクリックします。

   「プロファイルデータ」ウィンドウが開き、ここでプロファイルデータとして使うデータテーブルを指定します。
6. 「Potato Chip Profiles.jmp」を選択し、［**OK**］をクリックします。

   データテーブルの列が**「列の選択」**リストに表示されます。
7. 「列の選択」リストから**「プロファイル ID」**を選択し、**「役割変数の選択」**にある［**プロファイル ID**］をクリックします。
8. **「風味」**を選択し、**「プロファイル効果の作成」**にある［**追加**］をクリックします。

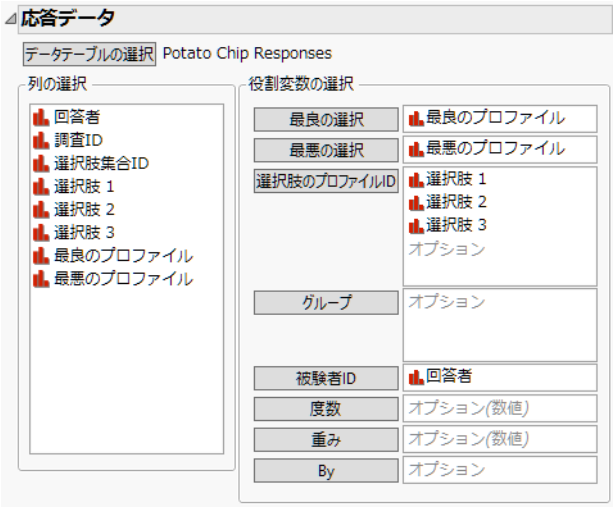
   なお、**「原産国」**も効果に含めてもよいですが、この例では省きます。

図5.5 入力後の「プロファイルデータ」アウトライン



9. 「応答データ」アウトラインを開きます。［**データテーブルの選択**］をクリックします。
10. 「Potato Chip Responses.jmp」を選択し、［**OK**］をクリックします。
11. 次の手順に従って、列に役割を割り当てます。入力後の起動ウィンドウは、図5.6のようになります。
    - 「**最良のプロファイル**」を選択し、［**最良の選択**］をクリックします。
    - 「**最悪のプロファイル**」を選択し、［**最悪の選択**］をクリックします。
    - 「**選択肢 1**」、「**選択肢 2**」、「**選択肢 3**」を選択し、［**選択肢のプロファイル ID**］をクリックします。
    - 「**回答者**」を選択し、［**被験者 ID**］をクリックします。

図 5.6 入力後の「応答データ」アウトライン



12.「被験者データ」アウトラインを開きます。［**データテーブルの選択**］をクリックします。

13.「Potato Chip Subjects.jmp」を選択し、［**OK**］をクリックします。

14.「**回答者**」を選択し、［**被験者 ID**］をクリックします。

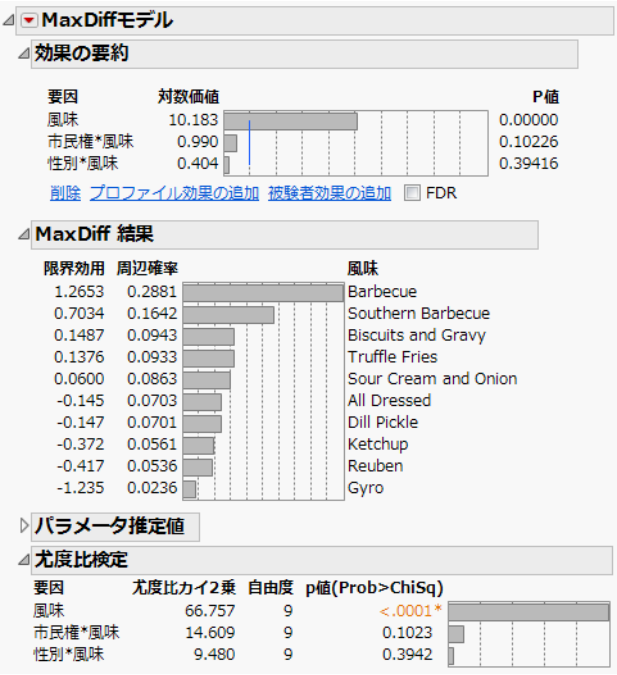
15.「**市民権**」と「**性別**」を選択し、「**モデル効果の作成**」にある［**追加**］をクリックします。

図 5.7 入力後の「被験者データ」アウトライン

モデルを調べる

1. ［**モデルの実行**］をクリックします。

図5.8「MaxDiff モデル」レポート



「効果の要約」レポートには、モデル項とその有意性を示す p 値が表示されます（図5.8）。**「風味」**はプロファイル効果です。**「市民権*風味」**と**「性別*風味」**は、被験者効果とプロファイル効果の交互作用です。

「尤度比検定」レポートは、**「風味」**が有意であることを示しています。

「MaxDiff」プラットフォームの起動

「MaxDiff」プラットフォームを起動するには、［**分析**］>［**消費者調査**］>［**MaxDiff**］を選択します。

「MaxDiff」プラットフォームで使用するデータは、1つのデータテーブルにまとめたものでも、2、3個のデータテーブルに分かれているものでもかまいません。「選択モデル」ウィンドウが開いたら、「データ形式」リストで、1つのデータテーブルを使うか、複数のデータテーブルを使うかを指定します。

1つのデータテーブル, 積み重ね

このデータ形式では、データを1つのデータテーブルにまとめてください。データテーブルの各行を、被験者に提示される各プロファイルに対応させてください。また、選択された最良と最悪のものがどれであるかを示す指示変数の列も含めてください。

データが1つのデータテーブルにまとめられている形式の例は、「1つのデータテーブル」（130ページ）を参照してください。詳細については、「［1つのデータテーブル, 積み重ね］の起動ウィンドウ」（137ページ）を参照してください。

複数のデータテーブル, 相互参照

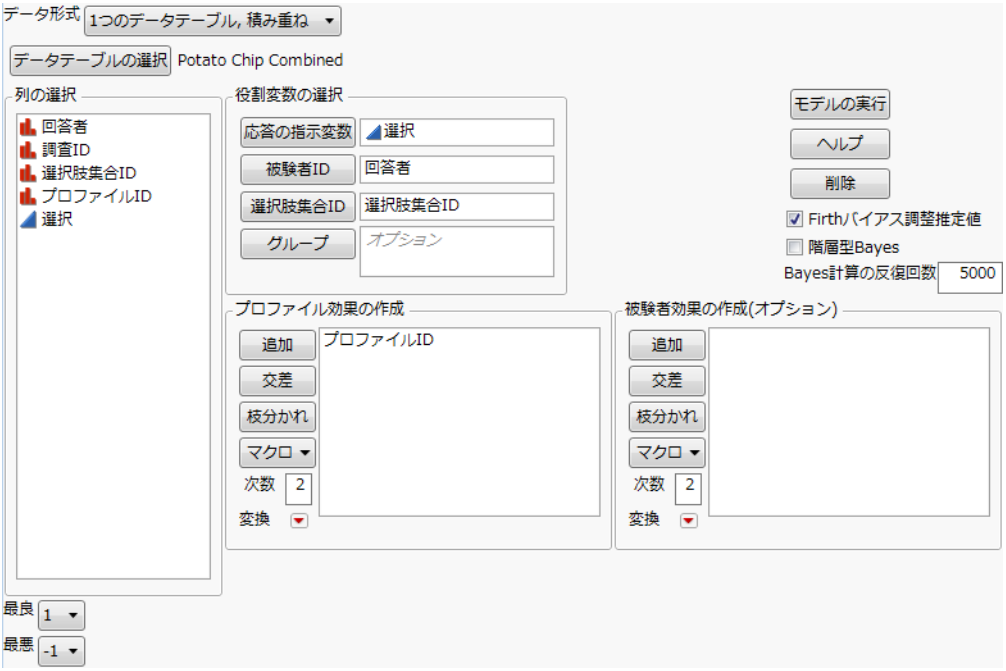
このデータ形式では、データを2つまたは3つのデータテーブルに分けてください。プロファイルデータと応答データは必須で、被験者データはオプションで使うことができます。「MaxDiff」起動ウィンドウは、各データテーブルに対応する3つのセクションで構成されています。各セクションは、必要に応じて開いたり閉じたりできます。

データが複数のデータテーブルに分かれている形式については、「複数のデータテーブル」(133ページ)を参照してください。詳細については、「[複数のデータテーブル, 相互参照] の起動ウィンドウ」(138ページ)を参照してください。

[1つのデータテーブル, 積み重ね] の起動ウィンドウ

図5.9は、1つのデータテーブル「Potato Chip Combined.jmp」だけを使って、列を指定した後の起動ウィンドウです。

図5.9 [1つのデータテーブル, 積み重ね] データ形式の起動ウィンドウ



データテーブルの選択 このボタンをクリックした後、データが保存されているデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、[その他...] を選択してください。

応答の指示変数 どれが「最良」や「最悪」であるかを示す列。「最良」と「最悪」に「1」、「-1」、「0」のうちのいずれか2つを指定すると、残りの1つが「最良」でも「最悪」でもないプロファイルとみなされます。ウィンドウ左下で「最良」と「最悪」を指定しない限り、「最良」には「1」、「最悪」には「-1」が使用されます。

被験者 ID 調査に参加した人のID。

選択肢集合 ID 被験者に提示されたプロファイルの集合を識別するためのID。これらの集合のなかから、回答者はいずれかのプロファイルを選択します。

グループ 「選択肢集合ID」と共に使用したときに、各選択肢集合を一意に示すことができる列。たとえば、**「選択肢集合 ID」**の値が1である選択肢集合が2つあったとします。データテーブルにおいて、一方の選択肢集合に対しては**「調査」**列の値がAで、他方の選択肢集合に対しては**「調査」**列がBであるなら、その**「調査」**列を［**グループ**］列に指定します。

プロファイル効果の作成 プロファイルの属性で構成された効果を追加します。

「プロファイル効果の作成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

被験者効果の作成(オプション) 被験者に関連する効果を追加します。

「被験者効果の作成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

Firth バイアス調整推定値 Firth法は、バイアス修正を伴う最尤推定であり、通常の最尤推定に比べ、推定や検定がより良い性質をもちます。さらに、ロジスティックモデルなどで生じる分離（separation）の問題も改善できます。ロジスティック回帰における分離問題については、Heinze and Schemper（2002）を参照してください。

JMP PRO **階層型 Bayes** Bayes流の方法で被験者ごとの係数値を求めます。詳細については、「Bayesパラメータ推定値」（147ページ）を参照してください。

JMP PRO **Bayes 計算の反復回数** （［階層型Bayes］を選択した場合のみ有効。）被験者ごとの係数値を計算する際に使う適応型Bayesアルゴリズムの反復回数。ここで指定した反復回数には、推定からは破棄されるバーンインの反復回数も含まれています。バーンインの反復回数は、起動ウィンドウで指定した「Bayes計算の反復回数」の半分です。

［複数のデータテーブル, 相互参照］の起動ウィンドウ

図5.10は、複数のデータテーブルを用いる例です。起動ウィンドウにおいて、「プロファイルデータ」に「Potato Chip Profile.jmp」を用いています。

図5.10 ［複数のデータテーブル, 相互参照］データ形式の起動ウィンドウ

［複数のデータテーブル, 相互参照］の場合、起動ウィンドウは3つのセクションに分かれています。

- 「プロファイルデータ」（139ページ）
- 「応答データ」（140ページ）
- 「被験者データ」（141ページ）

プロファイルデータ

プロファイルのデータテーブルには、それぞれの選択肢の属性を示すデータを保存してください。データテーブルの各列が1つの属性に対応するように、また、各行が1つのプロファイルに対応するように、データを作成してください。さらに、各プロファイルの**ID**を含んだ列を設けてください。

データテーブルの選択　このボタンをクリックした後、プロファイルデータが保存されているデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他...］を選択してください。

プロファイルID　属性の組み合わせ（プロファイル）を識別するためのID。［**プロファイルID**］によってプロファイルのデータテーブルにおける各行を一意に識別できない場合は、［**グループ**］変数も指定する必要があります。その場合、［**グループ**］列と［**プロファイルID**］列の組み合わせによって各行が一意に識別できるように、［**グループ**］列を追加してください。

グループ 「選択肢集合ID」列と共に使用したときに、各選択肢集合を一意に示すことができる列。たとえば、**「選択肢集合ID」**の値が1である選択肢集合が2つあったとします。データテーブルにおいて、一方の選択肢集合に対しては**「調査」**列の値がAで、他方の選択肢集合に対しては**「調査」**列がBであるなら、その**「調査」**列を［**グループ**］列に指定します。

プロファイル効果の作成 プロファイルの属性で構成された効果を追加します。

「プロファイル効果の作成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

Firthバイアス調整推定値 Firth法は、バイアス修正を伴う最尤推定であり、通常の最尤推定に比べ、推定や検定がより良い性質をもちます。さらに、ロジスティックモデルなどで生じる分離（separation）の問題も改善できます。ロジスティック回帰における分離問題については、Heinze and Schemper（2002）を参照してください。

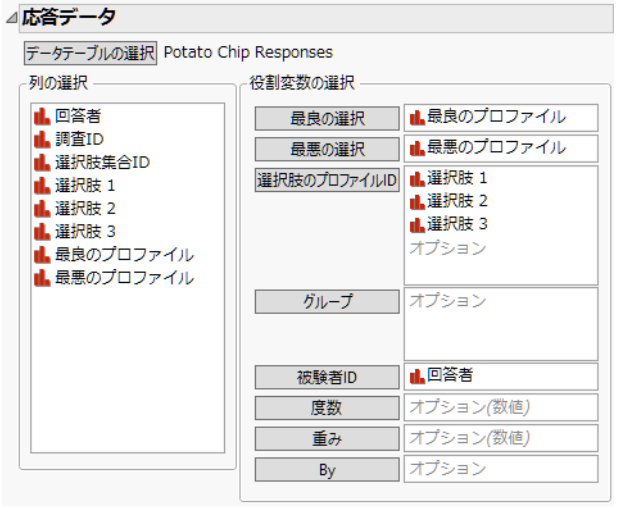
JMP PRO **階層型Bayes** Bayes流の方法で被験者ごとの係数値を求めます。詳細については、「Bayesパラメータ推定値」（147ページ）を参照してください。

JMP PRO **Bayes計算の反復回数** （［階層型Bayes］を選択した場合のみ有効。）被験者ごとの係数値を計算する際に使う適応型Bayesアルゴリズムの反復回数。ここで指定した反復回数には、推定からは破棄されるバーンインの反復回数も含まれています。バーンインの反復回数は、起動ウィンドウで指定した「Bayes計算の反復回数」の半分です。

応答データ

図5.11は、「Potato Chip Responses.jmp」を使って、列を指定した後の「応答データ」アウトラインです。

図5.11「応答データ」アウトライン



応答のデータテーブルは、調査した結果をまとめたものです。応答のデータテーブルには、各質問で提示された選択肢集合のIDと、回答者が選択した最良と最悪の選択肢を含めてください。このデータテーブルは、選択肢集合を示す列と、回答者が選択した選択肢を示す列により、プロファイルデータと関連付けられます。応答データとプロファイルデータにおける選択肢IDを対応させるのにグループ化が必要な場合には、その変数をグループ変数に指定してください。

データテーブルの選択　このボタンをクリックした後、応答のデータが保存されているデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他 ...］を選択してください。

最良の選択　被験者が「最良」であるとしたプロファイルのIDを含んだ列。

最悪の選択　被験者が「最悪」であるとしたプロファイルのIDを含んだ列。

選択肢のプロファイルID　各選択肢集合を構成する選択肢の**プロファイルID**を含んだ列。

グループ　**「選択されたプロファイルID」**列と共に使用したときに、各選択肢集合を一意に示すことができる列。

被験者ID　調査に参加した人を一意に示すID。

度数　度数を含んだ列。度数が n である行は、データにn回登場しているものとして計算に使用されます。度数が1未満である行や、欠測値である行は、分析に使用されません。

重み　データテーブルの各行に対する重みを含んだ列。0より大きい重みだけが分析に使われます。

By　この列の水準に従ってデータがグループ化され、それぞれ個別に分析されます。指定した列の水準ごとに個別に分析されます。分析結果は、水準ごとに個別のレポートにまとめられます。複数のBy変数を指定した場合は、By変数の水準のすべての組み合わせごとに分析が行われます。

被験者データ

図5.12は、「Potato Chip Subjects.jmp」を使って入力した「被験者データ」アウトラインです。

図 5.12「被験者データ」アウトライン

メモ: 被験者のデータテーブルは、被験者効果をモデルに含める場合のみ、必要になります。

このデータテーブルの列には、被験者 ID、および、被験者の属性（たとえば性別や年齢）を含めてください。このデータテーブルは、被験者数と同数の行で構成してください。また、応答のデータテーブルに、対応した被験者 ID の列を用意してください。

メモ: 被験者に関するデータは応答のデータテーブルに含めてもかまいませんが、その場合も、「被験者データ」アウトラインで被験者効果を指定する必要があります。

データテーブルの選択　このボタンをクリックした後、被験者のデータが保存されているデータテーブルを選択するか、まだデータテーブルが開かれていない場合にはファイルを開いてください。まだ開かれていないデータテーブルを開くには、［その他 ...］を選択してください。

被験者 ID　各被験者を一意に識別する ID。

モデル効果の構成　被験者に関する効果を追加します。

「モデル効果の構成」パネルについては、『基本的な回帰モデル』の「モデルの指定」章にある「モデル効果の構成」の節を参照してください。

「MaxDiff モデル」レポート

「MaxDiff モデル」ウィンドウには、起動ウィンドウで行った選択の内容に応じて以下のレポートのいくつかがデフォルトで表示されます。

- 「効果の要約」
- 「MaxDiff 結果」
- 「パラメータ推定値」
- 「Bayes パラメータ推定値」
- 「尤度比検定」

プラットフォームオプションについては、「「MaxDiff」プラットフォームのオプション」（148 ページ）を参照してください。

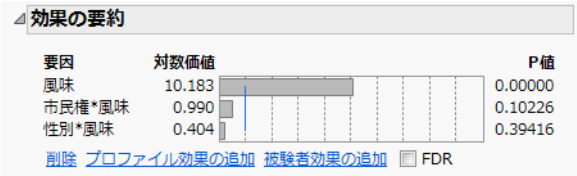
効果の要約

モデルに少なくとも 1 つの効果が含まれている場合は、「効果の要約」レポートが表示されます。モデルによって推定された効果がリストされ、それらの効果の対数価値（または FDR 対数価値）のプロットが表示されます。また、このレポートでは、モデルに効果を追加したり削除したりできます。このとき、モデルをあてはめた結果は、「効果の要約」レポートでの変更に合わせて自動的に更新されます。詳細については、『基本的な回帰モデル』の「標準最小2乗のレポートとオプション」章にある「効果の要約レポート」の節を参照してください。

起動ウィンドウで［Bayes 流の被験者効果］チェックボックスをオンにすると、「効果の要約」レポートは表示されません。Bayes 推定においては尤度比検定が行われないためです。

図 5.13 は、「Potato Chip Responses.jmp」の「**風味による MaxDiff**」スクリプトを実行したときに作成される「効果の要約」レポートです。

図 5.13 「効果の要約」レポート



「効果の要約」表の列

「効果の要約」表には次の列があります。

要因　モデル内の効果。*p* 値の小さい順に並べられます。

対数価値　各効果の対数価値（LogWorth）。$-\log_{10}(p\text{ 値})$ という式で計算されています。*p* 値を対数価値に変換すると、解釈がしやすくなります。対数価値が 2 以上あるものは、有意水準 0.01 で有意です（$-\log_{10}(0.01) = 2$）。

FDR 対数価値 各効果のFDR（False Discovery Rate; 偽発見率）の対数価値で、-$\log_{10}$(FDR 調整 p 値) という式で計算されています。対数価値やFDR 対数価値は、検定の有意性をグラフに表すのに適している統計量です。「対数価値」を「**FDR 対数価値**」に変更するには、［**FDR**］チェックボックスをオンにします。

棒グラフ 対数価値（またはFDR 対数価値）の棒グラフ。このグラフには整数値に縦の点線が、2の位置に青の参照線が引かれています。

p 値 各モデル効果の p 値。この p 値は、「尤度比検定」レポートにある有意性検定の p 値です。

FDR p 値 Benjamini-Hochberg 法で計算された、各効果のFDR 調整 p 値。FDR は、検定の多重性を考慮して、生の p 値を調整したものです。「**P 値**」列を「**FDR p 値**」に変更するには、［**FDR**］チェックボックスをオンにします。

FDR については、Benjamini and Hochberg（1995）を参照してください。偽発見率の詳細については、『予測モデルおよび発展的なモデル』の「応答のスクリーニング」章、または Westfall et al.（2011）を参照してください。

「効果の要約」表のオプション

要約表の下にあるオプションによって、効果を追加や削除できます。

削除 モデルから選択した効果を削除します。1つまたは複数の効果を削除するには、削除する効果に対応する列を選択し、［削除］ボタンをクリックします。

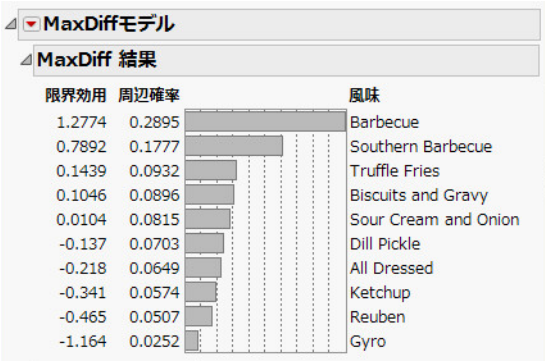
プロファイル効果の追加 パネルが開き、［1つのデータテーブル, 積み重ね］データ形式ならデータテーブルにあるすべての列、［複数のデータテーブル, 相互参照］データ形式ならプロファイルのデータテーブルにある列がリストされます。モデルに追加する列を選択し、リストの下にある［追加］ボタンをクリックすると、その効果がモデルに追加されます。［閉じる］をクリックすると、パネルが閉じます。

被験者効果の追加 パネルが開き、［1つのデータテーブル, 積み重ね］データ形式ならデータテーブルにあるすべての列、［複数のデータテーブル, 相互参照］データ形式なら被験者のデータテーブルにある列がリストされます。モデルに追加する列を選択し、リストの下にある［追加］ボタンをクリックすると、その効果がモデルに追加されます。［閉じる］をクリックすると、パネルが閉じます。

MaxDiff 結果

図5.14は、「Potato Chip Responses.jmp」の「**被験者効果のない MaxDiff**」スクリプトを実行したときに作成される「MaxDiff 結果」レポートです。

図5.14「MaxDiff 結果」レポート



起動ウィンドウで指定したプロファイル効果に対し、次のような値が表示されます。

限界効用　効用に与える影響を示した指標。限界効用が大きいほど、その機能は大きな価値があると考えられます。

周辺確率　他のすべての水準と一緒に提示されたときに、該当の水準を被験者が選択する確率の推定値。それぞれの効果において、周辺確率の和は1になります。

棒グラフ　周辺確率の棒グラフ。

効果の列　効果名の下に、水準がリストされます。水準ごとに、限界効用と周辺確率の推定値が表示されます。

パラメータ推定値

このレポートには、パラメータ推定値、モデル選択の規準、反復計算の結果といった詳細が表示されます。

図5.15は、「Potato Chip Responses.jmp」の「**風味による MaxDiff**」スクリプトを実行したときに作成される「パラメータ推定値」レポートです。

図 5.15「パラメータ推定値」レポート

パラメータ推定値

項	推定値	標準誤差
風味[All Dressed]	-0.15616234	0.2227718641
風味[Barbecue]	1.22210326	0.2978972284
風味[Biscuits and Gravy]	0.16398758	0.2249076601
風味[Dill Pickle]	-0.17356828	0.2151668801
風味[Gyro]	-1.11927509	0.2827016652
風味[Ketchup]	-0.47308393	0.2319900823
風味[Reuben]	-0.50927309	0.2294979936
風味[Sour Cream and Onion]	0.21115573	0.2450361368
風味[Southern Barbecue]	0.70149945	0.2695322771
市民権[Canadian]*風味[All Dressed]	-0.04368106	0.2239442590
市民権[Canadian]*風味[Barbecue]	-0.16180196	0.2978036763
市民権[Canadian]*風味[Biscuits and Gravy]	0.05734172	0.2233312021
市民権[Canadian]*風味[Dill Pickle]	-0.09824391	0.2188896544
市民権[Canadian]*風味[Gyro]	0.43257276	0.2907051874
市民権[Canadian]*風味[Ketchup]	-0.38035261	0.2349720398
市民権[Canadian]*風味[Reuben]	-0.34677939	0.2342643217
市民権[Canadian]*風味[Sour Cream and Onion]	0.56678250	0.2355493683
市民権[Canadian]*風味[Southern Barbecue]	-0.00712532	0.2720518372
性別[Female]*風味[All Dressed]	-0.26955535	0.2106170508
性別[Female]*風味[Barbecue]	0.39881410	0.2961368976
性別[Female]*風味[Biscuits and Gravy]	0.09184462	0.2245258406
性別[Female]*風味[Dill Pickle]	-0.11461679	0.2131955199
性別[Female]*風味[Gyro]	-0.40538827	0.2950786212
性別[Female]*風味[Ketchup]	-0.07317652	0.2135446256
性別[Female]*風味[Reuben]	0.20780054	0.2227994637
性別[Female]*風味[Sour Cream and Onion]	0.27353428	0.2337547848
性別[Female]*風味[Southern Barbecue]	0.07570279	0.2627935035

AICc	397.44423
BIC	456.27172
(-2)*対数尤度	327.00945
(-2)*Firth対数尤度	244.39836

勾配で収束しました

Firthバイアス調整推定値

項　モデル内の効果。

推定値　該当の項に対するパラメータの推定値。離散選択モデルでは、パラメータ推定値を「部分効用（part-worth）」と呼ぶことがあります。パラメータ推定値は、各項が効用に対してどれぐらい影響しているかを示す係数です。デフォルトでは、これらの推定値は、通常の最尤推定ではなく、Firth のバイアス調整最尤法で推定されます。

標準誤差　パラメータ推定値に対する標準誤差の推定値。

比較規準

適合度統計量として、AICc（修正された赤池情報量規準）、BIC（ベイズの情報量規準）、(– 2）× 対数尤度、(– 2）×Firth 対数尤度が表示されます。これらの統計量は、モデルを比較するのに用いられます。最初の 3 つの規準については、『基本的な回帰モデル』の付録「統計的詳細」に詳しい説明があります。

(– 2）×Firth 対数尤度は、起動ウィンドウで［Firth バイアス調整推定値］オプションをオンにした場合のみ求められます。このオプションは、デフォルトではオンになっています。

適合度統計量は、いずれも値が小さいほどよくあてはまっていることを示します。

JMP PRO Bayesパラメータ推定値

（この結果は、起動ウィンドウで［階層型Bayes］を選択した場合のみ表示されます。）「Bayesパラメータ推定値」レポートには、階層型Bayesモデルで求められた、モデル効果に関する結果が表示されます。階層型Bayesモデルでは、被験者による違いを尤度関数に組み込み、製品属性のパラメータにおける被験者の違いを推定します。Bayes推定のアルゴリズムには、適応型Metropolis-Hastings法（Train 2001）の一種が使われています。各モデル効果に対し、事後分布の平均と分散が計算されます。また、各モデル効果に関して、被験者ごとの係数値も出力できます。詳細については、「被験者推定値の保存」（150ページ）を参照してください。

このBayes流の推定では、被験者ごとの効果が確率変数（変量効果）とみなされます。製品の属性に関するパラメータが、被験者ごとに異なっており、特定の平均ベクトルと共分散行列を持つ多変量正規分布に従うと仮定されます。そして、ある被験者が複数の選択肢から選択することを表す尤度関数には、通常の最尤推定と同じように、多項ロジットモデルが仮定されます。（被験者ごとの効果が従う）多変量正規分布の平均ベクトルに対する事前分布には、平均ベクトルが0で、共分散行列が対角要素がすべて等しい対角行列である多変量正規分布が仮定されます。また、（被験者ごとの効果が従う）多変量正規分布の共分散行列が従う事前分布には、対角要素がすべて等しい対角行列を尺度行列とする逆Wishart分布が仮定されます。

反復計算では、最初にバーンイン期間が設けられており、最初のほうの反復の結果は破棄されます。デフォルトでは、そのバーンインの反復回数は、起動ウィンドウで指定した「Bayes計算の反復回数」の半分です。

図5.16 Bayesパラメータ推定値レポート

選択モデル: 選択

Bayesパラメータ推定値

項	事後 平均	事後 標準偏差	被験者 標準偏差
生地[Thick]	0.27529781	0.724891965	2.815958945
チーズ[Jack]	-7.01110292	4.697169974	2.593184268
トッピング[Pepperoni]	-1.06702410	0.941602303	2.260232580

全反復回数	5000
バーンイン反復回数	2500
回答者数	32
バーンイン後の平均対数尤度	-14.19342

項　モデルの項。

事後 平均　該当の項に対する事後分布を、被験者に関しても平均したもの。各反復において、各被験者に対する係数の値が生成されます。「事後 平均」は、その推定値を各反復で被験者に関して平均し、バーンイン反復後のそれらの全被験者平均をさらに平均したものです。

ヒント： 赤い三角ボタンのメニューから［Bayesチェーンの保存］を選択すると、反復ごとに生成された係数値を被験者に関して平均した値を確認できます。

事後 標準偏差　各反復において、生成された被験者ごとの係数値から平均が求められます。「事後 標準偏差」は、バーンイン反復後のそれらの全被験者平均から計算される標準偏差です。

被験者 標準偏差　まず、被験者ごとに、生成された係数値の平均をバーンイン後のものから計算します。「被験者 標準偏差」は、その標準偏差です。

ヒント: 赤い三角ボタンのメニューから［被験者推定値の保存］を選択すると、生成された係数値の、バーンイン後における、被験者ごとの平均を確認できます。

全反復回数　バーンイン反復を含め、実行された反復の合計数。

バーンイン反復回数　後で破棄されるバーンイン反復の回数。バーンインの反復回数は、起動ウィンドウで指定したBayes反復回数の半分に相当します。

回答者数　被験者の人数。

バーンイン後の平均対数尤度　バーンイン反復後の生成された値から求めた対数尤度関数の平均。

尤度比検定

図5.17は、「Potato Chip Responses.jmp」の**「風味によるMaxDiff」**スクリプトを実行したときに作成される「尤度比検定」レポートです。

図5.17 尤度比検定

尤度比検定

要因	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
風味	66.757	9	<.0001*
市民権*風味	14.609	9	0.1023
性別*風味	9.480	9	0.3942

要因　モデルに含まれる効果。

尤度比カイ2乗　該当する効果に対する検定の尤度比カイ2乗の値。

自由度　カイ2乗検定の自由度。

p値(Prob>ChiSq)　カイ2乗検定に対するp値。

棒グラフ　カイ2乗値の棒グラフ。

「MaxDiff」プラットフォームのオプション

JMP PRO **最尤推定値の表示**　（起動ウィンドウで［階層型Bayes］を選択した場合のみ表示。）モデル項の係数を、通常の最尤法で推定した結果（パラメータ推定値と標準誤差）。これらの値は、階層型Bayesアルゴリズムの開始値として使用されます。

複合因子検定 （起動ウィンドウで［階層型Bayes］を選択しなかった場合のみ表示。）モデル内の各因子を検定する目的で、その因子が関連しているすべての効果の尤度比検定を行います。複合因子検定について詳しくは、『基本的な回帰モデル』の「標準最小2乗のレポートとオプション」章を参照してください。

信頼区間 ［階層型Bayes］が選択されていないときには、「パラメータ推定値」レポートにパラメータの信頼区間を表示します。

［階層型Bayes］が選択されている場合、信頼区間は「Bayesパラメータ推定値］レポートに表示されます。このBayes流の信頼区間は、正規分布を仮定し、事後平均と事後標準偏差に基づいて求められます。

推定値の相関 ［階層型Bayes］が選択されていないときには、最尤法に基づいて、パラメータ推定値間の相関行列を表示します。

［階層型Bayes］が選択されている場合は、係数値の事後平均における相関行列を表示します。この相関行列は、バーンイン後の反復から求めたものです。バーンイン後、反復ごとに全被験者における事後平均が計算されます。これらの事後平均から求めた相関行列が、「推定値の相関」表に表示されます。

比較 設定された選択肢の比較を行います。比較したい因子と値を選択することができます。なお、［**すべて**］チェックボックスをチェックした場合、すべての水準が比較されます。［**すべて**］チェックボックスを複数の因子でチェックした場合、それらすべての因子の水準での比較ではなく、該当する1因子のすべての水準が、その他の因子を左側で選択されている水準に固定した上で、比較されます。詳細については、「「比較」レポート」（150ページ）を参照してください。

全水準の比較レポート 「全水準の比較レポート」には、2つのプロファイルを比較した結果が表示されます。モデルに被験者効果が含まれている場合は、被験者効果の水準を指定する必要があります。そのとき、被験者効果の指定された水準における比較だけが行われます。表の各セルには、効用の差、差の標準誤差、「差がない」ことを帰無仮説とするWald検定のp値が表示されます。

注意：「全水準の比較レポート」のp値は、多重比較に対する調整がまったく行われていません。これらのp値の結果は、大まかな目安として使用してください。

Wald検定のp値は、色付きで表示されます。行から列を引いた差が、正のときは青色、負のときは赤色で表示されます。青色または赤色の濃さは、有意性の高さを表します。濃い色ほど、差に対するp値が小さいことを示しています。

効用計算式の保存 複数のデータテーブルに対して分析を実行している場合は、効用の計算式列を含んだ、新しいデータテーブルを作成します。作成されたデータテーブルでは、各行がプロファイルと被験者の組み合わせに対応しており、プロファイルと被験者を示す列があります。1つのデータテーブルに対して分析を実行している場合は、「効用計算式」列が追加されます。

被験者ごとの勾配を保存 各パラメータの変化量を被験者ごとに含んだデータテーブルを作成します。パラメータの変化量は、ヘッセ行列で尺度化した各パラメータの傾きを、被験者ごとに平均して求められています。これは、該当する被験者と他の被験者との差異に対するラグランジュ乗数検定に対応します。作成されたデータテーブルを、付随しているスクリプトでクラスター分析すれば、マーケットセグメント（同じ選好構造をもつ消費者集団）を特定することができます。詳細については、「選択モデル」章の「セグメント化の例」（107ページ）を参照してください。

JMP PRO **被験者推定値の保存** （起動ウィンドウで［階層型Bayes］を選択した場合のみ表示。）各効果における被験者ごとのパラメータ推定値を含んだデータテーブルを新たに作成します。この値は、この被験者ごとに係数値を平均したものであり、「Bayesパラメータ推定値」レポートにある「事後 平均」です。「被験者受容率」は、Metropolis-Hastings法のステップにおいて、現在のステップとは異なる係数値に移動した割合（受容率）を示します。一般に、0.20ぐらいが良好な受容率とみなされています。詳細については、「Bayesパラメータ推定値」（147ページ）を参照してください。

JMP PRO **Bayesチェーンの保存** （起動ウィンドウで［階層型Bayes］を選択した場合のみ表示。）Bayes推定において使用されたチェーンの情報を含んだデータテーブルを、新たに作成します。詳細については、「Bayesチェーンの保存」（151ページ）を参照してください。

モデルダイアログ 現在の分析を実行した「MaxDiff」の起動ウィンドウを開きます。データテーブルやID、モデル効果を新しく指定できます。たとえば、現在のモデルから設定を少し変更して、別のモデルをあてはめることができます。

以下のオプションについて詳しくは、『JMPの使用法』の「JMPのレポート」章を参照してください。

やり直し 分析を繰り返したり、やり直したりするオプションを含みます。また、［自動再計算］オプションに対応しているプラットフォームにおいては、［自動再計算］オプションを選択すると、データテーブルに加えた変更が、該当するレポートに即座に反映されるようになります。

スクリプトの保存 レポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。

Byグループのスクリプトを保存 By変数の全水準に対するレポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。起動ウィンドウでBy変数を指定した場合のみ使用可能です。

「比較」レポート

「比較」レポートは、ペアごとの比較を指定したときに表示されます。レポートには、次のような列が含まれます。

因子 指定した被験者因子の水準。

比較対象1 比較対象となっている1つ目のプロファイルの因子と水準。

比較対象2 比較対象となっている2つ目のプロファイルの因子と水準。

効用1 「因子」列の被験者における、1つ目のプロファイルの効用。

効用2 「因子」列の被験者における、2つ目のプロファイルの効用。

確率1 「因子」列の被験者において、2つ目のプロファイルより1つ目のプロファイルを好む確率。

確率2 「因子」列の被験者において、1つ目のプロファイルより2つ目のプロファイルを好む確率。

オッズ比1 「確率1」を「確率2」で割った値。

オッズ比2 「確率2」を「確率1」で割った値。

比較の差 「効用1」から「効用2」を引いた値。

標準偏差 「比較の差」における標準誤差。

JMP PRO Bayesチェーンの保存

Bayesチェーンのデータテーブルを使うと、MCMCアルゴリズムで生成される係数値の安定性を調べることができます。作成されるデータテーブルには、(起動ウィンドウで指定した)「Bayes計算の反復回数」に1を足した数の行が含まれます。最初の行(反復1)は、開始値を示します。2行目以降には、生成された乱数が反復の順に保存されます。各列には、次のような情報が出力されます。

反復 反復の番号。最初の行は、開始値です。

対数尤度 その反復におけるモデルの対数尤度。各反復に対する対数尤度をプロットすれば、バーンイン期間(調整段階)での振る舞いを確認できます。

<モデル効果> 適応型シグマ 逆Wishart分布の尺度行列における対角要素の平方根。

<モデル効果> 受容率 MCMCアルゴリズムでの受容率。

<モデル効果> 平均 被験者ごとに生成された係数値、の平均。

<モデル効果> 分散 被験者ごとに生成された係数値、の分散。

第 6 章

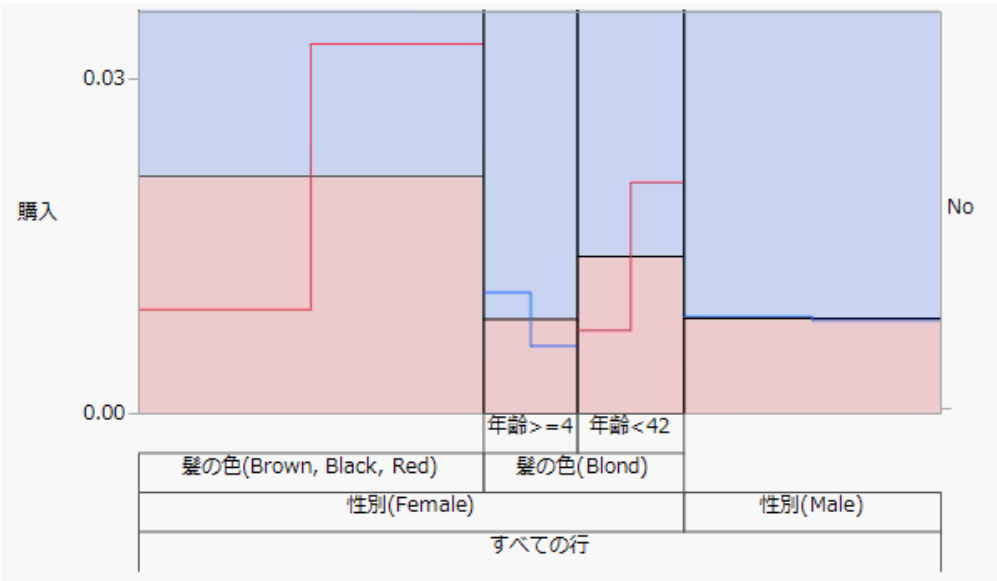
JMP PRO アップリフトモデル

施策が消費者行動に及ぼす効果をモデル化

JMP PRO この左のマークが付いている機能は、JMP Pro のみで提供されています。このプラットフォームは、JMP Pro のみで提供されています。

アップリフトモデルは、特定の施策に反応する可能性が高い個人の集団を特定するのに役立ちます。たとえば、マーケティングキャンペーンを最適化したり、個別化医療の方法を探し出したりすることなどが考えられます。アップリフトモデルは、増分モデル（incremental model）、真のリフトモデル（true lift model）、ネットモデル（net model）などとも呼ばれています。アップリフトモデルは、処置とその他の変数との間の交互作用を見つけていくという点で従来の手法とは異なります。アップリフトモデルは、施策（action）や処置（treatment；処理、治療）に対し、ポジティブに反応すると予測される集団をデータから探し出します。

図 6.1 ヘアケア製品の販売促進のためのアップリフトモデル例



項	分岐数	寄与	割合
性別	1	44.1799575	0.4691
年齢	1	25.474089	0.2705
髪の色	1	24.5280291	0.2604
米国地域	0	0	0.0000
居住地	0	0	0.0000

目次

「アップリフト」プラットフォームの概要 155
「アップリフト」プラットフォームの使用例 156
「アップリフト」プラットフォームの起動 157
「アップリフトモデル」レポート 158
アップリフトモデルのグラフ 159
アップリフトレポートのオプション 161

JMP PRO「アップリフト」プラットフォームの概要

「アップリフト」プラットフォームは、キャンペーンなどの処置（treatment；処理、治療）が消費者に及ぼす増分的な影響をモデル化します。アップリフトモデルは、処置に反応すると予測される消費者グループを特定するのに役立ちます。そのような消費者グループを特定できれば、リソースの割り当てや消費者に与える影響を最適化できるでしょう。（Radcliffe and Surry, 2011を参照）。

「アップリフト」プラットフォームでは、パーティションモデルをあてはめます。従来のパーティションモデルは予測が最良となるような分岐を見つけます。一方、アップリフトモデルは処置の差を最大化する分岐を見つけます。

パーティションによるアップリフトモデルでは、処置が強く影響するグループと、そうでないグループが存在していると仮定します。それらのグループを探し出すために、考えられるすべての二分岐で分けられたデータのそれぞれに、線形モデルやロジスティックモデルをあてはめます。応答変数が連続尺度の場合には、分岐、処置、および、分岐と処置の交互作用の線形関数としてモデル化されます。応答変数がカテゴリカルな場合には、同様に、分岐、処置、および、分岐と処置の交互作用のロジスティック関数としてモデル化されます。どちらの場合も、交互作用項は、分岐によって分かれた2つの消費者グループ間における、アップリフトの差に相当します。

「アップリフト」プラットフォームで分岐に使用される基準は、その交互作用項に対する検定のp値です。ただし、*p*値のみに基づいて説明変数を選択すると、水準が多い説明変数が有意になりやすいという多重性のバイアスが生じます。このため、JMPは、*p*値に水準数を考慮した調整を加えています（詳細については、JMP Webサイト「Monte Carlo Calibration of Distributions of Partition Statistics」を参照してください）。「アップリフト」プラットフォームの分岐は、交互作用効果に対する検定の調整済み*p*値を最大化することで決定されます。そして、それぞれの分岐に対しては、調整済み*p*値の対数価値（$-\log_{10}$）がレポートされます。

JMP PRO 「アップリフト」プラットフォームの使用例

「Hair Care Product.jmp」データテーブルは、毛染め製品に関する、仮想的なデータです。ある大手美容雑貨チェーンが、126,184名の男女のカード会員に対して、毛染め製品を購入するかどうかを追跡しました。まず、カード会員の中から約半数を無作為に抽出し、毛染め製品の広告を送付しました。そして、その後3か月にわたって、カード会員全員を対象に、製品を購入するかどうかを追跡しました。

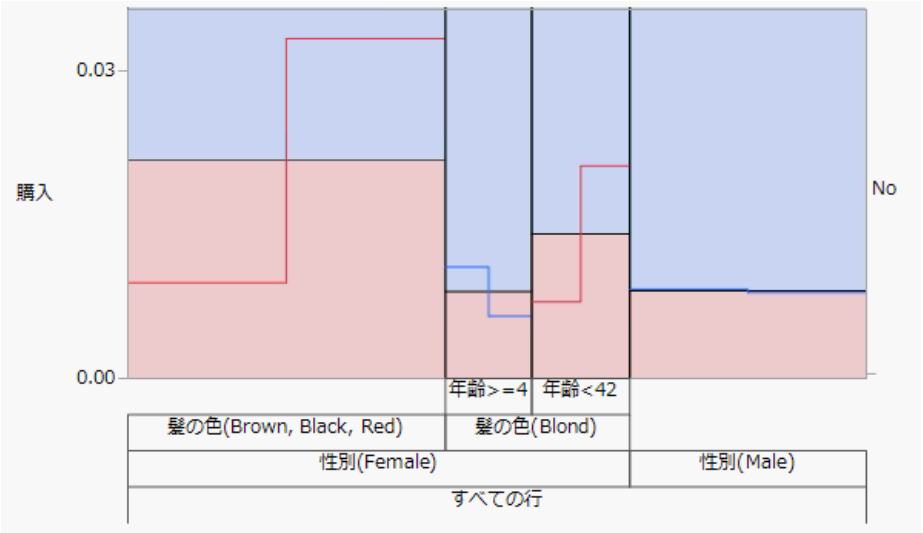
データテーブルの「**広告**」列は、広告が送付されたかどうかを示します。「**購入**」列は、製品を購入したかどうかを示します。各会員について、「**性別**」、「**年齢**」、「**髪の色**」、「**米国 地域**」、「**居住地**」（都市部かどうか）の情報が収集されています。また、「**検証**」列によって、カード会員の約33%が検証データとして使われるように設定されています。

JMPの「アップリフト」プラットフォームでは、応答変数がカテゴリカルである場合は、並び替えたときに順番が最初になっている水準が、関心のある水準として扱われます。「**購入**」列に「値の順序」列プロパティが設定されているのは、「Yes」を最初の水準とするためです。この列プロパティを設定しないと、「No」が最初の水準となり、分析において関心のある水準として扱われてしまいます。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Hair Care Product.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**アップリフト**］を選択します。
3. 「列の選択」リストから
   - 「**広告**」を選択し、「**処置**」をクリックします。
   - 「**購入**」を選択し、［**Y, 目的変数**］をクリックします。
   - 「**性別**」、「**年齢**」、「**髪の色**」、「**米国 地域**」、「**居住地**」を選択し、［**X, 説明変数**］をクリックします。
   - 「**検証**」列を選択し、［**検証**］ボタンをクリックします。
4. ［**OK**］をクリックします。
5. レポートのグラフの下にある［**実行**］をクリックします。

   検証セットに基づき、最適な分岐の数が3と判断されます。図6.2のようなグラフが作成されます。わかりやすいように、縦軸のスケールを変更しています。

図6.2 3回分岐した後のグラフ

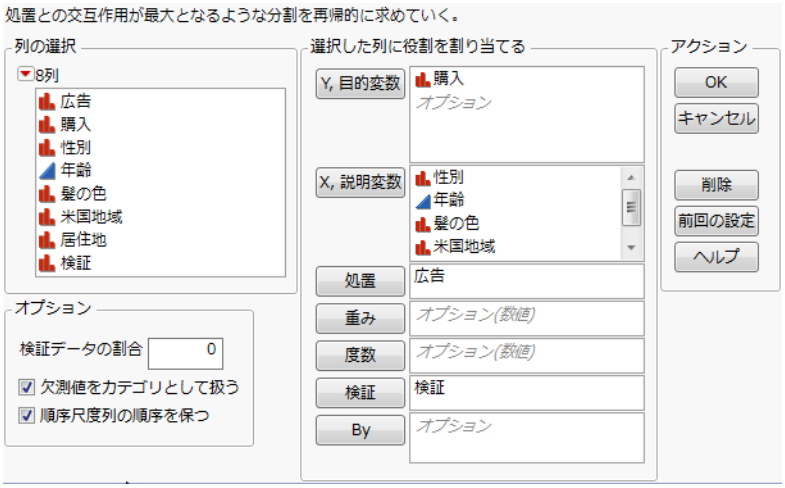


グラフから、広告によって購入にアップリフトが生じているのは、髪の色が黒（black）・赤（red）・茶色（brown）の女性、または低年齢層（**年齢** < 42）で髪の色がブロンドである女性のいずれかのグループだとわかります。高年齢層（**年齢** ≥ 42）でブロンドの女性の場合は、広告の効果がマイナスとなっています。

JMP PRO「アップリフト」プラットフォームの起動

「アップリフト」プラットフォームを起動するには、［**分析**］>［**消費者調査**］>［**アップリフト**］を選択します。図6.3は、「Hair Care Product.jmp」データテーブルの起動ウィンドウです。［Y, 目的変数］と［X, 説明変数］には、連続尺度の列とカテゴリカルな列のどちらでも指定できます。［処置］には通常、2水準のカテゴリカルな列を指定します。［処置］列に3つ以上の水準がある場合、最初の水準が「処置1」となり、残りのすべての水準は「処置2」としてまとめられます。

図6.3 アップリフトの起動ウィンドウ



検証データを用いるには、予め検証データを示すフラグの列を用意しておき、その列を［検証］列に指定します。もしくは、「検証データの割合」に割合を指定すると、その割合だけ無作為にデータが抽出されて、検証データとして使われます。「列の選択」リストから特定の列を選択しない状態で［検証］ボタンをクリックすると、データテーブルに検証列を追加できます。「検証列の作成」ユーティリティについての詳細は、『予測モデルおよび発展的なモデル』の「モデル化ユーティリティ」章を参照してください。

次のようなオプションも用意されています。

欠測値をカテゴリとして扱う　説明変数がカテゴリカルな場合、このチェックボックスをオンにすると、欠測値を個別のカテゴリとして処理できます。説明変数が連続尺度の場合は、欠測値が同一の数値をもつものとして扱われます。

順序尺度列の順序を保つ　このオプションを選択すると、順序が維持されるような分岐だけが考慮の対象となります。

JMP PRO 「アップリフトモデル」レポート

- 「アップリフトモデルのグラフ」
- 「アップリフトレポートのオプション」

JMP PRO アップリフトモデルのグラフ

アップリフトモデルのグラフにおける縦軸は、応答を示しています。そして、横軸は、分岐によって構成されたノードを表します。ノードごとに描かれている黒色の横線は、そのノードの応答の平均を示しています。各ノードには、処置による副分岐を表す赤色または青色の線も描かれています。これらの赤もしくは青の線は、各ノード内における処置による応答の違いを示しています。赤もしくは青の線は、各ノード内で処置で分岐したときに構成される2つのグループの平均応答を示しています。「処置」列におけるデータ値の並び順によって、どちらを「処置1」とするかが決まり、描かれる線が赤色になるか、青色になるかが決まります。分岐を進めていくにしたがって、分岐の条件式（条件式は横軸の下に表示されています）やグラフが更新されていきます。縦の線は、分岐の分かれ目を表しています。

グラフの下に、[**分岐**]、[**剪定**]、[**実行**] というボタンがあります。なお、このうち、[実行] ボタンは、検証データを使用している場合にのみ表示されます。また、「処置」列の名前と、その「処置」列の水準名も表示されます。「処置」列に3つ以上の水準がある場合、最初の水準以外はすべて「処置2」という1つの水準にまとめられ、2水準の応答として処理されます。

「処置」列情報の右側には、予測に関連する要約統計量がレポートされます（ただし、アップリフトモデルにおいては、最良の予測を行うことが分析目的ではない点に注意してください）。レポートは、分岐が実行されるたびに更新されます。検証データを使用している場合、学習データと検証データの両方に対する要約統計量が表示されます。

R2乗 ツリーに関連する回帰モデルのR2乗。この回帰モデルには、処置列との交互作用も含まれています。R2乗が1に近いほど、モデルのデータへのあてはまりが良いことを示します。

メモ: R2乗の値が0に近い場合、モデルに含まれていない変数の中に、説明されていない変動を説明するものがあると考えられます。ただし、データのばらつきがもともと大きなものであるときには、有効なアップリフトモデルであってもR2乗値が小さくなる場合もあります。

RMSE ツリーの回帰モデルにおけるRMSE（Root Mean Square Error; 平均平方誤差の平方根、誤差の標準偏差）。RMSEは、応答変数が連続尺度の場合にのみ算出されます。詳細については、『基本的な回帰モデル』を参照してください。

N オブザベーションの数（標本サイズ）。

分岐数 分岐が実行された回数。

AICc 回帰モデルを使って計算された、修正済みの赤池の情報量規準（AICc）。AICcは、応答変数が連続尺度の場合にのみ算出されます。詳細については、『基本的な回帰モデル』の付録「統計的詳細」を参照してください。

JMP PRO アップリフトのディシジョンツリー

アップリフトモデルで使用された分岐は、ツリーとして描画されます。「Hair Care Product.jmp」データテーブルを使った例は、図6.4を参照してください。描画されたツリーの各ノードには、次の情報が含まれます。

処置 「処置」列の列名と、その「処置」列に含まれている2つの水準が表示されます。

割合　応答変数が2水準のカテゴリカルの場合にのみ表示されます。各ノード内で、反応した消費者の割合が、処置の水準ごとに算出されています。

平均　応答変数が連続尺度の場合にのみ表示されます。各ノード内で処置の水準ごとに、応答の平均が表示されます。

度数　各ノード内で処置の水準ごとに、標本サイズ（たとえば、調べた消費者の人数）が表示されます。

t値　処置の水準間における差を検定するときに使われる *t* 値。なお、応答変数がカテゴリカルな場合は、この統計量は、連続尺度（0と1）に変換された後に算出されます。

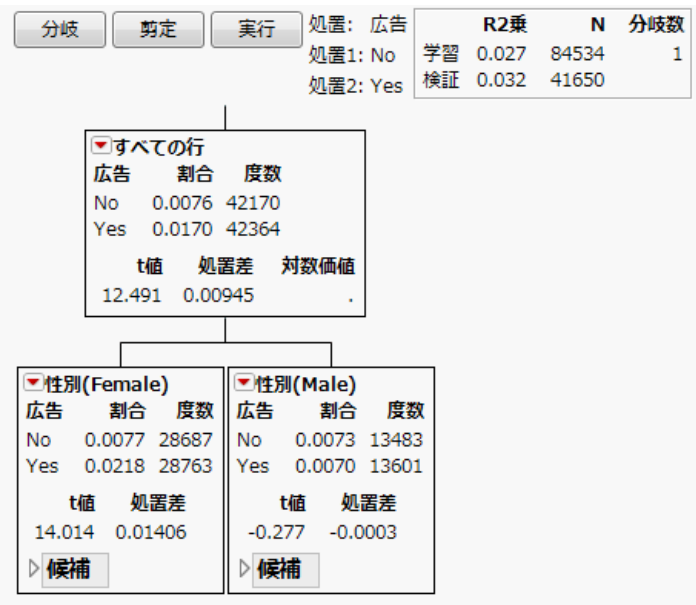
処置差　処置水準間の応答平均の差。これがアップリフトと呼ばれるものです。JMPの「アップリフト」プラットフォームでは、次のように仮定して計算されます。

– 「処置」列の値において、最初の水準が「処置群」を表し、それ以外の水準は「対照群」を表す。

– 応答変数の値は、値が大きいほど、ポジティブであることを示している。

対数価値　各ノードにおける、後続の分岐がもつ対数価値。

図6.4 最初の分岐のノード



「候補」レポート

各ノードには「候補」レポートも含まれます。レポートには、次のような情報が表示されます。

項　モデルの項。

対数価値　各項でのすべての分岐のうちで、分岐による統計量が最大となったときの対数価値。「対数価値」は、$-\log_{10}$（p）で計算される（ここで、pは、調整済み*p*値）

F値　応答変数が連続尺度の場合は、線形回帰モデルの交互作用項に対するF値。応答変数がカテゴリカルの場合は、名義ロジスティックモデルの交互作用項に対するカイ2乗値。いずれの回帰モデルも、処置、二分岐、および、それらの交互作用という3つの効果から構成されます。

ガンマ　応答変数が連続尺度の場合は、線形回帰モデルにおける交互作用の推定値。応答変数がカテゴリカルの場合は、名義ロジスティックモデルをFirthバイアス調整法で推定したときの、交互作用の推定値。

分岐点　該当の説明変数が連続尺度の場合は、分岐点が表示されます。該当の説明変数がカテゴリカルの場合は、最初（左）のノードに属するカテゴリが表示されます。

JMP PRO アップリフトレポートのオプション

「アップリフト」レポートの赤い三角ボタンのオプションについては、ここで紹介するオプションを除き、すべて「パーティション」プラットフォームのマニュアルで説明しています。ここで紹介しているもの以外のオプションは、『予測モデルおよび発展的なモデル』の「パーティションモデル」章を参照してください。

JMP PRO 分岐の最小サイズ

分岐を行う最小の標本サイズを、度数または標本全体に占める割合として指定するウィンドウが開きます。度数（標本サイズ）として指定する場合は、1以上の値を入力します。標本サイズ全体に対する割合として指定する場合は、1未満の値を入力します。アップリフトでは、Max（25, Floor（n/2000））がデフォルト値として設定されます。ここで、nはデータ全体の標本サイズです。Floor(x)はx以下の一番大きな整数、Max(a, b)はaとbのうちの大きい方の値です。

JMP PRO 列の寄与

この表とプロットは、作成されたツリーに対する各列の寄与を表わしています。列の寄与は、分岐のF値の合計として算出されます。F値は、回帰モデルにおける、処置と分岐の交互作用項に対する検定統計量です。

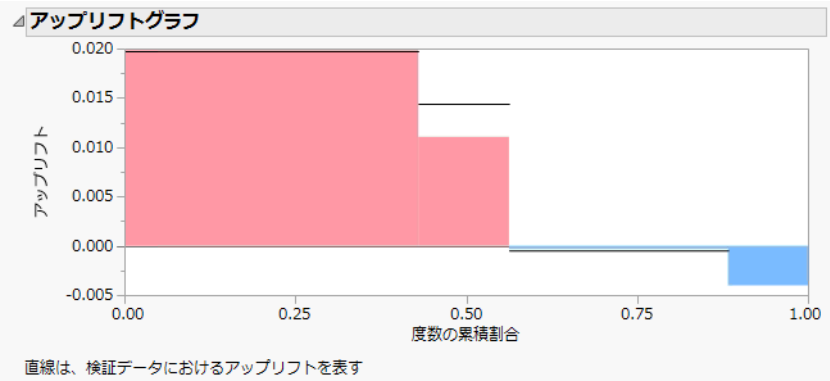
JMP PRO アップリフトグラフ

学習データにおいて、各ノードのアップリフト値を求めます。アップリフト値は、応答変数がカテゴリカルな場合は、商品を購入する確率の処置間における差です。応答変数が連続尺度の場合は、処置間における平均差です。その予測値が、最終的なツリーに基づいて、各ノードに対して算出されます。ノードごとのアップリフト値は、降順に並べて表示されます。アップリフトグラフでは、縦軸はアップリフト値です。横軸は、データの累積割合です。

図6.5は、「Hair Care Product.jmp」サンプルデータで、3回分岐した後のアップリフトグラフの例です。被験者のうち、2つのグループ（男性と、**年齢** ≥ 42のブロンドの女性）で、広告の効果がマイナスとなっています。

アップリフトグラフにある横線は、検証データにおけるアップリフト値です。学習データから推定されたツリーが、検証データに対して評価され、アップリフト値の推定値が算出されます。

図6.5 アップリフトグラフ



JMP PRO 列の保存

差の保存　各ノードに対して計算された、処置の水準間における平均差の推定値を保存します。この値が、アップリフト値の推定値です。

差の計算式を保存　差（アップリフト）の計算式を保存します。

差の計算式を発行　差の計算式を作成し、「計算式デポ」レポートに計算式列スクリプトとして保存します。「計算式デポ」レポートが開いていない場合は、このオプションを選択するとウィンドウが開かれます。『予測モデルおよび発展的なモデル』の「計算式デポ」章を参照してください。

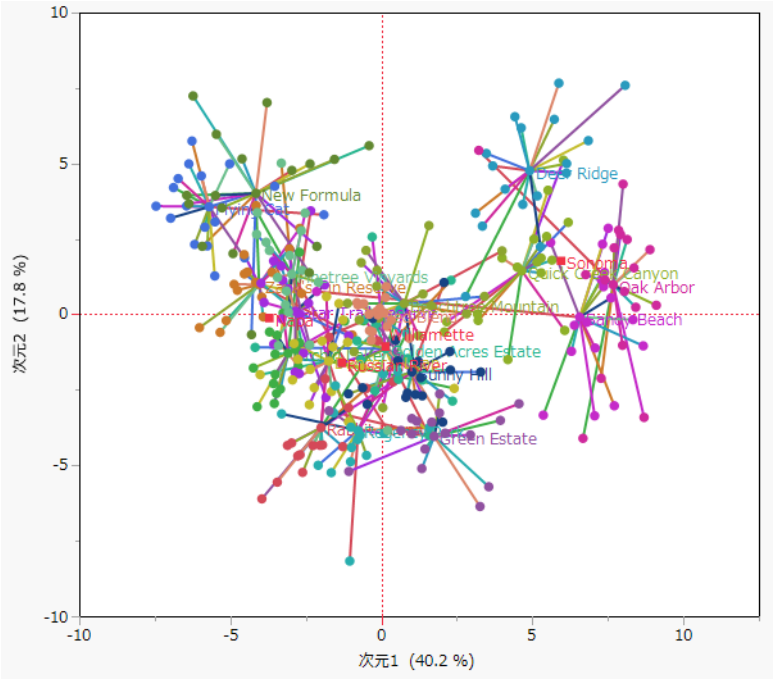
第 7 章

多重因子分析

複数の回答者における回答の類似性を視覚化する

多重因子分析（Multiple factor analysis; MFA）は、主成分分析（Principal Components Analysis: PCA）を発展させた手法です。多重因子分析は、主成分分析と同じように、多変量データを固有値分解によって互いに直交する主成分に変換します。いくつかの同じ商品や食品に対して異なった評価者から得たデータを分析する場面で、多重因子分析は特に役立ちます。多重因子分析で得られた主成分は、各商品が互いにどれだけ類似しているか、または異なっているかを判断するのに役立ちます。多重因子分析では、複数のデータに対する主成分分析を個別に行った後、全体に対する主成分分析を行います。

図 7.1 多重因子分析のコンセンサスマップ



目次

「多重因子分析」プラットフォームの概要 165
多重因子分析の例 166
「多重因子分析」プラットフォームの起動 167
データ形式 169
「多重因子分析」レポート 170
要約プロット 170
コンセンサスマップ 171
「多重因子分析」プラットフォームのオプション 171
「多重因子分析」プラットフォームの統計的詳細 174

「多重因子分析」プラットフォームの概要

多重因子分析（Multiple factor analysis; MFA）は、主成分分析（Principal Components Analysis: PCA）を発展させた手法です。主成分分析と異なるのは、多重因子分析は複数のデータセットを分析できるという点です。多重因子分析で分析される1つ1つのデータセットは、部分テーブル（sub-table）や部分行列（sub-matrix）と呼ばれています。多重因子分析が対象とする部分テーブルでは、含まれる行数は同じでなければいけません。各行は、商品（たとえば、評価をする製品や、官能検査をする食品）に対応しています。JMPの「多重因子分析」プラットフォームを実行するには、すべての部分テーブルが1つのJMPデータテーブルにまとめて保存されていなければいけません。そして、1つ1つの部分テーブルは、列のグループとして定義します。JMPの「多重因子分析」プラットフォームでは、この列のグループを「ブロック」と呼んでいます。ブロックについて、以下の点に注意してください。

- 各ブロックに含まれる列の数は同じでなくてもかまいません。たとえば、味や匂いなどを調べる官能検査においては、1つのブロックに含まれる各列は1名の検査者（パネリスト）が行った各評価に対応します。官能検査では、検査者が調べる評価項目が異なる場合もあります。
- 多重因子分析のブロックに含まれる評価項目やデータは、異なるものであってもかまいません。多重因子分析では、データの尺度（スケール）をブロックごとに調整した後、その尺度化された全データに対して主成分分析を適用します。

多重因子分析の主な目的は、商品や食品（データテーブルの行）における類似性をグラフで描き出すことです。もう1つの目的は、回答が大きく異なる回答者（反応が異なる検査者）を特定することです。なぜなら、ひとりだけ回答が大きく異なっている場合、そのデータは全体の結果にも影響するからです。多重因子分析では、結果を解釈しすくするために、追加変数（supplementary variable）を分析に追加することもできます。

多重因子分析では、測定や検査する商品属性が異なっていたり、実験器具・回答者・状況が異なっていたりしてもかまいません。多重因子分析は、特に官能検査データの分析で使用されることが多く、検査者間における違いを視覚化します。従来の官能検査では、複数の検査者が一貫した検査ができるようにするために、数時間に及ぶトレーニングが事前に必要でしょう。たとえば、ジュースの味覚に関する官能検査で、「フルーティ」・「甘い」・「さわやか」といった項目を各検査者は評価するとしましょう。従来の官能分析では、すべての検査者が同じような味の評価を行えるように事前にトレーニングしたり、検査者間で同じような評価になっているかどうか確認したりする必要があります。多重因子分析では、こうしたトレーニングを積んでいない検査者の評価データであっても分析を行えます。

多重因子分析が対象とするデータの典型例は、複数の同じ商品・食品・製品を、異なる検査者がいくつかの項目について評価したものです。この場合、1人の検査者から得られたデータを、1つのブロックとします。この官能検査データでは、データ行が測定対象の商品・食品・製品で、データ列が感覚によって判定された項目です。そして、1名の検査者が1つのブロックとして扱われます。なお、「多重因子分析」プラットフォームでは、欠測値は、列の平均値に置換されます。

多重因子分析の詳細については、Abdi et al.（2013）を参照してください。

多重因子分析の例

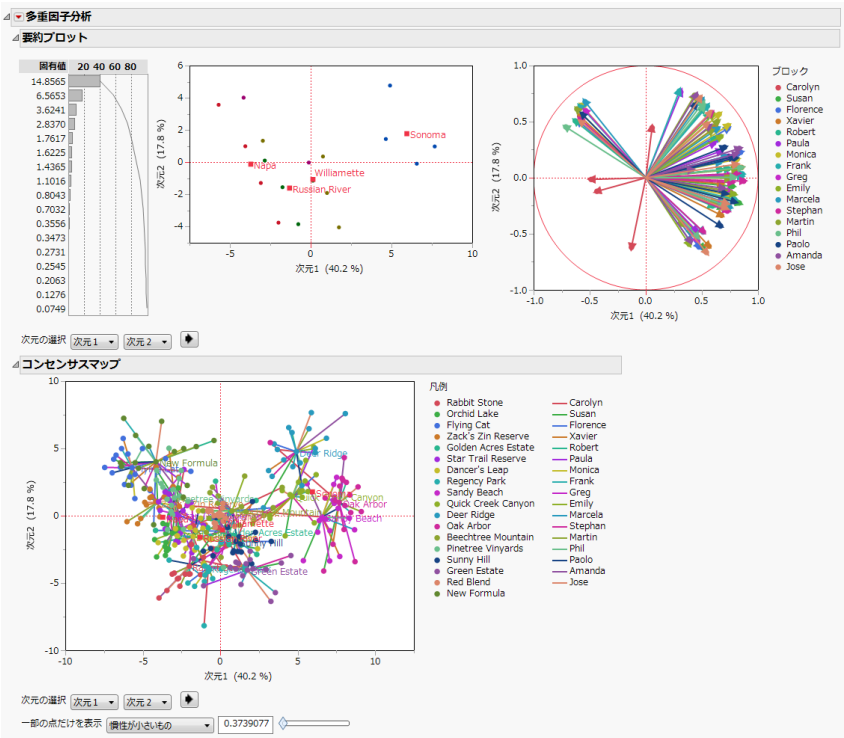
この例では、ワインの味や匂いを調べる官能検査のデータを使用します。回答者は、16種類のワインを対象に、複数の検査項目について、1（その特徴がない）から10（その特徴が顕著）で評価します。これらのワインの類似性や相違性を調べてみましょう。

1. ［**ヘルプ**］>［**サンプルデータライブラリ**］を選択し、「Wine Sensory Data.jmp」を開きます。
2. ［**分析**］>［**消費者調査**］>［**多重因子分析**］を選択します。
3. 「**ワインヤード**」を選択し、［**評価対象 ID**］をクリックします。
4. 「**地域**」を選択し、［**Z, 追加**］をクリックします。
5. 列グループをすべて選択し（「**Carolyn**」のグループから「**Jose**」のグループまで）、［**ブロックの追加**］をクリックします。

   メモ： このデータテーブルでは、回答者ごとに列がグループ化され、ブロックとしてまとめられています。列がグループ化されていない場合は、同じブロックにする列を選択して［**ブロックの追加**］をクリックするという操作を、各ブロックについて繰り返してください。

6. ［**モデルの実行**］をクリックします。

図7.2 最初の「多重因子分析」レポート



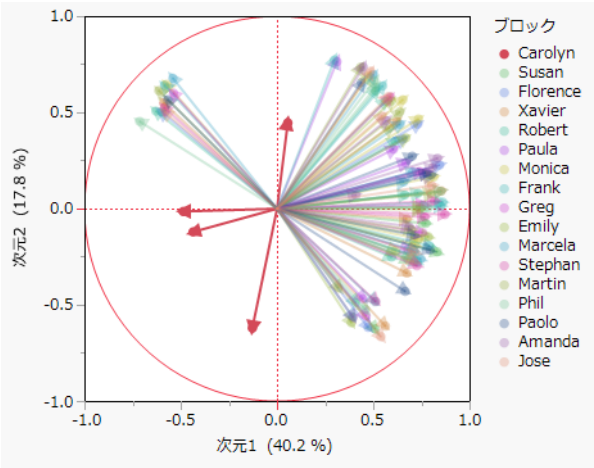
要約プロットで以下の点を確認してください。

– 最初の2次元上の因子スコアプロットを見ると、ワインの地域ごとにまとまっている傾向があります。

– 負荷量プロットの第3象限に位置している矢印は、検査者のうちキャロリン（Carolyn）に対応しています。これは、キャロリンと他の検査者には違いがあることを示唆しています。

7. 負荷量プロットの横にある凡例において「**Carolyn**」を選択してください。そうすると、キャロリンの結果が強調表示されます。

図7.3 負荷量プロットでキャロリンの結果を強調表示

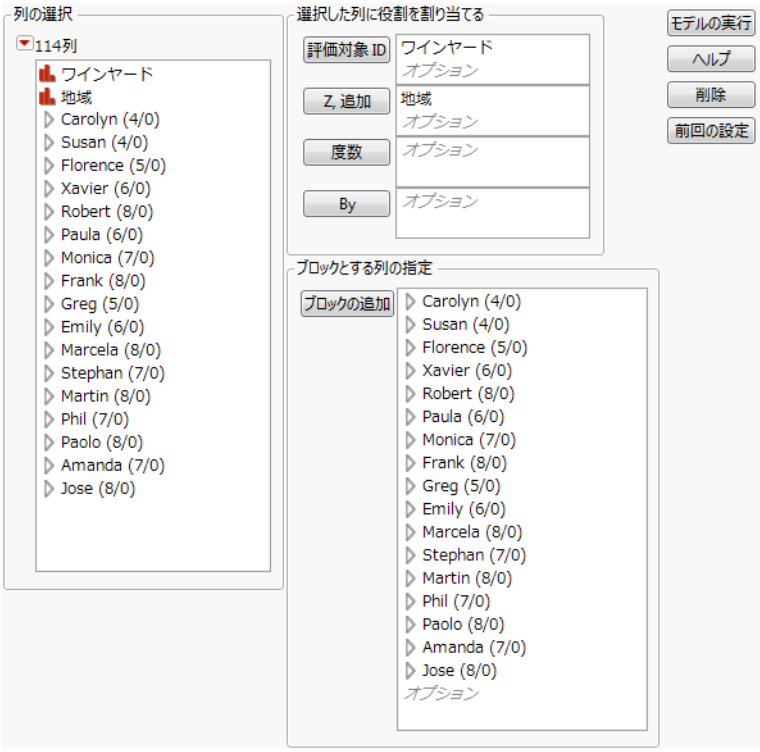


キャロリンの結果は、他の回答者と異なります。キャロリンの結果を除いて、分析をやり直すことを検討してもよいでしょう。キャロリンを除いた分析の結果については、図7.6（170ページ）を参照してください。

「多重因子分析」プラットフォームの起動

「多重因子分析」プラットフォームを起動するには、［**分析**］>［**消費者調査**］>［**多重因子分析**］を選択します。

図7.4「多重因子分析」起動ウィンドウ



評価対象 ID　評価をした対象物（製品・商品・食品）の列を指定します。

Z, 追加　追加変数（supplementary variable）として使用する列を指定します。追加変数は、分析者が関心をもっていて関係を調べたい変数ですが、空間を構成するときの計算には含めたくない変数のことです。

度数　この役割に割り当てられた列に含まれる数値は、分析において各行の度数として使用されます。

ブロックの追加　以下のいずれかの操作を行います。

– 複数の列を1つのブロックとして追加します。

– 列グループ（グループ化されている複数の列）をブロックとして追加します。

– 選択されているブロックに、個別の列を追加します。

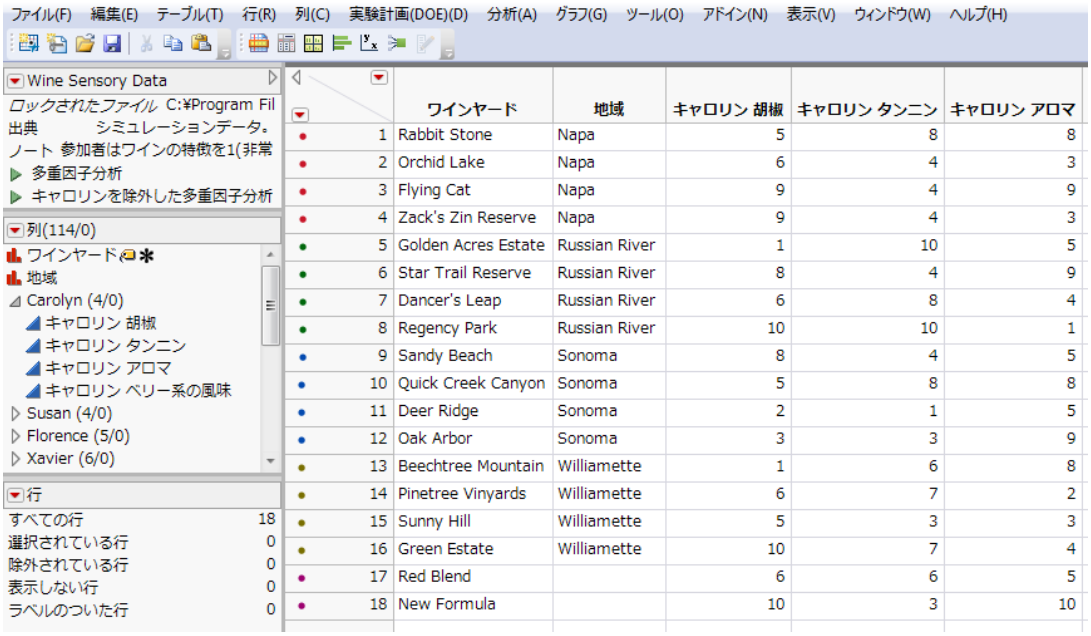
ヒント：プラットフォームを実行する前に複数の列をグループ化してブロックにまとめておいて、列グループのそれぞれを1つのブロックとして定義して、そのグループをまとめて追加するという操作を一度に行えます。そうでない場合は、同じグループにしたい列を選択して［ブロックの追加］をクリックするという操作を、ブロックごとに行ってください。なお、ブロック名は、ダブルクリックして変更できます。

データ形式

「多重因子分析」プラットフォームでは、部分テーブル（sub-tables）から構成される、1つのデータテーブルを使用します。個々の部分テーブルは「ブロック」と呼ばれています。各ブロックに含まれる列数は異なっていてもかまいません。また、各ブロックに含まれるデータの種類や内容は、互いに異なっていてもかまいません。ブロックは、JMPデータテーブルの列グループ（グループ化された列）として定義する必要はありませんが、JMPデータテーブルにて予めブロックごとに列をグループ化しておいたほうが、このプラットフォームを起動するときに便利です。

データテーブルの行は、評価する対象物（製品・商品・食品）を表しています。データテーブルでは、1つの対象物に対するデータが、1行に保存されているようにしてください。たとえば、図7.5は、異なるワインヤードで生産された16種類のワインの味や匂いを評価したデータテーブルです。このデータテーブルでは、評価者別に列がグループ化されていいます。それらの列グループは、「列」パネルで確認できます。

図7.5 多重因子分析用のデータテーブル（一部）



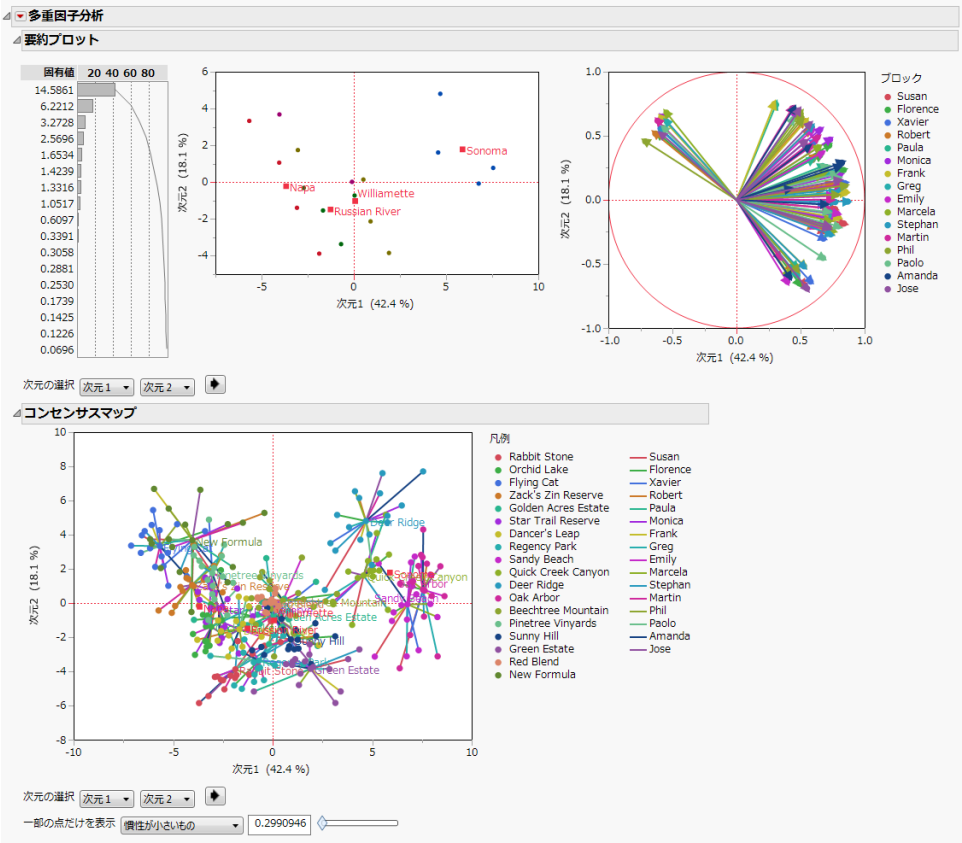
	ワインヤード	地域	キャロリン 胡椒	キャロリン タンニン	キャロリン アロマ
1	Rabbit Stone	Napa	5	8	8
2	Orchid Lake	Napa	6	4	3
3	Flying Cat	Napa	9	4	9
4	Zack's Zin Reserve	Napa	9	4	3
5	Golden Acres Estate	Russian River	1	10	5
6	Star Trail Reserve	Russian River	8	4	9
7	Dancer's Leap	Russian River	6	8	4
8	Regency Park	Russian River	10	10	1
9	Sandy Beach	Sonoma	8	4	5
10	Quick Creek Canyon	Sonoma	5	8	8
11	Deer Ridge	Sonoma	2	1	5
12	Oak Arbor	Sonoma	3	3	9
13	Beechtree Mountain	Williamette	1	6	8
14	Pinetree Vinyards	Williamette	6	7	2
15	Sunny Hill	Williamette	5	3	3
16	Green Estate	Williamette	10	7	4
17	Red Blend		6	6	5
18	New Formula		10	3	10

メモ：なお、「多重因子分析」プラットフォームでは、欠測値は列の平均値に置換されます。ただし、欠測値などによって、列の分散がゼロになる場合には、その列は分析から除外されます。欠測値に対する上記のルールは、追加変数を含むすべての変数に適用されます。

「多重因子分析」レポート

「多重因子分析」レポートには、初期状態では、固有値の表、要約プロット、コンセンサスマップが表示されます。

図7.6「多重因子分析」レポート



要約プロット

「要約プロット」は、以下の3つのセクションから構成されます。

- 1つ目は、コンセンサス主成分分析（consensus principal component analysis）の固有値です。各成分によって説明される分散の累積割合に関するプロットも表示されます。「コンセンサス主成分分析」とは、複数のブロックに共通の主成分を求めるための主成分分析であり、重み付けた部分テーブルをまとめ、データ全体に主成分分析したものです。
- 「要約プロット」の中央には、各データ行（各商品）に対する主成分スコアがプロットされています。カテゴリカルな追加変数を使用した場合は、その追加変数の水準がラベル付きマーカーで示されます。このプロットで近くに配置されている商品は、類似していると見なすことができます。

- 「要約プロット」の右には、主成分負荷量プロットが表示されます。この主成分負荷量は、ブロックごとに色分けされています。

ヒント: また、主成分負荷量プロットの凡例をクリックすると、該当するブロックが強調表示されます。

次元の選択　スコアプロットと負荷量プロットの次元を変更します。1つ目は横軸の次元を、2つ目は縦軸の次元をそれぞれ変更します。

コンセンサスマップ

「コンセンサスマップ」レポートは、商品ごとに、各評価者の評価、および、全評価者の平均をプロットしたものです。このマップは、各商品に関して、評価者の評価がどれだけ似ているかを見るのに役立ちます。たとえば、ある評価者の点が他の評価者の点よりも常に大きく異なっている場合は、その評価者を分析から除外することを検討してもよいでしょう。

次元の選択　コンセンサスマップの次元を変更します。1つ目は横軸の次元を、2つ目は縦軸の次元をそれぞれ変更します。

一部の点だけを表示　慣性（inertia; 分散）の値に基づき、各商品の点や線の濃淡を変更します。慣性の値が小さいものは、評価者間で評価が似ていることを示しています。慣性の値が大きいものは、評価者間で評価が異なっていることを示しています。

慣性が小さいもの　慣性がテキストボックスに指定した値以下になっている商品だけを強調表示します。慣性の閾値を変更するには、スライダを使用するか、またはテキストボックスに値を入力します。

慣性が大きいもの　慣性がテキストボックスに指定した値以上になっている商品だけを強調表示します。慣性の閾値を変更するには、スライダを使用するか、またはテキストボックスに値を入力します。

慣性が最小と最大のもの　慣性の値が最小になっている商品と最大となっている商品を強調表示します。このオプションを選択した場合は、テキストボックスの値やスライダの位置は結果に影響しません。

「多重因子分析」プラットフォームのオプション

「多重因子分析」の赤い三角ボタンのメニューには、以下のオプションが含まれています。

ブロックに対する重み　ブロックごとに計算された第1固有値、および、コンセンサス主成分分析を行うときに使われる各ブロックへの重みの表示／非表示を切り替えます。「重み」は、第1固有値の平方根の逆数です。

固有値　固有値表の表示／非表示を切り替えます。この表には、コンセンサス主成分分析の固有値が降順に表示されます。

固有ベクトル　固有ベクトル表の表示／非表示を切り替えます。この表には、コンセンサス主成分分析の固有ベクトルが表示されます。左から順に、第1次元目、第2次元目、… と表示されます。これらの固有ベクトルを使って元の変数から線形結合を計算したものが、コンセンサス主成分分析における主成分スコアで

す。各固有ベクトルは、ノルムが1になるように標準化されたものに、ブロックごとの重みが掛けられています。

変数の負荷量 コンセンサス主成分分析から計算された主成分負荷量の表示／非表示を切り替えます。表に表示されている数値に対する濃淡は、負荷量の絶対値が0からどの程度離れているかを表しており、負荷量の絶対値が0に近いほど薄く、0から離れているほど濃く表示されます。

変数の偏寄与率 変数の偏寄与率（partial contribution of a variable）の表示／非表示を切り替えます。この値は、コンセンサス主成分分析の各主成分に対して各変数がどの程度寄与しているかを、分散の割合でパーセントを単位として示したもの（該当の主成分における分散のうち、各変数がどれだけ説明しているかを表した寄与率）です。

変数の余弦2乗 「変数の余弦2乗」（squared cosine of a variable）の表示／非表示を切り替えます。変数ごとに、この余弦2乗をすべての主成分について足し合わせると、1になります。ある変数の余弦2乗は、その変数が該当の主成分によってどの程度表されているかを示したものです。

ヒント：変数の偏寄与率が0に近いほど、該当の主成分はその変数では構成されていないことを意味しています。また、変数の余弦2乗が0に近いほど、該当の主成分はその変数を説明していないことを意味しています。表に示されている数値の濃淡は、それらの大きさを表しています。

要約プロット 要約プロットの表示／非表示を切り替えます。要約プロットには、固有値プロット、スコアプロット、負荷量プロットが含まれます。

コンセンサスマップ コンセンサスマップの表示／非表示を切り替えます。詳細については、「コンセンサスマップ」（171ページ）を参照してください。

バイプロット スコアプロットと負荷量プロットを重ねたプロットの表示／非表示を切り替えます。コントロールを使用してプロットの2つの次元を選択できます。

部分主成分軸プロット 部分主成分軸プロット（partial axis plot）の表示／非表示を切り替えます。このグラフは、ブロックごとの主成分分析における主成分スコアと、コンセンサス主成分分析における主成分スコアとの相関係数をプロットしたものです。コントロールを使用してプロットの2つの次元を選択できます。凡例のブロックをクリックすると、該当するブロックがプロットで強調表示されます。

表示オプション

矢印線 負荷量プロットと部分主成分軸プロットの矢印の表示／非表示を切り替えます。矢印は、変数の総数が1000以下のときに表示されます。変数の総数が1000を超える場合は、矢印はデフォルトでは表示されません。

ラベルの表示 負荷量プロット、バイプロット、部分主成分軸プロットにおいて、点に対する列名ラベルの表示／非表示を切り替えます。

RV相関 RV相関を要素とする行列の表示／非表示を切り替えます。RV相関は、ブロック間における関係の大きさを、相関係数の2乗のようなもので表した指標です。

Lg係数　Lg係数を要素とする行列の表示／非表示を切り替えます。Lg係数は、ブロック間の関係の大きさを、共分散の2乗のようなもので表した指標です。

ブロック偏スコアと全体との相関　ブロックごとに求められた偏主成分スコアと、コンセンサス主成分の主成分スコアとの相関係数を要素とする行列の表示／非表示を切り替えます。同じ次元どうしの相関係数のみが表示されるため、行列は長方形になります。

ブロックの偏寄与率　「ブロックの偏寄与率」の表示／非表示を切り替えます。これは、ブロックが該当の次元にどれぐらい寄与しているかを示す寄与率です。ブロック内に含まれている変数の偏寄与率を合計したものです。

ブロックの偏慣性　尺度を変更したブロックの偏寄与率の表示／非表示を切り替えます。これは、「ブロックの偏寄与率」にコンセンサス主成分分析の固有値を掛け合わせ、100で割った値です。

ブロックの余弦2乗　「ブロックの余弦2乗」の表示／非表示を切り替えます。これは、ブロックごとに「ブロックの偏慣性」を2乗して、固有値を計算するのに用いた平方和で割ったものです。0から1の間の値をとり、各ブロックにおいて余弦2乗を全主成分について足し合わせると1になります。この値は、ブロックの分散のうち各主成分によって説明される割合です。

データ行のスコアを保存　コンセンサス主成分分析から計算される主成分スコアを、データテーブルの新しい列に保存します。カテゴリカルな追加変数を1つまたは複数使用した場合は、それらの変数の水準ごとの主成分スコアも新しいデータテーブルに保存されます。

データ行の余弦2乗を保存　各データ行の余弦2乗を、データテーブルの新しい列に保存します。カテゴリカルな追加変数を1つまたは複数使用した場合は、その余弦2乗も新しいデータテーブルに保存されます。

データ行の偏寄与率を保存　各データ行の偏寄与率を、データテーブルの新しい列に保存します。カテゴリカルな追加変数を1つまたは複数使用した場合は、その偏寄与率も新しいデータテーブルに保存されます。

ブロックの偏スコアを保存　ブロックの偏スコアを新しいデータテーブルに保存します。

部分主成分軸の座標を保存　部分主成分軸の座標を新しいデータテーブルに保存します。

以下のオプションについて詳しくは、『JMPの使用法』の「JMPのレポート」章を参照してください。

ローカルデータフィルタ　現在のレポートにおいて、データをフィルタリングするローカルデータフィルタの表示／非表示を切り替えます。

やり直し　分析を繰り返したり、やり直したりするオプションを含みます。また、［自動再計算］オプションに対応しているプラットフォームにおいては、［自動再計算］オプションを選択すると、データテーブルに加えた変更が、該当するレポートに即座に反映されるようになります。

スクリプトの保存　レポートを再現するためのスクリプトを保存するオプションが、保存先ごとに用意されています。

「多重因子分析」プラットフォームの統計的詳細

多重因子分析では、複数の部分テーブル（sub-tables）から、互いに直交する主成分スコアを求めます。基本的な手順は以下のとおりです。

- それぞれの部分テーブルに対する主成分分析を個別に行います。
- それぞれの部分テーブルの第1固有値から、重みを算出します。
- 部分テーブルを左右に連結し、まず、各列を平均0分散1に標準化します。
- 標準化した行列に対して、部分テーブルごとに重みを掛け合わせます。その重みを掛け合わされた行列に対して特異値分解を行います。これは、部分テーブルごとの重みを考慮した一般化特異値分解を標準化した行列に対して行うのと等価です。この計算を、「コンセンサス主成分分析」と言います。つまり、部分テーブルに対して重みを与えたものに主成分分析を適用します。

上記の手順によって、右特異ベクトル、左特異ベクトル、特異値という3つの行列が求められます。これらの行列から、コンセンサス主成分分析の主成分スコア、固有値、主成分負荷量が求められます。これらの3つの行列は、元のデータに含まれるさまざまな列を少数の次元にまで縮小したものになっています。最初の数次元を調べることにより、商品の類似度や相違性を判断することができます。

計算

多重因子分析における行列**X**の特異値分解は、以下のように定義することができます。

$$X = P\Delta Q^T \quad \text{ここで、} \quad P^T M P = Q^T A Q = I$$

各行列の定義は以下のとおりです。

Xは、部分テーブルから成る$n \times p$の行列で、平均0分散1に標準化されています。官能検査のデータでは、n個の商品（製品、食品）と、複数の評価者から得られたp種類の評価に該当します。

Qは、右特異ベクトルから成る$p \times q$の行列です。q個の主成分負荷量を計算するのに使用されます。

Δは、特異値から成る$q \times q$の対角行列です。主成分分析と同様に、特異値の2乗は固有値であり、各主成分の重要度を表しています。

Pは、左特異ベクトルから成る$n \times q$の行列です。q個の主成分スコアを求めるのに使用されます。

Mは、各データ行の重み（度数）を示す$n \times n$の対角行列です。

Aは、ブロックの重みから成る、$p \times p$の対角行列です。

多重因子分析の詳細については、Abdi et al.（2013）を参照してください。

慣性の計算について

JMPでは、慣性（分散）を計算する際に$N - 1$を使用しています。これらの計算は、データ行とブロックに対するスコアの計算に影響します。

付録 A

参考文献

Abdi, H., Williams, L. J., and Valentin, D. (2013). "Multiple factor analysis: principal component analysis for multitable and multiblock data sets." *WIREs Comp Stat*, 5:149-179. doi: 10.1002/wics.1246.

Benjamini, Y., and Hochberg, Y. (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B* 57:289-300.

Breslow, N. E., and Day, N. E. (1980). *The Analysis of Case-Control Studies*. Statistical Methods in Cancer Research, IARC Scientific Publications, vol. 1, no. 32. Lyon: International Agency for Research on Cancer.https://www.iarc.fr/en/publications/pdfs-online/stat/sp32/SP32.pdf. Retrieved 3/9/2017.

Firth, D. (1993). "Bias Reduction of Maximum Likelihood Estimates." Biometrika 80:27–38.

Heinze, G., and Schemper, M. (2002). "A Solution to the Problem of Separation in Logistic Regression." *Statistics in Medicine* 21:2409–2419.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

Lavassani, K. M., Movahedi, B., and Kumar, V. (2009). "Developments in Analysis of Multiple Response Survey Data in Categorical Data Analysis: The Case of Enterprise System Implementation in Large North American Firms." *Journal of Applied Quantitative Methods* 4:45–53.

Louviere, J. J., Flynn, T. N., and Marley, A. A. (2015). *Best-Worst Scaling: Theory, Methods and Applications*. Cambridge: Cambridge University Press.

McFadden, D. (1974). "Conditional Logit Analysis of Qualitative Choice Behavior." in *Frontiers in Econometrics* edited by P. Zarembka, 105–142. New York: Academic Press. See http://eml.berkeley.edu/reprints/mcfadden/zarembka.pdf. Retrieved 4/25/2016.

Radcliffe, N. J., and Surry, P. D. (2011). *Real-World Uplift Modeling with Significance-Based Uplift Trees*. Portrait Technical Report TR-2011-1, Stochastic Solutions. http://www.stochasticsolutions.com/pdf/sig-based-up-trees.pdf. Retrieved 1/13/2017.

Rossi, P. E., Allenby, G. M., and McCulloch, R. (2005). *Bayesian Statistics and Marketing*. Chichester, UK: John Wiley & Sons.

SAS Institute Inc. (2017). "The TTEST Procedure." In *SAS/STAT 14.3 User's Guide.* Cary NC: SAS Institute Inc. http://go.documentation.sas.com/?docsetId=statug&docsetTarget=statug_ttest_details14.htm&docsetVersion=14.3&locale=en. Retrieved 12/22/2017

Train, K. E. (2001). “A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit.” Department of Economics, University of California, Berkley. https://eml.berkeley.edu/~train/compare.pdf. Retrieved 1/13/2017.

Train, K. E. (2009). *Discrete Choice Methods and Simulation*. 2nd ed. Cambridge University Press.

Westfall, P. H., Tobias, R. D., and Wolfinger, R. D. (2011). *Multiple Comparisons and Multiple Tests Using SAS*. 2nd ed. Cary, NC: SAS Institute Inc.

索引
消費者調査

M
MaxDiff
- 応答データ 140
- 効用計算式の保存 149
- 最良 - 最悪スケーリング 129
- 推定値の相関 149
- 全水準の比較レポート 149
- 被験者ごとの勾配を保存 149
- ピザの選択の例 129–132
- プロファイルデータ 139

MaxDiff、複合因子検定 149

ア
アップリフトモデル 153
- プラットフォーム 157
- レポートオプション 161

オ
応答データ、MaxDiff 140
応答の等質性に対する検定 36

カ
各応答に対する検定 36
カテゴリカルプラットフォーム 27
- 起動ウィンドウ 30
- 自由回答 32
- 上位カテゴリ 45
- 表の構成 32
- レポート 34

コ
効用計算式の保存、MaxDiff 149
効用計算式の保存、選択モデルプラットフォーム 95

サ
最良 - 最悪スケーリング 129

シ
支払意思額、選択モデルプラットフォーム 95
信頼区間、MaxDiff 149

ス
推定値の相関、MaxDiff 149

セ
全水準の推定値、選択モデルプラットフォーム 93
全水準の比較レポート、MaxDiff 149
選択なしオプション、選択モデルプラットフォーム 83, 88
選択モデル 69–123
選択モデルプラットフォーム
- 効用計算式の保存 95
- 支払意思額 95
- セグメント化 72, 129
- 全水準の推定値 93
- 選択なしオプション 83, 88
- 被験者ごとの勾配を保存 72, 95, 129
- 被験者データ 88
- 複合因子検定 93
- プロファイルデータ 84

ヒ
比較、MaxDiff 149
被験者ごとの勾配を保存、MaxDiff 149
被験者ごとの勾配を保存、選択モデルプラットフォーム 72, 95, 129
被験者データ、選択モデルプラットフォーム 88
ピザの選択の例 107–110

ピザの選択の例、MaxDiff 129–132

フ

複合因子検定
maxdiff 149
複合因子検定、選択モデルプラットフォーム 93
プロファイルデータ、MaxDiff 139
プロファイルデータ、選択モデルプラットフォーム 84

モ

文字を使用した比較 38, 45

ユ

尤度比検定 93