

IBM SPSS Custom Tables

23

The IBM logo, consisting of the letters 'IBM' in a bold, black, sans-serif font. Each letter is composed of horizontal bars, with the 'I' having seven bars, the 'B' having eight bars, and the 'M' having nine bars.

お願い

本書および本書で紹介する製品をご使用になる前に、107 ページの『特記事項』に記載されている情報をお読みください。

本書は、IBM SPSS Statistics バージョン 23 リリース 0 モディフィケーション 0 および新しい版で明記されない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原典： IBM SPSS Custom Tables 23

発行： 日本アイ・ビー・エム株式会社

担当： トランスレーション・サービス・センター

目次

第 1 章 テーブル・ビルダー・インターフ

エース	1
テーブル・ビルダー・インターフェース	1
テーブルの作成	1
テーブルの作成	4
変数の積み重ね	4
変数のネスト	5
層	5
変数名や変数ラベルの表示と非表示	6
要約統計量	6
カテゴリと合計	11
計算対象カテゴリ	14
カテゴリを共有する変数のテーブル (比較テ ブル)	15
テーブル・ビルダーのカスタマイズ	16
カスタム・テーブル: 「オプション」タブ	16
カスタム・テーブル: 「表題」タブ	17
カスタム・テーブル: 「検定統計量」タブ	17

第 2 章 カテゴリ変数の単純なテーブル 19

カテゴリ変数の単純なテーブル	19
単一のカテゴリ変数	19
パーセント	20
合計	21
クロス集計	21
クロス集計表のパーセント	22
表示形式の制御	23
周辺合計	23
カテゴリのソートと除外	24

第 3 章 カテゴリ変数の積み重ね、ネ スト、層 27

カテゴリ変数の積み重ね	27
クロス集計を使用した積み重ね	27
カテゴリ変数のネスト	28
変数ラベルの抑止	29
ネストしたクロス集計表	30
層	32
2 つのカテゴリ層変数の積み重ね	33
ネストした 2 つのカテゴリ層変数	33

第 4 章 カテゴリ変数の合計と小計 35

単一変数の単純な合計	35
合計での WYSIWYG	36
合計の表示位置	36
ネストしたテーブルの合計	37
層変数の合計	38
小計	39
小計での WYSIWYG	39
小計カテゴリの非表示	40

層変数の小計	41
--------	----

第 5 章 カテゴリ変数の計算対象カテ ゴリ 43

単純な計算対象カテゴリ	43
計算対象カテゴリのカテゴリの非表示	44
計算対象カテゴリの小計の参照	45
計算対象カテゴリを使用した包括的で ない小計の表示	46

第 6 章 カテゴリを共有する変数のテ ブル 49

度数のテーブル	49
パーセントのテーブル	50
合計とカテゴリの制御	51
共有カテゴリのあるテーブルでのネ スト	52

第 7 章 要約統計量 53

要約統計量ソース変数	53
カテゴリ変数の要約統計量ソース	54
スケール変数の要約統計量ソース	54
積み重ね変数	55
カテゴリ変数のカスタム合計要約統計 量	56
カテゴリ値の表示	57

第 8 章 スケール変数の集計 59

スケール変数の集計	59
スケール変数の積み重ね	59
複数の要約統計量	60
度数、有効 N、および欠損値	60
各種変数の異なる要約	62
カテゴリ内でのグループ要約	63
複数のグループ化変数	63
スケール変数の中でのカテゴリ変 数のネスト	64

第 9 章 検定統計量 67

検定統計量	67
独立性の検定 (カイ 2 乗)	67
独立性の検定におけるネストと積 み重ねの影響	68
列の平均値の比較	70
列の平均値の検定におけるネスト と積み重ねの影響	71
列の比率の比較	72
列の比率の検定におけるネストと 積み重ねの影響	75
重み付けおよび多重回答セットに 関する注	77

第 10 章 多重回答セット 79

度数、回答数、パーセント、およ び合計	79
その他の変数との多重回答セットの 使用	81
統計量ソース変数と利用できる要 約統計量	82

多重カテゴリー・セットと重複回答	83	要約統計量の表示形式	91
多重回答セットでの有意差検定	84	要約統計量の表示ラベル	92
多重回答セットでの独立性の検定	84	列幅	93
多重回答セットでの列の平均値の比較	85	空のセルの表示値	94
		欠損統計量の表示値	94
第 11 章 欠損値	87	第 13 章 サンプル・ファイル	97
欠損値のないテーブル	87	特記事項	107
テーブルに欠損値を含める	88	商標	109
第 12 章 テーブルの書式設定とカスタマイズ	91	索引	111
テーブルの書式設定とカスタマイズ	91		

第 1 章 テーブル・ビルダー・インターフェース

テーブル・ビルダー・インターフェース

カスタム・テーブルでは、単純なドラッグ・アンド・ドロップ方式によるテーブル・ビルダー・インターフェースを使用して、テーブルをプレビューしながら、変数やオプションを選択できます。また、一般的なダイアログ・ボックスにはないレベルの柔軟性を備えており、例えばウィンドウのサイズやウィンドウ内のペインのサイズを変更することもできます。

テーブルの作成



図 1. 「カスタム・テーブル」ダイアログ・ボックスの「テーブル」タブ

テーブル・ビルダーの「テーブル」タブで、テーブルに表示される変数や要約統計量を選択します。

変数リスト: データ・ファイル内の変数は、ウィンドウの左上のペインに表示されます。「カスタム・テーブル」では、変数用の 2 つの異なる測定レベルが区別され、測定レベルに応じて、異なる方法で変数が処理されます。

カテゴリー: 限られた数の異なる値またはカテゴリーを持つデータ (例えば、性別や宗教)。カテゴリー変数には、文字列 (英数字) 変数、または数字コード (例えば、0 = 男性 および 1 = 女性) でカテゴリーを表す数値変数があります。これは、質的データとも呼ばれます。カテゴリー変数は、**名義変数**または**順序変数**のいずれかです。

- *Nominal* (名義データ). 本質的な順位を持たないカテゴリを表す値である場合 (従業員の勤務先企業での部署など)、変数を名義変数として取り扱うことができます。名義変数の例としては、地域、郵便番号、宗教上の所属などが挙げられます。
- *Ordinal* (順序データ). 値が本質的な順位を持つカテゴリを表す場合 (例えば、サービス満足度のレベルを「非常に不満」から「非常に満足」までの順位で評価する場合) は、変数を順序変数として扱うことができます。順序変数の例としては、満足度や信頼度を表す得点や嗜好得点などが挙げられます。

スケール: 間隔または比率スケールについて測定したデータ。この場合、データ値は、値の順序と値の間の距離の両方を示します。例えば、\$72,195 の給料は \$52,398 の給料より高く、この 2 つの値の間の距離は \$19,797 です。量的データ、連続型データとも呼ばれます。

カテゴリ変数は、テーブル内のカテゴリ (行、列、および層) を定義します。デフォルトの要約統計量は、度数 (各カテゴリのケース数) です。例えば、性別に関するカテゴリ変数のデフォルト・テーブルには、単純に男性の数と女性の数が表示されます。

通常、スケール変数は、カテゴリ変数のカテゴリ内で要約されます。デフォルトの要約統計量は平均値です。例えば、性別のカテゴリに含まれる収入のデフォルト・テーブルには、男性の平均収入と女性の平均収入が表示されます。

カテゴリ変数を使用してグループを定義することなく、スケール変数自体でスケール変数を要約することもできます。これは主に、複数のスケール変数の要約を**積み重ねる**場合に便利です。詳しくは、4 ページの『変数の積み重ね』のトピックを参照してください。

多重回答セット

カスタム・テーブルは、**多重回答セット**と呼ばれる特殊な種類の「変数」もサポートしています。多重回答セットは、実際には、一般的な意味での変数ではありません。多重回答セットはデータ・エディターには表示されず、他のプロシージャでも認識されません。多重回答セットは、複数の回答が可能な質問に対する回答を、複数の変数を使用して記録します。多重回答セットは、カテゴリ変数と同様に処理され、カテゴリ変数で実行できるほとんどの処理は、多重回答セットでも実行することができます。詳しくは、79 ページの『第 10 章 多重回答セット』のトピックを参照してください。

変数リストの各変数の隣にあるアイコンは、変数の型を識別します。

表 1. 測定レベル・アイコン









	数値	文字列	日付	時刻
スケール (連続)		なし		
順序				
名義				

表 2. 多重回答セットのアイコン

多重回答セット・タイプ	アイコン
多重回答セット、多重カテゴリ	

表 2. 多重回答セットのアイコン (続き)

多重回答セット・タイプ	アイコン
多重回答セット、二分法	

テーブル・ビルダーで変数の測定レベルを変更するには、変数リストで変数を右クリックし、ポップアップ・メニューから「**カテゴリ**」または「**スケール**」を選択します。データ・エディターの「変数ビュー」で、変数の測定レベルを永続的に変更することができます。「**名義**」または「**順序**」として定義された変数は、カスタム・テーブルでカテゴリ変数として処理されます。

カテゴリ: 変数リストでカテゴリ変数を選択すると、そのカテゴリ変数用に定義されたカテゴリが「カテゴリ」リストに表示されます。これらのカテゴリは、テーブルで変数を使用する際に、キャンバス・ペインにも表示されます。変数にカテゴリが定義されていない場合、「カテゴリ」リストとキャンバス・ペインには、プレースホルダー・カテゴリとして「**カテゴリ 1**」と「**カテゴリ 2**」の 2 つが表示されます。

テーブル・ビルダーに表示される定義済みカテゴリは、**値ラベル**に基づいています。これは説明的なラベルで、さまざまなデータ値に割り当てられます (例えば、数値 0 と 1 に、値ラベルの「**男性**」と「**女性**」が割り当てられるなど)。値ラベルを定義するには、データ・エディターの「変数ビュー」、またはデータ・エディター・ウィンドウの「データ」メニューにある「**変数プロパティの定義**」を使用します。

キャンバス・ペイン: テーブルを作成するには、キャンバス・ペインの行と列に変数をドラッグ・アンド・ドロップします。キャンバス・ペインには、作成されるテーブルのプレビューが表示されます。キャンバス・ペインのセルには、実際のデータ値は表示されませんが、キャンバス・ペインで、最終的なテーブルのレイアウトをかなり正確に確認することができます。カテゴリ変数の場合、値ラベルが定義されていない固有値がデータ・ファイル内に存在すると、プレビューに表示されているよりも多くのカテゴリが実際のテーブルに含まれることがあります。

- 「**標準**」ビューには、カテゴリ変数の要約統計量とカテゴリの行や列など、テーブルに含まれるすべての行と列が表示されます。
- 「**コンパクト**」ビューには、テーブル内の変数だけが表示され、テーブルに含まれる行と列のプレビューは表示されません。

テーブル作成時の基本的な規則と制約事項

- カテゴリ変数の場合、要約統計量は、統計ソース・ディメンション内で最も深いレベルの変数に基づきます。
- カテゴリ変数のデフォルトの統計ソース・ディメンション (行または列) は、変数をキャンバス・ペインにドラッグ・アンド・ドロップする順序に基づきます。例えば、最初に行トレイに変数をドラッグした場合、行ディメンションが、デフォルトの統計ソース・ディメンションになります。
- スケール変数は、行ディメンション内または列ディメンション内で最も深いレベルの変数のカテゴリ内でのみ要約できます (スケール変数は、テーブルの任意のレベルに配置できますが、要約は最も深いレベルで行われます)。
- スケール変数を他のスケール変数内で要約することはできません。ただし、複数のスケール変数の要約を積み重ねたり、カテゴリ変数のカテゴリ内でスケール変数を要約したりすることはできます。あるスケール変数を別のスケール変数内でネストしたり、あるスケール変数を行ディメンション内に配置し、別のスケール変数を列ディメンション内に配置したりすることはできません。

- アクティブ・データ・セット内のいずれかの変数に 12,000 を超える定義済み値ラベルが含まれている場合、テーブル・ビルダーを使用してテーブルを作成することはできません。この制限を超える変数をテーブルに含める必要がなければ、これらの変数を除外する変数セットを定義して適用することができます。12,000 を超える定義済み値ラベルを持つ変数を含める必要がある場合は、CTABLES コマンド構文を使用してテーブルを生成することができます。

テーブルの作成

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. 1 つ以上の変数を、キャンバス・ペインの行領域または列領域にドラッグ・アンド・ドロップするか、移動します。
3. 「OK」をクリックして、テーブルを作成します。
4. キャンバス・ペインで変数を選択 (クリック) します。
5. キャンバス・ペインの外側に変数をドラッグするか、Delete キーを押します。

変数の測定レベルを変更するには、以下の手順に従います。

6. 変数リストの変数を右クリックします (この操作は、変数リストでのみ行えます。キャンバス上で行うことはできません)。
7. ポップアップ・メニューから「カテゴリー」または「スケール」を選択します。

変数の積み重ね

積み重ねは、個別の複数のテーブルをまとめて同じ表示に貼り付けることとして考えられます。例えば、「性別」の情報と「年齢カテゴリー」の情報を、同じテーブルの別々のセクションに表示することができます。

変数を積み重ねる

1. 変数リストで、積み重ねたい変数をすべて選択してから、それらをまとめてキャンバス・ペインの行または列にドラッグ・アンド・ドロップします。

または

2. 個々の変数を、行にある既存の変数の上または下、あるいは列にある既存の変数の右または左にドラッグ・アンド・ドロップします。

表 3. 積み重ねたカテゴリー変数：

変数	カテゴリー	要約統計量
変数 1	カテゴリー 1	123
	カテゴリー 2	456
変数 2	カテゴリー 1	123
	カテゴリー 2	456
	カテゴリー 3	789

詳しくは、27 ページの『カテゴリー変数の積み重ね』のトピックを参照してください。

変数のネスト

クロス集計表と同様に、ネストでも 2 つのカテゴリー変数の間にある関係を示すことができます。ただし、同じディメンションで別の変数にネストされている変数は除きます。例えば、行ディメンションで「年齢カテゴリー」の中に「性別」をネストして、各年齢カテゴリーに属する男性と女性の数を表示できます。

カテゴリー変数の中にスケール変数をネストすることもできます。例えば、「性別」の中に「収入」をネストして、平均収入（または中央値や他の要約統計量）を男性と女性に分けて表示することができます。

変数をネストする

1. カテゴリー変数をキャンパス・ペインの行領域または列領域にドラッグ・アンド・ドロップします。
2. カテゴリー変数またはスケール変数を、カテゴリー行変数の左または右、あるいはカテゴリー列変数の上または下にドラッグ・アンド・ドロップします。

表 4. ネストしたカテゴリー変数：

変数 1	変数 2	要約統計量
カテゴリー 1	カテゴリー 1	12
	カテゴリー 2	34
	カテゴリー 3	56
カテゴリー 2	カテゴリー 1	12
	カテゴリー 2	34
	カテゴリー 3	56

詳しくは、28 ページの『カテゴリー変数のネスト』のトピックを参照してください。

注: カスタム・テーブルは、層化された分割ファイルの処理は行いません。層化分割ファイルと同じ結果を得るには、分割ファイル変数を、テーブルの最も浅いレベルにネストされた層に配置します。

層

層を使用してテーブルに深さのディメンションを追加し、3 次元の「立体」を作成することができます。層はネストや積み重ねと似ています。主な違いは、層カテゴリーは一度に 1 つしか表示できないことです。例えば、「年齢カテゴリー」を行変数、「性別」を層変数として使用すると、男性に関する情報と女性に関する情報をそれぞれ異なる層に表示するテーブルが作成されます。

層を作成する

1. テーブル・ビルダーの「テーブル」タブにある「層」をクリックして、「層」リストを表示します。
2. 層を定義するスケール変数またはカテゴリー変数を「層」リストにドラッグ・アンド・ドロップします。

「層」リスト内にスケール変数とカテゴリー変数を混在させることはできません。すべての変数は、同じタイプでなければなりません。多重回答セットは、「層」リストではカテゴリー変数として扱われます。層に含まれるスケール変数は、常に積み重ねられます。

複数のカテゴリー層変数がある場合、層を積み重ねるか、ネストすることができます。

- 「各カテゴリを層として表示」は、積み重ねに相当します。各層変数のカテゴリーごとに、個別の層が表示されます。層の合計数は、層変数ごとのカテゴリー数の合計です。例えば、3 つの層変数があり、それぞれの層変数に 3 つのカテゴリーがある場合、表には 9 つの層があることになります。

- 「各組み合せを層として表示」は、層のネストまたはクロス集計に相当します。層の合計数は、層変数ごとのカテゴリ数²の積です。例えば、3 つの変数があり、それぞれの変数に 3 つのカテゴリがある場合、表には 27 つの層があることとなります。

変数名や変数ラベルの表示と非表示

変数名および変数ラベルの表示には、次のオプションを使用できます。

- 変数ラベルだけを表示。定義済みの変数ラベルがない変数には、変数名が表示されます。これはデフォルト設定です。
- 変数名だけを表示。
- 変数ラベルと変数名の両方を表示。
- 変数名と変数ラベルを表示しない。変数ラベルや変数名が含まれる列または行は、キャンバス・ペインのテーブル・プレビューには表示されますが、実際のテーブルには表示されません。

変数ラベルや変数名を表示または非表示にするには、以下の手順に従います。

1. キャンバス・ペインのテーブル・プレビューで変数を右クリックします。
2. ポップアップ・メニューから「変数ラベルを表示」または「変数名を表示」を選択して、ラベルまたは名前の表示のオン/オフを切り替えます。選択項目の隣にあるチェック・マークは、その項目が表示されることを示します。

要約統計量

「要約統計量」ダイアログ・ボックスでは、次の処理を実行できます。

- テーブルの要約統計量の追加と削除を行う。
- 統計量のラベルを変更する。
- 統計量の順序を変更する。
- 統計量の書式を変更する (小数桁数を含む)。

このダイアログ・ボックスで選択可能な要約統計量 (およびその他のオプション) は、ダイアログ・ボックスの上部に表示される、要約統計量ソース変数の測定レベルに依存します。要約統計量のソース (要約統計量の計算に使用される変数) は、以下の基準で決まります。

- **測定レベル:** テーブル (または積み重ねテーブルのテーブル・セクション) にスケール変数がある場合、要約統計量は、そのスケール変数に基づきます。
- **変数の選択順序:** カテゴリ変数のデフォルトの統計ソース・ディメンション (行または列) は、変数をキャンバス・ペインにドラッグ・アンド・ドロップする順序に基づきます。例えば、最初に行領域に変数をドラッグした場合、行ディメンションがデフォルトの統計ソース・ディメンションになります。
- **ネスト:** カテゴリ変数の場合、要約統計量は、統計ソース・ディメンション内で最も深いレベルの変数に基づきます。

積み重ねテーブルには複数の要約統計量ソース変数 (スケール変数とカテゴリ変数の両方) が存在する場合がありますが、各テーブル・セクションには 1 つの要約統計量ソースしか存在しません。

要約統計量ソース・ディメンションを変更する

1. 「テーブル」タブの「要約統計量」グループにある「ソース」ドロップダウン・リストから、ディメンション (行、列、または層) を選択します。

テーブルに表示する要約統計量を制御する

1. 「テーブル」タブのキャンバス・ペインで、要約統計量ソース変数を選択（クリック）します。
2. 「テーブル」タブの「定義」グループで、「要約統計量」をクリックします。

または

3. キャンバス・ペインで要約統計量ソース変数を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
4. テーブルに含める要約統計量を選択します。矢印を使用して、選択した統計量を「統計」リストから「表示」リストに移動することができます。また、選択した統計量を「統計」リストから「表示」リストにドラッグ・アンド・ドロップすることもできます。
5. 現在選択している要約統計量の表示位置を変更するには、上矢印または下矢印をクリックします。
6. 「書式」ドロップダウン・リストから、選択した要約統計量の表示形式を選択します。
7. 選択した要約統計量の「小数部の桁数」セルに、表示する小数点以下の桁数を入力します。
8. キャンバス・ペインで現在選択されている変数に対して、選択した要約統計量を適用するには、「**選択項目に適用**」をクリックします。
9. キャンバス・ペインで積み重ねられた同じタイプのすべての変数に対して、選択した要約統計量を適用するには、「**すべてに適用**」をクリックします。

注: 「**すべてに適用**」を選択すると、既にキャンバス・ペイン上に配置されている同じタイプの積み重ね変数に対して適用されます。それ以外は、「**選択項目に適用**」と同じです。いずれの場合も、選択した要約統計量は、テーブルに追加した同じタイプの積み重ね変数に対して自動的に適用されます。

カテゴリー変数の要約統計量

カテゴリー変数で使用できる基本的な統計量は、度数とパーセントです。合計と小計に対して、カスタム要約統計量を指定することもできます。カスタム要約統計量には、一部の順序カテゴリー変数に適した中心傾向（平均値や中央値など）や散らばり（標準偏差など）の測定値などがあります。詳しくは、10 ページの『カテゴリー変数のカスタム合計要約統計量』のトピックを参照してください。

度数: テーブルの各セルに含まれるケースの数、または多重回答セットの回答者の数。

重み付けのない度数: テーブルの各セルに含まれる重み付けのないケースの数。重み付けが有効な場合のみ、度数とは異なって機能します。

列パーセント: 各列でのパーセント。サブテーブルの各列のパーセントを合計すると 100% になります（単純なパーセントの場合）。通常、列パーセントが役立つのは、カテゴリー行変数を使用されている場合だけです。

行パーセント: 各行でのパーセント。サブテーブルの各行のパーセントを合計すると 100% になります（単純なパーセントの場合）。通常、行パーセントが役立つのは、カテゴリー列変数を使用されている場合だけです。

層の行パーセントと層の列パーセント: ネストされたテーブル内のすべてのサブテーブルの行パーセントまたは列パーセントを合計すると 100% になります（単純なパーセントの場合）。テーブルに層が含まれている場合、それぞれの層において、ネストされたすべてのサブテーブルの行パーセントまたは列パーセントを合計すると 100% になります。

層パーセント: 各層でのパーセント。単純なパーセントの場合、現在表示されている層内のセルのパーセントを合計すると 100% になります。層変数を使用していない場合、これは表パーセントと同じになります。

表パーセント: 各セルのパーセントは、テーブル全体に基づきます。すべてのセルのパーセントは同じケース合計数に基づき、テーブル全体で合計すると 100% になります (単純なパーセントの場合)。

サブテーブル・パーセント: 各セルのパーセントは、サブテーブル全体に基づきます。サブテーブル内のすべてのセルのパーセントは同じケース合計数に基づき、サブテーブル内で合計すると 100% になります。ネストされたテーブルでは、最も深いネストレベルの前にある変数によってサブテーブルが定義されます。例えば、「年齢カテゴリー」の中に「性別」があり、その中に「婚姻状況」があるテーブルの場合、「性別」によってサブテーブルが定義されます。

多重回答セットでは、ケース数、回答数、度数に基づいてパーセントを計算することができます。詳しくは、『多重回答セットの要約統計量』のトピックを参照してください。

積み重ねテーブル

パーセントの計算では、積み重ね変数によって定義されたテーブルの各セクションが、個別のテーブルとして扱われます。層の行パーセント、層の列パーセント、および表パーセントは、積み重ねテーブルのセクションごとに合計すると 100% になります (単純なパーセントの場合)。さまざまなパーセントの計算で使用するパーセントの分母は、積み重ねテーブルの各セクションに含まれるケース数に基づきます。

パーセントの分母

パーセントの計算には 3 つの方法があります。どの方法で計算されるかは、計算の基準で欠損値をどのように扱うかによって決まります。

単純なパーセント: パーセントはテーブルで使用するケース数に基づいており、すべてのパーセントを合計すると常に 100% になります。カテゴリーをテーブルから除外した場合、そのカテゴリーのケースは分母となる数から除外されます。システム欠損値のあるケースは、常に分母となる数から除外されます。ユーザー欠損値のあるケースは、ユーザー欠損カテゴリーがテーブルから除外される場合 (デフォルト) には除外され、ユーザー欠損カテゴリーがテーブルに含まれる場合には含められます。名前に有効 N または合計 N が含まれないパーセントは、単純なパーセントです。

合計 N パーセント: システム欠損値およびユーザー欠損値があるケースが、単純なパーセントの分母に追加されます。パーセントの合計が、100% にならない場合があります。

有効 N パーセント: ユーザー欠損カテゴリーがテーブルに含まれている場合でも、ユーザーユーザー欠損値があるケースは、単純なパーセントの分母から除外されます。

注: ユーザー欠損カテゴリー以外のカテゴリーを手動で除外した場合、そのカテゴリーのケースは、常に分母となる数から除外されます。

多重回答セットの要約統計量

多重回答セットには、以下の要約統計量も使用できます。

列/層の応答 %: 応答数に基づくパーセントです。

列/層の応答 % (基準: 度数): 応答数が分子、合計度数が分母となります。

列/層の度数 % (基準: 応答数): 度数が分子、合計応答数が分母となります。

層の列/行の応答 %: サブテーブルでのパーセントです。応答数に基づくパーセントです。

層の列/行の応答 % (基準: 度数): サブテーブルでのパーセントです。応答数が分子、合計度数が分母となります。

層の列行の応答 % (基準: 応答数): サブテーブルでのパーセントです。度数が分子、合計応答数が分母となります。

応答: 応答数です。

サブテーブル/テーブルの応答 %: 応答数に基づくパーセントです。

サブテーブル/テーブルの応答 % (基準: 度数): 応答数が分子、合計度数が分母となります。

サブテーブル/テーブルの度数 % (基準: 応答数): 度数が分子、合計応答数が分母となります。

スケール変数とカテゴリ・カスタム合計の要約統計量

カテゴリ変数で使用できる度数とパーセントのほかに、スケール変数では以下の要約統計量も使用することができます。また、これらの要約統計量は、カテゴリ変数のカスタム合計や小計要約として使用することもできます。これらの要約統計量は、多重回答セットや文字列 (英数字) 変数では使用できません。

平均値: 算術平均 (合計をケース数で割った値)。

中央値: ケースの中央付近にある値 (50 パーセンタイル)。

最頻値: 最も頻繁に出現する値。値の出現頻度が同じである場合、最小の値が表示されます。

最小: 最小 (最低) 値。

最大: 最大 (最高) 値。

欠損値: 欠損値の度数 (ユーザー欠損値とシステム欠損値の両方)。

パーセンタイル: 5、25、75、95、99 のいずれかのパーセンタイル (あるいはこれらすべてのパーセンタイル) を指定できます。

範囲: 最大値と最小値の差分。

平均値の標準誤差: 同じ分布から抽出したサンプル間で平均値がどの程度異なるかについての測定値。これを使用して、観測された平均値と仮説の値のおおよその比較を行うことができます (すなわち、差と標準誤差の比率が -2 より小さいか +2 より大きい場合に、2 つの値は異なっていると結論付けることができます)。

標準偏差: 平均値を中心とした散らばりの測定値。正規分布では、68% のケースが平均値 ± 1 標準偏差に含まれ、95% のケースが平均値 ± 2 標準偏差に含まれます。例えば、平均年齢が 45 歳で標準偏差が 10 の場合、正規分布ではケースの 95% が 25 歳から 65 歳までに含まれます (分散の平方根)。

合計: 値の合計。

合計パーセント: 合計に基づくパーセント。行と列 (サブテーブル内)、行全体と列全体 (サブテーブル全体)、層、サブテーブル、テーブル全体で使用することができます。

合計 N: 非欠損値、ユーザー欠損値、およびシステム欠損値の度数。ユーザー欠損カテゴリを除き、手動で除外したカテゴリのケースは含まれません。

有効 N: 非欠損値の度数。ユーザー欠損カテゴリを除き、手動で除外したカテゴリのケースは含まれません。

分散: 平均値を中心とした散らばりの測定値。平均値からの偏差の平方和を、ケースの数から 1 を引いた値で割った結果に等しくなります。分散は、変数自体の平方を単位として測定されます (標準偏差の平方)。

積み重ねテーブル

積み重ね変数によって定義されたテーブルの各セクションは個別のテーブルとして扱われ、要約統計量もそれに従って計算されます。

カテゴリー変数のカスタム合計要約統計量

合計または小計を持つカテゴリー変数のテーブルでは、各カテゴリーに表示される要約とは異なる要約統計量を表示することができます。例えば、順序カテゴリー行変数には度数と列パーセントを表示し、「合計」統計量には中央値を表示することができます。

カスタム合計要約統計量を表示するカテゴリー変数のテーブルを作成するには、以下の手順を実行します。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

テーブル・ビルダーが開きます。

2. カテゴリー変数をキャンバスの行領域または列領域にドラッグ・アンド・ドロップします。
3. キャンバスで変数を右クリックし、ポップアップ・メニューから「**カテゴリと合計**」を選択します。
4. 「合計」チェック・ボックスをクリックして (チェック・マークを付け)、**「適用」**をクリックします。
5. キャンバスでもう一度変数を右クリックし、ポップアップ・メニューから「**要約統計量**」を選択します。
6. 「**合計と小計に対するユーザー指定の要約統計量**」をクリックして (チェック・マークを付け)、目的のカスタム要約統計量を選択します。

デフォルトでは、カスタム要約を含むすべての要約統計量は、カテゴリー変数が表示されるディメンションとは反対側のディメンションに表示されます。例えば、カテゴリー行変数を使用する場合、要約統計量により、テーブルの列が以下のように定義されます。

表 5. 行に表示された順序変数カテゴリーと、列に表示された要約統計量の度数と平均値：

変数	カテゴリー	度数	平均値
変数 1	1 賛成	196	2.29
	2 中立	936	
	3 反対	744	
	合計	1876	

要約統計量をカテゴリー変数と同じディメンションに表示するには、以下の手順を実行します。

7. テーブル・ビルダーの「テーブル」タブにある「要約統計量」グループで、「位置」ドロップダウン・リストからディメンションを選択します。

例えば、カテゴリー変数を行に表示する場合、ドロップダウン・リストから「行」を選択します。

要約統計量の表示形式

表示形式オプションには、以下の 2 つがあります。

nnnn: 単純な数値です。

nnnn%: 値の末尾にパーセント記号が追加されます。

自動: 定義済みの変数表示形式で、小数桁数を含みます。

N=nnnn: 値の前に $N=$ が表示されます。この形式は、要約統計量ラベルが表示されないテーブルでの度数、有効 N 、合計 N を表示する場合に役立つことがあります。

(nnnn): すべての値が括弧で囲まれます。

(nnnn)(負の値): 負の値だけが括弧で囲まれます。

(nnnn%): すべての値が括弧で囲まれ、値の末尾にパーセント記号が追加されます。

n,nnn,n: カンマ書式。ロケールの設定に関係なく、複数の桁数の区切り文字としてカンマが使用され、小数点としてピリオドが使用されます。

n.nnn,n: ドット書式。ロケールの設定に関係なく、複数の桁数の区切り文字としてピリオドが使用され、小数点としてカンマが使用されます。

\$n,nnn,n: ドル書式。値の前にドル記号が表示されます。ロケールの設定に関係なく、複数の桁数の区切り文字としてカンマが使用され、小数点としてピリオドが使用されます。

CCA/CCB/CCC/CCD/CCE: カスタム通貨フォーマット。カスタム通貨に対して現在定義されている書式がリストに表示されます。これらの書式は、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「通貨」タブで定義します。

一般的な規則と制限事項

- 「自動」を選択した場合を除き、小数の桁数は「小数部の桁数」列の設定によって決まります。
- カンマ書式、ドル書式、ドット書式を選択した場合を除き、Windows の「地域のオプション」コントロール・パネルで現在のロケールに対して定義されている小数点が表示されます。
- カンマ書式とドル書式の場合は、複数の桁数の区切り文字としてカンマが表示され、ドット書式の場合はピリオドが使用されますが、作成時に使用できる表示書式はないため、現在のロケール設定 (Windows の「地域のオプション」コントロール・パネルで定義) に基づく区切り文字が表示されます。

カテゴリーと合計

「カテゴリーと合計」ダイアログ・ボックスでは、次の処理を実行できます。

- カテゴリーの並べ替えと除外。
- 小計と合計の挿入。
- 計算対象カテゴリーの挿入。
- 空のカテゴリーの組み込みと除外。
- 欠損値の組み込みが定義されたカテゴリーの組み込みと除外。
- 定義済みの値ラベルを持たないカテゴリーの組み込みと除外。

- このダイアログ・ボックスは、カテゴリ変数と多重回答セットについてのみ使用することができます。スケール変数では使用できません。
- カテゴリが異なる複数の変数を選択した場合、小計の挿入、計算対象カテゴリの挿入、カテゴリの除外、手動によるカテゴリの並べ替えを行うことはできません。このような状況が発生するのは、キャンバス・プレビューで複数の変数を選択し、これらすべての変数を対象としてこのダイアログ・ボックスにアクセスした場合だけです。ただし、それぞれの変数に対して、個別にこれらの処理を実行することはできます。
- 定義済みの値ラベルがない変数の場合、カテゴリのソートと合計の挿入のみ、実行することができます。

「カテゴリと合計」ダイアログ・ボックスにアクセスする

1. カテゴリ変数または多重回答セットを、キャンバス・ペインにドラッグ・アンド・ドロップします。
2. キャンバス・ペインで変数を右クリックし、ポップアップ・メニューから「**カテゴリと合計**」を選択します。

または

3. キャンバス・ペインで変数を選択 (クリック) し、「テーブル」タブの「定義」グループで「**カテゴリと合計**」をクリックします。

キャンバス・ペインでは、以下の方法で、同じディメンション内の複数のカテゴリ変数を選択することもできます。

4. Ctrl キーを押しながら、キャンバス・ペインで各変数をクリックします。

または

5. キャンバス・ペインでテーブル・プレビューの外側をクリックし、クリックしたままドラッグして、選択したい変数が含まれている領域を選択します。

または

6. ディメンション内の任意の変数を右クリックし、「**全 [次元] 変数の選択**」を選択して、そのディメンション内の変数をすべて選択します。

カテゴリを並べ替える

手動でカテゴリの並べ替えを行うには、以下の手順を実行します。

1. リスト内のカテゴリを選択 (クリック) します。
2. 上矢印または下矢印をクリックして、カテゴリをリスト内で上下に移動します。

または

3. カテゴリの「値」列をクリックして、別の位置にドラッグ・アンド・ドロップします。

カテゴリを除外する

1. リスト内のカテゴリを選択 (クリック) します。
2. 「除外」リストの横にある矢印をクリックします。

または

3. カテゴリの「値」列をクリックして、リストの外側の任意の場所にドラッグ・アンド・ドロップします。

カテゴリーを除外すると、定義済みの値ラベルがないカテゴリーも除外されます。

カテゴリーをソートする

データ値、値ラベル、セル度数、要約統計量のいずれかを基準として、カテゴリーを昇順または降順にソートすることができます。

1. 「カテゴリーの並べ替え」グループで「並べ替え基準」ドロップダウン・リストをクリックして、使用するソート基準を選択します。ソート基準には、値、ラベル、度数、要約統計量（平均値、中央値、最頻値など）があります。ソートに使用できる要約統計量は、テーブルに表示するように選択した要約統計量によって異なります。
2. 「並び順」ドロップダウン・リストをクリックして、ソートの順序（昇順または降順）を選択します。

除外されているカテゴリーが 1 つでもあると、カテゴリーをソートすることはできません。

小計

1. リスト内で、小計に含めるカテゴリー範囲の最後のカテゴリーを選択（クリック）します。
2. 「小計を追加...」をクリックします。
3. 「小計の定義」ダイアログ・ボックスで、必要に応じて小計のラベル・テキストを変更します。
4. 小計だけを表示し、小計を定義するカテゴリーを非表示にするには、「小計されたカテゴリをテーブルから隠す」を選択します。
5. 「続行」をクリックして、小計を追加します。

合計

1. 「合計」チェック・ボックスをクリックします。合計のラベル・テキストを変更することもできます。

選択した変数が別の変数内でネストされている場合、各サブテーブルに合計が挿入されます。

合計と小計の表示位置

合計と小計は、各合計に含まれるカテゴリーの上または下に表示することができます。

- 「合計と小計の表示」グループで「下」を選択した場合、合計は各サブテーブルの上に表示され、選択したカテゴリーを含め、その上にあるすべてのカテゴリー（ただし、先行する小計の下にあるカテゴリー）が、各小計に含まれます。
- 「合計と小計の表示」グループで「上」を選択した場合、合計は各サブテーブルの下に表示され、選択したカテゴリーを含め、その下にあるすべてのカテゴリー（ただし、先行する小計の上にあるカテゴリー）が、各小計に含まれます。

重要: 小計を定義する前に、小計の表示位置を選択する必要があります。表示位置を変更すると、現在選択されている小計だけでなくすべての小計に影響し、小計に含まれるカテゴリーも変更されます。

計算対象カテゴリー

要約統計量、合計、小計、または定数（あるいはこれらの組み合わせ）から計算されたカテゴリーを表示することができます。詳しくは、14 ページの『計算対象カテゴリー』のトピックを参照してください。

カスタム・テーブルと小計の要約統計量

「要約統計量」ダイアログ・ボックスを使用して、テーブルの合計領域や小計領域に「合計」以外の統計量を表示することができます。詳しくは、7 ページの『カテゴリー変数の要約統計量』のトピックを参照してください。

注: テーブルの本体にも表示されるカスタムの合計統計量を複数選択し、統計ラベルを非表示にすると、テーブル内での順序と同じになるように、合計が再ソートされます。ただし、ラベルは表示されないため、それぞれの合計統計量が実際に何を表しているかがわからない場合があります。そのため、複数の統計量を選択して統計ラベルを非表示にすることは、通常はお勧めしません。

合計、小計、および除外されたカテゴリー

除外されたカテゴリーのケースは、合計の計算には含まれません。

欠損値、空のカテゴリー、および値ラベルのない値

欠損値: これは、**ユーザー欠損値**、または欠損値を含むものとして定義された値 (例えば、男性の妊娠率は「適用外」であることを表すコード 99 など) の表示を制御します。デフォルトでは、ユーザー欠損値は除外されます。テーブルにユーザー欠損カテゴリーを含めるには、このオプションを選択します (チェック・マークを付けます)。変数に複数の欠損値カテゴリーを含めることができますが、キャンバス上のテーブル・プレビューには、1 つの総称的な欠損値カテゴリーだけが表示されます。定義済みユーザー欠損カテゴリーは、すべてテーブルに含まれます。**システム欠損値** (データ・エディター内の数値変数のセル) は、常に除外されます。

空のカテゴリー: 空のカテゴリーとは、定義済みの値ラベルを持っているが、特定のテーブルやサブテーブルについてはケースが存在しないカテゴリーを指します。デフォルトでは、空のカテゴリーはテーブルに含まれます。テーブルから欠損カテゴリーを除外するには、このオプションの選択を解除します。

データのスキャン時にその他の値: デフォルトでは、定義済み値ラベルがないデータ・ファイル内のカテゴリー値は、自動的にテーブルに含まれます。定義済み値ラベルがない値をテーブルから除外するには、このオプションの選択を解除します。定義済み値ラベルを持つカテゴリーを除外すると、定義済み値ラベルがないカテゴリーも除外されます。

計算対象カテゴリー

テーブルには、要約統計量の集計結果を表示できるだけでなく、これらの集計結果、定数値、小計と合計、またはこれらの組み合わせから計算された 1 つ以上のカテゴリーを表示することもできます。これらの計算結果は、計算対象カテゴリーまたは計算後と呼ばれます。計算対象カテゴリーは、単一変数のカテゴリーと同様に機能しますが、以下に示す類似点と相違点があります。

- 計算対象カテゴリーは、他のカテゴリーと同様に配置されます。
- 計算対象カテゴリーは、他のカテゴリーと同じ統計上で機能します。
- 計算対象カテゴリーは、小計、合計、有意差検定には影響しません。
- 計算対象カテゴリーの値は、デフォルトで、他のカテゴリーの要約統計量と同じ書式を使用します。この書式は、計算対象カテゴリーを定義する際に上書きすることができます。

計算対象カテゴリーを使用して集計結果を合計できるため、小計と同じように計算対象カテゴリーを使用することができます。ただし、計算対象カテゴリーは、以下の点で小計よりも便利です。

- 計算対象カテゴリーは、他の小計の結果から計算することができます。
- 計算対象カテゴリーは、相互にオーバーラップさせ、同じカテゴリー (または同じカテゴリーの一部) で機能させることができます。
- 計算対象カテゴリーには、計算対象カテゴリーの上下にある他のすべてのカテゴリーの値を含める必要はありません。つまり、計算対象カテゴリーは包括的ではありません。
- 計算対象カテゴリーには、隣接していないカテゴリーの値を含めることができます。

合計や小計とは異なり、計算対象カテゴリーは、元のデータではなく、集計したデータから計算されます。したがって、計算対象カテゴリーの値は、合計と小計の結果とは一致しない場合があります。また、計算対象カテゴリーを定義する際に、ソース・カテゴリーを非表示にするオプションを指定できるため、生成されたテーブルで小計の内容を解釈するのが難しくなる場合があります。計算対象カテゴリーを使用する場合は、小計に対してカスタム・ラベルを指定することをお勧めします。

計算対象カテゴリーを定義する

計算対象カテゴリーは、「カテゴリと合計」ダイアログ・ボックスから追加します。このダイアログ・ボックスにアクセスする方法については、11 ページの『カテゴリと合計』のトピックを参照してください。

1. 「カテゴリと合計」ダイアログ・ボックスで、「**カテゴリを追加...**」をクリックします。
2. 「**計算対象カテゴリーのラベル**」で、計算対象カテゴリーのラベルを指定します。「カテゴリー」リストから該当するカテゴリーをドラッグして、そのカテゴリーのラベルを含めることができます。
3. カテゴリーまたは合計と小計 (あるいはその両方) を選択し、演算子を使用して式を作成することにより、計算対象カテゴリーを定義します。500 などの定数値を入力して式に含めることもできます。
4. 計算対象カテゴリーだけを表示し、その計算対象カテゴリーを定義するカテゴリーを非表示にするには、「**式で使用されているカテゴリをテーブルから隠す**」を選択します。
5. 「**表示形式**」タブをクリックして、計算対象カテゴリーの表示形式と小数の桁数を変更します。詳しくは、『計算対象カテゴリーの表示形式』のトピックを参照してください。
6. 「**続行**」をクリックして、計算対象カテゴリーを追加します。

計算対象カテゴリーの表示形式

デフォルトでは、計算対象カテゴリーは、変数の他のカテゴリーと同じ表示形式および小数桁数を使用します。これらの設定は、「計算対象カテゴリー」ダイアログ・ボックスの「表示形式」タブで指定変更できます。「表示形式」タブには、計算対象カテゴリーが処理する現在の要約統計量と、これらの要約統計量の表示形式および小数桁数がリストされます。

要約統計量ごとに、以下の操作を実行できます。

1. 「書式」ドロップダウン・リストから、要約統計量の表示形式を選択します。表示形式の完全なリストについては、11 ページの『要約統計量の表示形式』のトピックを参照してください。
2. 選択した要約統計量の「小数部の桁数」セルに、表示する小数点以下の桁数を入力します。

カテゴリーを共有する変数のテーブル (比較テーブル)

多くの場合、アンケート調査の多数の質問には、予想される回答項目の共通セットが用意されています。積み重ねを使用すれば、これらの関連する変数を同じテーブルに表示し、共有される応答カテゴリーをテーブルの列に表示できます。

カテゴリーを共有する複数の変数のテーブルを作成する

表 6. 列に積み重ねられた共有応答カテゴリーのある変数

変数	カテゴリー 1	カテゴリー 2	カテゴリー 3
変数 1	12	34	56
変数 2	56	12	34
変数 3	34	56	12

詳しくは、49 ページの『第 6 章 カテゴリーを共有する変数のテーブル』のトピックを参照してください。

テーブル・ビルダーのカスタマイズ

標準的なダイアログ・ボックスとは異なり、テーブル・ビルダーのサイズは、標準的なウィンドウと同じ方法で変更できます。

1. テーブル・ビルダーのサイズを大きくするか、小さくするには、テーブル・ビルダーの上下左右の端または角をクリックしたままドラッグします。

「テーブル」タブでは、変数リスト、「カテゴリー」リスト、およびキャンバス・ペインのサイズを変更することもできます。

2. 変数リストと「カテゴリー」リストの間にある水平バーをクリックしたままドラッグすると、リストが長くなったり、短くなったりします。下に動かすと、変数リストが長くなり、「カテゴリー」リストが短くなります。上に動かすと、逆の結果になります。
3. 変数リストと「カテゴリー」リストをキャンバス・ペインから区切っている垂直バーをクリックしたままドラッグすると、リストの幅が広くなったり、狭くなったりします。キャンバスは、残りのスペースを埋めるように自動的にサイズ変更されます。

カスタム・テーブル: 「オプション」タブ

「オプション」タブでは、次の処理を実行できます。

- 空のセルと統計量を計算できないセルに表示する内容を指定する。
- スケール変数統計量の計算で欠損値を処理する方法を制御する。
- データ列幅の最小値と最大値を設定する。
- 多重カテゴリー・セットに含まれる重複回答の処理を制御する。

データ・セルの表示: 空のセルと統計量を計算できないセルに表示する内容を制御します。

- **空のセル:** ケースが含まれていないテーブル・セル (セルの度数が 0) について、3 つの表示オプション (0、空白、指定するテキスト値) のいずれかを選択することができます。テキスト値の最大文字数は 255 文字です。
- **計算できない統計量:** 統計量を計算できない場合 (例えば、ケースのないカテゴリーの平均値など) に表示されるテキスト。テキスト値の最大文字数は 255 文字です。デフォルト値はピリオド (.) です。

データ列の幅: データ列の最小幅と最大幅を制御します。この設定は、行ラベルの列幅には影響しません。

- **テーブルルックの設定:** 現在のデフォルトのテーブルルックに指定されたデータ列幅を使用します。新規テーブルの作成時に使用する独自のカスタム・デフォルト・テーブルルックを作成したり、テーブルルックで行ラベル列幅とデータ列幅の両方を制御したりすることができます。
- **カスタム:** データ列幅に対するデフォルト・テーブルルックの設定を上書きします。テーブルの最小データ列幅と最大データ列幅を指定し、測定の単位 (ポイント、インチ、またはセンチメートル) を指定してください。

スケール変数の欠損値: 複数のスケール変数を持つテーブルについて、スケール変数統計量の欠落データの処理を制御します。

- **使用可能なデータを最大限に使用 (変数ごとの削除):** 各スケール変数に対して有効な値を持つすべてのケースが、そのスケール変数の要約統計量に含まれます。
- **スケール変数間で一貫したケース基準を使う (リストワイズ除去):** テーブル内のいずれかのスケール変数について欠損値があるケースは、テーブル内のすべてのスケール変数の要約統計量から除外されます。

多重カテゴリ グループに対して重複した回答を数える: 重複回答とは、多重カテゴリ・セット内の複数の変数に対する同じ回答のことです。デフォルトでは、重複回答はカウントされませんが、重複回答が、重複回答をカウントに含めたい場合もあります (アンケート調査の回答者が最後に購入した 3 台の車のメーカーを表す多重カテゴリ・セットの場合など)。

小さい度数を非表示にする: ここで指定した整数よりも小さな度数を非表示にすることができます。非表示の値は、 $<N$ として表示されます (N は、指定する整数)。2 以上の整数を指定する必要があります。

カスタム・テーブル: 「表題」タブ

「表題」タブでは、タイトル、キャプション、およびコーナー・ラベルの表示を制御します。

タイトル: テーブルの上に表示されるテキストです。

キャプション: テーブルの下の脚注の上に表示されるテキストです。

コーナー: テーブルの左上コーナーに表示されるテキストです。コーナー・テキストが表示されるのは、テーブルに行変数が含まれる場合、およびピボット・テーブルの行ディメンションのラベル・プロパティが「入れ子」に設定されている場合のみです。これは、デフォルトのテーブルルックの設定ではありません。

テーブルのタイトル、キャプション、またはコーナー・ラベルには、以下の自動的に生成される値を組み込むことができます。

日付: Windows の現在の「地域のオプション」設定に基づく形式で表示される現在の年、月、日付です。

時刻: Windows の現在の「地域のオプション」設定に基づく形式で表示される現在の時間、分、秒です。

テーブル式: テーブルで使用される変数であり、テーブルでのこれらの変数の使用方法です。変数に定義済みの変数ラベルがある場合、そのラベルが表示されます。生成されるテーブルでは、以下の記号が、テーブルで変数がどのように使用されているかを示します。

- 「+」は、変数が積み重ねられていることを示します。
- 「>」は、ネストしていることを示します。
- 「BY」は、クロス集計または層を示します。

カスタム・テーブル: 「検定統計量」タブ

「検定統計量」タブでは、カスタム・テーブルに対して、以下のような各種の有意差検定を要求することができます。

- 独立性のカイ 2 乗検定。
- 列平均の等質性検定。
- 列比率の等質性検定。
- 多重回答セットと小計の有意差検定。(多重回答セットの有意差検定については、84 ページの『多重回答セットでの有意差検定』を参照してください)。

カテゴリ ラベルがデフォルトのテーブル ディメンションから移動されたテーブルおよび計算対象カテゴリでは、これらの検定を使用することはできません。

列の平均値を比較 (t 検定): このオプションを選択すると、列に 1 つ以上のカテゴリ変数があり、行に 1 つ以上のスケール変数があるテーブルについて、列の平均値の等質性がペアごとに検定されます。テーブルには、要約統計量として平均値が含まれている必要があります。Bonferroni 法を使用して検定の p 値を

調整するかどうかを選択できます。また、検定のアルファ・レベルも指定することができます。これは、0 よりも大きな 1 未満の値でなければなりません。さらに、多重回答検定の場合、平均値の検定の分散は、常に比較されるカテゴリーだけに基づくのに対し、通常のカテゴリー変数の場合は、比較されるカテゴリーだけを使用して推定することも、すべてのカテゴリーから推定することもできます。

列の比率を比較 (z 検定): このオプションを選択すると、列と行の両方に 1 つ以上のカテゴリー変数があるテーブルで、列の比率の等質性がペアごとに検定されます。テーブルには、度数または単純な列パーセントが含まれている必要があります。Bonferroni 法を使用して検定の p 値を調整するかどうかを選択できます。また、検定のアルファ・レベルも指定できます。これは、0 よりも大きな 1 未満の値でなければなりません。

有意差の特定: 「列の平均値を比較」または「列の割合を比較」を選択した場合、有意差をどのように示すかを選択することができます。

- **個別テーブル:** 有意差検定の結果は、個別テーブルに表示されます。2 つの値が大きく異なる場合、大きいほうの値に対応するセルには、小さいほうの値の列を示すキーが表示されます。以下に例を示します。
- **APA スタイルのサブスクリプトを使用するメイン テーブル:** メイン・テーブル自体が、下付き文字による APA スタイル形式を使用して有意差を特定します。2 つの値が大きく異なる場合、これらの値には別の下付き文字が表示されます。これらの下付き文字は、脚注ではありません。このオプションを有効にすると、現在のテーブルルックで定義されている脚注のスタイルが上書きされ、上付き文字の番号として表示されます。以下に例を示します。

これらのテーブルを作成する方法と解釈する方法を具体的に説明する詳しい例については、72 ページの『列の比率の比較』を参照してください。

注: APA スタイルのテーブルには、テーブル内の下付き文字の解釈方法を説明するキャプションが組み込まれます。「表題」タブでキャプションを指定すると、下付き文字を説明するキャプションの代わりに、「表題」タブで指定されたキャプションが表示されます。

独立性の検定 (カイ 2 乗): このオプションを選択すると、行と列の両方に 1 つ以上のカテゴリー変数があるテーブルで、独立性のカイ 2 乗検定が実行されます。また、検定のアルファ・レベルも指定できます。これは、0 よりも大きな 1 未満の値でなければなりません。

小計カテゴリーの代わりに小計を使用: このオプションを選択すると、小計のカテゴリーではなく、それぞれの小計が有意差検定で使用されます。このオプションを選択しなかった場合は、小計カテゴリーが非表示になっている小計だけが、小計のカテゴリーに代わって検定で使用されます。

検定に多重回答変数を含める: このオプションを選択すると、多重回答セットのカテゴリーを使用して検定が実行されます。このオプションを選択しなかった場合は、検定の実行時に多重回答セットが無視されます。

第 2 章 カテゴリー変数の単純なテーブル












カテゴリー変数の単純なテーブル

作成するほとんどのテーブルには、おそらく少なくとも 1 つの**カテゴリー変数** が含まれることとなります。カテゴリー変数とは、限られた数の異なる数値またはカテゴリーを持つ変数のことです (例えば、性別や宗教)。カテゴリー変数は、**名義変数**または**順序変数**のいずれかです。

- *Nominal* (名義データ). 本質的な順位を持たないカテゴリーを表す値である場合 (従業員の勤務先企業での部署など)、変数を名義変数として取り扱うことができます。名義変数の例としては、地域、郵便番号、宗教上の所属などが挙げられます。
- *Ordinal* (順序データ). 値が本質的な順位を持つカテゴリーを表す場合 (例えば、サービス満足度のレベルを「非常に不満」から「非常に満足」までの順位で評価する場合) は、変数を順序変数として扱うことができます。順序変数の例としては、満足度や信頼度を表す得点や嗜好得点などが挙げられます。

変数リストで各変数の隣にあるアイコンは、変数の型を識別します。

表 7. 測定レベル・アイコン

	数値	文字列	日付	時刻
スケール (連続)		なし		
順序				
名義				

カスタム・テーブルは、定義済みの**値ラベル**を持つカテゴリー変数用に最適化されています。詳しくは、1 ページの『テーブルの作成』のトピックを参照してください。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス (「編集」メニューから「オプション」を選択) の「全般」タブで設定します。

単一のカテゴリー変数

単一のカテゴリー変数のテーブルは、簡単に作成できる最も単純なテーブルの 1 つですが、ほとんどの場合はこのテーブルで十分です。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。

キャンバス・ペインにテーブルのプレビューが表示されます。プレビューには実際のデータ値は表示されません。データが表示される場所に、プレースホルダーが表示されるだけです。

3. 「OK」をクリックして、テーブルを作成します。

「ビューアー」ウィンドウにテーブルが表示されます。

		度数
年齢カテゴリー	25歳未満	242
	25 - 34	627
	35 - 44	679
	45 - 54	481
	55 - 64	320
	65歳以上	479

図 2. 行にある単一のカテゴリー変数

この単純なテーブルでは、「度数」という列見出しは実際のところ必要ないので、この列見出しのないテーブルを作成できます。

4. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
5. 「要約統計量」グループで、「位置」に対して「非表示」を選択（クリック）します。
6. 「OK」をクリックして、テーブルを作成します。

年齢カテゴリー	25歳未満	242
	25 - 34	627
	35 - 44	679
	45 - 54	481
	55 - 64	320
	65歳以上	479

図 3. 要約統計量の列ラベルがない単一のカテゴリー変数

パーセント

度数に加え、パーセントを表示することもできます。単一のカテゴリー変数からなる単純なテーブルでは、変数を行に表示する場合、おそらく列パーセントを知りたいと思うはずですが。逆に、変数を列に表示する場合には、行パーセントを知りたいと思うはずですが。

1. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 「要約統計量」グループで、「位置」に対する「非表示」を選択解除します。このテーブルには 2 つの列があるため、各列が何を表しているのかがわかるように列ラベルを表示します。
3. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
4. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。

- 「表示」リストの「ラベル」セルで、デフォルトのラベルを削除して「パーセント」と入力します。
- 「**選択項目に適用**」をクリックし、テーブル・ビルダーで「**OK**」をクリックしてテーブルを作成します。

		度数	パーセント
年齢カ	25歳未満	242	8.6%
テコリ	25 - 34	627	22.2%
	35 - 44	679	24.0%
	45 - 54	481	17.0%
	55 - 64	320	11.3%
	65歳以上	479	16.9%

図4. 度数と列パーセント

合計

合計は、カスタム・テーブルに自動的に組み込まれませんが、テーブルに合計を追加するのは簡単です。

- テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
- キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「**カテゴリと合計**」を選択します。
- 「カテゴリと合計」ダイアログ・ボックスで、「**合計**」を選択（クリック）します。
- 「**適用**」をクリックし、テーブル・ビルダーで「**OK**」をクリックしてテーブルを作成します。

		度数	パーセント
年齢カ	25歳未満	242	8.6%
テコリ	25 - 34	627	22.2%
	35 - 44	679	24.0%
	45 - 54	481	17.0%
	55 - 64	320	11.3%
	65歳以上	479	16.9%
	合計	2828	100.0%

図5. 度数、列パーセント、および合計

詳しくは、35ページの『第4章 カテゴリー変数の合計と小計』のトピックを参照してください。

クロス集計

クロス集計表は、2つのカテゴリー変数の間にある関係を調べる基本的な手法です。例えば、「年齢カテゴリー」を行変数、「性別」を列変数として使用して、各年齢カテゴリーに属する男性と女性の数を示す2次元クロス集計表を作成できます。

- テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
- 「戻す」をクリックして、テーブル・ビルダーでの前の選択項目を削除します。
- テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。

- 変数リストからキャンバス・ペインの列領域に「性別」をドラッグ・アンド・ドロップします。(この変数を見つけるために、変数リストをスクロールダウンしなければならない場合もあります)。
- 「OK」をクリックして、テーブルを作成します。

		性別	
		男性	女性
		度数	度数
年齢カ	25歳未満	108	134
テコリ	25 - 34	276	351
	35 - 44	309	370
	45 - 54	221	260
	55 - 64	136	184
	65歳以上	178	301

図 6. 年齢カテゴリーと性別のクロス集計表

クロス集計表のパーセント

2次元のクロス集計表では、行パーセントと列パーセントの両方が、有用な情報を提供する場合があります。

- テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
- キャンバス・ペインで「性別」をクリックします。

ポップアップ・メニューで「要約統計量」が無効になっていることがわかります。これは、要約統計量は、統計ソース・ディメンションで最も深いレベルにある変数に対してのみ選択できるためです。カテゴリー変数のデフォルトの統計ソース・ディメンション (行または列) は、変数をキャンバス・ペインにドラッグ・アンド・ドロップする順序に基づきます。この例では、最初に「年齢カテゴリー」を行ディメンションにドラッグしています。行ディメンションには他の変数がないため「年齢カテゴリー」が統計量ソース変数となっています。統計ソース・ディメンションを変更することはできますが、この例の場合、その必要はありません。詳しくは、6ページの『要約統計量』のトピックを参照してください。

- キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
- 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
- 「統計量」リストで「行 N %」を選択し、矢印をクリックして「表示」リストに追加します。
- 「選択項目に適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

		性別					
		男性			女性		
		度数	列%	行%	度数	列%	行%
年齢カ テゴリ	25歳未満	108	8.8%	44.6%	134	8.4%	55.4%
	25 - 34	276	22.5%	44.0%	351	21.9%	56.0%
	35 - 44	309	25.2%	45.5%	370	23.1%	54.5%
	45 - 54	221	18.0%	45.9%	260	16.3%	54.1%
	55 - 64	136	11.1%	42.5%	184	11.5%	57.5%
	65歳以上	178	14.5%	37.2%	301	18.8%	62.8%

図7. 行パーセントと列パーセントを表示するクロス集計表

表示形式の制御

要約統計量で表示する小数桁数などの表示形式を制御することができます。例えば、デフォルトでは、パーセントは小数点以下 1 桁とパーセント記号で表示されます。ただし、セルの値を小数点以下 2 桁で表示し、パーセント記号を表示しないようにするとします。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
3. 要約統計量から選択した 2 つのパーセント (「列 N %」と「行 N %」) の両方について、「書式」ドロップダウン・リストから「nnnn.n」を選択し、「小数部の桁数」セルに「2」と入力します。
4. 「OK」をクリックして、テーブルを作成します。

		性別					
		男性			女性		
		度数	列%	行%	度数	列%	行%
年齢カ テゴリ	25歳未満	108	8.79%	44.63%	134	8.38%	55.37%
	25 - 34	276	22.48%	44.02%	351	21.94%	55.98%
	35 - 44	309	25.16%	45.51%	370	23.13%	54.49%
	45 - 54	221	18.00%	45.95%	260	16.25%	54.05%
	55 - 64	136	11.07%	42.50%	184	11.50%	57.50%
	65歳以上	178	14.50%	37.16%	301	18.81%	62.84%

図8. 行と列のパーセントに対して書式設定したセルの表示

周辺合計

クロス集計表では、各行や各列の合計である**周辺合計**を表示するのがごく一般的です。デフォルトでは、周辺合計はカスタム・テーブルに組み込まれないため、明示的にテーブルに追加する必要があります。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 「戻す」をクリックして、テーブル・ビルダーでの前の選択項目を削除します。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
4. 変数リストからキャンバス・ペインの列領域に「性別」をドラッグ・アンド・ドロップします。(この変数を見つけるために、変数リストをスクロールダウンしなければならない場合もあります)。

5. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
6. 「カテゴリと合計」ダイアログ・ボックスで、「合計」を選択 (クリック) し、「適用」をクリックします。
7. キャンバス・ペインで「性別」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
8. 「カテゴリと合計」ダイアログ・ボックスで、「合計」を選択 (クリック) し、「適用」をクリックします。
9. 「要約統計量」グループで、「位置」に対して「非表示」を選択 (クリック) します。度数だけを表示しているため、テーブルのデータ・セルに表示する「統計量」を識別する必要はありません。
10. 「OK」をクリックして、テーブルを作成します。

		性別		
		男性	女性	合計
年齢カ テコリ	25歳未満	108	134	242
	25 - 34	276	351	627
	35 - 44	309	370	679
	45 - 54	221	260	481
	55 - 64	136	184	320
	65歳以上	178	301	479
	合計	1228	1600	2828

図9. 周辺合計を表示するクロス集計表

カテゴリーのソートと除外

デフォルトでは、カテゴリーは、カテゴリーの値ラベルが表すデータ値の昇順で表示されます。例えば、年齢カテゴリーとして「Less than 25」、「25 - 34」、「35 - 44」などの値ラベルが表示されていても、実際の基礎となるデータ値は 1、2、3 などであり、これらの基礎となるデータ値によって、カテゴリーのデフォルトの表示順序が制御されます。

カテゴリーの順序は簡単に変更できます。また、テーブルに表示したくないカテゴリーを除外することもできます。

カテゴリーのソート

次の項目について、カテゴリーを手動で再配置することも、カテゴリーを昇順や降順でソートすることもできます。

- データ値。
 - 値ラベル。
 - セル度数。
 - 要約統計量。ソートに使用できる要約統計量は、テーブルに表示するように選択した要約統計量によって異なります。
1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
 2. キャンバス・ペインの行領域に「年齢カテゴリー」が表示されていない場合は、このカテゴリーを行領域にドラッグ・アンド・ドロップします。

3. キャンパス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。

データ値とそれに関連付けられた値ラベルは、現在のソート順で表示されます。この例では、ソート順はデータ値の昇順のままになっています。

4. 「カテゴリのソート」グループで、「順序」ドロップダウン・リストから「降順」を選択します。

これで、ソート順が逆になります。

5. 「基準」ドロップダウン・リストから「ラベル」を選択します。

これで、カテゴリは値ラベルの降順のアルファベット順でソートされます。

この例では、「*Less than 25*」というラベルの付いたカテゴリがリストの先頭に表示されることに注意してください。アルファベット順のソートでは、文字よりも数字が先になります。このラベルは文字で始まる唯一のラベルであり、リストは降順（逆向き）にソートされるため、このカテゴリがリストの先頭に表示されます。

特定のカテゴリをリスト内の別の場所に表示させたい場合は、そのカテゴリを簡単に移動することができます。

6. 「ラベル」リストで「*Less than 25*」のラベルが付いたカテゴリをクリックします。
7. リストの右側にある下矢印をクリックします。カテゴリが、リストの 1 行下に移動します。
8. カテゴリがリストの最後にくるまで、下矢印をクリックし続けます。

カテゴリの除外

テーブル内に表示したくないカテゴリがある場合は、そのカテゴリを除外することができます。

1. 「ラベル」リストで「*Less than 25*」のラベルが付いたカテゴリをクリックします。
2. 「除外」リストの左側にある矢印キーをクリックします。
3. 「ラベル」リストで「*65 or older*」のラベルが付いたカテゴリをクリックします。
4. 「除外」リストの左側にある矢印キーをもう一度クリックします。

2 つのカテゴリが「表示」リストから「除外」リストに移動されます。元に戻したい場合は、簡単に「表示」リストに戻すことができます。

5. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

		性別		
		男性	女性	合計
年齢カゴリ	55 - 64	136	184	320
	45 - 54	221	260	481
	35 - 44	309	370	679
	25 - 34	276	351	627
	合計	942	1165	2107

図 10. 値ラベルの降順でソートされ、一部のカテゴリが除外されているテーブル

2 つのカテゴリを除外する前よりも、合計の値が小さくなっていることに注意してください。これは、テーブルに含まれるカテゴリによって合計が計算されるためです。除外されたカテゴリは、合計の計算からも除外されます。詳しくは、35 ページの『第 4 章 カテゴリ変数の合計と小計』のトピックを参照してください。

第 3 章 カテゴリー変数の積み重ね、ネスト、層

積み重ね、ネスト、層は、いずれも複数の変数を同じテーブルに表示するための方法です。この章では、これらの手法をカテゴリー変数で使用する場合を中心に説明しますが、スケール変数でも同じ手法を使用できます。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「全般」タブで設定します。

カテゴリー変数の積み重ね

積み重ねは、個別の複数のテーブルをまとめて同じ表示に貼り付けることとして考えられます。例えば、「性別」の情報と「年齢カテゴリー」の情報を、同じテーブルの別々のセクションに表示することができます。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「性別」をドラッグ・アンド・ドロップします。
3. 変数リストの「年齢カテゴリー」を、行領域の「性別」の下にドラッグ・アンド・ドロップします。

これで、2 つの変数が行ディメンションに積み重ねられます。

4. 「OK」をクリックして、テーブルを作成します。

		度数
性別	男性	1232
	女性	1600
年齢カテゴリー	25歳未満	242
	25 - 34	627
	35 - 44	679
	45 - 54	481
	55 - 64	320
	65歳以上	479

図 11. 行にカテゴリー変数を積み重ねたテーブル

同じようにして、列に変数を積み重ねることもできます。

クロス集計を使用した積み重ね

積み重ねテーブルには、別のディメンションに異なる変数を含めることができます。例えば、行に積み重ねた 2 つの変数を、列ディメンションに表示される 3 番目の変数でクロス集計することができます。

1. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 行で「年齢カテゴリー」と「性別」をまだ積み重ねていない場合は、前の手順の説明に従って積み重ねます。
3. 変数リストからキャンバス・ペインの列領域に「インターネットから情報を得る」をドラッグ・アンド・ドロップします。
4. 「OK」をクリックして、テーブルを作成します。

		インターネットから情報を得る	
		いいえ	はい
		度数	度数
性別	男性	873	359
	女性	1092	508
年齢カテゴリー	25歳未満	146	96
	25 - 34	368	259
	35 - 44	435	244
	45 - 54	346	135
	55 - 64	252	68
	65歳以上	416	63

図 12. 列変数でクロス集計された 2 つの積み重ねられた行変数

注: 「... から情報を得る」というラベルを持つ変数はいくつかあるため、変数リストでこれらの変数を区別するのが難しい場合があります（ラベルの幅が広すぎて、変数リストに完全に表示されないためです）。変数ラベル全体を表示するには、次の 2 つの方法があります。

- リスト内の変数にマウス・ポインターを合わせて、ポップアップ・ツールチップにラベル全体を表示します。
- 変数と「カテゴリー」リストをキャンバス・ペインから区切っている縦線をクリックしたままドラッグして、リストの幅を広げます。

カテゴリー変数のネスト

クロス集計表と同様に、ネストでも 2 つのカテゴリー変数の間にある関係を示すことができます。ただし、同じディメンションで別の変数にネストされている変数は除きます。例えば、行ディメンションで「年齢カテゴリー」の中に「性別」をネストして、各年齢カテゴリーに属する男性と女性の数を表示できます。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 「戻す」をクリックして、テーブル・ビルダーでの前の選択項目を削除します。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
4. 変数リストの「性別」を、行領域の「年齢カテゴリー」の右側にドラッグ・アンド・ドロップします。

キャンバス・ペインのプレビューから、このネストされたテーブルには度数の列が 1 つあり、その列の各セルに、年齢カテゴリーごとの男性の数と女性の数が示されることがわかります。

「性別」という変数ラベルが、それぞれの年齢カテゴリーに繰り返し表示されていることに気付くはずですが。このような繰り返しを最小限に抑えるには、カテゴリー数が最も少ない変数を、ネストの最も浅いレベルに配置します。

5. キャンバス・ペインで、変数ラベル「性別」をクリックします。

6. 変数を行領域の左端にドラッグ・アンド・ドロップします。

これで、「性別」が 6 回繰り返される代わりに、「年齢カテゴリー」が 2 回繰り返されることとなります。これで、テーブルは基本的に同じ結果を生成しながらも、より見やすくなります。

7. 「OK」をクリックして、テーブルを作成します。

				度数
性別	男性	年齢カテゴリー	25歳未満	108
			25 - 34	276
		35 - 44	309	
		45 - 54	221	
		55 - 64	136	
		65歳以上	178	
	女性	年齢カテゴリー	25歳未満	134
			25 - 34	351
		35 - 44	370	
		45 - 54	260	
		55 - 64	184	
		65歳以上	301	

図 13. 「性別」の中に「年齢カテゴリー」をネストしたテーブル

注: カスタム・テーブルは、層化された分割ファイルの処理は行いません。層化分割ファイルと同じ結果を得るには、分割ファイル変数を、テーブルの最も浅いレベルにネストされた層に配置します。

変数ラベルの抑止

ネストしたテーブルでの変数ラベルの重複に対処するもう 1 つの方法は、単純に変数名またはラベルの表示を抑止することです。「性別」と「年齢カテゴリー」は、どちらも変数ラベルがなくても値ラベルだけで十分理解できるため、この両方の変数についてはラベルを除去できます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューにある「変数ラベルを表示」を選択解除します。
3. 「性別」についても同じ操作を行います。

テーブル・プレビューでは変数ラベルが表示されたままになっていますが、テーブルには変数ラベルが表示されなくなります。

4. 「OK」をクリックして、テーブルを作成します。

		度数
男性	25歳未満	108
	25 - 34	276
	35 - 44	309
	45 - 54	221
	55 - 64	136
	65歳以上	178
女性	25歳未満	134
	25 - 34	351
	35 - 44	370
	45 - 54	260
	55 - 64	184
	65歳以上	301

図 14. 変数ラベルがない状態のネストしたテーブル

テーブル本体に変数ラベルを何度も表示することなく、テーブルのどこかに変数ラベルを残したい場合は、テーブル・タイトルまたはコーナー・ラベルに含めることができます。

5. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
6. 「表題」タブをクリックします。
7. 「タイトル」テキスト・ボックス内の任意の場所をクリックします。
8. 「テーブル式」をクリックします。「タイトル」テキスト・ボックスに `&[テーブル式]` というテキストが表示されます。これで、テーブル・タイトルに、テーブル内で使用される変数の変数ラベルが含まれるテーブル・タイトルが生成されることとなります。
9. 「OK」をクリックして、テーブルを作成します。

性別 > 年齢カテゴリ

		度数
男性	25歳未満	108
	25 - 34	276
	35 - 44	309
	45 - 54	221
	55 - 64	136
	65歳以上	178
女性	25歳未満	134
	25 - 34	351
	35 - 44	370
	45 - 54	260
	55 - 64	184
	65歳以上	301

図 15. テーブル・タイトル内の変数ラベル

タイトル内にある、より大きい記号 (>) は、「年齢カテゴリ」が「性別」の中にネストされていることを示します。

ネストしたクロス集計表

ネストしたテーブルでは、別のディメンションに異なる変数を含めることができます。例えば、行で「年齢カテゴリ」を「性別」の中にネストし、このネストした行を、列ディメンションにある 3 番目の変数でクロス集計することができます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。

- 行で「年齢カテゴリー」をまだ「性別」の中にネストしていない場合は、前の手順の説明に従ってネストします。
- 変数リストからキャンバス・ペインの列領域に「インターネットから情報を得る」をドラッグ・アンド・ドロップします。

テーブルが大きすぎて、キャンバス・ペインに完全に表示できないことに気付くはずですが、テーブル・プレビューの表示されていない部分を見るには、キャンバス・ペインで上下左右にスクロールするか、あるいは以下の方法を使用できます。

- テーブル・ビルダーで「コンパクト」をクリックして、コンパクト・ビューを表示します。この場合、変数ラベルだけが表示され、テーブルに含まれるカテゴリーや要約統計量の情報は表示されません。
- テーブル・ビルダーの上下左右の端または角をクリックしたままドラッグして、テーブル・ビルダーのサイズを大きくします。

- 「OK」をクリックして、テーブルを作成します。

				インターネットから情報を得る	
				いいえ	はい
				度数	度数
性別	男性	年齢カテゴリー	25歳未満	59	49
			25 - 34	159	117
			35 - 44	217	92
			45 - 54	169	52
			55 - 64	112	24
		65歳以上	155	23	
	女性	年齢カテゴリー	25歳未満	87	47
			25 - 34	209	142
			35 - 44	218	152
			45 - 54	177	83
			55 - 64	140	44
		65歳以上	261	40	

図 16. ネストしたクロス集計表

行と列の交換

長い時間をかけて複雑なテーブルをセットアップした後、テーブルの向きを逆にして、すべての行変数を列に移動し、すべての列変数を行に移動しなければ、完璧なテーブルにはならないと判断した場合を考えてみてください。例えば、行に「年齢カテゴリー」と「性別」カテゴリーをネストしたクロス集計表を作成した後、この 2 つの人口統計的変数を代わりに列にネストするとします。

- テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
- キャンバス・ペインの任意の場所を右クリックし、ポップアップ・メニューから「行と列を入れ替える」を選択します。

これで、行変数と列変数が入れ替わりました。

テーブルを作成する前に、多少の変更を行って、表示の複雑さを軽減します。

- 「非表示」を選択して、要約統計量の列ラベルの表示を抑制します。
- キャンバス・ペインで「性別」を右クリックし、「変数ラベルを表示」を選択解除します。

5. 「OK」をクリックして、テーブルを作成します。

		男性						女性					
		年齢カテゴリ						年齢カテゴリ					
		25歳未満	25 - 34	35 - 44	45 - 54	55 - 64	65歳以上	25歳未満	25 - 34	35 - 44	45 - 54	55 - 64	65歳以上
インターネットから情報を得る	いいえ	59	159	217	169	112	155	87	209	218	177	140	261
	はい	49	117	92	52	24	23	47	142	152	83	44	40

図 17. 列で人口統計的変数をネストしたクロス集計表

層

層を使用してテーブルに深さのディメンションを追加し、3次元の「立体」を作成することができます。実際に、層はネストや積み重ねと非常に似ています。主な違いは、層カテゴリーは一度に1つしか表示できないことです。例えば、「年齢カテゴリー」を行変数、「性別」を層変数として使用すると、男性に関する情報と女性に関する情報をそれぞれ異なる層に表示するテーブルが作成されます。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 「戻す」をクリックして、テーブル・ビルダーでの前の選択項目を削除します。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
4. テーブル・ビルダーの上部にある「層」をクリックして、「層」リストを表示します。
5. 変数リストから「層」リストに「性別」をドラッグ・アンド・ドロップします。

この時点で、層変数を追加しても、キャンバス・ペインに表示されるプレビューには視覚的な影響がないことがわかります。層変数が統計量ソース変数で、要約統計量を変更しない限り、層変数がキャンバス・ペインのプレビューに影響を与えることはありません。

6. 「OK」をクリックして、テーブルを作成します。

性別 男性

年齢カテゴリ	度数
25歳未満	108
25 - 34	276
35 - 44	309
45 - 54	221
55 - 64	136
65歳以上	178

図 18. 層を使用する単純なテーブル

一見したところ、このテーブルは、単一のカテゴリー変数を使用する単純なテーブルと変わらないように見えます。唯一の違いは、テーブルの先頭に「性別 男性」というラベルが表示されていることです。

7. 「ビューアー」ウィンドウ内のテーブルをダブルクリックしてアクティブにします。
8. 「性別 男性」というラベルが、実際にはドロップダウン・リストの選択項目であることがわかります。

9. ドロップダウン・リストの下矢印をクリックして、層のリスト全体を表示します。

このテーブルでは、リスト内にもう 1 つの選択項目しかありません。

10. ドロップダウン・リストから「性別 女性」を選択します。

性別 女性		度数
年齢カテゴリー	25歳未満	134
	25 - 34	351
	35 - 44	370
	45 - 54	260
	55 - 64	184
	65歳以上	301

図 19. 単純な層を使用するテーブルでの異なる層の表示

2 つのカテゴリー層変数の積み重ね

層に複数のカテゴリー変数がある場合、層変数を積み重ねるか、ネストすることができます。デフォルトでは、層変数は積み重ねられます。(注: スケール層変数がない場合は、層変数の積み重ねしかできません)。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 行に「年齢カテゴリー」がなく、層に「性別」がない場合は、前の手順の説明に従って層を使ったテーブルを作成します。
3. 変数リストの「最終学歴」を、「層」リストの「性別」の下にドラッグ・アンド・ドロップします。

これで、「層」リストの下にある「層出力」グループの 2 つのラジオ・ボタンがアクティブになります。デフォルトでは、「各カテゴリーを層として表示」が選択されます。これは、積み重ねに相当します。

4. 「OK」をクリックして、テーブルを作成します。
5. 「ビューアー」ウィンドウ内のテーブルをダブルクリックしてアクティブにします。
6. ドロップダウン・リストの下矢印をクリックして、層のリスト全体を表示します。

テーブルには、2 つの「性別」カテゴリーと 5 つの「最終学歴」カテゴリーのそれぞれに対応する、合計 7 つの層があります。積み重ねた層の場合、層の合計数は、層変数のカテゴリー数の合計です (これには、層変数に要求した合計や小計のカテゴリーも含まれます)。

ネストした 2 つのカテゴリー層変数

カテゴリー層変数をネストすると、層変数カテゴリーの組み合わせごとに個別の層が作成されます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 層を積み重ねたテーブルをまだ作成していない場合は、前の手順の説明に従って作成します。
3. 「層出力」グループで、「各組み合わせを層として表示」を選択します。これは、ネストに相当します。
4. 「OK」をクリックして、テーブルを作成します。
5. 「ビューアー」ウィンドウ内のテーブルをダブルクリックしてアクティブにします。

6. ドロップダウン・リストの下矢印をクリックして、層のリスト全体を表示します。

テーブルには、「性別」と「最終学歴」の各組み合わせに 1 つずつ、合計 10 の層があります (すべての層を確認するには、リストをスクロールする必要があります)。ネストされた層の場合、層の合計数は、層変数ごとのカテゴリー数の積 です (この例では、 $5 \times 2 = 10$)。

層にしたテーブルの印刷

デフォルトでは、現在表示されている層だけが印刷されます。テーブルのすべての層を印刷するには、以下の手順に従います。

1. 「ビューアー」ウィンドウ内のテーブルをダブルクリックしてアクティブにします。
2. 「ビューアー」ウィンドウのメニューから、次の項目を選択します。

「書式」 > 「テーブル プロパティ...」

3. 「印刷」タブをクリックします。
4. 「すべての層を印刷」を選択します。

デフォルトのテーブルルックを含めて、この設定をテーブルルックの一部として保存することもできます。

第 4 章 カテゴリー変数の合計と小計

カスタム・テーブルには、合計と小計の両方を表示することができます。合計と小計は、行、列、層といったディメンションやネストのレベルに関わらず、カテゴリー変数に適用できます。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「全般」タブで設定します。

単一変数の単純な合計

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
3. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
4. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
5. 「表示」リストの「ラベル」セルで、デフォルトのラベルを削除して「パーセント」と入力します。
6. 「選択項目に適用」をクリックします。
7. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
8. 「カテゴリと合計」ダイアログ・ボックスで、「合計」を選択（クリック）します。
9. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

	度数	パーセント
年齢カ テコリ		
25歳未満	242	8.6%
25 - 34	627	22.2%
35 - 44	679	24.0%
45 - 54	481	17.0%
55 - 64	320	11.3%
65歳以上	479	16.9%
合計	2828	100.0%

図 20. 単一カテゴリー変数の単純な合計

合計での WYSIWYG

合計は、テーブルに表示されたカテゴリに基づきます。テーブルから一部のカテゴリを除外すると、それらのカテゴリのケースは合計の計算に含まれません。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインで「年齢カテゴリ」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
3. 「ラベル」リストで「25 歳未満」のラベルが付いたカテゴリをクリックします。
4. 「除外」リストの左側にある矢印キーをクリックします。
5. 「ラベル」リストで「65 歳以上」のラベルが付いたカテゴリをクリックします。
6. 「除外」リストの左側にある矢印キーをもう一度クリックします。

2 つのカテゴリが「表示」リストから「除外」リストに移動されます。

7. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

	度数	パーセント	
年齢カ テゴリ	25 - 34	627	29.8%
	35 - 44	679	32.2%
	45 - 54	481	22.8%
	55 - 64	320	15.2%
	合計	2107	100.0%

図 21. 除外したカテゴリのあるテーブルでの合計

このテーブルの合計度数は、すべてのカテゴリを含めた場合は 2,828 であるのに対し、2,107 しかありません。合計には、テーブルで使用されているカテゴリのみが含まれます。(パーセントは、データ・ファイル内の合計ケース数ではなく、テーブルで使用されているケースの合計数に基づいているため、パーセント合計は 100% になっています)。

合計の表示位置

合計は、デフォルトでは、合計されるカテゴリの下に表示されます。合計の表示位置を変更して、合計されるカテゴリの上に合計を表示することもできます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインで「年齢カテゴリ」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
3. 「合計と小計の表示」グループで、「上記のカテゴリに適用」を選択します。
4. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

		度数	パーセント
年齢カ テゴリ	合計	2107	100.0%
	25 - 34	627	29.8%
	35 - 44	679	32.2%
	45 - 54	481	22.8%
	55 - 64	320	15.2%

図 22. 合計したカテゴリーの上に表示された合計

ネストしたテーブルの合計

合計は任意のネスト・レベルにあるカテゴリー変数に適用できるため、複数のネスト・レベルでグループ合計を表示するテーブルを作成することができます。

グループ合計

他のカテゴリー変数の中にネストされたカテゴリー変数の合計は、グループ合計を表します。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 「性別」を、キャンバス・ペインにある「年齢カテゴリー」の左側にドラッグ・アンド・ドロップします。
3. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。

テーブルを作成する前に、合計されるカテゴリーの下に合計を戻します。

4. 「合計と小計の表示」グループで、「下記のカテゴリに適用」を選択します。
5. 「適用」をクリックして設定を保存し、テーブル・ビルダーに戻ります。
6. 「OK」をクリックして、テーブルを作成します。

			度数	パーセント	
性別	男性	年齢カ テゴリ	25 - 34	276	29.3%
		35 - 44	309	32.8%	
		45 - 54	221	23.5%	
		55 - 64	136	14.4%	
		合計	942	100.0%	
女性	年齢カ テゴリ	25 - 34	351	30.1%	
		35 - 44	370	31.8%	
		45 - 54	260	22.3%	
		55 - 64	184	15.8%	
		合計	1165	100.0%	

図 23. 「性別」カテゴリーの中にネストされた「年齢カテゴリー」の合計

これで、テーブルには 2 つのグループ合計が表示されるようになりました。1 つは男性のグループ合計、もう 1 つは女性のグループ合計です。

総計

ネストした変数に適用される合計は、総計ではなく、常にグループ合計です。テーブル全体での合計が必要な場合には、最も浅いネスト・レベルにある変数に合計を適用できます。

1. テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインで「性別」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
3. 「カテゴリと合計」ダイアログ・ボックスで、「合計」を選択 (クリック) します。
4. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

				度数	パーセント
性別	男性	年齢カ	25 - 34	276	29.3%
		テゴリ	35 - 44	309	32.8%
			45 - 54	221	23.5%
			55 - 64	136	14.4%
			合計	942	100.0%
女性	年齢カ	テゴリ	25 - 34	351	30.1%
			35 - 44	370	31.8%
			45 - 54	260	22.3%
			55 - 64	184	15.8%
			合計	1165	100.0%
合計	年齢カ	テゴリ	25 - 34	627	29.8%
			35 - 44	679	32.2%
			45 - 54	481	22.8%
			55 - 64	320	15.2%
			合計	2107	100.0%

図 24. ネストしたテーブルの総計

総計は 2,828 ではなく、2,107 であることに注意してください。2 つの年齢カテゴリーはテーブルから除外されたままなので、これらのカテゴリーに含まれるケースが総計から除外されています。

層変数の合計

層変数の合計は、テーブルに含まれる個別の層として表示されます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. テーブル・ビルダーで「層」をクリックして、「層」リストを表示します。
3. キャンバス・ペインの行領域から「層」リストに「性別」をドラッグ・アンド・ドロップします。

注: 「性別」にはすでに合計が指定されているため、ここで合計を指定する必要はありません。ディメンション間で変数を移動しても、その変数の設定には影響しません。

4. 「OK」をクリックして、テーブルを作成します。
5. 「ビューアー」でテーブルをダブルクリックしてアクティブにします。
6. 「層」ドロップダウン・リストの下矢印をクリックして、テーブル内のすべての層のリストを表示します。

テーブルには、「性別 男性」、「性別 女性」、および「性別 合計」の 3 つの層があります。

層合計の表示位置

層変数の合計の場合、合計の表示位置 (上または下) によって、合計の層の位置が決まります。例えば、層変数の合計に「上記のカテゴリに適用」を指定すると、合計の層が最初に表示される層になります。

小計

変数のカテゴリーのサブセットに対して小計を含めることができます。例えば、サンプル調査の 45 歳未満のすべての回答者を表す年齢カテゴリーと 45 歳以上のすべての回答者を表す年齢カテゴリーのそれぞれに対して小計を含めることができます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
4. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
5. 「値」リストで「3.00」を選択します。
6. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
7. 「ラベル」テキスト・フィールドに、「45 歳未満の小計」と入力します。
8. 「続行」をクリックします。

これにより、最初の 3 つの年齢カテゴリーの小計を含む行が挿入されます。

9. 「値」リストで「6.00」を選択します。
10. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
11. 「ラベル」テキスト・フィールドに、「45 歳以上の小計」と入力します。
12. 「続行」をクリックします。

重要: 小計を定義する前に、合計と小計の表示位置 (「上記のカテゴリに適用」または「下記のカテゴリに適用」) を選択する必要があります。表示位置を変更すると、現在選択されている小計だけでなくすべての小計に影響し、小計に含まれるカテゴリーも変更されます。

13. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

	度数
年齢カテゴリー	
25歳未満	242
25 - 34	627
35 - 44	679
45歳未満の小計	1548
45 - 54	481
55 - 64	320
65歳以上	479
45歳以上の小計	1280

図 25. 年齢カテゴリーの小計

小計での WYSIWYG

合計と同じく、小計はテーブルに含まれるカテゴリーに基づきます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。

注: 最初の小計として表示される (ラベルではなく) 値は、「1.00...3.00」です。これは、小計には、1 から 3 のリストに含まれるすべての値が含まれることを意味します。

3. 「値」リストで「1.00」を選択します (または、ラベル「25 歳未満」をクリックします)。
4. 「除外」リストの左側にある矢印キーをクリックします。

最初の年齢カテゴリーが除外され、最初の小計として表示される値が「2.00...3.00」に変更されます。小計はテーブルに含まれるカテゴリーに基づくため、この値は、除外されたカテゴリーが合計に含まれていないことを示します。カテゴリーを除外すると、そのカテゴリーが自動的に小計から除外されます。したがって、例えば、小計が基にするカテゴリーを除外して、小計だけを表示することはできません。

小計カテゴリーの非表示

小計を定義するカテゴリーの表示を抑止し、基本となるデータに影響を与えることなく効果的にカテゴリーを「省略」して、小計だけを表示できます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
4. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
5. 「値」リストで「3.00」を選択します。
6. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
7. 「ラベル」テキスト・フィールドに、「45 歳未満」と入力します。
8. 「小計されたカテゴリをテーブルから隠す」を選択します (チェック・マークを付けます)。
9. 「続行」をクリックします。

これにより、最初の 3 つの年齢カテゴリーの小計を含む行が挿入されます。

10. 「値」リストで「6.00」を選択します。
11. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
12. 「ラベル」テキスト・フィールドに、「45 歳以上」と入力します。
13. 「小計されたカテゴリをテーブルから隠す」を選択します (チェック・マークを付けます)。
14. 「続行」をクリックします。
15. 小計を含む合計を表示するには、「表示」グループの「合計」を選択します (チェック・マークを付けます)。
16. 「適用」をクリックします。

キャンバスに、小計を表示する一方、小計を定義するカテゴリーを表示から除外する設定が反映されません。

17. 「OK」をクリックして、テーブルを生成します。

		度数
年齢	45 歳未満	1548
カテゴリ	45 歳以上	1280
	合計	2828

図 26. 小計と合計のみを表示するテーブル

層変数の小計

合計と同じく、層変数の小計もテーブルに含まれる個別の層として表示されます。基本的に、小計はカテゴリとして扱われます。各カテゴリはテーブルに含まれる個別の層であり、小計カテゴリの表示位置を含め、層カテゴリの表示順序は、「カテゴリと合計」ダイアログ・ボックスで指定されたカテゴリの並び順によって決まります。

第 5 章 カテゴリー変数の計算対象カテゴリー

カスタム・テーブルには、計算対象カテゴリーを含めることができます。これらのカテゴリーは、行、列、層といったディメンションやネストのレベルに関わらず、同じ変数のカテゴリーから計算される新しいカテゴリーです。例えば、2 つのカテゴリーの間の違いを示す計算対象カテゴリーを含めることができます。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

単純な計算対象カテゴリー

1. メニューから次の項目を選択します。
「分析」 > 「テーブル」 > 「カスタム・テーブル...」
2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
3. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
4. 「値」リストで「3.00」を選択します。
5. 「カテゴリを追加」をクリックして「計算対象カテゴリーの定義」ダイアログ・ボックスを表示します。
6. 「計算対象カテゴリのラベル」テキスト・フィールドに、「45 歳未満」と入力します。
7. 「カテゴリ」リストの「25 歳未満 (1.00)」を選択し、矢印ボタンをクリックして「計算対象カテゴリの式」テキスト・ボックスにコピーします。[1] が式に表示されます。
8. ダイアログ・ボックスでプラス (+) 演算子ボタンをクリックします (または、キーボードでプラス (+) キーを押します)。
9. 「カテゴリ」リストの「25 - 34 (2.00)」を選択し、矢印ボタンをクリックして「計算対象カテゴリの式」テキスト・ボックスにコピーします。
10. ダイアログ・ボックスでプラス (+) 演算子ボタンをクリックします (または、キーボードでプラス (+) キーを押します)。
11. 「カテゴリ」リストの「35 - 44 (3.00)」を選択し、矢印ボタンをクリックして「計算対象カテゴリの式」テキスト・ボックスにコピーします。
12. 「続行」をクリックします。
これにより、最初の 3 つの年齢カテゴリーの小計を含む行が挿入されます。
13. 「値」リストで「5.00」を選択します。
14. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
15. 「ラベル」テキスト・フィールドに、「65 歳未満」と入力します。
16. 「続行」をクリックします。
これにより、最初の 5 つの年齢カテゴリーの小計を含む行が挿入されます。

17. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

	度数
年齢がゴリ 25歳未満	242
25 - 34	627
35 - 44	679
45歳未満	1548
45 - 54	481
55 - 64	320
65歳未満	2349
65歳以上	479

図 27. 小計を含む計算対象カテゴリ

このテーブルには、計算対象カテゴリ (45 歳未満) と小計 (65 歳未満) が含まれています。小計には、計算対象カテゴリにも含まれているカテゴリが含まれます。複数の小計が同じカテゴリを共有することはできないため、小計だけを使用して、これと同じテーブルを作成することはできません。

計算対象カテゴリのカテゴリの非表示

小計と同じく、計算対象カテゴリの式で使用されるカテゴリの表示を抑止し、計算対象カテゴリのみを表示することができます。以下の例は、前の例に基づいて作成します。

1. メニューから次の項目を選択します。
「分析」 > 「テーブル」 > 「カスタム・テーブル...」
2. キャンバス・ペインで「年齢カテゴリ」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
3. 「値」リストで「45 歳未満」の計算対象カテゴリを選択します。
4. 「編集」をクリックして「計算対象カテゴリの定義」ダイアログ・ボックスを表示します。
5. 「式で使用されているカテゴリをテーブルから隠す」を選択します。
6. 「続行」をクリックします。
7. 「値」リストで「65 歳未満」の小計を選択します。
8. 「編集」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
9. 「小計されたカテゴリをテーブルから隠す」を選択します。
10. 「続行」をクリックします。
11. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

	度数
年齢がゴリ 45歳未満	1548
65歳未満	2349
65歳以上	479

図 28. 小計を含む計算カテゴリと非表示のカテゴリ

前の例と同様、テーブルには計算対象カテゴリと小計が含まれます。ただしこの場合、それぞれのカテゴリは非表示にされているため、小計だけが表示されます。

計算対象カテゴリーの小計の参照

計算対象カテゴリーの式には、小計を含めることができます。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「就業状況」をドラッグ・アンド・ドロップします。
4. 変数リストから列領域に「婚姻状況」をドラッグ・アンド・ドロップします。
5. キャンバス・ペインで「就業状況」を右クリックし、ポップアップ・メニューから「カテゴリーと合計」を選択します。
6. 「値」リストで「2」を選択します。
7. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
8. 「ラベル」テキスト・フィールドに、「就業中」と入力します。
9. 「小計されたカテゴリーをテーブルから隠す」を選択します。
10. 「続行」をクリックします。

これにより、最初の 2 つの就業状況カテゴリーの小計を示す行が挿入されます。

11. 「値」リストで「8」を選択します。
12. 「小計を追加」をクリックして「小計の定義」ダイアログ・ボックスを表示します。
13. 「ラベル」テキスト・フィールドに、「就業中ではない」と入力します。
14. 「小計されたカテゴリーをテーブルから隠す」を選択します。
15. 「続行」をクリックします。

これにより、他の就業状況カテゴリーの小計を示す行が挿入されます。

16. 「値」リストで「就業中ではない」の小計を選択します。
17. 「カテゴリーを追加」をクリックして「計算対象カテゴリーの定義」ダイアログ・ボックスを表示します。
18. 「計算対象カテゴリーのラベル」テキスト・フィールドに、「就業中 / 就業中ではない」と入力します。
19. 「合計と小計」リストで「就業中 (就業中 #1)」を選択し、矢印ボタンをクリックして「計算対象カテゴリーの式」テキスト・ボックスにコピーします。
20. ダイアログ・ボックスで除算 (/) 演算子ボタンをクリックします (またはキーボードで「/」キーを押します)。
21. 「合計と小計」リストで「就業中ではない (就業中ではない #2)」を選択し、矢印ボタンをクリックして「計算対象カテゴリーの式」テキスト・ボックスにコピーします。

デフォルトでは、変数の統計と同じ書式 (この場合は度数) が計算対象カテゴリーで使用されます。計算対象カテゴリーの式の除算によって算出される小数部を表示する必要がありますが、度数のデフォルトの書式には小数部が含まれていないため、この書式を変更する必要があります。

22. 「表示形式」タブをクリックします。
23. 「度数」の「小数部の桁数」設定を「2」に変更します。
24. 「続行」をクリックします。

25. 「適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

		婚姻状況				
		既婚	死別	離婚	別離	未婚
		度数	度数	度数	度数	度数
就業状況	働いています	916	64	330	67	494
	働いていません	429	219	116	26	169
	働いています / 働いていません	2.14	.29	2.84	2.58	2.92

図 29. 小計の比率を示す計算対象カテゴリー

このテーブルには、2 つの小計と 1 つの計算対象カテゴリーがあります。計算対象カテゴリーには小計の比率が表示されるため、各小計で示されるグループを容易に比較することができます。就業中の回答者と就業中ではない回答者の比率は、配偶者と死別した回答者のグループにおいて、他のグループよりもかなり低くなっています。また、既婚の回答者グループでの比率がわずかに低くなっているのは、家で子供と一緒に過ごさなければならないことが理由であると考えられます。

計算対象カテゴリーを使用した包括的でない小計の表示

小計は包括的です。つまり、テーブル内のすべての小計には、テーブル内でその小計が位置する場所の上下にあるすべての値が含まれます。一方、計算対象カテゴリーは包括的でないため、テーブル内の混在したカテゴリーを合計することができます。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

- 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
- テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「自分のことをリベラルまたは保守的であると思う」をドラッグ・アンド・ドロップします。
- キャンバス・ペインで「自分のことをリベラルまたは保守的であると思う」を右クリックし、ポップアップ・コンテキスト・メニューから「カテゴリと合計」を選択します。
- 「値」リストで「3」を選択します。
- 「カテゴリを追加」をクリックして「計算対象カテゴリーの定義」ダイアログ・ボックスを表示します。
- 「計算対象カテゴリのラベル」テキスト・フィールドに、「 リベラル小計」と入力します。テキストの前に 4 つのスペースがあることに注意してください。このスペースは、生成されるテーブルのインデントに使用されます。
- 「カテゴリ」リストの「極めてリベラル (1)」を選択し、矢印ボタンをクリックして「計算対象カテゴリの式」テキスト・ボックスにコピーします。
- ダイアログ・ボックスでプラス (+) 演算子ボタンをクリックします (または、キーボードでプラス (+) キーを押します)。
- 「カテゴリ」リストの「リベラル (2)」を選択し、矢印ボタンをクリックして「計算対象カテゴリの式」テキスト・ボックスにコピーします。
- ダイアログ・ボックスでプラス (+) 演算子ボタンをクリックします (または、キーボードでプラス (+) キーを押します)。
- 「カテゴリ」リストの「ややリベラル (3)」を選択し、矢印ボタンをクリックして「計算対象カテゴリの式」テキスト・ボックスにコピーします。

13. 「**続行**」をクリックします。

これにより、リベラル・カテゴリーの小計を含む行が挿入されます。

14. 「**値**」リストで「**7**」を選択します。

15. 「**カテゴリを追加**」をクリックして「**計算対象カテゴリーの定義**」ダイアログ・ボックスを表示します。

16. 「**計算対象カテゴリーのラベル**」テキスト・フィールドに、「**保守小計**」と入力します。テキストの前に 4 つのスペースがあることに注意してください。このスペースは、生成されるテーブルのインデントに使用されます。

17. 「**カテゴリー**」リストの「**やや保守的 (5)**」を選択し、矢印ボタンをクリックして「**計算対象カテゴリーの式**」テキスト・ボックスにコピーします。

18. ダイアログ・ボックスで**プラス (+)** 演算子ボタンをクリックします (または、キーボードで**プラス (+)** キーを押します)。

19. 「**カテゴリー**」リストの「**保守的 (6)**」を選択し、矢印ボタンをクリックして「**計算対象カテゴリーの式**」テキスト・ボックスにコピーします。

20. ダイアログ・ボックスで**プラス (+)** 演算子ボタンをクリックします (または、キーボードで**プラス (+)** キーを押します)。

21. 「**カテゴリー**」リストの「**極めて保守的 (7)**」を選択し、矢印ボタンをクリックして「**計算対象カテゴリーの式**」テキスト・ボックスにコピーします。

22. 「**続行**」をクリックします。

これにより、保守カテゴリーの小計を含む行が挿入されます。

23. 「**適用**」をクリックし、**テーブル・ビルダー**で「**OK**」をクリックしてテーブルを作成します。

	度数	
自分自身自由党派か保守党派か?	非常に自由党	64
	自由党	357
	どちらかという自由党	351
	自由党小計	772
	中間	986
	どちらかという保守党	432
	保守党	415
	非常に保守党	86
	保守党小計	933

図 30. 包括的でない小計を表示する計算対象カテゴリー

このテーブルには 2 つの計算対象カテゴリーがありますが、テーブル内に表示されているすべてのカテゴリーがこれらの計算対象カテゴリーに含まれているわけではありません。「**中間**」カテゴリーは、いずれの計算対象カテゴリーにも含まれていません。小計は包括的であるため、小計を使用して、これと同じテーブルを作成することはできません。

第 6 章 カテゴリーを共有する変数のテーブル

多くの場合、アンケート調査の多数の質問には、予想される回答項目の共通セットが用意されています。例えば、サンプル調査には、さまざまな公共と民間の機関およびサービスへの信頼度に関するいくつかの変数があります。これらの変数にはすべて、共通の応答カテゴリー・セットがあります (1 = とても信頼している、2 = 信頼している、3 = 少し信頼している)。積み重ねを使用すれば、これらの関連する変数を同じテーブルに表示し、共有される応答カテゴリーをテーブルの列に表示できます。計算対象カテゴリーを使用している場合、そのすべての計算対象カテゴリーのラベルと式がすべての変数で同じであるという規定に従っていれば、これらの機能を計算対象に使用することもできます。

	とても信頼している	信頼している	少し信頼している
銀行、金融機関に対する信頼度	490	1068	306
教育に対する信頼度	511	1055	315
大手会社に対する信頼度	500	1078	243
薬に対する信頼度	844	864	167
新聞、雑誌、出版物に対する信頼度	176	878	808
TV (の情報) に対する信頼度	196	936	744

図 31. カテゴリーを共有する変数のテーブル

注: 前のバージョンのカスタム・テーブルでは、これは「度数表」として知られていました。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス (「編集」メニューから「オプション」を選択) の「全般」タブで設定します。

度数のテーブル

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーの変数リストで、「銀行、金融機関に対する信頼度」をクリックし、Shift キーを押しながら「TV (の情報) に対する信頼度」をクリックして、すべての「信頼度」の変数を選択します (注: 変数リスト内の変数ラベルが、ファイル順ではなくアルファベット順で表示されていることが前提です)。
3. 信頼度に関する 6 つの変数を、キャンバス・ペインの行領域にドラッグ・アンド・ドロップします。

この操作により、これらの変数が行ディメンションに積み重ねられます。デフォルトでは、各変数のカテゴリー・ラベルも行に表示されるため、非常に長く幅の狭いテーブルになります (6 つの変数 x 3 つのカテゴリー = 18 行)。ただし、6 つの変数すべてが、同じ定義済みカテゴリー・ラベル (値ラベル) を共有するため、カテゴリー・ラベルは列ディメンションに配置することができます。

4. 「カテゴリーの位置」ドロップダウン・リストから、「列内の行ラベル」を選択します。

これで、テーブルの行は 6 つ (積み重ね変数ごとに 1 つの行) だけになり、定義済みカテゴリーがテーブルの列になります。

- 要約統計量ラベルの「度数」は必須ではないため、テーブルを作成する前に、「要約統計量」グループの「位置」に対して「非表示」を選択 (クリック) します。
- 「OK」をクリックして、テーブルを作成します。

	とても信頼している	信頼している	少し信頼している
銀行、金融機関に対する信頼度	490	1068	306
教育に対する信頼度	511	1055	315
大手会社に対する信頼度	500	1078	243
薬に対する信頼度	844	864	167
新聞、雑誌、出版物に対する信頼度	176	878	808
TV(の情報)に対する信頼度	196	936	744

図 32. 列に共有カテゴリー・ラベルを配置して積み重ねられた行変数のテーブル

変数を行に表示し、列にカテゴリーを表示する代わりに、列に変数を積み重ねて、行にカテゴリーを表示するテーブルを作成することもできます。この例では、カテゴリーよりも変数のほうが多くなっていますが、変数よりもカテゴリーのほうが多い場合には、この方法が適している場合があります。

パーセントのテーブル

変数を行に積み重ねてカテゴリーを列に表示するテーブルの場合、表示すべき最も重要な (または、少なくとも最も理解しやすい) パーセントは、行パーセントです (変数を列に積み重ねてカテゴリーを行に表示するテーブルの場合は、列パーセントが必要になります)。

- テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
- キャンバス・ペインのテーブル・プレビューで、信頼度変数のいずれかを右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
- 「統計量」リストで「行 N %」を選択し、矢印ボタンをクリックして「表示」リストに移動します。
- 「表示」リストの「度数」行で任意のセルを選択し、矢印ボタンをクリックして「統計」リストに戻します。この操作により、「表示」リストからこのセルが削除されます。
- 「すべてに適用」をクリックし、要約統計量の変更をテーブル内のすべての積み重ね変数に適用します。

注: テーブル・プレビューがこの図のように表示されていない場合、「すべてに適用」ではなく「選択項目に適用」をクリックした可能性があります。この場合、新しい要約統計量は選択した変数だけに適用されます。この例の場合、その操作によってカテゴリーごとに 2 つの列が作成され、一方の列には、他のすべての変数の度数のプレースホルダーが表示され、もう一方の列には、選択した変数の行パーセントのプレースホルダーが表示されることになります。こうしたテーブルが作成されたとしても、それは、この例で目的とするテーブルではありません。

- 「OK」をクリックして、テーブルを作成します。

	とても信頼している	信頼している	少し信頼している
銀行、金融機関に対する信頼度	490	1068	306
大手会社に対する信頼度	500	1078	243
教育に対する信頼度	511	1055	315
新聞、雑誌、出版物に対する信頼度	176	878	808
薬に対する信頼度	844	864	167
TV (の情報) に対する信頼度	196	936	744

図 33. 行に変数が積み重ねられ、列にカテゴリーが表示された行パーセントのテーブル

注: 共有カテゴリーを持つ変数のテーブルには、任意の数の要約統計量を含めることができます。本書の例では、わかりやすくするために、一度に 1 つだけ示しています。

合計とカテゴリーの制御

変数とは反対のディメンションにカテゴリーが表示されるテーブルを作成できるのは、テーブル内のすべての変数が同じカテゴリーになっていて、同じ順序で表示される場合だけです。このテーブルを作成するには、合計、小計、その他のカテゴリーの調整を行う必要があります。つまり、「カテゴリと合計」ダイアログ・ボックスで行う変更は、テーブル内でカテゴリーを共有するすべての変数に対して行う必要があるということです。

1. テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインのテーブル・プレビューで最初の信頼度変数を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
3. 「カテゴリと合計」ダイアログ・ボックスで、「合計」を選択し (チェック・マークを付け)、「適用」をクリックします。

最初に、カテゴリー・ラベルが列から行に戻ったことがわかります。また、「カテゴリーの位置」コントロールが無効になっていることもわかります。これは、この変数が同じ「カテゴリー」のセットを共有しなくなったためです。また、変数の 1 つに合計カテゴリーが表示されています。

4. キャンバス・ペインでいずれかの信頼度変数を右クリックし、ポップアップ・メニューから「すべての行変数を選択」を選択します。または、キャンバス・ペインで Ctrl キーを押しながら各積み重ね変数をクリックし、すべての積み重ね変数を選択します (ペインをスクロールダウンするか、テーブル・ビルダーのウィンドウを拡張しなければならない場合があります)。
5. 「定義」グループで「カテゴリと合計」をクリックします。
6. 「カテゴリと合計」ダイアログ・ボックスで「合計」がまだ選択されていない場合 (チェック・マークが付いていない場合) は、選択してから「適用」をクリックします。
7. すべての変数に合計カテゴリーが追加されたため、「カテゴリーの位置」ドロップダウン・リストが再び有効になります。このドロップダウン・リストから、「列内の行ラベル」を選択します。
8. 「OK」をクリックして、テーブルを作成します。

	とても信頼 している	信頼して いる	少し信頼 している	合計
銀行、金融機関に対しての 信頼度	490	1068	306	1864
大手会社に対しての信頼度	500	1078	243	1821
教育に対しての信頼度	511	1055	315	1881
新聞、雑誌、出版物に対 しての信頼度	176	878	808	1862
薬に対しての信頼度	844	864	167	1875
TV (の情報) に対しての信 頼度	196	936	744	1876

図 34. 行に変数が積み重ねられ、列にカテゴリーと合計が表示された行パーセントのテーブル

共有カテゴリーのあるテーブルでのネスト

ネストされたテーブルでは、共有カテゴリーを持つ積み重ね変数の反対側のディメンションにカテゴリー・ラベルを表示させる場合、これらの変数を現在のディメンションで最も深いネスト・レベルに位置させる必要があります。

1. テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 変数リストから行領域の左側に「性別」をドラッグ・アンド・ドロップします。

これで、共有カテゴリーを持つ積み重ね変数が、テーブル・プレビューで「性別」カテゴリーの中にネストされました。

3. テーブル・プレビューで、積み重ねられたいずれかの信頼度変数の右側に、「性別」をドラッグ・アンド・ドロップします。

ここで再び、カテゴリー・ラベルが行ディメンションに戻されて、「カテゴリーの位置」コントロールが無効になります。この時点で、積み重ね変数の 1 つの中に「性別」がネストされている一方で、他の積み重ね変数にはネストされた変数がありません。積み重ね変数のそれぞれに「性別」をネストした変数として追加することもできますが、その場合、行ラベルを列に移動すると、列には共有カテゴリーを持つ積み重ね変数のカテゴリー・ラベルではなく、「性別」のカテゴリー・ラベルが表示されることになります。その理由は、「性別」が最も深いレベルにネストされた変数になり、カテゴリーの位置の変更は、常に最も深いレベルにネストされた変数に適用されるためです。

第 7 章 要約統計量

「要約統計量」には、カテゴリー変数の単純な度数から、スケール変数の平均値の標準誤差などの分散の測定までのあらゆるものが含まれます。ただし、「カスタム・テーブル」ダイアログ・ボックスの「検定統計量」タブで選択できる有意差検定は含まれません。詳しくは、67 ページの『検定統計量』のトピックを参照してください。

カテゴリー変数と多重回答セットの要約統計量には、度数と、以下のさまざまなパーセントの計算が含まれます。

- 行パーセント
- 列パーセント
- サブテーブル・パーセント
- 表パーセント
- 有効数パーセント

スケール変数の要約統計量とカテゴリー変数のカスタム合計要約には、カテゴリー変数に使用できる要約統計量に加え、以下のものがあります。

- 平均値
- 中央値
- パーセンタイル
- 合計
- 標準偏差
- 範囲
- 最小値と最大値

さらに、多重回答セットには他の要約統計量も使用できます。要約統計量の完全なリストもあります。詳しくは、6 ページの『要約統計量』のトピックを参照してください。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「全般」タブで設定します。

要約統計量ソース変数

使用可能な要約統計量は、要約統計量ソース変数の測定レベルに依存します。要約統計量のソース（要約統計量の計算に使用する変数）は、以下の基準で決まります。

- **測定レベル:** テーブル（または積み重ねテーブルのテーブル・セクション）にスケール変数がある場合、要約統計量は、そのスケール変数に基づきます。

- **変数の選択順序:** カテゴリー変数のデフォルトの統計ソース・ディメンション (行または列) は、変数をキャンバス・ペインにドラッグ・アンド・ドロップする順序に基づきます。例えば、最初に行領域に変数をドラッグした場合、行ディメンションがデフォルトの統計ソース・ディメンションになります。
- **ネスト:** カテゴリー変数の場合、要約統計量は、統計ソース・ディメンション内で最も深いレベルの変数に基づきます。

積み重ねテーブルには複数の要約統計量ソース変数 (スケール変数とカテゴリー変数の両方) が存在する場合がありますが、各テーブル・セクションには 1 つの要約統計量ソースしか存在しません。

カテゴリー変数の要約統計量ソース

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...

2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
3. キャンバス・ペインで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します (これがテーブルの唯一の変数であるため、この変数が統計量ソース変数になります)。
4. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
5. 「選択項目に適用」をクリックします。
6. テーブル・ビルダーで、変数リストの「インターネットから情報を得る」をキャンバス・ペインの「年齢カテゴリー」の右側にドラッグ・アンド・ドロップします。
7. キャンバス・ペインでもう一度「年齢カテゴリー」をクリックします。これにより、ポップアップ・メニューの「要約統計量」項目が無効になります。これは、「年齢カテゴリー」が統計ソース・ディメンションで最も深いレベルにネストされた変数ではなくなったためです。
8. キャンバス・ペインで「インターネットから情報を得る」を右クリックします。これにより、「要約統計量」項目が有効になります。これは、この変数が統計ソース・ディメンションで最も深いレベルにネストされており、要約統計量ソース変数になっているためです (テーブルには 1 つのディメンション (行) しかないため、行が統計ソース・ディメンションになります)。
9. キャンバス・ペインの行領域から列領域に「インターネットから情報を得る」をドラッグ・アンド・ドロップします。
10. キャンバス・ペインで「インターネットから情報を得る」をもう一度右クリックします。これにより、ポップアップ・メニューの「要約統計量」項目が無効になります。これは、この変数が統計ソース・ディメンション上に存在しなくなったためです。

カテゴリー変数のデフォルトの統計ソース・ディメンションは、テーブルの作成時に変数を配置した最初のディメンションであるため、「年齢カテゴリー」が再び統計量ソース変数になります。この例で最初に行った作業は、変数を行ディメンションに配置することでした。したがって、行ディメンションがデフォルトの統計ソース・ディメンションになっています。現在、このディメンション上に存在する変数は「年齢カテゴリー」だけであるため、これが統計量ソース変数になります。

スケール変数の要約統計量ソース

1. スケール変数「1日にTVを見る時間」を、キャンバス・ペインの行領域にある「年齢カテゴリー」の左側にドラッグ・アンド・ドロップします。

まず、「度数」と「列 N %」の集計が「平均値」に置き換えられていることに気付くはずですが。さらに、キャンバス・ペインで「1日にTVを見る時間」を右クリックすると、これが要約統計量ソース変数になっていることがわかります。スケール変数を使用するテーブルでは、スケール変数のネスト・レベルやディメンションに関わらず、常にスケール変数が統計量ソース変数になります。スケール変数のデフォルトの要約統計量は平均値です。

- 「1日にTVを見る時間」を、行領域から列領域の「インターネットから情報を得る」の上にドラッグ・アンド・ドロップします。
- 「1日にTVを見る時間」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。(別のディメンションに移動しても、この変数は統計量ソース変数のままです)。
- 「要約統計量」ダイアログ・ボックスの「表示」リストで平均値の「書式」セルをクリックし、「書式」ドロップダウン・リストから「nnnn」を選択します(この選択項目を探すには、リストをスクロールアップしなければならない場合があります)。
- 「小数部の桁数」セルに、2 と入力します。
- 「選択項目に適用」をクリックします。

これで、キャンバス・ペインのテーブル・プレビューには、小数点以下 2 桁の平均値が表示されます。

- 「OK」をクリックして、テーブルを作成します。

		1日にTVを見る時間	
		インターネットから情報を得る	
		いいえ	はい
		平均値	平均値
年齢カテゴリー	25歳未満	3.54	2.12
	25 - 34	3.42	2.14
	35 - 44	3.00	2.01
	45 - 54	2.83	2.06
	55 - 64	3.24	2.37
	65歳以上	3.82	2.33

図 35. クロス集計表のカテゴリー変数で集計されたスケール変数

積み重ね変数

積み重ねテーブルでは、複数の統計量ソース変数を組み込み、これらの統計量ソース変数ごとに異なる要約統計量を指定できるため、積み重ねテーブルで要約統計量を指定する際には、いくつかの事項を特別に考慮する必要があります。

- テーブル・ビルダーを開きます(「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
- 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
- 変数リストで「インターネットから情報を得る」をクリックし、Shift キーを押しながら変数リスト内の「TV から情報を得る」をクリックして、「ニュース」関連のすべての変数を選択します(注: 変数リスト内の変数ラベルが、ファイル順ではなくアルファベット順で表示されていることが前提です)。
- 5 つのニュース関連の変数を、キャンバス・ペインの行領域にドラッグ・アンド・ドロップします。
- 5 つのニュース関連の変数が、行ディメンション内に積み重ねられます。

5. キャンバス・ペインで「インターネットから情報を得る」をクリックし、この変数だけが選択されている状態にします。
6. 「インターネットから情報を得る」を右クリックし、ポップアップ・メニューから「**要約統計量**」を選択します。
7. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します (矢印を使用して、選択した統計量を「統計」リストから「表示」リストに移動することも、選択した統計量を「統計」リストから「表示」リストにドラッグ・アンド・ドロップすることもできます)。
8. 「**選択項目に適用**」をクリックします。

列パーセントに 1 つの列が追加されます。ただし、キャンバス・ペインのテーブル・プレビューを確認すると、列パーセントは 1 つの変数に対してだけ表示されることがわかります。これは、積み重ねテーブルに複数の統計量ソース変数があり、それぞれの統計量ソースで異なる要約統計量を表示できるためです。ただし、この例では、すべての変数について同じ要約統計量を表示します。

9. キャンバス・ペインで「新聞から情報を得る」を右クリックし、ポップアップ・メニューから「**要約統計量**」を選択します。
10. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
11. 「**すべてに適用**」をクリックします。

テーブル・プレビューを確認すると、すべての積み重ね変数について列パーセントが表示されることがわかります。

カテゴリー変数のカスタム合計要約統計量

カテゴリー統計量ソース変数の場合、変数のカテゴリーに対して表示される統計量とは異なるカスタム合計要約統計量を含めることができます。例えば、順序変数の場合、各カテゴリーにはパーセントを表示し、カスタム合計要約統計量には平均値または中央値を表示することができます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
3. 変数リストで「新聞に対する信頼度」をクリックし、Ctrl キーを押しながら「TV に対する信頼度」をクリックして、この両方の変数を選択します。
4. 2 つの変数を、キャンバス・ペインの行領域にドラッグ・アンド・ドロップします。これにより、2 つの変数が行ディメンションに積み重ねられます。
5. キャンバス・ペインでいずれかの変数を右クリックし、ポップアップ・メニューから「**すべての行変数を選択**」を選択します (既にどちらも選択されているかもしれませんが、ここでは確実に選択します)。
6. 変数をもう一度右クリックし、ポップアップ・メニューから「**カテゴリと合計**」を選択します。
7. 「カテゴリと合計」ダイアログ・ボックスで、「合計」をクリックしてから「**適用**」をクリックします。

これで、キャンバス・ペインのテーブル・プレビューに、両方の変数の合計行が表示されます。カスタム合計要約統計量を表示するには、テーブルに対して合計または小計 (あるいはその両方) が指定されている必要があります。

8. キャンバス・ペインでいずれかの変数を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
9. 「要約統計量」ダイアログ・ボックスの「表示」リストで「度数」をクリックし、矢印をクリックして「統計」リストに移動します。この操作により、「表示」リストから度数が削除されます。
10. 「統計量」リストで「列 N %」をクリックし、矢印キーをクリックして「表示」リストに移動します。
11. 「合計と小計に対するユーザー指定の要約統計量」をクリックします (チェック・マークを付けます)。
12. カスタム要約の「表示」リストで「度数」をクリックし、矢印キーをクリックしてカスタム要約の「統計量」リストに移動します。この操作により、「表示」リストから度数が削除されます。
13. カスタム要約の「統計量」リストで「平均値」をクリックし、矢印をクリックしてカスタム要約の「表示」リストに移動します。
14. 「表示」リストで、平均値の「書式」セルをクリックし、書式のドロップダウン・リストから「nnnn」を選択します (この選択項目を探すには、リストをスクロールアップしなければならない場合があります)。
15. 「小数部の桁数」セルに、2 と入力します。
16. 「すべてに適用」をクリックして、これらの設定をテーブルの両方の変数に適用します。

カスタム合計要約統計量について新しい列が追加されますが、期待していたものとは違っているかもしれません。キャンバス・ペインのプレビューを確認すると、テーブルに多数の空のセルが含まれていることがわかります。

17. テーブル・ビルダーで、「要約統計量」グループの「位置」ドロップダウン・リストから「行」を選択します。

この操作により、すべての要約統計量が行ディメンションに移動され、テーブルの 1 つの列にすべての要約統計量が表示されます。

18. 「OK」をクリックして、テーブルを作成します。

新聞、雑誌、出版物に対する信頼度	とても信頼している	列 %	9.5%
	信頼している	列 %	47.2%
	少し信頼している	列 %	43.4%
	合計	平均値	2.34
TV の情報) に対する信頼度	とても信頼している	列 %	10.4%
	信頼している	列 %	49.9%
	少し信頼している	列 %	39.7%
	合計	平均値	2.29

図 36. カスタム合計要約統計量を使用したカテゴリー変数

カテゴリー値の表示

前述のテーブルには、小さな問題が 1 つだけあります。それは、テーブルが基礎とするカテゴリー値を知らなければ、平均値を解釈するのが難しいことです。平均値 2.34 が「とても信頼している」と「信頼している」の間にあるのか、「信頼している」と「少し信頼している」の間にあるのかがわかりません。

「カスタム・テーブル」で直接この問題に対処することはできませんが、より一般的な方法で対処することはできます。

1. メニューから次の項目を選択します。

「編集」 > 「オプション...」

2. 「オプション」ダイアログ・ボックスで、「出力ラベル」タブをクリックします。
3. 「ピボット・テーブルのラベル付け」グループで、「ラベル中の変数値の表示」ドロップダウン・リストから「値とラベル」を選択します。
4. 「OK」をクリックして、この設定を保存します。
5. テーブル・ビルダーを開き（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）、「OK」をクリックしてテーブルをもう一度作成します。

新聞、雑誌、出版物に対する信頼度	1	とても信頼している	列 %	9.5%
	2	信頼している	列 %	47.2%
	3	少し信頼している	列 %	43.4%
		合計	平均値	2.34
TV の情報) に対する信頼度	1	とても信頼している	列 %	10.4%
	2	信頼している	列 %	49.9%
	3	少し信頼している	列 %	39.7%
		合計	平均値	2.29

図 37. 変数カテゴリーに表示された値とラベル

カテゴリー変数により、平均値 2.34 が「信頼している」と「少し信頼している」の間にあることが明らかになります。テーブルにカテゴリー値を表示すると、平均値などのカスタム合計要約統計量の値を遥かに容易に解釈できるようになります。

この表示設定は、すべての手続きによって出力されるピボット・テーブルのすべてに影響するグローバル設定であり、設定を変更するまでは、セッションが変わってもそのまま維持されます。値ラベルだけを表示する元の設定に変更するには、以下のようにします。

6. メニューから次の項目を選択します。

「編集」 > 「オプション...」

7. 「オプション」ダイアログ・ボックスで、「出力ラベル」タブをクリックします。
8. 「ピボット・テーブルのラベル付け」グループで、「ラベル中の変数値の表示」ドロップダウン・リストから「ラベル」を選択します。
9. 「OK」をクリックして、この設定を保存します。

第 8 章 スケール変数の集計

スケール変数の集計

スケール変数では、さまざまな要約統計量を利用できます。カテゴリ変数で利用できる度数やパーセントに加え、スケール変数では次の要約統計量も利用できます。

- 平均値
- 中央値
- パーセンタイル
- 合計
- 標準偏差
- 範囲
- 最小値と最大値

詳しくは、9 ページの『スケール変数とカテゴリ・カスタム合計の要約統計量』のトピックを参照してください。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「全般」タブで指定します。

スケール変数の積み重ね

テーブル内でスケール変数を積み重ねることで、同じテーブル内にある複数のスケール変数を集計できます。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーで、変数リストの「回答者の年齢」をクリックし、Ctrl キーを押しながら「就学年数」をクリックし、さらに Ctrl キーを押しながら「1 日に TV を見る時間」をクリックして、3 つすべてのスケール変数を選択します。
3. 選択した 3 つの変数を、キャンバス・ペインの行領域にドラッグ・アンド・ドロップします。

3 つの変数が、行ディメンションに積み重ねられます。3 つの変数はすべてスケール変数であるため、カテゴリは表示されません。デフォルトの要約統計量は平均値です。

4. 「OK」をクリックして、テーブルを作成します。

	平均値
回答者の年齢	46
就学年数	13
1日にTVを見る時間	3

図 38. 積み重ねられたスケール変数の平均値を示すテーブル

複数の要約統計量

スケール変数に対しては、デフォルトでは平均値が表示されますが、スケール値に他の要約統計量を選択したり、複数の要約統計量を表示したりすることもできます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインのテーブル・プレビューで、3 つのスケール変数のいずれかを右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
3. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「中央値」を選択し、矢印をクリックして「表示」リストに追加します。(矢印を使用して、選択した統計量を「統計」リストから「表示」リストに移動できます。また、選択した統計量を「統計」リストから「表示」リストにドラッグ・アンド・ドロップすることもできます)。
4. 「表示」リストで、中央値の「書式」セルをクリックし、書式のドロップダウン・リストから「nnnn」を選択します。
5. 「小数部の桁数」セルに、1 と入力します。
6. 「表示」リストの平均値について、同じ変更を行います。
7. 「すべてに適用」をクリックして、これらの変更を 3 つすべてのスケール変数に適用します。
8. テーブル・ビルダーで「OK」をクリックして、テーブルを作成します。

	平均値	中央値
回答者の年齢	45.6	42.0
就学年数	13.3	13.0
1日にTVを見る時間	2.9	2.0

図 39. スケール変数を積み重ねたテーブルに表示された平均値と中央値

度数、有効 N、および欠損値

多くの場合、平均値などの要約統計量の計算に使用されるケース数を表示すると役立ちます。この情報は、要約統計量の「度数」によって提供されると考えても問題ありません。ただし、欠損値がある場合、計算に使用する分母となる正確なケース数は表示されません。分母となる正確なケース数を取得するには、「有効 N」を使用してください。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインのテーブル・プレビューで、3 つのスケール変数のいずれかを右クリックし、ポップアップ・メニューから「要約統計量」を選択します。

- 「要約統計量」ダイアログ・ボックスの「統計量」リストで「度数」を選択し、矢印をクリックして「表示」リストに追加します。
- 「統計量」リストで「有効 N」を選択し、矢印をクリックして「表示」リストに追加します。
- 「すべてに適用」をクリックして、これらの変更を 3 つすべてのスケール変数に適用します。
- テーブル・ビルダーで「OK」をクリックして、テーブルを作成します。

	平均値	中央値	度数	有効な N
回答者の年齢	45.6	42.0	2832	2828
就学年数	13.3	13.0	2832	2820
1日にTVを見る時間	2.9	2.0	2832	2337

図 40. 度数と有効 N の対比

3 つすべての変数で、「度数」が同じ値 (2,832) になっています。これは偶然ではなく、この値はデータ・ファイル内の合計ケース数であるためです。スケール変数はカテゴリー変数内でネストされないため、「度数」は、データ・ファイル内の合計ケース数をそのまま示しています。

一方、「有効な N」は変数によって異なり、「1日にTVを見る時間」では「度数」との差がかなりあります。これは、この変数には多数の「欠損値」があるためです。つまり、この変数について記録された値が存在しないケースや、欠落データとして定義された値 (例えば、男性の妊娠率は「該当しない」であることを表すコード 99 など) を持つケースがあるということです。

7. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
8. キャンパス・ペインのテーブル・プレビューで、3 つのスケール変数のいずれかを右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
9. 「要約統計量」ダイアログ・ボックスの「表示」リストで「有効 N」を選択し、矢印キーをクリックして「統計」リストに戻します。この操作により、「表示」リストから「有効 N」が削除されます。
10. 「表示」リストで「度数」を選択し、矢印キーをクリックして「統計量」リストに戻します。この操作により、「表示」リストから「度数」が削除されます。
11. 「統計量」リストで「欠損値」を選択し、矢印キーをクリックして「表示」リストに追加します。
12. 「すべてに適用」をクリックして、これらの変更を 3 つすべてのスケール変数に適用します。
13. テーブル・ビルダーで「OK」をクリックして、テーブルを作成します。

	平均値	中央値	欠損
回答者の年齢	45.6	42.0	4
就学年数	13.3	13.0	12
1日にTVを見る時間	2.9	2.0	495

図 41. スケール要約統計量のテーブルに表示された欠損値の数

これで、テーブルには各スケール変数の欠損値の数が表示されるようになりました。これにより、「1日にTVを見る時間」には多数の欠損値がある一方、他の 2 つの変数には欠損値がほとんどないことが明らかにわかります。この変数の要約値を信頼する前に、このことを考慮する必要があります。

各種変数の異なる要約

複数の要約統計を表示するだけでなく、積み重ねテーブルでは、各種のスケール変数についてさまざまな要約統計を表示することができます。例えば、前のセクションで使用したテーブルには、3 つの変数の 1 つだけに多数の欠損値があることがわかりました。そこで、この 1 つの変数についてのみ、欠損値の数を表示することにします。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインのテーブル・プレビューで「回答者の年齢」をクリックし、Ctrl キーを押しながら「就学年数」をクリックして両方の変数を選択します。
3. 選択した 2 つの変数のいずれかを右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
4. 「要約統計量」ダイアログ・ボックスの「表示」リストで「欠損値」を選択し、矢印キーをクリックして「統計」リストに戻します。この操作により、「表示」リストから「欠損値」が削除されます。
5. 「選択項目に適用」をクリックして、選択した 2 つの変数だけに変更を適用します。

テーブルのデータ・セル内のプレースホルダーは、「1 日に TV を見る時間」についてのみ欠損値の数が表示されることを示しています。

6. 「OK」をクリックして、テーブルを作成します。

	平均値	中央値	欠損
回答者の年齢	45.6	42.0	
就学年数	13.3	13.0	
1日にTVを見る時間	2.9	2.0	495

図 42. 異なる変数について異なる要約統計量を表示するテーブル

このテーブルには必要な情報が表示されていますが、このレイアウトでは、テーブルの内容を解釈するのが難しい場合があります。このテーブルを見た人の中には、「欠損」列の空白セルが、その変数に欠損値がないことを表していると考えられる人がいるかもしれません。

7. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
8. テーブル・ビルダーで、「要約統計量」グループの「位置」ドロップダウン・リストから「行」を選択します。
9. 「OK」をクリックして、テーブルを作成します。

回答者の年齢	平均値	45.6
	中央値	42.0
就学年数	平均値	13.3
	中央値	13.0
1日にTVを見る時間	平均値	2.9
	中央値	2.0
	欠損	495

図 43. 行ディメンションで表示された要約統計量と変数

これで、このテーブルには、1 つの変数についてのみ欠損値の数が表示されていることが簡単にわかるようになります。

カテゴリー内でのグループ要約

カテゴリー変数をグループ化変数として使用して、カテゴリー変数で定義したグループに含まれるスケール変数の要約を表示することができます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 変数リストからキャンバス・ペインの列領域に「性別」をドラッグ・アンド・ドロップします。

キャンバス・ペインのテーブル・プレビューで「性別」を右クリックすると、ポップアップメニューで「要約統計量」が無効になっていることがわかります。これは、スケール変数のあるテーブルでは、スケール変数が常に統計量ソース変数となるためです。

3. 「OK」をクリックして、テーブルを作成します。

		性別	
		男性	女性
回答者の年齢	平均値	44.6	46.3
	中央値	42.0	43.0
就学年数	平均値	13.4	13.2
	中央値	13.0	13.0
1日にTVを見る時間	平均値	2.8	2.9
	中央値	2.0	2.0
	欠損	213	282

図 44. カテゴリー列変数を使用してグループ化された要約

このテーブルでは、男性と女性の平均 (平均値と中央値) を簡単に比較でき、男性と女性の間には大きな差がないことが明らかにわかります。これは特に興味深い結果ではありませんが、役立つ情報になるかもしれません。

複数のグループ化変数

カテゴリー・グループ化変数をネストしたり、行と列の両方のカテゴリー・グループ化変数を使用したりすることで、グループをさらに細分化できます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 変数リストの「インターネットから情報を得る」を、キャンバス・ペインの行領域の左端にドラッグ・アンド・ドロップします。スケール変数のいずれか 1 つだけでなく、3 つすべてのスケール変数がある中にネストされるように配置してください。

上記の 2 番目の例のようにならない場合もありますが、この例には該当しません。

3. 「OK」をクリックして、テーブルを作成します。

				性別	
				男性	女性
インターネットから情報を得る	いいえ	回答者の年齢	平均値	47.0	48.8
			中央値	45.0	46.0
		就学年数	平均値	13.4	13.1
			中央値	13.0	12.0
		1日にTVを見る時間	平均値	3.2	3.4
		中央値	2.0	3.0	
		欠損	213	282	
	はい	回答者の年齢	平均値	38.7	41.1
			中央値	35.0	38.0
		就学年数	平均値	13.2	13.3
		中央値	13.0	13.0	
1日にTVを見る時間		平均値	2.1	2.1	
	中央値	2.0	2.0		
	欠損	0	0		

図 45. 行および列のカテゴリ変数でグループ化されたスケールの要約

スケール変数の中でのカテゴリ変数のネスト

前のテーブルには必要な情報が示されていますが、その表示形式は、最も簡単に解釈できるようにはなっていません。例えば、インターネットから情報を得る男性の平均年齢と、そうでない男性の平均年齢を比較することはできますが、それぞれの値が離れているよりも、隣合わせに並んでいたほうが比較しやすくなります。2つの行変数の位置を交換し、3つのスケール変数の中にカテゴリ・グループ化変数をネストすることで、理解しやすいテーブルに改善できる可能性があります。スケール変数の場合、ネスト・レベルは統計量ソース変数に影響しません。ネスト・レベルに関係なく、スケール変数は常に統計量ソース変数になります。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. キャンバス・ペインのテーブル・プレビューで「回答者の年齢」をクリックし、Ctrl キーを押しながら「就学年数」をクリックし、さらに Ctrl キーを押しながら「1日にTVを見る時間」をクリックして、3つすべてのスケール変数を選択します。
3. スケール変数を行領域の左端にドラッグ・アンド・ドロップして、3つのスケール変数のそれぞれの中にカテゴリ変数「インターネットから情報を得る」をネストします。
4. 「OK」をクリックして、テーブルを作成します。

				性別	
				男性	女性
回答者の年齢	インター	いい	平均値	47.0	48.8
	ネットから情報を得る	え	中央値	45.0	46.0
		はい	平均値	38.7	41.1
			中央値	35.0	38.0
就学年数	インター	いい	平均値	13.4	13.1
	ネットから情報を得る	え	中央値	13.0	12.0
		はい	平均値	13.2	13.3
			中央値	13.0	13.0
1日にTVを見る時間	インター	いい	平均値	3.2	3.4
	ネットから情報を得る	え	中央値	2.0	3.0
			欠損	213	282
		はい	平均値	2.1	2.1
		中央値	2.0	2.0	
		欠損	0	0	

図46. 積み重ねたスケール変数の中にネストされたカテゴリー行変数

ネストの順番は、テーブル内で強調したい関係や比較によって選択します。スケール変数のネストの順番を変更しても、要約統計量の値は変わりません。テーブル内でのスケール変数の相対位置が変わるだけです。

第 9 章 検定統計量

検定統計量

行変数と列変数の間の関係を調べるには、3 つの異なる有意差検定を使用できます。この章では、ネストと積み重ねの影響に特に注意しながら、これらの検定のそれぞれの出力について説明します。詳しくは、27 ページの『第 3 章 カテゴリー変数の積み重ね、ネスト、層』のトピックを参照してください。

サンプル・データ・ファイル

この章の例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

独立性の検定 (カイ 2 乗)

独立性のカイ 2 乗検定を使用して、2 つのカテゴリー変数の間に関係があるかどうかを判別します。例えば、「就業状況」が「婚姻状況」に関係しているかどうかを判別するとします。

1. メニューから次の項目を選択します。
「分析」 > 「テーブル」 > 「カスタム・テーブル...」
2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「就業状況」をドラッグ・アンド・ドロップします。
3. 変数リストから列領域に「婚姻状況」をドラッグ・アンド・ドロップします。
4. 要約統計量の位置として「行」を選択します。
5. 「就業状況」を選択し、「定義」グループで「要約統計量」をクリックします。
6. 「統計量」リストで「列 N %」を選択し、「表示」リストに追加します。
7. 「選択項目に適用」をクリックします。
8. 「カスタム・テーブル」ダイアログ・ボックスで、「検定統計量」タブをクリックします。
9. 「独立性の検定 (カイ 2 乗)」を選択します。
10. 「OK」をクリックしてテーブルを作成し、カイ 2 乗検定結果を取得します。

就業状況		婚姻状況				
		既婚	死別	離婚	別離	未婚
正社員	度数	778	44	295	58	392
	列%	57.8%	15.5%	66.1%	62.4%	59.1%
パート	度数	138	20	35	9	102
	列%	10.3%	7.1%	7.8%	9.7%	15.4%
契約社員	度数	23	2	9	1	11
	列%	1.7%	.7%	2.0%	1.1%	1.7%
失業中	度数	13	3	10	0	32
	列%	1.0%	1.1%	2.2%	.0%	4.8%
退職	度数	168	150	53	6	17
	列%	12.5%	53.0%	11.9%	6.5%	2.6%
学生	度数	9	1	7	2	60
	列%	.7%	.4%	1.6%	2.2%	9.0%
主婦	度数	200	55	25	13	35
	列%	14.9%	19.4%	5.6%	14.0%	5.3%
その他	度数	16	8	12	4	14
	列%	1.2%	2.8%	2.7%	4.3%	2.1%

図 47. 「就業状況」と「婚姻状況」を使用した集計

このテーブルは、「就業状況」と「婚姻状況」を集計したクロス集計表で、度数と列の比率が要約統計量として示されています。列の比率は、各列での合計が 100% 以内になるように計算されています。この 2 つの変数に関係がない場合は、列全体を通して各行の比率が近接した値になります。このテーブルの比率には差があるように見えますが、カイ 2 乗検定でそれを確認することができます。

		婚姻状況
就業状況	カイ 2 乗	729.242
	df	28
	Sig.	.000*

結果はそれぞれ最も内部の副表中の空でない行と列に基づいて生成されます。

*. カイ 2 乗検定は 0.05 水準において有意です。

図 48. Pearson のカイ 2 乗検定

独立性の検定は、「就業状況」と「婚姻状況」との間に関係がないという仮説を立てます。つまり、列比率が列全体で同じであり、観測される違いは偶然の変動によるものであるという仮説です。カイ 2 乗統計は、観測されるセルの度数と、列の比率が列全体で同じである場合に予測される度数との全体的な違いを測定します。カイ 2 乗統計量で示される違いが大きいほど、観測されたセル度数と予測されるセル度数の差が大きいことになります。これは、列比率が等しくないこと、独立性の仮説が正しくないこと、したがって「就業状況」と「婚姻状況」の間に関係があることの強い証拠となります。

計算されたカイ 2 乗統計量の値は 729.242 です。この値が独立性の仮説を否定するのに十分な証拠になるかどうかを判断するために、統計量の有意確率値が計算されます。有意確率値とは、自由度 28 のカイ 2 乗分布から取り出された乱数変量が 729.242 よりも大きくなる確率です。この値は「検定統計量」タブで指定されたアルファ・レベルを下回っているため、0.05 レベルで独立性の仮説を否定することができます。したがって、「就業状況」と「婚姻状況」の間には、実際に関係があることになります。

独立性の検定におけるネストと積み重ねの影響

独立性の検定の規則に従い、最も深いサブテーブルごとに個別の検定が実行されます。この検定に対するネストの影響を調べるために、前のセクションで使用した例をここでも使用します。ただし今回は、「性別」のレベルで「婚姻状況」をネストします。

1. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 変数リストの「性別」を、キャンパス・ペインの列領域の「婚姻状況」の上にドラッグ・アンド・ドロップします。

3. 「OK」をクリックして、テーブルを作成します。

		性別	
		男性	女性
		婚姻状況	婚姻状況
就業 状況	カイ2乗	246.637	542.589
	df	28	28
	Sig.	.000 ^{a,b}	.000 ^{a,b}

結果はそれぞれ最も内部の副表中の空でない行と列に基づいて生成されます。

*. カイ2乗検定は 0.05 水準において有意です。

- a. この副表のセルの 20% 以上が5未満の期待セル度数を持ちます。カイ2乗検定結果は不正である可能性があります。
- b. この副表中の最小期待セル度数は 1 未満です。カイ2乗検定結果は不正である可能性があります

図 49. Pearson のカイ 2 乗検定

「婚姻状況」を「性別」のレベル内でネストした状態で、「性別」のレベルごとに 1 つずつ、合計 2 つの検定が実行されています。各テストの有意確率値を確認すると、男性と女性の両方について、「婚姻状況」と「就業状況」との間には関係がないという仮説を否定することができます。ただし、テーブルの注には、各テーブルのセルの 20% を超えるセルの期待セル度数が 5 未満であることと、最小期待セル度数が 1 未満であることが記されています。この 2 つの注は、これらのテーブルがカイ 2 乗検定の前提を満たしていない可能性があるため、検定の結果は信頼できないものであることを示しています。

注: セル境界によって、脚注がビューから切り離される場合があります。「セル プロパティ」ダイアログ・ボックスでセルの配置を変更することにより、脚注を表示することができます。

積み重ねが検定に与える影響を確認するには、次の手順を実行してください。

4. テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
5. 変数リストの「最終学歴」を、行領域の「就業状況」の下にドラッグ・アンド・ドロップします。
6. 「OK」をクリックして、テーブルを作成します。

		性別	
		男性	女性
		婚姻状況	婚姻状況
就業 状況	カイ2乗	246.637	542.589
	df	28	28
	Sig.	.000 ^{a,b}	.000 ^{a,b}
最終 学歴	カイ2乗	43.844	105.506
	df	16	16
	Sig.	.000*	.000*

結果はそれぞれ最も内部の副表中の空でない行と列に基づいて生成されます。

*. カイ2乗検定は 0.05 水準において有意です。

- a. この副表のセルの 20% 以上が5未満の期待セル度数を持ちます。カイ2乗検定結果は不正である可能性があります。
- b. この副表中の最小期待セル度数は 1 未満です。カイ2乗検定結果は不正である可能性があります

図 50. Pearson のカイ 2 乗検定

「最終学歴」と「就業状況」を積み重ねた状態で、4 つの検定が実行されています (「性別」の各レベルでの「婚姻状況」と「就業状況」の独立性の検定および「婚姻状況」と「最終学歴」の独立性の検定)。「婚姻状況」と「就業状況」の検定結果は、前のセクションと同じです。「婚姻状況」と「最終学歴」の検定結果からは、これらの変数の間に関係があることがわかります。

列の平均値の比較

列の平均値の検定を使用して、列内のカテゴリ変数と行内の連続型変数との間に関係があるかどうかを判別します。さらに、この検定結果を使用し、連続型変数の平均値を基準として、カテゴリ変数のカテゴリの相対的な順序を決定することもできます。例えば、「1日にTVを見る時間」が「新聞から情報を得る」と関係しているかどうかを判別するとします。

1. メニューから次の項目を選択します。
「分析」 > 「テーブル」 > 「カスタム・テーブル...」
2. 「戻す」をクリックして、すべてのタブをデフォルト設定に戻します。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「1日にTVを見る時間」をドラッグ・アンド・ドロップします。
4. 変数リストから列領域に「新聞から情報を得る」をドラッグ・アンド・ドロップします。
5. 「1日にTVを見る時間」を選択し、「定義」グループで「要約統計量」をクリックします。
6. 書式として「nnnn」を選択します。
7. 表示する小数部の桁数として「2」を選択します。これによって書式が「nnnn.nn」に変わったことに注意してください。
8. 「選択項目に適用」をクリックします。
9. 「カスタム・テーブル」ダイアログ・ボックスで、「検定統計量」タブをクリックします。
10. 「列の平均値を比較 (t 検定)」を選択します。
11. 「OK」をクリックしてテーブルを作成し、列の平均値の検定結果を取得します。

	新聞から情報を得る	
	いいえ	はい
	平均値	平均値
1日にTVを見る時間	2.92	2.74

図 51. 「1日にTVを見る時間」と「新聞から情報を得る」を使用した集計

このテーブルには、新聞から情報を得る人とそうでない人の「1日にTVを見る時間」の平均値が示されています。これらの平均値で認められる差から、新聞から情報を得ない人がTVを見る時間は、新聞から情報を得る人より約 0.18 時間多くなっていることがわかります。この差が偶然の変動によるものかどうかを確認するには、列の平均値の検定結果を調べます。

	新聞から情報を得る	
	いいえ	はい
	(A)	(B)
1日にTVを見る時間		

図 52. 列の平均値の比較

列の平均値の検定テーブルでは、列変数の各カテゴリに文字キーが割り当てられます。「新聞から情報を得る」の「いいえ」カテゴリには文字 A が割り当てられ、「はい」カテゴリには文字 B が割り当てられています。列の各ペアについて、t 検定を使用して列の平均値が比較されます。列は 2 つしかないため、1 つの検定だけが実行されます。それぞれの有意確率ペアについて、平均値が小さいほうのカテゴリのキーが、平均値が大きいほうのカテゴリの下に配置されます。テーブルのセルにはキーが示されていないため、列の平均値には、統計上の大きな差異はないことがわかります。

APA スタイル表記での有意性の結果

有意性の結果を別個のテーブルに表示したくない場合は、これらの結果をメイン・テーブルに表示することができます。有意性の結果は、下付き文字による APA スタイル表記を使用して識別されます。ここでも、列の平均値を比較する上記の手順を実行しますが、「検定統計量」タブで以下の変更を行う必要があります。

1. 「有意差の特定」領域で、「**APA スタイルのサブスクリプトを使用するメイン テーブル**」を選択します。
2. 「**OK**」をクリックしてテーブルを作成し、APA スタイル表記を使用した列の平均値の検定結果を取得します。

	新聞から情報を得る	
	いいえ	はい
	平均値	平均値
1日にTVを見る時間	2.92 _a	2.74 _a

図 53. APA スタイル表記を使用した列の平均値の比較

列の平均値の検定テーブルでは、列変数のカテゴリーに下付き文字が割り当てられます。列の各ペアについて、 t 検定を使用して列の平均値が比較されます。値のペアに大きな差がある場合、これらの値には異なる下付き文字が割り当てられます。列は 2 つしかないため、1 つの検定だけが実行されます。この例での列の平均値には同じ下付き文字が示されているため、列の平均値には、統計上の大きな差異はないことがわかります。

列の平均値の検定におけるネストと積み重ねの影響

列の平均値の検定の規則に従い、最も深い各サブテーブルについて、個別セットのペアごとの検定が実行されます。ここでは、この検定にネストが与える影響を調べるために、前のセクションで使用した例をもう一度使用します。ただし今回は、「就業状況」のレベルで「1日にTVを見る時間」をネストします。

1. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 変数リストからキャンバス・ペインの行領域に「就業状況」をドラッグ・アンド・ドロップします。
3. 「**OK**」をクリックして、テーブルを作成します。

			新聞から情報を得る	
			いいえ	はい
			(A)	(B)
就業 状況	正社員	1日にTVを見る時間	B	
	パートタイム	1日にTVを見る時間		
	契約社員	1日にTVを見る時間		
	失業中	1日にTVを見る時間		
	退職	1日にTVを見る時間		
	学生	1日にTVを見る時間		
	主婦	1日にTVを見る時間		
	その他	1日にTVを見る時間		

図 54. 列の平均値の比較

ここでは、「1日にTVを見る時間」が「就業状況」のレベルにネストされた状態で、「就業状況」のレベルごとに 1 セットずつ、合計 7 セットの列の平均値の検定が実行されます。「新聞から情報を得る」のカテゴリーには、同じ文字キーが割り当てられています。「フルタイムで勤務」の回答者には、A 列に B キーが示されています。これは、フルタイムで勤務している人の「1日にTVを見る時間」

の平均値が、新聞から情報を得る人よりも小さいことを意味します。列に他のキーは示されていないため、統計上、列の平均値には他に大きな差はないと結論付けることができます。

Bonferroni の調整: 複数の検定を実行すると、Bonferroni の調整が各列の平均値の検定に適用されません。これにより、「検定統計量」タブで指定されたアルファ・レベル (真偽比率) が検定の各セットに適用されます。この例では 7 セットの検定が実行されますが、各セットで 1 つの列ペアしか比較されないため、このテーブルでは Bonferroni の調整が適用されていません。

積み重ねが検定に与える影響を確認するには、次の手順を実行してください。

4. テーブル・ビルダーをもう一度開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
5. 変数リストの「インターネットから情報を得る」を、列領域の「新聞から情報を得る」の左側にドラッグ・アンド・ドロップします。
6. 「OK」をクリックして、テーブルを作成します。

		インターネットから情報を得る		新聞から情報を得る	
		いいえ (A)	はい (B)	いいえ (A)	はい (B)
就業 状況	正社員	1日にTVを見る時間	B		B
	パートタイム	1日にTVを見る時間	B		
	契約社員	1日にTVを見る時間			
	失業中	1日にTVを見る時間	B		
	退職	1日にTVを見る時間	B		
	学生	1日にTVを見る時間	B		
	主婦	1日にTVを見る時間	B		
その他	1日にTVを見る時間	B			

図 55. 列の平均値の比較

ここでは、「インターネットから情報を得る」と「新聞から情報を得る」を積み重ねた状態で、「インターネットから情報を得る」と「新聞から情報を得る」に対して「就業状況」のレベルごとに 1 セットずつ、合計 14 セットの列の平均値の検定が実行されます。この場合も、各セットで 1 つの列ペアしか比較されないため、Bonferroni の調整は適用されません。「新聞から情報を得る」の検定結果は前の例と同じです。

「インターネットから情報を得る」の「いいえ」カテゴリーには文字 A が割り当てられ、カテゴリー「はい」カテゴリーには文字 B が割り当てられています。一時的に失業している回答者を除き、列の平均値の検定の各セットの A 列には B キーが示されています。これは、「1 日に TV を見る時間」の平均値が、新聞から情報を得ない人に比べ、インターネットから情報を得る人のほうが小さいことを意味します。「一時的に失業」セットにはキーが示されていないため、これらの回答者の場合、列の平均値について統計上の大きな差異はないことがわかります。

列の比率の比較

列の比率の検定は、行カテゴリー変数のカテゴリー比率を基準として、列カテゴリー変数のカテゴリーの相対的な並び順を決める場合に使用します。例えば、カイ 2 乗検定を使用して「就業状況」と「婚姻状況」が互いに独立していないことを調べてから、どの行と列がこの関係に関与しているのかを確認することができます。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. 「戻す」をクリックして、すべてのタブをデフォルト設定に戻します。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「就業状況」をドラッグ・アンド・ドロップします。

- 変数リストから列領域に「婚姻状況」をドラッグ・アンド・ドロップします。
- 「就業状況」を選択し、「定義」グループで「要約統計量」をクリックします。
- 「統計量」リストで「列 N %」を選択し、「表示」リストに追加します。
- 「表示」リストで「度数」の選択を解除します。
- 「選択項目に適用」をクリックします。
- 「カスタム・テーブル」ダイアログ・ボックスで、「検定統計量」タブをクリックします。
- 「列の比率を比較 (z 検定)」を選択します。
- 「OK」をクリックしてテーブルを作成し、列の比率の検定結果を取得します。

		婚姻状況				
		既婚	死別	離婚	別離	未婚
		列 %	列 %	列 %	列 %	列 %
就業 状況	正社員	57.8%	15.5%	66.1%	62.4%	59.1%
	パートタイム	10.3%	7.1%	7.8%	9.7%	15.4%
	契約社員	1.7%	.7%	2.0%	1.1%	1.7%
	失業中	1.0%	1.1%	2.2%	.0%	4.8%
	退職	12.5%	53.0%	11.9%	6.5%	2.6%
	学生	.7%	.4%	1.6%	2.2%	9.0%
	主婦	14.9%	19.4%	5.6%	14.0%	5.3%
	その他	1.2%	2.8%	2.7%	4.3%	2.1%

図 56. 「就業状況」と「婚姻状況」を使用した集計

このテーブルは「就業状況」と「婚姻状況」を集計したクロス集計表で、列の比率が要約統計量として示されています。

		婚姻状況				
		既婚	死別	離婚	別離	未婚
		(A)	(B)	(C)	(D)	(E)
就業 状況	正社員	B		A B	B	B
	パートタイム					A B C
	契約社員				a	A B
	失業中					
	退職	E	A C D E	E		
	学生					A B C
	主婦	C E	C E		C E	
	その他					

図 57. 列の比率の比較

列の比率の検定テーブルでは、列変数の各カテゴリーに文字キーが割り当てられます。「婚姻状況」の「既婚」カテゴリーには文字 A が割り当てられ、「死別」カテゴリーには文字 B が割り当てられます。このようにして、最後の「未婚」カテゴリーには文字 E が割り当てられます。列の各ペアについて、z 検定を使用して列の比率が比較されます。「就業状況」のレベルごとに 1 セットずつ、合計 7 セットの列の比率の検定が実行されます。「婚姻状況」には 5 つのレベルがあるため、検定の各セットでは $(5*4)/2 = 10$ ペアの列が比較され、Bonferroni の調整を使用して有意確率値が調整されます。各有意確率のペアについて、比率が小さいほうのカテゴリーのキーが、比率が大きいほうのカテゴリーの下に配置されます。

「フルタイムで勤務」に関連する検定のセットでは、B 列以外の各列に B キーが示されています。また、C 列には A キーも示されています。これらのキー以外に、他の列で示されているキーはありません。したがって、フルタイムで勤務している離婚経験者の比率は、フルタイムで勤務している既婚者の比率よりも高く、フルタイムで勤務している既婚者の比率は、フルタイムで勤務している配偶者と死別した人の比率よりも高いと結論付けることができます。フルタイムで勤務している別居中の人または未婚者の比率と、フルタ

イムで勤務している離婚経験者または既婚者の比率との差を確認することはできませんが、これらの比率は、フルタイムで勤務している配偶者と死別した人の比率よりも高くなっています。

「パートタイム」または「学生」に関連する検定では、E 列に A キー、B キー、C キーが示されています。これらのキー以外に、他の列で示されているキーはありません。したがって、学校に通っているか、パートタイムで勤務している未婚者の比率は、学校に通っているか、パートタイムで勤務している既婚者、配偶者と死別した人、離婚経験者の比率よりも高いこととなります。

「一時的に失業」または「その他」の就業状況に関連する検定では、どの列にもキーが示されていません。したがって、既婚者、配偶者と死別した人、離婚経験者、別居中の人、未婚の人で、一時的に失業している人、あるいはカテゴリーにない雇用状況にある人の比率では、それぞれに目に見える違いはありません。

「退職」に関連する検定では、配偶者と死別した退職者の比率は、他のすべての婚姻状況のカテゴリーに属する退職者の比率よりも高くなっています。さらに、既婚者または離婚経験者で退職した人の比率は、未婚で退職した人の比率よりも高くなっています。

既婚者、配偶者と死別した人、別居中の人で家事をしている人の比率は、離婚経験者と、未婚の人で家事をしている人の比率よりも高くなっています。

未婚で「失業中」の人の比率は、失業中の既婚者と、配偶者と死別した人の比率よりも高くなっています。また、「別居中」列に「。」のマークが示されていることに注意してください。これは、「失業中」の行で測定された別居中の人の比率が 0 または 1 であるため、失業中の回答者の「別居中」列を使用して比較を行うことはできないことを意味しています。

APA スタイル表記での有意性の結果

有意性の結果を別個のテーブルに表示したくない場合は、これらの結果をメイン・テーブルに表示することができます。有意性の結果は、下付き文字による APA スタイル表記を使用して識別されます。ここでも、列の比率を比較する上記の手順を実行しますが、「検定統計量」タブで以下の変更を行う必要があります。

1. 「有意差の特定」領域で、「**APA スタイルのサブスクリプトを使用するメイン テーブル**」を選択します。
2. 「**OK**」をクリックしてテーブルを作成し、APA スタイル表記を使用した列の平均値の検定結果を取得します。

		婚姻状況				
		既婚	死別	離婚	別離	未婚
		列の N %	列の N %	列の N %	列の N %	列の N %
就業状況	正社員	57.8% _a	15.5% _b	66.1% _c	62.4% _{a,c}	59.1% _{a,c}
	パートタイム	10.3% _a	7.1% _a	7.8% _a	9.7% _{a,b}	15.4% _b
	契約社員	1.7% _a	.7% _a	2.0% _a	1.1% _a	1.7% _a
	失業中	1.0% _a	1.1% _a	2.2% _{a,b}	.0% ¹	4.8% _b
	退職	12.5% _a	53.0% _b	11.9% _a	6.5% _{a,c}	2.6% _c
	学生	.7% _a	.4% _a	1.6% _a	2.2% _{a,b}	9.0% _b
	主婦	14.9% _a	19.4% _a	5.6% _b	14.0% _a	5.3% _b
	その他	1.2% _a	2.8% _a	2.7% _a	4.3% _a	2.1% _a

図 58. APA スタイル表記を使用した列の比率の比較

列の比率の検定テーブルでは、列変数のカテゴリーに下付き文字が割り当てられます。列の各ペアについて、z 検定を使用して列の比率が比較されます。値のペアに大きな差がある場合、これらの値には異なる下付き文字が割り当てられます。

「正社員」に関連する検定のセットでは、「死別」カテゴリーに他の列では使用されていない下付き文字が示されていて、「別離」と「未婚」カテゴリーでは 2 つの下付き文字が同じです。これらの下付き文字と、テーブルに示されている実際の比率を検討することにより、前の例で別のテーブルを使用して説明したものと同一分析を行うことができます。したがって、フルタイムで勤務している離婚経験者の比率は、フルタイムで勤務している既婚者の比率よりも高く、フルタイムで勤務している既婚者の比率は、フルタイムで勤務している配偶者と死別した人の比率よりも高いと結論付けることができます。フルタイムで勤務している別居中の人または未婚者の比率と、フルタイムで勤務している離婚経験者または既婚者の比率との差を確認することはできませんが、これらの比率は、フルタイムで勤務している配偶者と死別した人の比率よりも高くなっています。残りについても、前の例の分析が当てはまります。

列の比率の検定におけるネストと積み重ねの影響

列の比率の検定の規則に従い、最も深い各サブテーブルについて、個別セットのペアごとの検定が実行されます。ここでは、この検定にネストが与える影響を調べるために、前のセクションで使用した例をもう一度使用します。ただし今回は、「性別」のレベルで「就業状況」をネストします。

1. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 変数リストからキャンバス・ペインの行領域に「性別」をドラッグ・アンド・ドロップします。
3. 「OK」をクリックして、テーブルを作成します。

				婚姻状況				
				既婚 (A)	死別 (B)	離婚 (C)	別離 (D)	未婚 (E)
性別	男性	就業 状況	正社員	B		B	B	B
			パートタイム					A
			契約社員		a			
			失業中		a		a	A
			退職	E	A C D E	E		
	女性	就業 状況	学生		a			A C
			主婦				A	
			その他			A B	A B	B
			正社員	B				B
			パートタイム	B			a	B
			契約社員				a	A
			失業中				a	
			退職	E	A C D E	E		
			学生					A B C
			主婦	C E	C E		C	
		その他						

図 59. 列の比率の比較

「就業状況」を「性別」のレベル内でネストした状態で、「性別」の各レベルの「就業状況」のレベルごとに 1 セットずつ、合計 14 セットの列の比率の検定が実行されています。「婚姻状況」の Kategorii には、同じ文字キーが割り当てられています。

このテーブルの結果には、注目すべき点がいくつかあります。

- 検定の数が増えると、列の比率が 0 になる列が増えてきます。これは、別居中の回答者と、配偶者と死別した男性において最も顕著です。
- 「家事」の回答者において、以前の例で見られた列の差は、すべて女性が原因で発生したものと思われる。

積み重ねが検定に与える影響を確認するには、次の手順を実行してください。

4. テーブル・ビルダーをもう一度開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
5. 変数リストの「最終学歴」を、行領域の「性別」の下にドラッグ・アンド・ドロップします。
6. 「OK」をクリックして、テーブルを作成します。

				婚姻状況				
				既婚 (A)	死別 (B)	離婚 (C)	別離 (D)	未婚 (E)
性別	男性	就業 状況	正社員	B		B	B	B
			パートタイム					A
			契約社員		a			
			失業中		a		a	A
			退職	E	A C D E	E		A C
	学生		a					
	主婦							
	女性	就業 状況	正社員	B		A B	A	B
			パートタイム	B			B	B
			契約社員				a	
失業中						a	A	
退職			E	A C D E	E		A B C	
学生								
主婦	C E	C E		C				
その他								
最終学歴	中卒未満			A C E				
	高卒							
	短大卒	B			B		B	
	4大卒	B					B	
大学院卒	B							

図 60. 列の比率の比較

「最終学歴」と「性別」を積み重ねた状態で、前述の 14 セットのほかに、「最終学歴」のレベルごとに 1 セットずつ、合計 19 セットの列の平均値の検定が実行されています。「婚姻状況」のカテゴリーには、同じ文字キーが割り当てられています。

このテーブルの結果には、いくつかの注目すべき点があります。

- 上で実行した 14 セットの検定については、同じ検定結果となっています。
- 既婚者、離婚経験者、未婚者の回答者に比べ、配偶者と死別した男性において、高卒未満の学歴を持つ人の比率が高くなっています。
- 配偶者と死別した男性よりも、既婚者、離婚経験者、未婚者の回答者において、高卒より上の学歴を持つ人の比率が高くなる傾向があります。

重み付けおよび多重回答セットに関する注

ケースの重み付けは、変数のいずれかが多重回答変数である場合でも、回答数ではなく度数に基づいています。

第 10 章 多重回答セット

カスタム・テーブル および図表ビルダーは、**多重回答セット**という特殊な種類の「変数」をサポートしています。多重回答セットは、実際には一般的な意味での「変数」ではありません。これらはデータ・エディターには表示されず、その他の手続きにも認識されません。多重回答セットは、複数の回答が可能な質問に対する回答を、複数の変数を使用して記録します。多重回答セットは、カテゴリ変数と同様に処理され、カテゴリ変数で実行できるほとんどの処理は、多重回答セットでも実行することができます。

多重回答セットは、データ・ファイルに複数の変数を使用して構成されます。データ・ファイル内では、多重回答セット特殊な構成要素です。IBM® SPSS® Statistics データ・ファイル内には複数の多重回答セットを定義して保存できますが、他のファイル・フォーマットに対して多重回答セットをインポートまたはエクスポートすることはできません。他の IBM SPSS Statistics データ・ファイルから多重回答セットをコピーすることはできます。それには、データ・エディター・ウィンドウの「データ」メニューからアクセスする「「データ・プロパティのコピー」を使用します。

サンプル・データ・ファイル

この章の例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「全般」タブで指定します。

度数、回答数、パーセント、および合計

カテゴリ変数で使用できるすべての要約統計量は、多重回答セットでも使用することができます。さらに、他の統計量も多重回答セットで使用することができます。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. 変数リストからキャンバス・ペインの行領域に「ニュース・ソース」（これは、多重回答セット *\$mltnews* を説明するラベルです）をドラッグ・アンド・ドロップします。

変数リスト内の「変数」の隣にあるアイコンは、その変数が多重二値セットであることを示します。



図 61. 多重二値セットのアイコン

多重二値セットの場合、各「カテゴリ」は、実際には個別の変数であり、カテゴリ・ラベルは変数ラベル（または、定義済み変数ラベルがない変数の場合は変数名）になります。この例で表示される度数は、セット内の各変数に「はい」の回答が含まれるケースの数を表します。

3. キャンバス・ペインのテーブル・プレビューで「ニュース・ソース」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。

- 「カテゴリと合計」ダイアログ・ボックスで「合計」を選択 (クリック) し、次に「適用」をクリックします。
- 「ニュース・ソース」をもう一度右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
- 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
- 「選択項目に適用」をクリックし、「OK」をクリックしてテーブルを作成します。

		度数	列 %
ニュース ソース	新聞から情報を得る	805	38.7%
	雑誌から情報を得る	294	14.1%
	TVから情報を得る	1077	51.8%
	ラジオから情報を得る	551	26.5%
	インターネットから情報を得る	867	41.7%
	合計	2081	100.0%

図 62. 多重二値の度数と列パーセント

加算されない合計

このテーブルの数値を見ると、「合計」の数値と、それぞれの値を合計した値との間に大きな差異があることがわかります (具体的には、「合計」の数値が、実際の合計値よりも大幅に小さくなっています)。これは、テーブル内の各「カテゴリー」の度数は、その変数の値が 1 (「はい」の回答) となっているケースの数であるためです。多重二値セットに含まれる 5 つすべての変数に対する「はい」の回答を合計すると、データ・ファイル内のケースの総数を大幅に上回ることがあります。

ただし、合計の「度数」は、セット内の 1 つ以上の変数に対して「はい」と回答されたケースの総数であるため、この値がデータ・ファイル内のケースの総数を上回ることはありません。この例の合計度数 2,081 は、データ・ファイル内のケースの総数よりも 800 ほど小さな値です。これらの変数のいずれも欠損値を持っていない場合、約 800 人のアンケート回答者は、情報を得る際に、これらの情報源のいずれも利用していないということになります。合計度数は、列パーセントの計算で分母となる値です。したがって、この例の列パーセントを合計すると、列パーセントの合計に表示されている 100% よりも大きな数値になります。

加算される合計

通常、「度数」というのはほとんど曖昧さのない用語ですが、上記の例における多重回答セットでの合計というコンテキスト (つまり、通常は実際に必要な要約統計量が「回答数」となるコンテキスト) では、この用語が混乱を招く可能性があるということがわかります。

- テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
- キャンパス・ペインのテーブル・プレビューで「ニュース・ソース」を右クリックし、ポップアップメニューから「要約統計量」を選択します。
- 「要約統計量」ダイアログ・ボックスの「統計量」リストで「回答数」を選択し、矢印をクリックして「表示」リストに追加します。
- 「統計量」リストで「列の回答 %」を選択し、矢印をクリックして「表示」リストに追加します。
- 「選択項目に適用」をクリックし、「OK」をクリックしてテーブルを作成します。

		度数	列%	回答	列の回答%
ニュース ソース	新聞から情報を得る	805	38.7%	805	22.4%
	雑誌から情報を得る	294	14.1%	294	8.2%
	TVから情報を得る	1077	51.8%	1077	30.0%
	ラジオから情報を得る	551	26.5%	551	15.3%
	インターネットから情報を得る	867	41.7%	867	24.1%
	合計	2081	100.0%	3594	100.0%

図 63. 多重二値の回答数と列の回答パーセント

多重二値セットに含まれる各「カテゴリー」では、「回答数」と「度数」が同じになります。多重二値セットでは、常にこうなります。ただし、合計についてはかなり異なります。合計回答数は 3,594 で、合計度数との差は 1,500 を超え、データ・ファイル内の合計ケース数との差は 700 を超えています。

パーセントについては、「列 N %」と「列の回答 %」はどちらも 100% になりますが、多重二値セットに含まれる各カテゴリーのパーセントは、列の回答パーセントより大幅に小さな値になります。これは、列の回答パーセントの計算で、パーセントの分母となる値が合計回答数（この例では 3,594）であるため、列のパーセントの分母である 2,081 より大幅に小さくなるからです。

100% 以上の合計パーセント

列のパーセントと列の回答パーセントは、どちらも合計が 100% になっていますが、この例の場合、「列 N %」列の個々の値を合計すると、明らかに 100% よりも大きな値になります。合計回答数ではなく合計度数でパーセントを計算する一方で、個々のカテゴリーのパーセントの合計を正確に反映した「合計」パーセントを表示するには、以下の手順を実行します。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. キャンバス・ペインのテーブル・プレビューで「ニュース・ソース」を右クリックし、ポップアップメニューから「要約統計量」を選択します。
3. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列の回答 % (基準: 度数)」を選択し、矢印をクリックして「表示」リストに追加します。
4. 「選択項目に適用」をクリックし、「OK」をクリックしてテーブルを作成します。

		度数	列%	応答	列の回答%	列の回答% (基準: 度数)
ニュース ソース	新聞から情報を得る	805	38.7%	805	22.4%	38.7%
	雑誌から情報を得る	294	14.1%	294	8.2%	14.1%
	TVから情報を得る	1077	51.8%	1077	30.0%	51.8%
	ラジオから情報を得る	551	26.5%	551	15.3%	26.5%
	インターネットから情報を得る	867	41.7%	867	24.1%	41.7%
	合計	2081	100.0%	3594	100.0%	172.7%

図 64. 度数をパーセントの分母として使用した列の回答パーセント

その他の変数との多重回答セットの使用

一般に、多重回答セットはカテゴリー変数と同じように使用できます。例えば、カテゴリー変数を使用して多重回答セットをクロス集計したり、カテゴリー変数の中に多重回答セットをネストしたりできます。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。

- 変数リストからプレビュー・ペインの行領域の左側に「性別」をドラッグ・アンド・ドロップし、多重回答セット「ニュース・ソース」を性別カテゴリーの中にネストします。
- キャンバス・ペインのテーブル・プレビューで「性別」を右クリックし、ポップアップ・メニューで「変数ラベルを表示」を選択解除します。
- 「ニュース・ソース」についても同じ操作を行います。

これにより、テーブルから変数ラベルの付いた列が除去されます（これらの列は、この例の場合には必要ないためです）。

- 「OK」をクリックして、テーブルを作成します。

		度数	列のN%	応答	列の回答%	列の回答% (基準: 度数)
男性	新聞から情報を得る	375	41.9%	375	24.4%	41.9%
	雑誌から情報を得る	121	13.5%	121	7.9%	13.5%
	TVから情報を得る	451	50.3%	451	29.3%	50.3%
	ラジオから情報を得る	233	26.0%	233	15.1%	26.0%
	インターネットから情報を得る	359	40.1%	359	23.3%	40.1%
	合計	896	100.0%	1539	100.0%	171.8%
女性	新聞から情報を得る	430	36.3%	430	20.9%	36.3%
	雑誌から情報を得る	173	14.6%	173	8.4%	14.6%
	TVから情報を得る	626	52.8%	626	30.5%	52.8%
	ラジオから情報を得る	318	26.8%	318	15.5%	26.8%
	インターネットから情報を得る	508	42.9%	508	24.7%	42.9%
	合計	1185	100.0%	2055	100.0%	173.4%

図 65. カテゴリー変数の中にネストされた多重回答セット

統計量ソース変数と利用できる要約統計量

テーブルにスケール変数がない場合、カテゴリー変数と多重回答セットは、統計量ソース変数と同じ方法で扱われます。統計量ソース変数は、統計ソース・ディメンションで最も深いレベルにネストされている変数です。一部の要約統計量は多重回答セットだけに割り当てることができるので、このような特殊な多重回答の要約統計量を求める場合は、多重回答セットは統計ソース・ディメンションで最も深くネストされた変数にする必要があります。

- テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
- キャンバス・ペインのテーブル・プレビューで、「ニュース・ソース」を「性別」の左側にドラッグ・アンド・ドロップして、ネストの順序を変更します。

回答数、列の回答パーセントなど、特殊な多重回答のすべての要約統計量はテーブル・プレビューから取り除かれます。カテゴリー変数「性別」は、最も深いレベルにネストされた変数であり、これが統計量ソース変数になっているためです。

幸い、テーブル・ビルダーはこれらの設定を「記憶」します。「ニュース・ソース」を前の位置に戻して「性別」の中にネストすると、回答関連のすべての要約統計量がテーブル・プレビューに復元されます。

多重カテゴリー・セットと重複回答

多重カテゴリー・セットには、多重二値セットでは使用できない機能が 1 つあります。それは、重複した回答をカウントする機能です。多くの場合、多重カテゴリー・セットに含まれる重複した回答は、コーディング・エラーを表しています。例えば、「最も優れた自動車を製造していると思う国を 3 つ挙げてください」という質問に対し、「スウェーデン、ドイツ、スウェーデン」という回答は正しくありません。

ただし、重複した回答が完全に正しい場合もあります。例えば、「最後に購入した 3 台の自動車はどこで製造されましたか?」という質問の場合、「スウェーデン、ドイツ、スウェーデン」という回答は、完全に意味が通った回答です。

「カスタム・テーブル」には、多重カテゴリー・セットでの重複回答に対する選択肢が用意されています。デフォルトでは、重複回答はカウントされませんが、重複回答をカウントするように指定することもできます。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 「戻す」をクリックして、以前の設定をすべてクリアします。
3. 変数リストからキャンバス・ペインの行領域に「自動車メーカー (最新)」をドラッグ・アンド・ドロップします。

変数リスト内の「変数」の隣にアイコンが表示されている場合、その変数は多重カテゴリー・セットです。



図 66. 多重カテゴリー・セットのアイコン

多重カテゴリー・セットの場合、表示されるカテゴリーは、そのセット内のすべての変数に対して定義されている値ラベルの共通セットを表します（一方、多重二値セットでの「カテゴリー」は、実際には、セット内の各変数の変数ラベルです）。

4. キャンバス・ペインのテーブル・プレビューで「自動車メーカー (最新)」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
5. 「カテゴリと合計」ダイアログ・ボックスで「合計」を選択（クリック）し、次に「適用」をクリックします。
6. 「自動車メーカー (最新)」をもう一度右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
7. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「回答数」を選択し、矢印をクリックして「表示」リストに追加します。
8. 「選択項目に適用」をクリックし、「OK」をクリックしてテーブルを作成します。

		度数	回答
自動車	アメリカ	1938	1938
車	日本	1327	1327
メーカー	韓国	695	695
(最新)	ドイツ	693	693
	スウェーデン	360	360
	その他	343	343
	合計	2832	5356

図 67. 多重カテゴリー・セット: 重複のない度数と回答

デフォルトでは、重複回答はカウントされません。したがって、このテーブルの「度数」列と「回答」列に示されている各カテゴリーの値は同じになります。異なるのは合計だけです。

9. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
10. 「オプション」タブをクリックします。
11. 「多重カテゴリー グループに対して重複した回答を数える」をクリックします (チェック・マークを付けます)。
12. 「OK」をクリックして、テーブルを作成します。

		度数	応答
自動車	アメリカ	1938	2797
車	日本	1327	1717
メーカー	韓国	695	760
(最新)	ドイツ	693	754
	スウェーデン	360	383
	その他	343	359
	合計	2832	6770

図 68. 重複回答が含まれる複数カテゴリー・セット

このテーブルでは、特にアメリカ製の自動車について、「度数」列と「回答数」列に示されている値に大きな違いがあります。これは、多くの回答者が複数のアメリカ製の自動車を所有していることを示しています。

多重回答セットでの有意差検定

有意差検定では、本質的にカテゴリー変数を使用する場合と同じように、多重回答セットを使用できます。

- 独立性の検定 (カイ 2 乗) や列の比率の比較 (Z 検定) では、度数を基準に検定が行われるため、度数が、テーブルに表示される要約統計量のうちの 1 つになっていなければなりません。
- 多重カテゴリー・セットの場合、「オプション」タブで「多重カテゴリー グループに対して重複した回答を数える」が選択されていると、列の比率または列の平均値を比較する検定 (t 検定) は実行されません。詳しくは、16 ページの『カスタム・テーブル: 「オプション」タブ』のトピックを参照してください。

多重回答セットでの独立性の検定

この例では、カテゴリー変数と多重回答セットのクロス集計表を作成し、クロス集計表で独立性のカイ 2 乗検定を行います。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. 「戻す」をクリックして、以前の設定をすべてクリアします。

- 変数リストからキャンバス・ペインの列領域に「ニュース・ソース」(これは、多重二値セット *\$mltnews* を説明するラベルです) をドラッグ・アンド・ドロップします。
- 変数リストからキャンバス・ペインの行領域に「性別」をドラッグ・アンド・ドロップします。
- 「検定統計量」タブをクリックします。
- 「独立性の検定 (カイ 2 乗)」を選択します (チェック・マークを付けます)。
- 「検定に多重応答変数を含める」がまだ選択されていない場合は、これを選択します。
- 「OK」をクリックして手続きを実行します。

		ニュースソース
性別	Chi-square	10.266
	df	5
	Sig.	.068

図 69. カイ 2 乗の結果

カイ 2 乗検定での有意水準 0.068 は、ニュース・ソースの選択において女性と男性の間におそらく大きな差があることを意味します (統計的な有意性を判断する基準として、有意確率値 0.05 未満を使用していることが前提です)。

多重回答セットでの列の平均値の比較

この例では、多重回答セットで定義されたカテゴリー内でスケール変数の平均値を計算し、各カテゴリーの平均値を他のすべてのカテゴリーの平均値と比較して有意差を求めます。

- テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
- 「戻す」をクリックして、以前の設定をすべてクリアします。
- 変数リストからキャンバス・ペインの列領域に「ニュース・ソース」(これは、多重二値セット *\$mltnews* を説明するラベルです) をドラッグ・アンド・ドロップします。
- 「回答者の年齢」をキャンバス・ペインの行領域にドラッグ・アンド・ドロップします。
- 「検定統計量」タブをクリックします。
- 「列の平均値を比較 (t 検定)」を選択します (チェック・マークを付けます)。
- 「検定に多重応答変数を含める」がまだ選択されていない場合は、これを選択します。
- 「OK」をクリックして手続きを実行します。

	ニュースソース				
	新聞から 情報を得 る	雑誌から 情報を得 る	TVから情 報を得る	ラジオから 情報を得 る	インター ネット から情報 を得る
	平均値	平均値	平均値	平均値	平均値
回答者の年齢	52	40	48	40	40

列平均の比較*

	ニュースソース				
	新聞から 情報を得 る	雑誌から 情報を得 る	TVから情 報を得る	ラジオから 情報を得 る	インター ネット から情報 を得る
	(A)	(B)	(C)	(D)	(E)
回答者の年齢	B C D E		B D E		

結果は有意水準 0.05 で等分散を仮定する2群の差の検定に基づいています。各有意確率 α について、より小さいカテゴリのキーがより大きい平均を持ったカテゴリの下に現れます。

図 70. 有意差検定の結果

- 多重回答セットの各カテゴリは、文字 (A、B、C、D、E) で識別され、別のカテゴリの平均値がそのカテゴリの平均値と比較して低く、差が有意である場合、平均値が低いほうのカテゴリを表す文字が表示されます。
- 「新聞から情報を得る」(A) の平均年齢は最も高く、他のすべてのカテゴリの平均値とは有意な差があります。
- 「TV から情報を得る」(C) の平均年齢は 2 番目に高く、残りのすべてのカテゴリの平均値 (B、D、および E) とは有意な差があります。(前述のとおり、C も A との有意な差があります)。
- 「雑誌から情報を得る」(B)、 「ラジオから情報を得る」(D)、 および 「インターネットから情報を得る」(E) の平均年齢は、互いに有意な差がありません。

第 11 章 欠損値

多くのデータ・ファイルには、ある程度の欠落データがあります。欠落データは、さまざまな原因で発生します。例えば、アンケート調査の回答者が一部の質問に答えない場合、特定の変数が一部のケースに適用されない場合、コーディング・エラーによって一部の値が消失された場合などです。

IBM SPSS Statistics には、次の 2 種類の欠損値があります。

- **ユーザー欠損値:** 欠落データを含むと定義された値。これらの値には、データ欠損の理由を識別する値ラベルを割り当てることができます (例えば、男性の妊娠率に対するコード 99 および「適用外」という値ラベル)。
- **システム欠損値:** 数値変数に値がない場合、その変数にはシステム欠損値が割り当てられます。これは、データ・エディターの「データ・ビュー」でピリオドによって表されます。

欠損値による影響を補い、さらに欠落データのパターン分析にも使用できる数々の機能があります。ただし、この章では単純に、単純に「カスタム・テーブル」が欠落データを処理する方法、そして欠落データが要約統計量の計算に与える影響を説明することを目的とします。

サンプル・データ・ファイル

この章に記載する例では、データ・ファイル *missing_values.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。これは、欠損値の基本概念を説明することを目的として、非常にシンプルに、完全に人工的に作成されたデータ・ファイルであり、1 つの変数と 10 のケースしか含まれていません。

欠損値のないテーブル

デフォルトでは、ユーザー欠損カテゴリはカスタム・テーブルには表示されません (システム欠損値も表示されません)。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「欠損値のある変数」(ファイル内にある唯一の変数) をドラッグ・アンド・ドロップします。
3. キャンバス・ペインで変数を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
4. 「カテゴリと合計」ダイアログ・ボックスで、「合計」をクリックして (チェック・マークを付けて) 「適用」をクリックします。
5. キャンバス・ペインのテーブル・プレビューでもう一度「欠損値のある変数」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
6. 「要約統計量」ダイアログ・ボックスの「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
7. 「選択項目に適用」をクリックします。

キャンバス・ペインのテーブル・プレビューに表示されるカテゴリと、「カテゴリ」リスト (テーブル・ビルダー左側の変数リストの下) に表示されるカテゴリには、わずかな違いがあることがわか

ります。「カテゴリ」リスト内の「欠損値」というラベルの付いたカテゴリは、テーブル・プレビューには表示されません。これは、欠損値カテゴリはデフォルトで除外されるためです。

8. 「OK」をクリックして、テーブルを作成します。

		度数	列%
欠損 値の ある 変数	低	2	28.6%
	中	3	42.9%
	高	2	28.6%
	合計	7	100.0%

図 71. 欠損値のないテーブル

このテーブルには、まったく問題がありません。カテゴリ値を加算して合計が計算され、パーセント値には、合計度数をパーセントの分母として使用して算出された値 (例えば、 $3/7 = 0.429$ 、つまり 42.9%) が正確に反映されています。ただし、合計度数はデータ・ファイルに含まれるケースの合計数ではありません。これは、欠損値がないケース、つまり該当する変数のユーザー欠損値またはシステム欠損値がないケースの合計数です。

テーブルに欠損値を含める

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択)。
2. キャンバス・ペインのテーブル・プレビューで「欠損値のある変数」を右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
3. 「カテゴリと合計」ダイアログ・ボックスで、「欠損値」をクリックして (チェック・マークを付けて)「適用」をクリックします。

これで、テーブル・プレビューに「欠損値」カテゴリが組み込まれます。テーブル・プレビューには、欠損値用に 1 つのカテゴリしか表示されませんが、テーブルにはすべてのユーザー欠損カテゴリが表示されます。

4. キャンバス・ペインのテーブル・プレビューでもう一度「欠損値のある変数」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
5. 「要約統計量」ダイアログ・ボックスで、「合計と小計に対するユーザー指定の要約統計量」をクリックします (チェック・マークを付けます)。
6. カスタム要約統計量のリストで「有効 N」を選択し、矢印をクリックして「表示」リストに追加します。
7. 「合計 N」についても同じ操作を行います。
8. 「選択項目に適用」をクリックし、テーブル・ビルダーで「OK」をクリックしてテーブルを作成します。

		度数	列 %	有効 N	合計 N
欠損 値の ある 変数	低	2	22.2%		
	中	3	33.3%		
	高	2	22.2%		
	未知	1	11.1%		
	不適格	1	11.1%		
	合計	9	100.0%	7	10

図 72. 欠損値のあるテーブル

これで、2 つの定義済みユーザー欠損カテゴリ（「わからない」と「該当なし」）がテーブルに表示されるようになります。ユーザー欠損値を持つ 2 つのケース（各ユーザー欠損カテゴリに 1 つずつ）が追加されたため、合計度数は 7 から 9 に変わります。また、非欠損値とユーザー欠損値の数に基づいて列パーセントが計算されるため、列パーセントの値も変わります。パーセントの計算に含まれないのは、システム欠損値だけです。

「有効 N」には、欠損値がないケースの合計数 (7) が表示され、「合計 N」には、ユーザー欠損値とシステム欠損値の両方を含めたケースの合計数が表示されます。ケースの合計数は 10 ですが、これは、「度数」列に合計として表示されている非欠損値とユーザー欠損値の度数よりも 1 だけ大きい値です。これは、システム欠損値を持っているケースが 1 つあるためです。

9. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
10. キャンパス・ペインのテーブル・プレビューで「欠損値のある変数」を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
11. 「統計量」リスト（合計と小計のカスタム統計量ではありません）で「列の有効 N %」をクリックし、矢印をクリックして「表示」リストに追加します。
12. 「列の合計 N %」についても同じ操作を行います。
13. また、この両方の項目を、合計と小計のカスタム要約統計量のリストに追加することもできます。
14. 「選択項目に適用」をクリックし、「OK」をクリックしてテーブルを作成します。

		度数	列 %	列の有効 N %	列の合計 N %	有効な N	合計 N
欠損 値の ある 変数	低	2	22.2%	28.6%	20.0%		
	中	3	33.3%	42.9%	30.0%		
	高	2	22.2%	28.6%	20.0%		
	未知	1	11.1%	.0%	10.0%		
	不適格	1	11.1%	.0%	10.0%		
	合計	9	100.0%	100.0%	100.0%	7	10

図 73. 欠損値、有効パーセント、合計パーセントがあるテーブル

- 「列 N %」は、非欠損とユーザー欠損値の数に基づく各カテゴリのパーセントです（テーブルに、ユーザー欠損値が明示的に含まれているため）。
- 「列の有効 N %」は、欠損値のない有効なケースだけにに基づく各カテゴリのパーセントです。これらの値は、ユーザー欠損値が含まれていない元のテーブルでの列パーセントと同じです。
- 「列の合計 N %」は、ユーザー欠損値とシステム欠損値の両方を含む、すべてのケースに基づく各カテゴリのパーセントです。このカテゴリに含まれる個々のカテゴリのパーセントを合計すると、90% にしかならないことがわかります。これは、合計 10 ケースのうち、1 つのケース (10%) にシステム欠損値があるためです。このケースは、パーセンテージ計算の分母には含まれていますが、システム欠損値のあるケースに対するカテゴリは、テーブル内には提供されません。

第 12 章 テーブルの書式設定とカスタマイズ

テーブルの書式設定とカスタマイズ

「カスタム・テーブル」では、テーブル作成プロセスの一環として、以下を含む多くのテーブル書式設定プロパティを制御できます。

- 要約統計量の表示形式とラベル
- データ列幅の最小値と最大値
- 空のセルに表示するテキストまたは値

これらの設定は、テーブル・ビルダー・インターフェース内で維持されるので（これらの設定を変更するか、テーブル・ビルダーの設定をリセットするか、他のデータ・ファイルを開くまで）、テーブルを作成してから手動で設定を編集することなく、同じ書式プロパティを使用した複数のテーブルを作成できます。また、テーブル・ビルダー・インターフェースの「貼り付け」ボタンを使用してコマンド構文をシンタックス・ウィンドウに貼り付け、これをファイルとして保存すれば、これらの書式設定を他のすべてのテーブル・パラメーターと併せて保存することができます。

テーブルの多くの書式プロパティは、テーブルを作成した後に変更することもできます。この場合、ピボット・テーブルのビューアーに用意されているすべての書式設定機能を利用できます。ただし、この章では、テーブルを作成する前にテーブル書式プロパティを制御する方法を中心に説明します。ピボット・テーブルについて詳しくは、ヘルプ・システムの「キーワード」タブを使用して、キーワードとして「ピボット テーブル」と入力してください。

サンプル・データ・ファイル

この章の例では、データ・ファイル *survey_sample.sav* を使用します。詳しくは、データ・ファイルのトピックを参照してください。

ここに記載するすべての例でダイアログ・ボックスに表示される変数ラベルは、アルファベット順にソートされています。変数リストの表示プロパティは、「オプション」ダイアログ・ボックス（「編集」メニューから「オプション」を選択）の「全般」タブで設定します。

要約統計量の表示形式

「カスタム・テーブル」では、要約統計量に対して、比較的インテリジェントなデフォルト形式が適用されますが、これらのデフォルトを上書きしたい場合もあります。

1. メニューから次の項目を選択します。

「分析」 > 「テーブル」 > 「カスタム・テーブル...」

2. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
3. 行領域の「年齢カテゴリー」の下に「TV (の情報) に対する信頼度」をドラッグ・アンド・ドロップし、行ディメンションに 2 つの変数を積み重ねます。
4. キャンバス・ペインのテーブル・プレビューで「年齢カテゴリー」を右クリックし、ポップアップ・メニューから「すべての行変数を選択」を選択します。

5. 「年齢カテゴリー」をもう一度右クリックし、ポップアップ・メニューから「カテゴリと合計」を選択します。
6. 「カテゴリと合計」ダイアログ・ボックスで、「合計」を選択して (チェック・マークを付けて)、「適用」をクリックします。
7. キャンバス・ペインのテーブル・プレビューでいずれかの変数を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。
8. 「統計量」リストで「列 N %」を選択し、矢印キーをクリックして「表示」リストに追加します。
9. 「合計と小計に対するユーザー指定の要約統計量」を選択します (チェック・マークを付けます)。
10. カスタム要約統計量の「統計量」リストで「列 N %」を選択し、矢印をクリックして「表示」リストに追加します。
11. 「平均値」についても同じ操作を行います。
12. 「すべてに適用」をクリックします。

テーブル・プレビューのプレースホルダー値は、各要約統計量のデフォルト形式を表しています。

- 度数のデフォルト表示形式は、小数部のない整数値を表す **nnnn** です。
- パーセントのデフォルト表示形式は、小数点以下 1 桁の数値にパーセント記号が続く **nnnn.n%** です。
- 平均値のデフォルト表示形式は、これら 2 つの変数で異なります。

度数形式 (「有効 N」と「合計 N」を含む) でもパーセント形式でもない要約統計量の場合、データ・エディターで変数用に定義された表示形式がデフォルトの表示形式になります。データ・エディターの変数ビューで変数を確認すると、「年齢カテゴリー」(変数 *agecat*) は 2 桁の小数部を持つように定義され、「TV (の情報) に対しての信頼度」(変数 *contv*) は小数部を持たないように定義されていることがわかります。

両方の平均値が同じ小数桁数で表示されたほうが見やすいため、これは、デフォルト形式が望ましい形式ではない場合の一例です。

13. キャンバス・ペインのテーブル・プレビューでいずれかの変数を右クリックし、ポップアップ・メニューから「要約統計量」を選択します。

「表示」リストの「書式」セルには、平均値の形式が「自動」として表示されます。この場合、変数の定義済みの表示形式が使用され、「小数部の桁数」セルは無効になります。小数部の桁数を指定するには、最初に別の書式を選択する必要があります。

14. カスタム要約統計量の「表示」リストで平均値の「書式」セルをクリックし、書式のドロップダウン・リストから「nnnn」を選択します。
15. 「小数部の桁数」セルに、値として 1 を入力します。
16. 「すべてに適用」をクリックして、この設定を両方の変数に適用します。

これで、テーブル・プレビューには、両方の平均値が小数点以下 1 桁の値として表示されるようになります (ここで、このテーブルを作成することもできますが、「年齢カテゴリー」の「平均」値は多少解釈しにくいかもしれません。この変数の実際の数値コードは、1 から 6 までの範囲しかないので)。

要約統計量の表示ラベル

要約統計量の表示形式だけでなく、各要約統計量の内容を説明するラベルも制御できます。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 「戻す」をクリックして、テーブル・ビルダーでの以前の設定をすべてクリアします。
3. テーブル・ビルダーで、変数リストからキャンバス・ペインの行領域に「年齢カテゴリー」をドラッグ・アンド・ドロップします。
4. 変数リストからキャンバス・ペインの列領域に「給料体制」をドラッグ・アンド・ドロップします。
5. キャンバス・ペインのテーブル・プレビューで「年齢カテゴリー」を右クリックし、ポップアップ・コンテキスト・メニューから「要約統計量」を選択します。
6. 「統計量」リストで「列 N %」を選択し、矢印キーをクリックして「表示」リストに追加します。
7. 「表示」リストの「ラベル」セルで「列」という単語をダブルクリックして、セルの内容を編集します。ラベルから単語「列」を削除し、「%」だけのラベルにします。
8. 「度数」の「ラベル」セルも同じように編集し、「N」だけのラベルにします。

この画面を表示している間に、「列 N %」統計量の書式を変更して、不要なパーセント記号を削除します（この列にパーセントが表示されることは、列ラベルからわかるためです）。

9. 「列 N %」の「書式」セルをクリックし、書式のドロップダウン・リストから「nnnn.n」を選択します。
10. 「選択項目に適用」をクリックします。

テーブル・プレビューに変更後の書式と変更後のラベルが表示されます。

11. 「OK」をクリックして、テーブルを作成します。

		給料体制											
		時給制		日給制		週給制		月給制		年俸制		その他	
		N	%	N	%	N	%	N	%	N	%	N	%
年齢カテゴリー	25歳未満	91	14.0%	0	0%	12	9.7%	3	2.0%	7	3.1%	14	7.7%
	25 - 34	175	26.9%	5	29.4%	33	26.6%	37	24.8%	63	28.0%	31	17.1%
	35 - 44	185	28.5%	5	29.4%	42	33.9%	45	30.2%	66	29.3%	61	33.7%
	45 - 54	124	19.1%	5	29.4%	25	20.2%	38	25.5%	58	25.8%	41	22.7%
	55 - 64	52	8.0%	0	0%	10	8.1%	23	15.4%	29	12.9%	19	10.5%
	65歳以上	23	3.5%	2	11.8%	2	1.6%	3	2.0%	2	.9%	15	8.3%

図 74. 要約統計量のラベルが変更されたテーブル

列幅

前の例では、テーブルの幅がかなり広がったことに気付いたはずですが、この問題を解決する 1 つの方法は、単純に行変数と列変数を交換することです。あるいは、列の幅が必要以上に広いことから、列の幅を狭くするという解決方法もあります。（実際のところ、要約統計量ラベルを短くした理由は、列の幅を狭くできるようにするためです）。

1. テーブル・ビルダーを開きます（「分析」メニューから「テーブル」、「カスタム・テーブル」の順に選択）。
2. 「オプション」タブをクリックします。
3. 「データ列の幅」グループで、「カスタム」をクリックします。
4. 「最大値」として 36 と入力します。（「単位」の設定が「ポイント」であることを確認してください）。
5. 「OK」をクリックして、テーブルを作成します。

		給料体制											
		時給制		日給制		週給制		月給制		年俸制		その他	
		N	%	N	%	N	%	N	%	N	%	N	%
年齢カ テコリ	25歳未満	91	14.0%	0	.0%	12	9.7%	3	2.0%	7	3.1%	14	7.7%
	25 - 34	175	26.9%	5	29.4%	33	26.6%	37	24.8%	63	28.0%	31	17.1%
	35 - 44	185	28.5%	5	29.4%	42	33.9%	45	30.2%	66	29.3%	61	33.7%
	45 - 54	124	19.1%	5	29.4%	25	20.2%	38	25.5%	58	25.8%	41	22.7%
	55 - 64	52	8.0%	0	.0%	10	8.1%	23	15.4%	29	12.9%	19	10.5%
	65歳以上	23	3.5%	2	11.8%	2	1.6%	3	2.0%	2	.9%	15	8.3%

図 75. 列幅を狭くしたテーブル

これで、テーブルはかなりコンパクトになりました。

空のセルの表示値

デフォルトでは、空のセル (ケースが含まれないセル) には 0 が表示されます。代わりに、これらのセルに何も表示されないようにするか (空白のままにします)、あるいは空のセルに表示するテキスト文字列を指定できます。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、 「カスタム・テーブル」の順に選択)。
2. 「オプション」タブをクリックします。
3. 「データ・セルの表示」グループで、「空のセル」に対して「テキスト」を選択し、「なし」と入力します。
4. 「OK」をクリックして、テーブルを作成します。

		給料体制											
		時給制		日給制		週給制		月給制		年俸制		その他	
		N	%	N	%	N	%	N	%	N	%	N	%
年齢カ テコリ	25歳未満	91	14.0%	なし	なし	12	9.7%	3	2.0%	7	3.1%	14	7.7%
	25 - 34	175	26.9%	5	29.4%	33	26.6%	37	24.8%	63	28.0%	31	17.1%
	35 - 44	185	28.5%	5	29.4%	42	33.9%	45	30.2%	66	29.3%	61	33.7%
	45 - 54	124	19.1%	5	29.4%	25	20.2%	38	25.5%	58	25.8%	41	22.7%
	55 - 64	52	8.0%	なし	なし	10	8.1%	23	15.4%	29	12.9%	19	10.5%
	65歳以上	23	3.5%	2	11.8%	2	1.6%	3	2.0%	2	.9%	15	8.3%

図 76. 空のセルに「なし」と表示されたテーブル

これで、テーブル内の 4 つの空のセルに、値 0 ではなく「なし」というテキストが表示されるようになりました。

欠損統計量の表示値

統計量を計算できない場合は、デフォルトの表示値としてピリオド (.) が表示されます。これは、システム欠損値を示すために使用される記号です。欠損統計量は「空の」セルとは異なるため、欠損統計量の表示値は、ケースがないセルの表示値とは別に制御します。

1. テーブル・ビルダーを開きます (「分析」メニューから「テーブル」、 「カスタム・テーブル」の順に選択)。
2. 変数リストからキャンパスの列領域の一番上 (つまり「給料体制」の上) に、「1 日に TV を見る時間」をドラッグ・アンド・ドロップします。

「1 日に TV を見る時間」はスケール変数であるため、自動的に統計量ソース変数になり、要約統計量は平均値に変更されます。

3. キャンパス・ペインのテーブル・プレビューで「1日にTVを見る時間」を右クリックし、ポップアップ・コンテキスト・メニューから「要約統計量」を選択します。
4. 「統計量」リストで「有効 N」を選択し、矢印キーをクリックして「表示」リストに追加します。
5. 「選択項目に適用」をクリックします。
6. 「オプション」タブをクリックします。
7. 「計算できない統計量」のフィールドに NA と入力します。
8. 「OK」をクリックして、テーブルを作成します。

		1日にTVを見る時間											
		給料体制											
		時給制		日給制		週給制		月給制		年俸制		その他	
		平均値	有効な N	平均値	有効な N	平均値	有効な N	平均値	有効な N	平均値	有効な N	平均値	有効な N
年齢 カテゴリ	25歳未満	3	71	なし	なし	3	10	2	3	2	6	2	8
	25 - 34	3	134	5	2	2	30	2	29	2	52	2	22
	35 - 44	3	136	2	5	3	30	2	34	2	47	3	46
	45 - 54	2	90	2	4	2	22	2	36	2	45	2	34
	55 - 64	3	40	なし	なし	3	7	2	15	2	23	3	15
	65歳以上	3	18	2	2	1	1	なし	0	1	2	3	11

図 77. 欠損統計量として「NA」が表示されているテーブル

テーブル内の 3 つのセルで、平均値として「NA」というテキストが表示されています。各ケースにおいて、対応する「有効 N」の値がその理由を示しています。つまり、平均値の計算に使用できるケースがないということです。

ただし、表示されている内容には、わずかな違いがあることがわかります。それは、3 つの「有効 N」の値の 1 つは、ケースがないセルについて表示されるラベル「なし」ではなく、「0」が表示されていることです。これは、平均値の計算に使用できる有効なケースはないが、カテゴリは実際には空ではないことを示しています。カテゴリ変数が 2 つしかない元のテーブルに戻ると、このクロス集計されたカテゴリに、実際には 3 つのケースがあることがわかります。ただし、この 3 つケースは、いずれもスケール変数「1日にTVを見る時間」に対して欠損値を持っているため、有効なケースではありません。

第 13 章 サンプル・ファイル

製品とともにインストールされるサンプル・ファイルは、インストール・ディレクトリーの *Samples* サブディレクトリーにあります。Samples サブディレクトリー内には、英語、フランス語、ドイツ語、イタリア語、日本語、韓国語、ポーランド語、ロシア語、中国語 (簡体字)、スペイン語、中国語 (繁体字) の言語ごとに個別のフォルダーがあります。

すべてのサンプル・ファイルが、すべての言語で使用できるわけではありません。あるサンプル・ファイルが特定の言語で使用できない場合、その言語のフォルダーには、サンプル・ファイルの英語バージョンが含まれています。

説明

このドキュメントのさまざまな例で使用されているサンプル・ファイルの簡単な説明を以下に示します。

- **accidents.sav:** これは、特定の地域における自動車事故の危険因子を年齢と性別ごとに調査している保険会社に関する架空のデータ・ファイルです。各ケースが、年齢カテゴリーと性別のクロス分類に対応しています。
- **adl.sav:** これは、脳卒中患者に対して提案されるタイプの治療の効果を特定する取り組みに関する架空のデータ・ファイルです。医師団は、女性の脳卒中患者たちを、2つのグループのいずれかにランダムに割り当てました。最初のグループは標準的な理学療法を受け、2番目のグループはさらに心理療法を追加で受けました。治療の3カ月後に、各患者が日常生活における一般的な活動をどの程度行うことができるかを、順序変数として得点付けしました。
- **advert.sav:** これは、広告費とその売上成果との関係を調べることを目的とした、小売業者の取り組みに関する架空のデータ・ファイルです。この目的のために、過去の売上高の数値と、それに関する広告費のデータが収集されています。
- **aflatoxin.sav:** これは、収穫物によって濃度が大きく異なる毒物であるアフラトキシンを、トウモロコシの収穫物で検定することに関する架空のデータ・ファイルです。ある穀物加工業者は、8つの収穫物から16のサンプルを受け取り、10億分の1単位でアフラトキシン・レベルを測定しました。
- **anorectic.sav:** 拒食行動と過食行動の症状の標準化を目的として、調査員¹が摂食障害を持つ55人の若者について調査を行いました。各患者が4年間で4回診察を受け、観察記録の合計は220件となっています。毎回の観察では、16種類の症状それぞれについて患者にスコアが付けられました。患者71(2回目)、患者76(2回目)、患者47(3回目)の症状のスコアは欠落しているため、有効な観察記録は217件ということになります。
- **bankloan.sav:** これは、債務不履行率を低減させるための銀行の取り組みに関する架空のデータ・ファイルです。このファイルには、過去の顧客と見込み客850人に関する財務情報と人口統計情報が含まれています。最初の700件のケースは、以前に貸付を行った顧客です。残りの150件のケースは見込み顧客です。銀行は、これらの顧客に関して信用リスクの良し悪しを分類する必要があります。
- **bankloan_binning.sav:** これは、過去の5,000人の顧客に関する財務情報と人口統計情報が含まれている架空のデータ・ファイルです。

1. Van der Ham, T., J. J. Meulman, D. C. Van Strien, and H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, 363-368.

- **behavior.sav:** 典型的な例²として、52人の学生に対し、15種類の状況と15種類の行動の組み合わせについて、0 = 「非常に適切」から9 = 「非常に不適切」までの10段階でランク付けするよう依頼しました。個人間で平均を取り、非類似度として値を取りました。
- **behavior_ini.sav:** このデータ・ファイルには、*behavior.sav* の2次元の解の初期構成が含まれています。
- **brakes.sav:** これは、高性能自動車のディスク・ブレーキを生産している工場における品質管理に関する架空のデータ・ファイルです。このデータ・ファイルには、8台の機械で生産した16個のディスクの直径測定値が含まれています。ブレーキの目標の直径は322ミリメートルです。
- **breakfast.sav:** 典型的な調査³として、21人のWharton School MBAの学生とその配偶者に対し、15種類の朝食を好みの順に1 = 「最も好き」から15 = 「最も嫌い」までランク付けするよう依頼しました。調査対象者の嗜好は、「全体的な好み」から「スナックとドリンクのみ」まで、6つの異なるシナリオで記録されています。
- **breakfast-overall.sav:** このデータ・ファイルには、最初のシナリオ（「全体的な好み」）における朝食の品目についての好みだけが含まれています。
- **broadband_1.sav:** これは、全国規模のプロードバンド・サービスの地域ごとの加入者数が含まれている架空のデータ・ファイルです。このデータ・ファイルには、85地域の月々の加入者数が4年間分含まれています。
- **broadband_2.sav:** このデータ・ファイルは *broadband_1.sav* と同じですが、3カ月分のデータが追加されています。
- **car_insurance_claims.sav:** これは、他の場所⁴で表示および分析される、自動車の損害請求に関するデータ・セットです。逆リンク関数を使用して、従属変数の平均値を保険契約者の年齢、車種、製造年の線型結合に関連付けることにより、平均請求数をガンマ分布としてモデル化することができます。申請された請求の数は、尺度重み付けとして使用することができます。
- **car_sales.sav:** このデータ・ファイルには、自動車のさまざまな車種やモデルの架空の売上推定値、表示価格、物理的仕様が含まれています。表示価格と物理的仕様は、それぞれ *edmunds.com* と製造元のサイトから入手したものです。
- **car_sales_uprepared.sav:** これは、フィールドの変換バージョンが含まれていない、*car_sales.sav* の修正バージョンです。
- **carpet.sav:** 一般的な例⁵として、新しいカーペット専用洗剤の市販を計画している企業が、消費者の嗜好に関する5つの因子（パッケージのデザイン、ブランド名、価格、「Good Housekeeping」のシール、返金保証）の影響について調べようとしています。パッケージのデザインには3つの因子レベルがあり、それぞれのレベルによって塗布用ブラシの位置が異なります。また、3つのブランド名（*K2R*、*Glory*、*Bissell*）と3つの価格水準があり、残り2つの因子のそれぞれについて2つのレベル（「なし」と「あり」）があります。10人の消費者が、これらの因子によって定義された22個のプロファイルに順位を付けます。「嗜好」という変数には、各プロファイルの平均順位のランキングが格納されます。ランキングが低いほど、嗜好度は高くなります。この変数は、各プロファイルの全体的な嗜好測定値を反映しています。
- **carpet_prefs.sav:** このデータ・ファイルは *carpet.sav* で説明したものと例に基づいていますが、10人の消費者それぞれから収集した実際のランキングが含まれています。これらの消費者は、22種類の製

2. Price, R. H., and D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579-586.

3. Green, P. E., and V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

4. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

5. Green, P. E., and Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

品プロファイルを、一番好きなものから一番嫌いなものまで順位付けすることを依頼されています。変数 *PREF1* から *PREF22* には、*carpet_plan.sav* で定義された、関連するプロファイルの ID が格納されます。

- **catalog.sav:** このデータ・ファイルには、あるカタログ会社が販売した 3 つの製品の、架空の月間売上高の数値が含まれています。また、5 つの予測変数のデータも含まれています。
- **catalog_seasfac.sav:** このデータ・ファイルは *catalog.sav* と同じですが、季節性の分解プロシージャとそれに付随する日付変数から計算した一連の季節因子が追加されています。
- **cellular.sav:** これは、チャーン (契約/解約を繰り返す顧客の流動現象) を削減するための携帯電話会社の取り組みに関する架空のデータ・ファイルです。チャーンの傾向スコアは、0 から 100 の範囲でアカウントに適用されます。スコアリングが 50 以上のアカウントは、プロバイダーを変更しようとしていると考えられます。
- **ceramics.sav:** これは、新しい上質の合金は、標準的な合金より高い耐熱性を持っているかどうかを判別するための、ある製造業者の取り組みに関する架空のデータ・ファイルです。各ケースが 1 つの合金の個別のテストを表し、耐熱性の限界温度が記録されます。
- **cereal.sav:** これは、880 人を対象に、朝食の好みについて、年齢、性別、婚姻状況、ライフスタイルが活動的かどうか (週 2 回以上運動するか) を含めて調査した、架空のデータ・ファイルです。各ケースが個別の回答者を表しています。
- **clothing_defects.sav:** これは、ある衣料品工場での品質管理工程に関する架空のデータ・ファイルです。工場で生産される各ロットから、検査官が衣料品のサンプルを取り出し、不良品の数を数えます。
- **coffee.sav:** これは、6 つのアイス・コーヒー・ブランド⁶ の印象に関連するデータ・ファイルです。回答者は、アイス・コーヒーに関する 23 の各印象属性について、その属性が表現していると思われるすべてのブランドを選択しました。機密保持のため、6 つのブランドを AA、BB、CC、DD、EE、FF で表しています。
- **contacts.sav:** これは、企業のコンピューター営業担当者グループの連絡先リストに関する架空のデータ・ファイルです。それぞれの連絡先は、所属する会社の部門と会社のランクによって分類されています。また、最後に販売した金額、最後の販売以降の経過時間、連絡先の会社の規模も記録されています。
- **creditpromo.sav:** これは、最近のクレジットカード・プロモーションの有効性を評価することを目的とした、あるデパートの取り組みに関する架空のデータ・ファイルです。この目的のために、500 人のカード所有者がランダムに選択されました。そのうちの半分に対して、今後 3 カ月間の買い物に関して利率を下げることを知らせる広告を送付しました。残りの半分には、通常どおりの定期的な広告を送付しました。
- **customer_dbase.sav:** これは、自社のデータウェアハウスにある情報を使用して、最も反応がありそうな顧客に対して特典を提供するための、ある会社の取り組みに関する架空のデータ・ファイルです。顧客ベースのサブセットをランダムに選択して特典を提供し、それに対する顧客の反応が記録されています。
- **customer_information.sav:** これは、名前や住所など、顧客の連絡先情報が含まれている架空のデータ・ファイルです。
- **customer_subset.sav:** これは、*customer_dbase.sav* の 80 件のケースのサブセットです。
- **debate.sav:** これは、政治討論会の出席者から得た、討論前と討論後の調査に対する回答のペアに関する架空のデータ・ファイルです。各ケースが個々の回答者に対応しています。

6. Kennedy, R., C. Riquier, and B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56-70.

- **debate_aggregate.sav:** これは、*debate.sav* 内の回答が集計されている架空のデータ・ファイルです。各ケースが、討論前後の好みのクロス分類に対応しています。
- **demo.sav:** これは、毎月の特典を送付することを目的とした、購入顧客のデータベースに関する架空のデータ・ファイルです。顧客が特典に反応したかどうか、さまざまな人口統計情報と共に記録されています。
- **demo_cs_1.sav:** これは、調査情報のデータベースを作成するための、ある会社の取り組みの最初のステップに関する架空のデータ・ファイルです。各ケースが個別の都市に対応し、地域、地方、地区、および都市の ID が記録されています。
- **demo_cs_2.sav:** これは、調査情報のデータベースを作成するための、ある会社の取り組みの第 2 のステップに関する架空のデータ・ファイルです。各ケースが、最初のステップで選択した都市の個別の世帯単位に対応し、地域、地方、地区、都市、区画、および単位の ID が記録されます。計画の最初の 2 段階からのサンプリング情報も含まれています。
- **demo_cs.sav:** これは、複合サンプリング計画を使用して収集された調査情報を含む架空のデータ・ファイルです。各ケースが個別の世帯単位に対応し、さまざまな人口統計情報と抽出情報が記録されています。
- **dmdata.sav:** これは、ダイレクト・マーケティング企業の人口統計情報と購入情報が含まれている架空のデータ・ファイルです。*dmdata2.sav* には、テスト・メールを受け取った連絡先のサブセットの情報が含まれ、*dmdata3.sav* には、テスト・メールを受け取らなかった残りの連絡先に関する情報が含まれています。
- **dietstudy.sav:** この架空のデータ・ファイルには、「Stillman diet」⁷ の研究結果が含まれています。各ケースが個々の被験者に対応し、被験者のダイエット前後の体重 (ポンド単位) と、トリグルセリド・レベル (mg/100 ml 単位) が記録されています。
- **dvdplayer.sav:** これは、新しい DVD プレイヤーの開発に関する架空のデータ・ファイルです。マーケティング・チームはプロトタイプを使用して、フォーカス・グループ・データを収集しました。各ケースが個々の調査対象ユーザーに対応し、ユーザーの人口統計情報と、プロトタイプに関する質問への回答が記録されています。
- **german_credit.sav:** このデータ・ファイルは、カリフォルニア大学アーバイン校の Repository of Machine Learning Databases⁸ に含まれている「German credit」データ・セットから抜粋したものです。
- **grocery_1month.sav:** これは、*grocery_coupons.sav* データ・ファイルの週ごとの購入を「ロールアップ」して、各ケースが個々の顧客に対応するようにした架空のデータ・ファイルです。そのため、週ごとに変っていた変数の一部が表示されなくなり、記録された購入額が、調査を行った 4 週間の購入額の合計になっています。
- **grocery_coupons.sav:** これは、顧客の購買習慣に関心を持っている食料雑貨店チェーンが収集した調査データが含まれている架空のデータ・ファイルです。各顧客を 4 週間にわたって追跡し、各ケースが個々の顧客の週に対応しています。その週の食料品購入金額など、顧客がいつどこで買物をするかに関する情報が記録されています。
- **guttman.sav:** Bell⁹ は、予想される社会グループを示す表を作成しました。Guttman¹⁰ は、この表の一部を使用しました。この表では、社会的相互作用、グループへの帰属感、メンバーとの物理的な近接

7. Rickman, R., N. Mitchell, J. Dingman, and J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228:, 54-58.

8. Blake, C. L., and C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

9. Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

10. Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469-506.

性、関係の形式化などを表す 5 つの変数が、7 つの理論上の社会グループと交差しています。このグループには、観衆 (フットボールの試合の観戦者など)、視聴者 (映画館や授業の参加者など)、公衆 (新聞やテレビの読者や視聴者など)、群衆 (観衆に似ているが、より強い相互の関連がある)、第 1 次集団 (親密な関係)、第 2 次集団 (自発的な集団)、および近代コミュニティ (物理的に密接した近接性と特化されたサービスの必要性による緩い同盟関係) があります。

- **health_funding.sav:** これは、医療用資金 (人口 100 人あたりの金額)、罹患率 (人口 10,000 人あたりの人数)、医療サービス機関の受診率 (人口 10,000 人あたりの人数) に関するデータが含まれている架空のデータ・ファイルです。各ケースが、個別の都市を表しています。
- **hivassay.sav:** これは、HIV 感染を発見する迅速な分析方法を開発することを目的とした、ある製薬研究所の取り組みに関する架空のデータ・ファイルです。分析の結果は 8 段階の濃さの赤で表現され、色が濃いほど感染の可能性が高くなります。研究所では 2,000 件の血液サンプルに関して試験を行い、その半数が HIV に感染しており、残りの半分は感染していませんでした。
- **hourlywagedata.sav:** これは、さまざまな経験レベルを持つ管理職から現場担当までの看護師の時給に関する架空のデータ・ファイルです。
- **insurance_claims.sav:** これは、不正請求の恐れがある疑わしい請求にフラグを立てるためのモデルを作成したいと考えているある保険会社に関する架空のデータ・ファイルです。各ケースが個々の請求を表しています。
- **insure.sav:** これは、10 年満期の生命保険契約に対し、顧客が請求を行うかどうかを示す危険因子を調査している保険会社に関する架空のデータ・ファイルです。データ・ファイルの各ケースは、年齢と性別が一致する、請求が記録された契約と記録されなかった契約のペアを表しています。
- **judges.sav:** これは、訓練を受けた審判 (および 1 人の熱心なファン) が、300 件の体操の演技に対して付けた得点に関する架空のデータ・ファイルです。各行が個々の演技を表しています。この審判たちは、同じ演技を見ました。
- **kinship_dat.sav:** Rosenberg と Kim ¹¹ は、15 種類の親族関係用語 (叔母、兄弟、いとこ、娘、父、孫娘、祖父、祖母、孫息子、母、甥、姪、姉妹、息子、叔父) の分析を行いました。Rosenberg と Kim は、大学生の 4 つのグループ (女性 2 組、男性 2 組) に、類似性に基づいて上記の用語を並べ替えるよう依頼しました。2 つのグループ (女性 1 組、男性 1 組) には、この最初の条件とは異なる条件に基づいて、2 回目の並べ替えをするように依頼しました。したがって、合計で 6 つの「ソース」が取得されたこととなります。各ソースは、15×15 の近接行列に対応しています。この近接行列のセルの数は、ソース内の人数から、ソース内でオブジェクトを分割した回数を引いたものと同じとなります。
- **kinship_ini.sav:** このデータ・ファイルには、*kinship_dat.sav* の 3 次元の解の初期配置が含まれています。
- **kinship_var.sav:** このデータ・ファイルには、*gender* (性別)、*gener* (世代)、*degree* (親等) という独立変数が含まれています。これらの変数を使用して、*kinship_dat.sav* の解のディメンションを解釈することができます。具体的には、これらの変数を使用して、解の空間をこれらの変数の線型結合に制限することができます。
- **marketvalues.sav:** これは、1999 年から 2000 年までの、イリノイ州アルゴンキンの新興住宅地開発における住宅売上に関するデータ・ファイルです。これらの売上は、公開レコードの対象となっています。
- **nhis2000_subset.sav:** National Health Interview Survey (NHIS) は、米国国民を対象とした人口ベースの大規模な調査です。全国の代表的な世帯サンプルで、対面式の調査が行われます。各世帯のメンバーについて、人口統計情報と、健康に関する行動と状態の観測値が収集されます。このデータ・ファイルには、2000 年の調査から取得された情報のサブセットが含まれています (National Center for Health

11. Rosenberg, S., and M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

Statistics. National Health Interview Survey, 2000. 一般使用のデータ・ファイルとドキュメント。
ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/。2003 年にアクセス)。

- **ozone.sav**: このデータには、残りの変数からオゾン濃度を予測するための、6 つの気象変数に対する 330 の観測値が含まれています。これまでの研究者¹²、¹³が、他の研究者と共に、これらの変数間に非線型性を確認しています。この場合、標準的な回帰アプローチは使用できません。
- **pain_medication.sav**: この架空のデータ・ファイルには、慢性関節炎を治療する抗炎症薬の臨床試験の結果が含まれています。特に重要なのは、薬の効果が出るまでの時間と、既存の薬剤との比較です。
- **patient_los.sav**: この架空のデータ・ファイルには、心筋梗塞 (MI、または「心臓発作」) の疑いで入院した患者の治療記録が含まれています。各ケースは個々の患者に対応し、入院に関連する多くの変数が記録されています。
- **patlos_sample.sav**: この架空のデータ・ファイルには、心筋梗塞 (MI、または「心臓発作」) の治療中に血栓溶解剤を投薬された患者のサンプルの治療記録が含まれています。各ケースは個々の患者に対応し、入院に関連する多くの変数が記録されています。
- **poll_cs.sav**: これは、市民の法案支持率を議会開会前に特定するための、世論調査員の取り組みに関する架空のデータ・ファイルです。各ケースは、登録されている有権者に対応しています。ケースごとに、有権者が居住している郡、町、区域が記録されています。
- **poll_cs_sample.sav**: この架空のデータ・ファイルには、*poll_cs.sav* の有権者のサンプルが含まれています。このサンプルは、*poll.csplan* 計画ファイルで指定されている計画に従って抽出され、このデータ・ファイルには包含確率とサンプル重み付けが記録されています。ただし、サンプリング計画では確率比例 (PPS) 法が使用されるため、結合選択確率が含まれているファイル (*poll_jointprob.sav*) もあります。サンプルの抽出後、有権者の人口統計と法案に関する意見に対応する追加の変数が収集され、データ・ファイルに追加されました。
- **property_assess.sav**: これは、限られたリソースで資産価値の評価を最新に保つための、郡の評価担当者の取り組みに関する架空のデータ・ファイルです。各ケースは、前年に郡内で売却された資産に対応しています。データ・ファイル内の各ケースには、その資産が存在する町、最後にその資産を訪問した評価担当者、その評価からの経過時間、そのときに行われた評価、その資産の売却価値が記録されています。
- **property_assess_cs.sav**: これは、限られたリソースで資産価値の評価を最新に保つための、州の評価担当者の取り組みに関する架空のデータ・ファイルです。各ケースは、州内の資産に対応しています。データ・ファイル内の各ケースには、その資産が存在する郡、町、区域、最後の評価からの経過時間、そのときに行われた評価が記録されています。
- **property_assess_cs_sample.sav**: この架空のデータ・ファイルには、*property_assess_cs.sav* にリストされている資産のサンプルが含まれています。このサンプルは、*property_assess.csplan* 計画ファイルで指定されている計画に従って抽出され、このデータ・ファイルには包含確率とサンプル重み付けが記録されています。サンプルの抽出後、現在の価値 変数が収集され、データ・ファイルに追加されました。
- **recidivism.sav**: これは、管轄地域での累犯率を把握するための、政府の法執行機関の取り組みに関する架空のデータ・ファイルです。各ケースは元犯罪者に対応し、この犯罪者の人口統計情報、最初の犯罪の詳細、2 回目の逮捕までの期間 (初犯から 2 年以内の場合) が記録されています。
- **recidivism_cs_sample.sav**: これは、管轄地域での累犯率を把握するための、政府の法執行機関の取り組みに関する架空のデータ・ファイルです。各ケースは、2003 年の 7 月中に最初の逮捕から釈放された元犯罪者に対応し、この犯罪者の人口統計情報、最初の犯罪の詳細、2 回目の逮捕のデータ (2006 年 6

12. Breiman, L., and J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580-598.

13. Hastie, T., and R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

月末日までに逮捕された場合) が記録されています。犯罪者は、*recidivism_cs.csplan* で指定されたサンプリング計画に従って抽出された部門から選択されています。サンプリング計画では確率比例 (PPS) 法が使用されるため、結合選択確率が含まれているファイル (*recidivism_cs_jointprob.sav*) もあります。

- **rfm_transactions.sav:** これは、購入日、購入品目、各取引の金額など、購買取引データが含まれている架空のデータ・ファイルです。
- **salesperformance.sav:** これは、2 つの新しい販売トレーニング・コースの評価に関する架空のデータ・ファイルです。60 人の従業員を 3 つのグループに分け、全員が標準のトレーニングを受けます。さらに、グループ 2 は技術的なトレーニングを受け、グループ 3 は実践的なチュートリアルを受けます。トレーニング・コースの最後にすべての従業員がテストを受け、そのスコアが記録されました。データ・ファイルの各ケースは個々の受講生を表し、その受講生が割り当てられたグループと、テストのスコアが記録されています。
- **satisf.sav:** これは、ある小売業者が 4 箇所の店舗で行った満足度調査に関する架空のデータ・ファイルです。合計で 582 人の顧客について調査が行われ、各ケースは 1 人の顧客からの回答を表しています。
- **screws.sav:** このデータ・ファイルには、ねじ、ボルト、ナット、鉋 (びょう) ¹⁴ の特性に関するデータが含まれています。
- **shampoo_ph.sav:** これは、あるヘアケア製品工場での品質管理に関する架空のデータ・ファイルです。定期的に、6 つの異なる生産バッチが測定され、その pH が記録されます。目標の範囲は 4.5 から 5.5 です。
- **ships.sav:** これは、他の場所 ¹⁵ で表示および分析される、波が貨物船に与える損害に関するデータ・セットです。事故カウントは、船舶の種類、建造期間、サービス期間により、ポワゾン率で発生するものとしてモデル化することができます。各因子のクロス分類によって形成されたテーブルの各セルのサービス月数の集計により、危険にさらされる確率の値が得られます。
- **site.sav:** これは、業務拡大に向けて新たな敷地を選定することを目的とした、ある会社の取り組みに関する架空のデータ・ファイルです。2 人のコンサルタントを雇い、敷地を別々に評価させました。これらのコンサルタントは、詳細なレポートのほかに、各敷地を「良い」、「普通」、「悪い」のいずれかで集計しました。
- **smokers.sav:** このデータ・ファイルは、1998 年の National Household Survey of Drug Abuse から抽出した、アメリカの世帯の確率サンプルです (<http://dx.doi.org/10.3886/ICPSR02934>)。したがって、このデータ・ファイルを分析するには、人口の傾向を反映させるために、最初にデータの重み付けを行う必要があります。
- **stocks.sav:** この架空のデータ・ファイルには、1 年分の在庫の価格と量が含まれています。
- **stroke_clean.sav:** この架空のデータ・ファイルには、Data Preparation オプションのプロシーチャーを使用して整理した後の、医療データベースの状態が含まれています。
- **stroke_invalid.sav:** この架空のデータ・ファイルには、医療データベースの初期状態と、いくつかのデータ入力エラーが含まれています。
- **stroke_survival:** これは、虚血性脳卒中後のリハビリ・プログラムを終えた後に数回の困難に直面した患者の生存時間に関する架空のデータ・ファイルです。脳卒中後の心筋梗塞、虚血性脳卒中、または出血性脳卒中の発症と、その発症時刻が記録されています。このサンプルには、脳卒中後に実施されたりハビリ・プログラムの最後まで生存した患者だけが含まれているため、サンプルの左側は切り捨てられています。

14. Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.

15. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

- **stroke_valid.sav:** この架空のデータ・ファイルには、データの検証プロセスによる確認後の、医療データベースの状態が含まれています。異常ケースの可能性のあるケースも含まれています。
- **survey_sample.sav:** このデータ・ファイルには、人口統計データや各種の態度指標などの調査データが含まれています。これは「1998 NORC General Social Survey」の変数のサブセットに基づいていますが、一部のデータ値が変更され、いくつかの架空の変数がデモの目的で追加されています。
- **telco.sav:** これは、顧客ベースにおけるチャーン（契約/解約を繰り返す顧客の流動現象）を削減するための電気通信会社の取り組みに関する架空のデータ・ファイルです。各ケースが個々の顧客に対応し、人口統計やサービス利用状況などのさまざまな情報が記録されています。
- **telco_extra.sav:** このデータ・ファイルは *telco.sav* データ・ファイルと似ていますが、「期間」変数と対数変換された顧客支出変数が削除され、標準化された対数変換後の顧客支出変数に置き換えられています。
- **telco_missing.sav:** このデータ・ファイルは *telco.sav* データ・ファイルのサブセットですが、一部の人口統計データ値が欠損値で置き換えられています。
- **testmarket.sav:** この架空のデータ・ファイルは、ファースト・フード・チェーンでの新しいメニューの追加計画に関連しています。新製品をプロモーションするためのキャンペーンには 3 つの候補があるため、新メニューは、ランダムに選択されたマーケットのいくつかの場所で導入されます。場所ごとに別々のプロモーションを展開し、最初の 4 週間について、新メニューの週間売上高が記録されます。各ケースが、個々の場所と週に対応しています。
- **testmarket_1month.sav:** この架空のデータ・ファイルは、*testmarket.sav* データ・ファイルの週ごとの売上を「ロールアップ」して、各ケースを個々の場所に対応させたものです。そのため、週ごとに変わっていた変数の一部が表示されなくなり、記録された売上高が、調査を行った 4 週間の売上高の合計になっています。
- **tree_car.sav:** これは、人口統計と自動車購入価格のデータが含まれている架空のデータ・ファイルです。
- **tree_credit.sav:** これは、人口統計と銀行ローン履歴のデータが含まれている架空のデータ・ファイルです。
- **tree_missing_data.sav:** これは、多数の欠損値とともに、人口統計と銀行ローン履歴のデータが含まれている架空のデータ・ファイルです。
- **tree_score_car.sav:** これは、人口統計と自動車購入価格のデータが含まれている架空のデータ・ファイルです。
- **tree_textdata.sav:** これは、測定レベルと値ラベルを割り当てる前の変数のデフォルトの状態を示すことを主な目的とする変数が 2 つだけ含まれている単純なデータ・ファイルです。
- **tv-survey.sav:** これは、テレビ・スタジオで実施された、ヒットした番組の放送期間を延長するかどうかを検討する調査に関する架空のデータ・ファイルです。906 人の回答者に対して、さまざまな条件下でこの番組を視聴するかどうかを質問しました。各行は個々の回答者を表し、各列は個々の条件を表しています。
- **ulcer_recurrence.sav:** このファイルには、潰瘍の再発を防ぐための 2 つの治療の有効性を比較するための調査から得られた情報の一部が含まれています。これは区間打ち切りデータの好例で、他の場所¹⁶ で表示と分析が行われています。
- **ulcer_recurrence_recoded.sav:** このファイルでは、単に調査終了時のイベント確率ではなく調査の間隔ごとのイベント確率をモデル化できるように、*ulcer_recurrence.sav* の情報が再編成されています。これは、他の場所¹⁷ で表示と分析が行われています。

16. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

17. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

- **verd1985.sav:** このデータ・ファイルは、調査¹⁸に関連しています。8 つの変数に対する 15 人の被験者の回答が記録されています。対象となる変数は 3 つのグループに分類されています。グループ 1 には「年齢」と「婚姻」、グループ 2 には「ペット」と「ニュース」、グループ 3 には「音楽」と「居住地域」がそれぞれ含まれています。「ペット」は多重名義として尺度化され、「年齢」は順序として尺度化されます。他のすべての変数は、単一名義として尺度化されます。
- **virus.sav:** これは、自社のネットワーク上のウィルスの影響を特定するための、インターネット・サービス・プロバイダー (ISP) の取り組みに関する架空のデータ・ファイルです。ネットワーク上の感染した E メール・トラフィックの (およその) パーセンテージが、検出時点から脅威が阻止されるまでの間、追跡されています。
- **wheeze_stuebenville.sav:** これは、子供に対する大気汚染の健康上の影響に関する長期的な調査¹⁹ から得られたサブセットです。このデータには、オハイオ州スチューベンビルに住む 7 歳、8 歳、9 歳、10 歳の子供を対象に行った喘鳴の状態の反復 2 値測定と、調査の初年に母親が喫煙していたかどうかの固定的な記録が含まれています。
- **workprog.sav:** これは、体の不自由な人をより条件のよい仕事に就かせようとする政府の事業プログラムに関する架空のデータ・ファイルです。このプログラムの参加候補者のサンプルが追跡され、その中には、ランダムに選択されてプログラムに参加した人と、そうでない人がいます。各ケースが、個々のプログラム参加者を表しています。
- **worldsales.sav:** この架空のデータ・ファイルには、大陸別と製品別の販売収益が含まれています。

18. Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

19. Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, and B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, 366-374.

特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒103-8510

東京都中央区日本橋箱崎町19番21号

日本アイ・ビー・エム株式会社

法務・知的財産

知的財産権ライセンス渉外

以下の保証は、国または地域の法律に沿わない場合は、適用されません。IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態を提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Software Group

ATTN: Licensing

200 W. Madison St.

Chicago, IL; 60606

U.S.A.

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができますが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

この文書に含まれるいかなるパフォーマンス・データも、管理環境下で決定されたものです。そのため、他の操作環境で得られた結果は、異なる可能性があります。一部の測定が、開発レベルのシステムで行われた可能性があります。その測定値が、一般に利用可能なシステムのものと同じである保証はありません。さらに、一部の測定値が、推定値である可能性があります。実際の結果は、異なる可能性があります。お客様は、お客様の特定の環境に適したデータを確かめる必要があります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者にお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があります、単に目標を示しているものです。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

著作権使用許諾:

本書には、様々なオペレーティング・プラットフォームでのプログラミング手法を例示するサンプル・アプリケーション・プログラムがソース言語で掲載されています。お客様は、サンプル・プログラムが書かれているオペレーティング・プラットフォームのアプリケーション・プログラミング・インターフェースに準拠したアプリケーション・プログラムの開発、使用、販売、配布を目的として、いかなる形式においても、IBM に対価を支払うことなくこれを複製し、改変し、配布することができます。このサンプル・プログラムは、あらゆる条件下における完全なテストを経ていません。従って IBM は、これらのサンプル・プログラムについて信頼性、利便性もしくは機能性があることをほのめかしたり、保証することはできません。これらのサンプル・プログラムは特定物として現存するままの状態を提供されるものであり、いかなる保証も提供されません。IBM は、お客様の当該サンプル・プログラムの使用から生ずるいかなる損害に対しても一切の責任を負いません。

それぞれの複製物、サンプル・プログラムのいかなる部分、またはすべての派生的創作物にも、次のように、著作権表示を入れていただく必要があります。

© (お客様の会社名) (西暦年). このコードの一部は、IBM Corp. のサンプル・プログラムから取られています。

© Copyright IBM Corp. _年を入れる_. All rights reserved.

商標

IBM、IBM ロゴおよび ibm.com は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

Adobe、Adobe ロゴ、PostScript、PostScript ロゴは、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

インテル、Intel、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Centrino、Intel Centrino ロゴ、Celeron、Xeon、Intel SpeedStep、Itanium、および Pentium は、Intel Corporation または子会社の米国およびその他の国における商標または登録商標です。

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

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

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

Java およびすべての Java 関連の商標およびロゴは Oracle やその関連会社の米国およびその他の国における商標または登録商標です。

索引

日本語, 数字, 英字, 特殊文字の順に配列されています。なお, 濁音と半濁音は清音と同等に扱われています。

[ア行]

値

 カテゴリのラベルと値の表示 57
 値と値ラベル 57

[カ行]

カイ 2 乗

 カスタム・テーブル 67
各種変数の異なる要約統計量
 積み重ねテーブル 62
カスタム合計要約統計量 56
カスタム・テーブル
 カスタム合計 10
 カテゴリの省略 40
 カテゴリの除外 11, 24
 カテゴリのソート 24
 カテゴリの並べ替え 11
 カテゴリ変数 1
 カテゴリ変数の値ラベル 1
 カテゴリ変数の単純なテーブル 19
 カテゴリを共有する変数のテーブル 15, 49
 空のセル 16
 キャプション 17
 行パーセントと列パーセント 20
 行変数と列変数の交換 31
 クロス集計 21
 計算後のカテゴリ 14, 43
 計算対象カテゴリ 11, 14, 43
 検定統計量 17, 67
 コーナー・ラベル 17
 合計 11, 21, 35
 コンパクト・ビュー 30
 周辺合計 23
 小計 11, 35
 小計カテゴリの非表示 40
 除外されたカテゴリがあるテーブルでの合計 24
 スケール変数 1
 スケール変数のための欠損値の除外 16
 層にしたテーブルの印刷 34
 層変数 32, 33

カスタム・テーブル (続き)

 層変数のネスト 33
 測定レベルの変更 1
 タイトル 17
 多重回答セット 1, 79
 多重回答セットのパーセント 8
 多重カテゴリ・セット 16
 テーブルの作成 4
 統計ソース・ディメンション 22
 統計量のラベルの非表示 19
 度数表 15, 49
 パーセント 7, 8, 20, 22
 比較テーブル 15, 49
 表示形式 6
 表示する小数桁数の制御 6
 ファイルの分割処理 5
 平均度数分布表 10
 変数の積み重ね 27
 変数のネスト 28, 30
 変数名とラベルの表示と非表示 6
 有意差検定と多重回答 77
 要約統計量 7, 8, 9
 要約統計量ディメンションの変更 10
 要約統計量の表示形式 11
 要約統計量のラベルの変更 20
 列幅 16
カスタム・テーブル内の統計量のラベルの非表示 19
カテゴリ値の表示 57
カテゴリの削除
 カスタム・テーブル 11, 24
カテゴリの省略
 カスタム・テーブル 24, 40
カテゴリの除外
 カスタム・テーブル 11, 24
カテゴリのソート
 カスタム・テーブル 24
カテゴリの並べ替え
 カスタム・テーブル 11
空のセル
 カスタム・テーブルに表示される値 16, 94
キャプション
 カスタム・テーブル 17
グループ化した要約
 スケール変数 63
グループ合計 37
クロス集計
 カスタム・テーブル 21
計算後のカテゴリ
 カスタム・テーブル 14, 43

計算対象カテゴリ

 カスタム・テーブル 14, 43
 式に含まれるカテゴリの非表示 44
 小計から 45
 表示形式 15
欠損値 60, 87
 カスタム・テーブルに含める 88
 パーセント計算に与える影響 88
検定統計量
 カスタム・テーブル 17, 67
コーナー・ラベル
 カスタム・テーブル 17
合計
 カスタム・テーブル 9, 11, 21, 35
 カスタム・テーブルの周辺合計 23
 グループ合計 37
 除外したカテゴリ 36
 層 38
 ネストしたテーブル 37
 表示位置 36
合計 N 88

[サ行]

最小

 カスタム・テーブル 9
最大
 カスタム・テーブル 9
最頻値
 カスタム・テーブル 9
サブグループ合計 37
サンプル・ファイル
 場所 97
時刻
 カスタム・テーブルに現在の時刻を含める 17
システム欠損値 87
小計 39
 カスタム・テーブル 11, 35
 小計カテゴリの非表示 40
小数桁数
 カスタム・テーブルで表示する小数桁数の制御 6, 23, 91
スケール変数
 行および列カテゴリ変数でグループ化された要約 63
 グループ化した要約 63
 積み重ね 59
 ネスト 64
 複数の要約統計量 60
 要約統計量 59

層変数

- カスタム・テーブル 32, 33
- 層にしたテーブルの印刷 34
- 層変数の積み重ね 33
- 層変数のネスト 33
- 層を使用したテーブルの印刷 34
- 測定レベル
 - カスタム・テーブルでの変更 1

[タ行]

タイトル

- カスタム・テーブル 17
- 多重回答セット 79
 - 多重カテゴリ・セットの重複回答 16
- パーセント 8
- 有意差検定 77, 84, 85

中央値 60

- カスタム・テーブル 9

テーブル

- カスタム・テーブル 1

統計量

- カスタム合計要約統計量 56
- 積み重ねテーブル 55
- 要約統計量 53

度数

- 有効 N との対比 60

度数表

- カスタム・テーブル 15, 49

[ハ行]

パーセント

- カスタム・テーブル 7, 8, 20, 22
- 欠損値 88
- 多重回答セット 8

範囲

- カスタム・テーブル 9

比較テーブル 15, 49

日付

- カスタム・テーブルに現在の日付を含める 17

表示形式 23

- カスタム・テーブル内の要約統計量 11, 91

表示する小数桁数の制御 23

標準偏差

- カスタム・テーブル 9

ファイルの分割処理

- カスタム・テーブル 5

分散

- カスタム・テーブル 9

平均 59

平均値

- カスタム・テーブル 9

平均度数分布表 10, 56

変数の積み重ね

- 各種変数の異なる要約統計量 62
- カスタム・テーブル 27
- スケール変数 59
- 層変数の積み重ね 33
- 複数の要約統計量ソース変数 55

変数のネスト

- カスタム・テーブル 28, 30
- スケール変数 64

変数ラベル

- カスタム・テーブルでの表示の抑止 6

[ヤ行]

ユーザー欠損値 87

有意差検定

- カスタム・テーブル 17
- 多重回答セット 84, 85

有効 N 60, 88

- カスタム・テーブル 9

要約統計量 53

- カスタム合計要約統計量 56
- ソース変数 53
- ソース・ディメンション 53

積み重ねテーブル 55

- 積み重ねテーブルの各種変数の異なる要約 62

表示形式 91

- ラベル・テキストの変更 92

要約統計量ソース変数

- スケール変数 64

[ラ行]

ラベル

- 要約統計量のラベル・テキストの変更 92

列の比率の統計量

- カスタム・テーブル 72

列の平均値の統計量

- カスタム・テーブル 70

列幅

- カスタム・テーブルでの制御 16, 93



Printed in Japan