



0 はじめに

テキスト

芳賀敏郎（2011）医薬品開発のための統計解析
第1部 基礎 改訂版、サイエンティスト社、p.275



第1部 基礎

- 1. 統計の基礎
 - 1.1 宝くじの期待値と分散、1.2 サイコロの目の数の期待値と分散
 - 1.3 分散の加法性・中心極限定理・正規分布、1.4 統計的推測、1.5 モデル
- 2. 1組のデータの解析 .
 - 2.1 データの特徴の記述、2.2 データのグラフ表示と外れ値
 - 2.3 対数変換と対数正規分布、2.4 平均に関する推測（母標準偏差 σ 既知）
 - 2.5 分散に関する推測、2.6 平均に関する推測（母標準偏差 σ 未知）
- 3. 2組のデータの解析 .
 - 3.1 データのグラフ化、3.2 平均値の差の t 検定、3.3 分散の違いの検定
 - 3.4 分散が異なる場合の平均値の差の比較、
 - 3.5 対応のある場合の平均値の差の t 検定、3.6 検出力と n の決め方
 - 3.7 ノンパラメトリック検定
- 4. 相関・回帰
 - 4.1 散布図、4.2 相関係数、4.3 回帰モデルとモデルの推定
 - 4.4 誤差を考慮した推定、4.5 回帰分析適用上の諸問題



0 はじめに

p.1

0.1 統計解析用ソフトウェア

0.2 学び方

0.3 統計解析とデータ解析

テキストの
該当ページ

補足 「R を使った学習」

使用するファイル

Excel ファイル「基本改1.xls」

サイエンティスト社ホームページからダウンロード

★プレゼンテーションの
スピーカーノートを、
PDF の注釈に変換してあります



0.1 統計解析用ソフトウェア

このテキストで使用する統計解析用ソフトウェア



●Excel（エクセル）

表計算プログラム（事務計算～技術計算）

セルに文字・数値や計算式を入力し、四則演算から高度な計算までを実行できる

統計解析の計算過程を追いかけて、理解を深めることができる

実際に、統計の実務で利用するには問題がある（間違いの修正、公的な申請での制約等）

●JMP（ジャンプ）

統計解析専用のプログラムの一つ、SAS Institute Japan から提供（日本語に対応）

データのグラフ化機能が豊富（グラフを見て考え、解析方法を選択、さらに解析を進展）

試行錯誤を反復する解析（探索的解析）に適している

実務家が持つ固有技術をフルに生かせる

●R（アール）

統計解析専用のプログラムの一つ（巻末で補足）



●Excel（エクセル）

表計算プログラム（事務計算～技術計算）

セルに文字・数値や計算式を入力し、四則演算から高度な計算までを実行できる

統計解析の計算過程を追いかけて、理解を深めることができる

実際に、統計の実務で利用するには問題がある（間違いの修正、公的な申請での制約等）

●JMP（ジャンプ）

統計解析専用のプログラムの一つ、SAS Institute Japan から提供（日本語に対応）

データのグラフ化機能が豊富（グラフを見て考え、解析方法を選択、さらに解析を進展）

試行錯誤を反復する解析（探索的解析）に適している

実務家が持つ固有技術をフルに生かせる

●R（アール）

統計解析専用のプログラムの一つ（巻末で補足）



使用するソフトウェア

●Excel（エクセル）

表計算プログラム（事務計算～技術計算）

セルに文字・数値や計算式を入力し、四則演算から高度な計算までを実行できる

統計解析の計算過程を追いかけて、理解を深めることができる

実際に、統計の実務で利用するには問題がある（間違いの修正、公的な申請での制約等）

●JMP（ジャンプ）

統計解析専用のプログラムの一つ、SAS Institute Japan から提供（日本語に対応）

データのグラフ化機能が豊富（グラフを見て考え、解析方法を選択、さらに解析を進展）

試行錯誤を反復する解析（探索的解析）に適している

実務家が持つ固有技術をフルに生かせる

●R（アール）

統計解析専用のプログラムの一つ（巻末で補足）

ものを作ったり、サービスを提供するときに必要な技術
それぞれの分野における専門技術、専門知識



●Excel の有用な機能

分析ツール	一連の統計計算を自動的に実行
ゴールシーク	先に計算結果を指定して、数値を逆算する機能
ソルバー	同上、機能がより豊富
データテーブル	同じ計算を繰り返して、計算結果をテーブル形式で表示
BVA	専用のプログラミング言語（Visual Basic for Application） ユーザー定義関数を作成、マクロ機能の作成、複雑な計算、自動化

●Excel の関数

多様な機能をもつ統計関数が整備（p.68, p.114 参照）

Excel 2007 以前の古い関数、Excel 2010 以降の新しい関数の 2 種類

●Excel ヒント

目次の最期のページ viii に一覧表

適宜、今後の学習の中で身に着ける



●Excel の有用な機能

分析ツール	一連の統計計算を自動的に実行
ゴールシーク	先に計算結果を指定して、数値を逆算する機能
ソルバー	同上、機能がより豊富
データテーブル	同じ計算を繰り返して、計算結果をテーブル形式で表示
BVA	専用のプログラミング言語（Visual Basic for Application） ユーザー定義関数を作成、マクロ機能の作成、複雑な計算、自動化

●Excel の関数

多様な機能をもつ統計関数が整備（p.68, p.114 参照）

Excel 2007 以前の古い関数、Excel 2010 以降の新しい関数の 2 種類

●Excel ヒント

目次の最期のページ viii に一覧表

適宜、今後の学習の中で身に着ける



- 数年ごとにバージョンアップ

テキストはJMP 9.02（2011年）を使用、スライドはJMP 10.02（2012）を使用
JMP 11（2013年）、JMP 12（2015年）、JMP 13（2016年）・・・

- JMP ヒント

目次の最期のページ viii に一覧表

- 変数（実験の要因と結果）を3つの尺度に分類

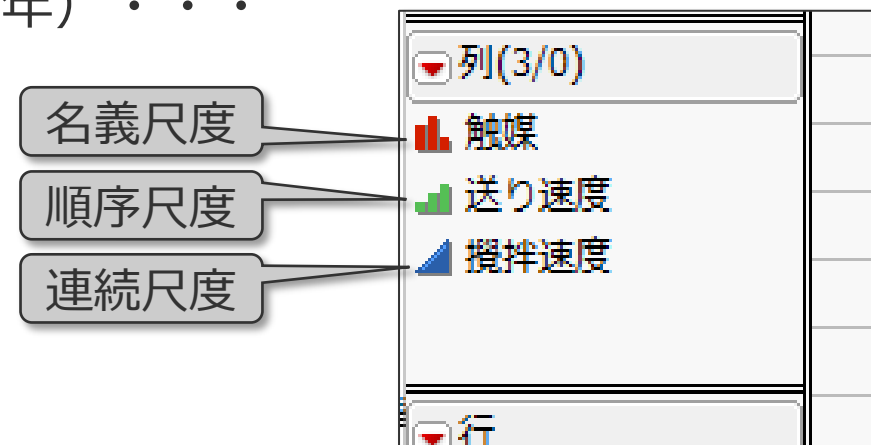
変数	尺度	例
量的変数	連続尺度 (Continuous)	身長、体重、温度、時間、濃度、年齢
質的変数 (カテゴリカル変数)	名義尺度 (Nominal)	反応 (-, +)、処理区 (A, B, C)
	順序尺度 (Ordinal)	反応 (-, +, ++)、サイズ (大、中、小)

●数年ごとにバージョンアップ

テキストは JMP 9.02 (2011年) を使用、スライドは JMP 10.02 (2012) を使用
 JMP 11 (2013年)、JMP 12 (2015年)、JMP 13 (2016年) . . .

●JMP ヒント

目次の最期のページ viii に一覧表



●変数 (実験の要因と結果) を3つの尺度に分類

変数	尺度	例
量的変数	連続尺度 (Continuous)	身長、体重、温度、時間、濃度、年齢
質的変数 (カテゴリカル変数)	名義尺度 (Nominal)	反応 (-, +)、処理区 (A、B、C)
	順序尺度 (Ordinal)	反応 (-, +, ++)、サイズ (大、中、小)

質的変数 名義尺度

量的変数 連続尺度

薬剤	体重
薬剤A	10.7
薬剤A	11.0
薬剤A	10.8
薬剤B	11.4
薬剤B	10.7
薬剤B	10.9
薬剤B	11.3

量的変数 連続尺度

量的変数 連続尺度

濃度	体重
10	10.7
10	11.0
10	10.8
20	13.2
20	12.4
20	13.8
30	15.9
30	14.6
30	16.8
40	20.1
40	22.1
40	25.1

変数の尺度と、その組み合わせに応じて、JMPが適切なグラフや解析方法を自動的に選択し、結果を表示してくれる。尺度の設定は非常に重要。



0.2 学び方

本テキストを使った学び方のポイント



- 「テキスト」を読むだけでは十分な理解は困難
- 学習用の「Excel ファイル」と「JMP ファイル」を使う
 - サイエンティスト社の[ホームページ](#)からファイル入手（無料）
 - なお、ExcelとJMPが使用できる環境であることを前提（JMPは体験版も利用可）
 - Excel ファイルで解析の考え方、手順を理解する
 - 同じデータをJMPで解析し、両者の結果を見比べて、JMPの出力の意味を理解
 - JMPの解析結果を見る
 - どのような考え方で求められたかをExcelで理解
- Excelによる「シミュレーション」を使って統計の基本的事項を理解
 - シミュレーションの面白さを体験
- 「演習問題」を自分で解く



- 「テキスト」を読むだけでは十分な理解は困難
- 学習用の「Excel ファイル」と「JMP ファイル」を使う
 - サイエンティスト社の[ホームページ](#)からファイル入手（無料）
 - なお、ExcelとJMPが使用できる環境であることを前提（JMPは体験版も利用可）
 - Excel ファイルを操作して、解析の考え方、手順を理解する
 - 同じデータをJMPで解析し、両者の結果を見比べて、JMPの出力の意味を理解
 - JMPの解析結果を見る
 - どのような考え方で求められたかをExcelで理解
- Excelによる「シミュレーション」を使って統計の基本的事項を理解
 - シミュレーションの面白さを体験
- 「演習問題」を自分で解く

Rでの代替が可能



入手する Excel ファイルと JMP ファイル

第 1 部

Excelファイル	JMPファイル
基礎改1.xls	2-1群1.JMP
基礎改2.xls	2-演習.JMP
基礎改3.xls	3-2群1.JMP
基礎改4.xls	3-2群2.JMP
基礎Macros.xls	3-2群ノンパラ.JMP
	3-ROC.jsp
	3-演習2.JMP
	4-RankCorr.JMP
	4-ビッグ クラス.JMP
	4-演習8.JMP
	4-演習データ.JMP
	4-相関1.JMP
	4-相関2.JMP
	4-相関3.JMP
	永田3_5_7.JMP
	永田4_6.JMP
	永田本.pdf
	永田本.xls

第 2 部

Excelファイル	JMPファイル	JMPファイル
DE改1- 1 因子(質).xlsm	1- 1 因子1.JMP	7-枝分れs.jsp
DE改2- 1 因子(量).xlsm	1- 1 因子2.JMP	7-変量.jsp
DE改3-乱塊法.xlsm	1- 1 因子3.JMP	7-乱塊法.jsp
DE改4-共分散.xlsm	2- 1 因子1.JMP	
DE改5-2因子.xlsm	2- 1 因子2.jsp	
DE改6-多因子.xlsm	3-乱塊法.JMP	
DE改7-変量.xlsm	4-共分散1.JMP	
DEマ如.xlsm	4-共分散2.JMP	
DE改0-はじめに.xlsx	4-共分散3.JMP	
	4-群分け.jsp	
	5-2因子1.JMP	
	5-2因子2.JMP	
	5-2因子4.JMP	
	6-応答曲面.jsp	
	6-直交表.JMP	
	7-経時.jsp	
	7-交差2.JMP	
	7-交差3.JMP	

第 3 部

Excelファイル	JMPファイル	JMPファイル	JMPファイル
改1非線形.xlsm	12-逆推定.JMP	34-4x4.jsp	演13-2.jsp
改2非線形.xlsm	12-単回帰.JMP	35-4x4.jsp	演14-1.jsp
改3計数値.xlsm	13-指数曲線.JMP	35-4x4Ex.jsp	演22-3.jsp
改4Log回帰.xlsx	13-指数曲線2.jsp	36-(2).jsp	演23-3.jsp
	14-Logist1.JMP	36-虫歯.jsp	演33-2.jsp
	14-Logist2.JMP	36-複数.JMP	演36-1.jsp
	21-指数曲線.jsp	41-二変量.jsp	演41-1.jsp
	21-不等分散.jsp	42-2次式A.jsp	演42-1.jsp
	21-不等分散2.jsp	42-2次式B.jsp	演42-2.jsp
	22-効力比.jsp	42-Log回帰A.JMP	
	23-併用効果.jsp	42-Log回帰B.jsp	
	24-複数曲線.jsp	43-a.jsp	
	25-PK1.jsp	43-Log個々.jsp	
	25-PK2.jsp	43-効力比.jsp	
	32-割合.JMP	43-車の調査.jsp	
	33-(11).jsp	43-多項ロジ.jsp	
	33-割合の差.jsp	44-ロバスト回帰.jsp	
	34-2xb.jsp	演12-2.jsp	



テキストに掲載されている図表

R での代替が可能

- テキストの図表の内容を Excel と JMP で確認

テキストの図表のほとんどは、Excel ファイルと JMP ファイルで提供される

図表の元になっているデータを確認し、計算過程を理解し、計算の意味を追いかける

- 図と表 → 表示

テキストの図表は、「表示 0.0.0」のように通し番号で識別

- 「テキスト」、「Excelファイル」、「JMPファイル」をフルに使って学習

Excelファイルと JMP ファイルを使い、手を動かしてテキストの内容を理解する

R スクリプトファイルで代替も可



●Excel ファイルの読み込み

入手したExcel ファイルには VBA マクロが含まれる

Excel では、「マクロを有効にする」ための設定を行う

Excel のセキュリティ管理の設定が必要になる場合がある（ウィルスの対策への対応 p.52）

●Excel ファイル・シートのコピー

学習を進めていくうえで、データや数式を書き換える場合があるため、元の内容を保存

Excel のファイルを、学習用として別名で保存して使用

Excel のシートを、学習用として別名でコピーして使用

（シート同士で参照している場合があるので、一部のシートのみでの保存はしない）

●名前ボックスの利用

テキストの「表示 0.0.0」に対応したExcel の該当箇所へ移動する

●Excelファイルの読み込み

Excelファイル「基本改1.xls」

名前ボックスから「表示1.3.3」 (Fig13_03) を選択

テキストの図と表
→ 表示 0.0.0



名前ボックスの
▼をクリック

Fig13_03

ファイル ホーム 挿入 ページレイアウト

MS 明朝

B I U

クリップボード

N31 ▼

Fig11_01

Fig11_02

Fig11_03

Fig11_04

Fig11_05

Fig12_01

Fig12_02

Fig13_01

Fig13_02

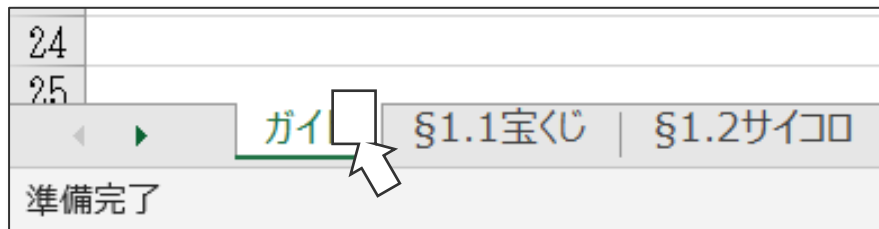
Fig13_03

Fig13_04

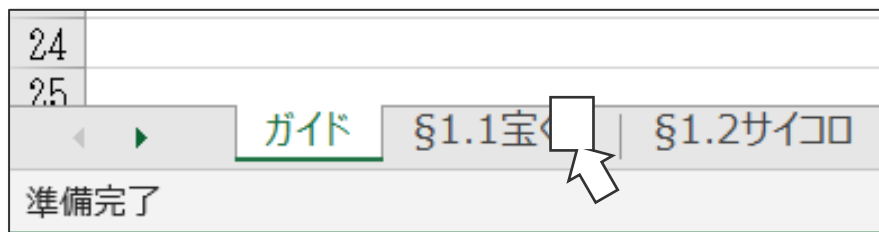
	E	F	G
1 ~ 6 の目の出る確			
0	400	500	600
3	0.188	0.186	0.177
7	0.155	0.164	0.166
7	0.143	0.146	0.155
3	0.170	0.170	0.166
7	0.163	0.156	0.155
3	0.183	0.178	0.183

●シートの移動とコピー (p.6)

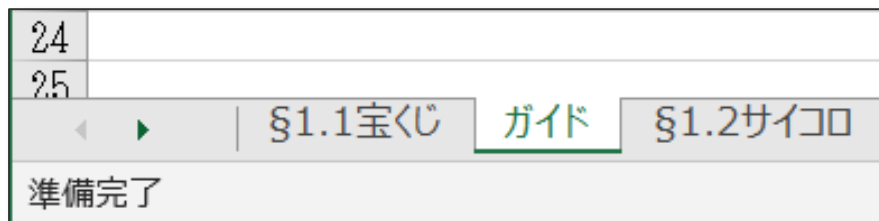
移動



シートをクリック

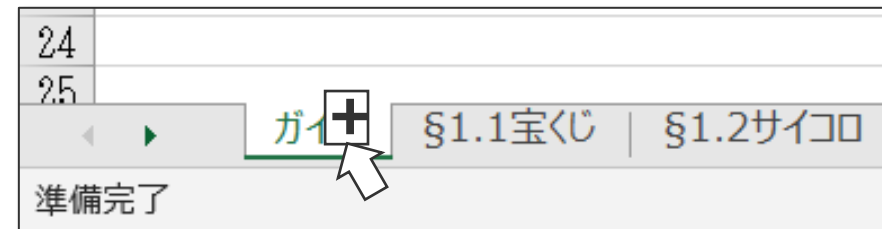


目的の位置にドラッグ

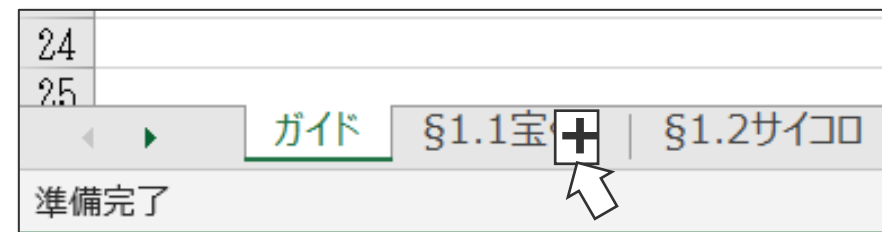


シートが移動

コピー



Ctrlキーを押しながらシートをクリック



目的の位置にドラッグ



シートがコピー (こちらのシートで学習)

●ワークシート分析（セルの参照元と参照先を表示）

表示 1.1.4 のセル014を選択

トップメニュー[数式]
> [ワークシート分析]

[参照元のトレース]
[参照先のトレース]
を選択して表示

[トレース矢印の削除]
を選択して消去

表示1.1.4

	I	J	K	L	M	N	O
12	等	¥	本数	偏差	確率		
13	i	x_i	n_i	d_i	π_i	$d_i * \pi_i$	$d_i^2 * \pi_i$
14	1等	10,000	10	9,930	0.0010	9.93	98,605
15	2等	1,000	200	930	0.0200	18.60	17,298
16	3等	100	4,000	30	0.4000	12.00	360
17	はずれ	0	5,790	-70	0.5790	-40.53	2,837
18	合計	0	10,000		1.0000	0.00	119,100

セル014を選択

数式 データ 校閲 表示 ヘルプ 実行したい作業を入力してください

名前
の管理

名前の定義
数式で使用
選択範囲から作成
定義された名前

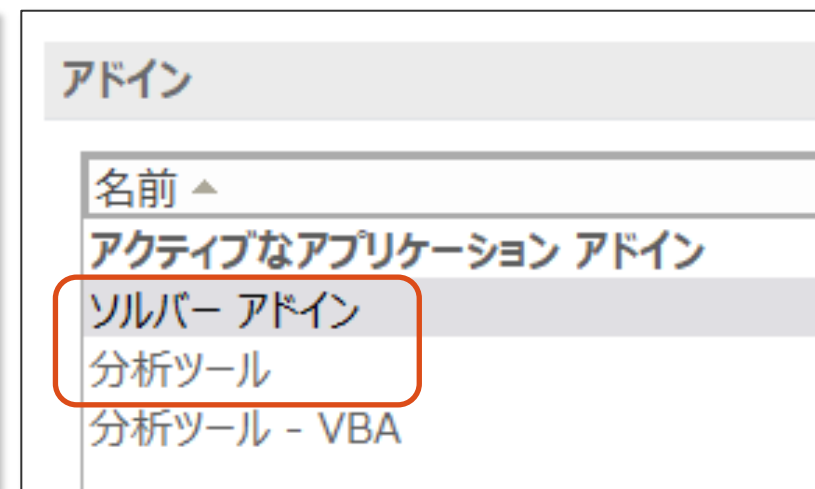
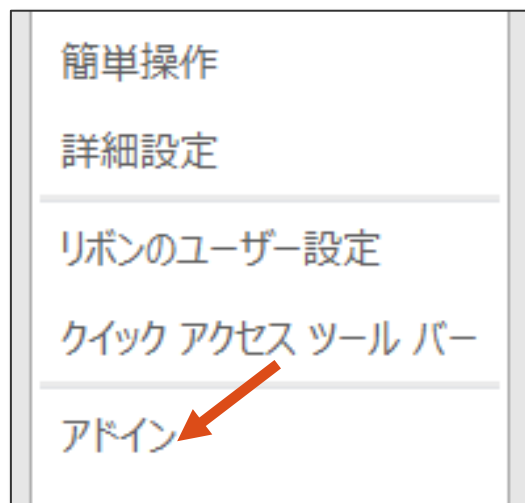
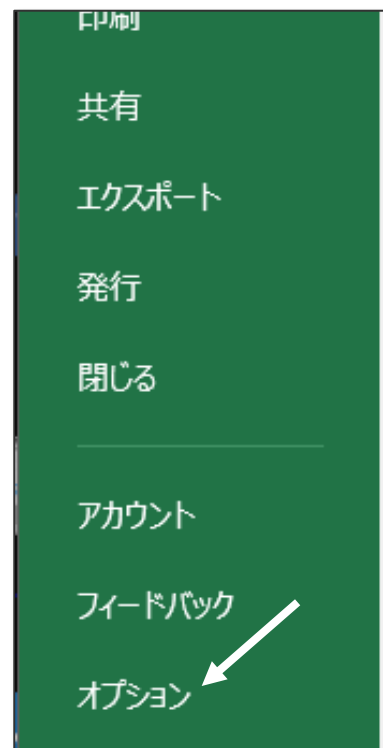
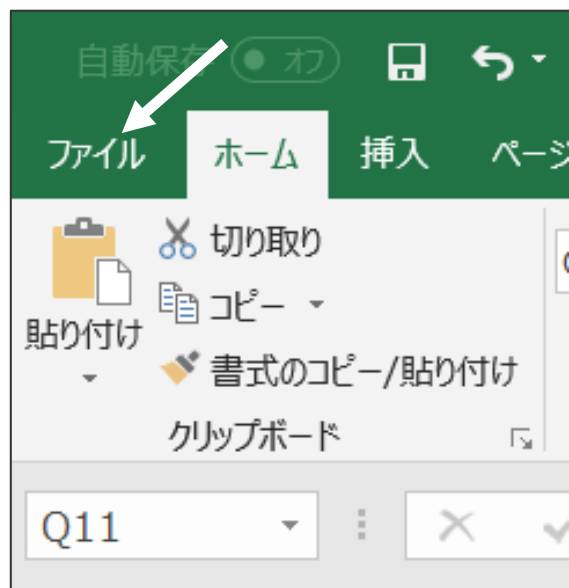
参照元のトレース
参照先のトレース
トレース矢印の削除
ワークシート分析

数式の表示
エラー チェック
数式の検証

●ソルバー、分析ツールの確認

[ファイル] > [オプション] > [アドイン]

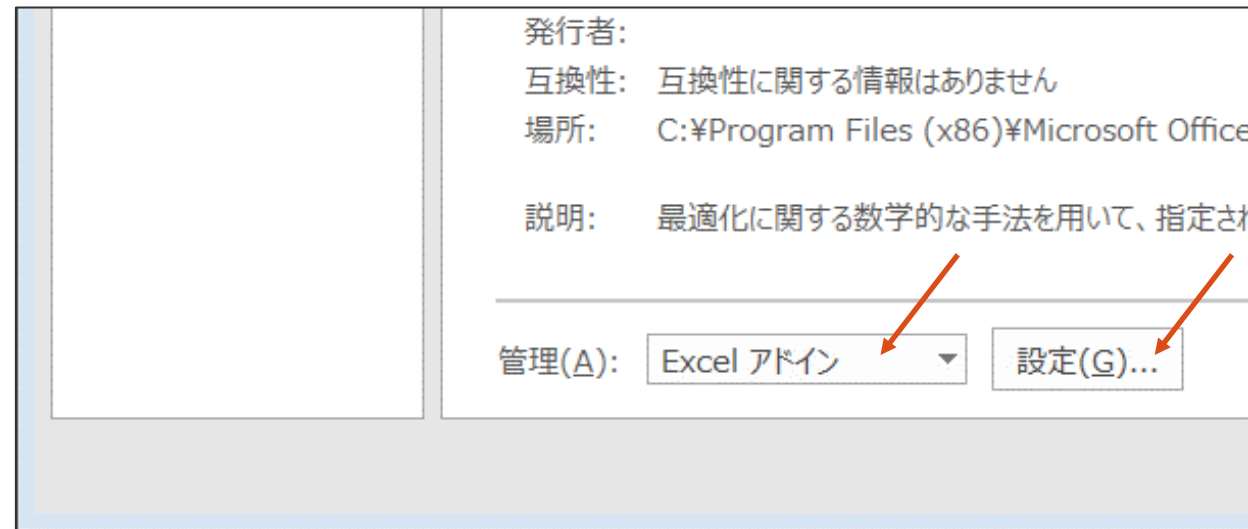
p.87, p.235参照



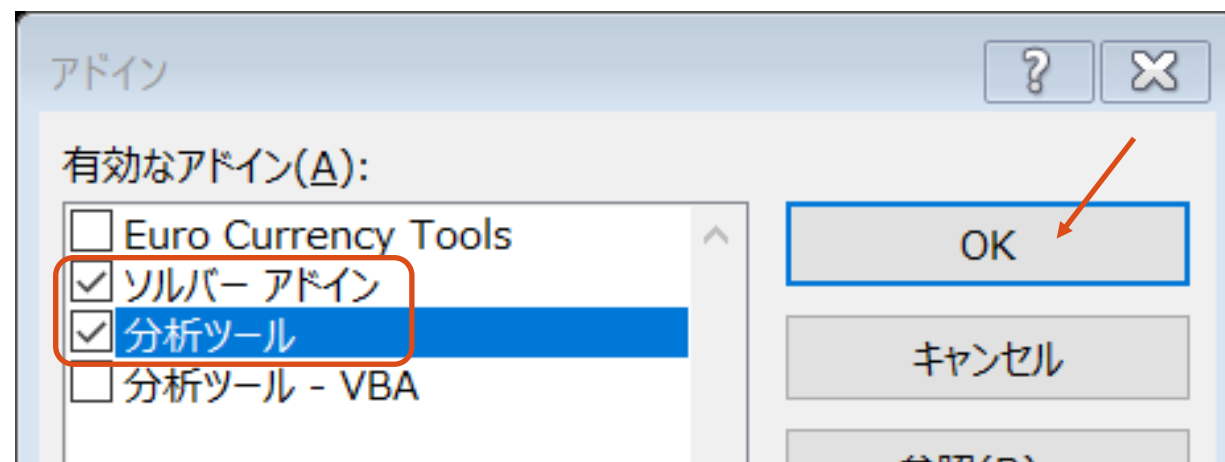
[アクティブなアプリケーション アドイン] に
[ソルバーアドイン]、[分析ツール] が
あれば設定不用
無ければ、Excelアドインをインストール

●ソルバー、分析ツールのインストール

[Excelアドイン] + [設定]



[ソルバーアドイン]
[分析ツール] に
チェックを入れて [OK]



●テキストと Excel ファイルの数値を確認する上での留意点

テキストでは、一連の計算を進めていく中で、計算過程の値を適当な桁に四捨五入して表示

実際は、Excel は表示されない桁の数値も使って次の計算を実行

テキストの計算過程の値を使ってその先を計算すると、値が一致しない場合がある

(a/b) × c を求める

a	12.8
b	2.3
c	9.1

テキスト	a / b =	5.57
	(a / b) × c =	50.64

Excel	a / b =	5.57
	(a / b) × c =	50.64

手計算	5.57 × c =	50.69
-----	------------	-------

Excel が保持している値

a / b = 5.56521739130435

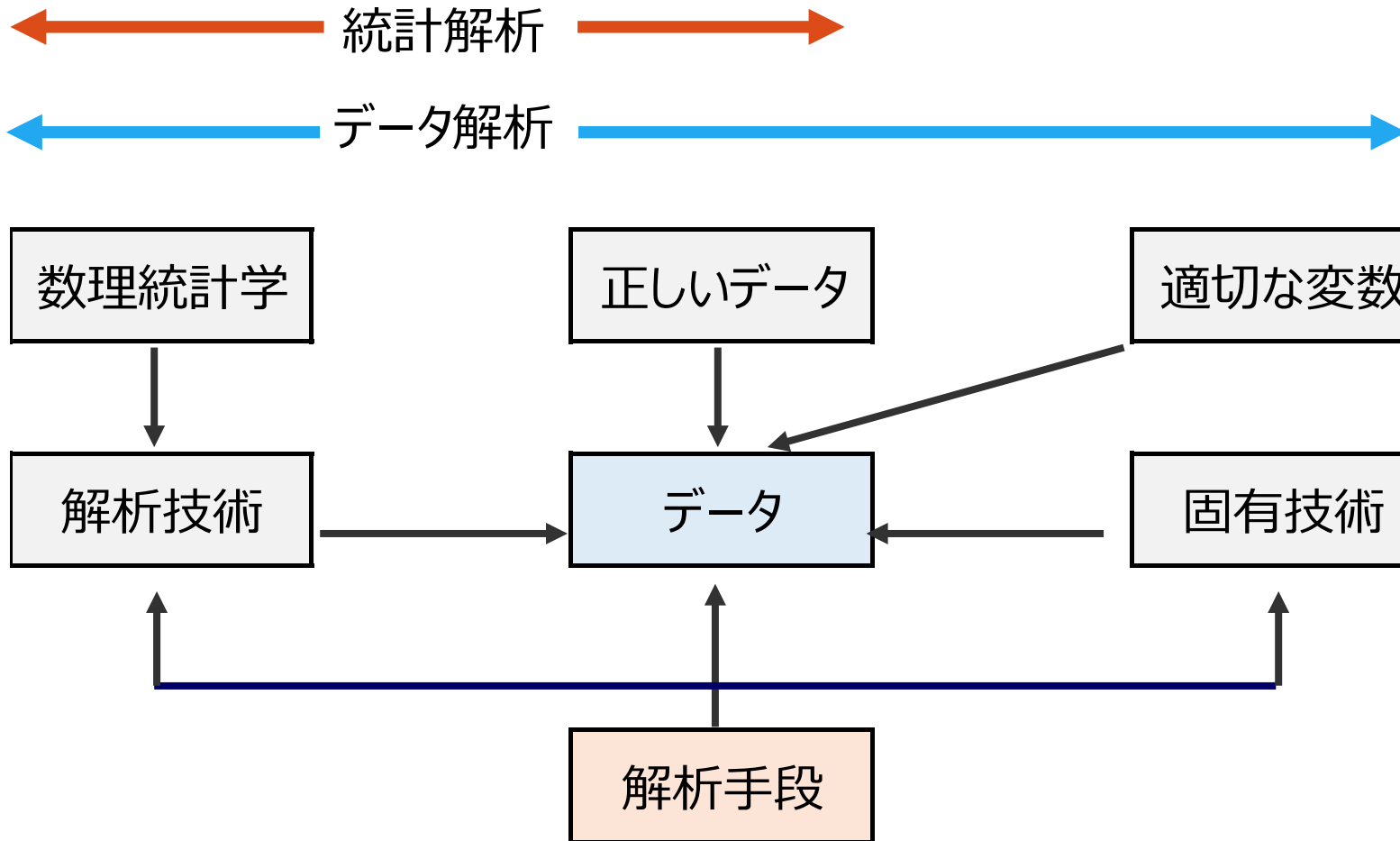
50.64 と一致しない

5.570000



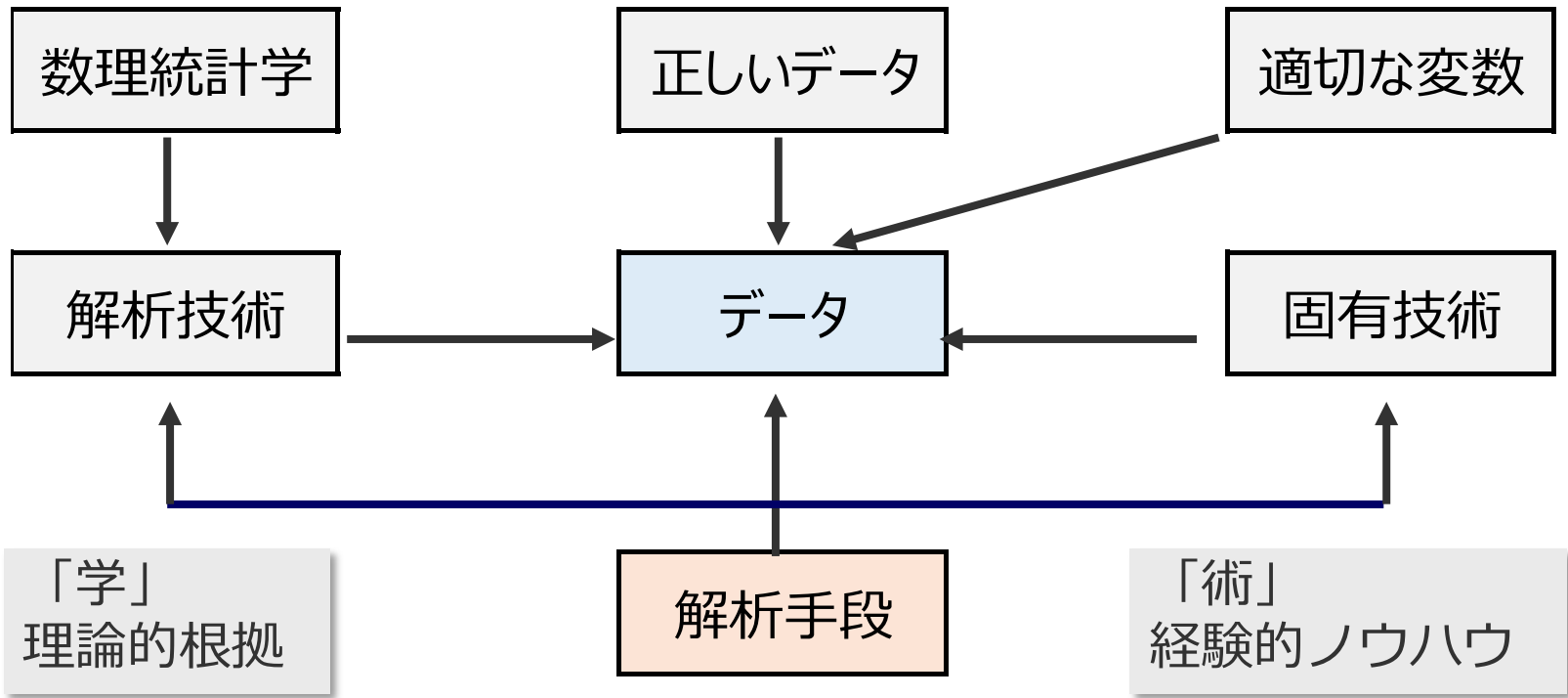
0.3 統計解析とデータ解析

表示0.3.1 統計解析とデータ解析



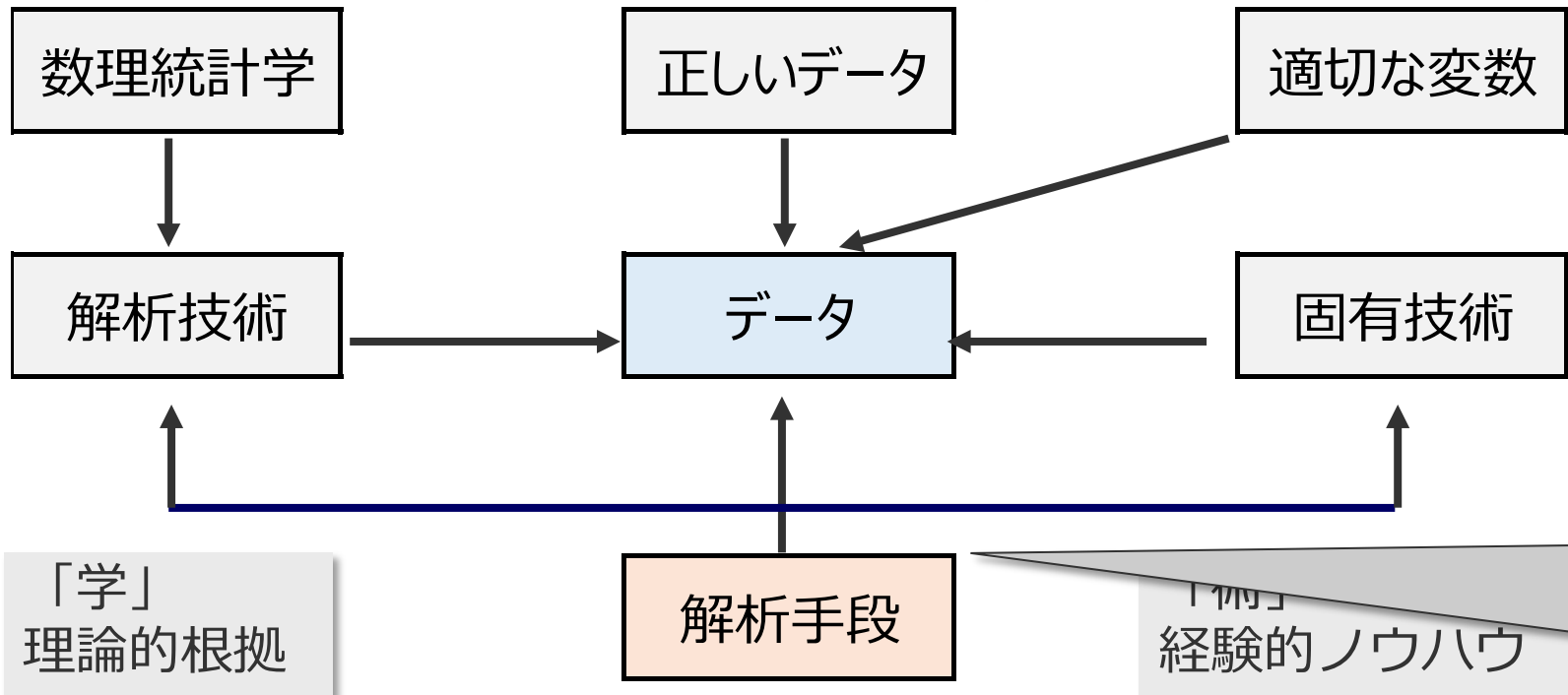
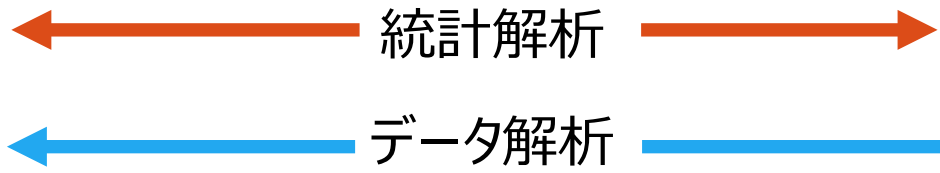
- 統計解析
データは x_{ij} ,
 i 番目の変数の
 j 番目のサンプルの観測値
と抽象化して解析
- データ解析
 i 番目の変数が何であり,
他の変数とどのような
関連があるか
 j 番目のサンプルは
どのような状態で観測
されたものであるか
イメージをもって解析

表示0.3.1 統計解析とデータ解析



解析手順のマニュアル化は不可能
経験と工夫を積み重ねて身に付けられる

表示0.3.1 統計解析とデータ解析



データが正しくなければ (誤りが含まれていれば), 誤った解析結果が得られる データが正しいか、確認

解析目的に対して 適切な変数が含まれていなければならない
たいへん困難な問題であり 豊富で深い固有技術が必要

データを 解析技術と固有技術から 追及するために使う道具 使いこなせることが必須
ここでは、Excel、JMP、R を利用



補足 Rを使った学習

JMP の代わりに R を使って学習を進める



補足：R を使った学習

●R と RStudio による統計解析

- R : 統計処理に特化したプログラミング言語、統計解析環境、フリーソフトウェア
- RStudio : R を利用するための IDE (Integrated Development Environment、統合開発環境)

●R と RStudio の利用方法

以下の PDF ファイルで使い方を習得できる

「R と RStudioの使い方入門－ 1 」 [PDF ファイル](#) (6 MB)

「R と RStudioの使い方入門－ 2 」 [PDF ファイル](#) (6 MB)

R と RStudio の初期設定と操作方法に限定した解説

R と RStudio を使うために、これらを十分学習することは必須

補足：Rを使った学習

●提供する2種類のPDFファイルとRスクリプトファイル

本サイトでは、

「テキストを解説したPDF」、「Rに関するPDF」、「Rスクリプトファイル」を提供します。

インターネットに表示される本サイト（テキストの目次）

0 はじめに

1 統計の基礎

1.1 宝くじの期待値と分散 Rの補足

1.2 サイコロの目の数の期待値と分散 Rの補足

1.3 分散の加法性、中心極限定理、正規分布 Rの補足

1.4 統計的推測 Rの補足

1.5 モデル

Rスクリプトのダウンロード

目次の項目をクリック

「テキストを解説したPDF」をダウンロード
(Excel と JMP によるテキストの解説)

「Rの補足」をクリック

「Rに関するPDF」をダウンロード
(その目次に対応したRの解説)

「Rスクリプトのダウンロード」をクリック

Rスクリプトファイルとデータファイルを
一括してダウンロード

補足：R を使った学習

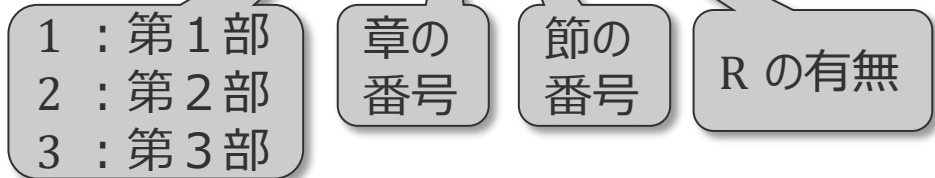
●提供する2種類のPDFファイル

「テキストを解説したPDF」の例
(Excel と JMP を使用)

Green1-1-1.pdf

「Rに関するPDF」の例

Green1-1-1R.pdf



ファイル名の番号は共通

「R」の文字の有無で区別

テキストの目次	提供するPDF	
	テキストを解説したPDF (Excel と JMP)	Rに関するPDF
1. 統計の基礎		
1.1 宝くじの期待値と分散	Green1-1-1.pdf	Green1-1-1R.pdf
.....
2. 1組のデータの解析		
2.1 データの特徴の期熟	Green1-2-1.pdf	Green1-2-1R.pdf
2.2 データのグラフ表示と外れ値	Green1-2-2.pdf	Green1-2-2R.pdf
.....
3. 2組のデータの解析		
3.1 データのグラフ化	Green1-3-1.pdf	Green1-3-1R.pdf
3.2 平均値の差のt検定	Green1-3-2.pdf	Green1-3-2R.pdf
.....

補足：R を使った学習

●提供する2種類のPDFファイル

テキストを解説したPDF (Excel と JMP)

テキストの
該当ページ

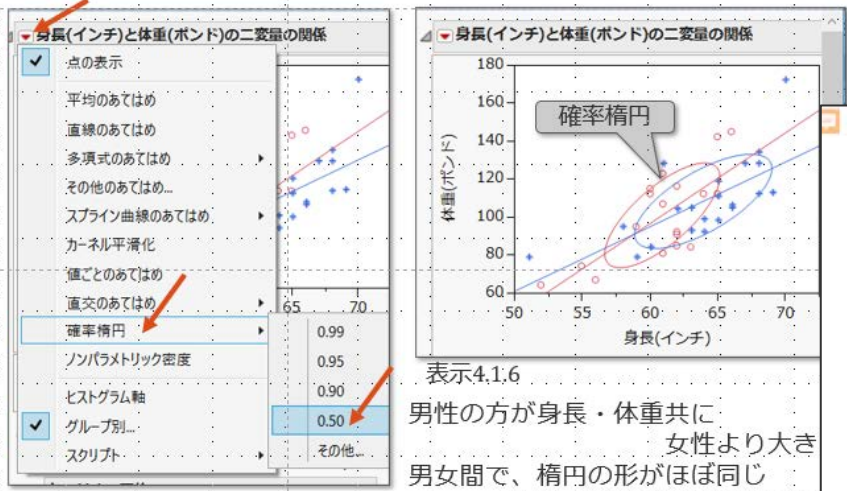
自身でRを操作して出力結果を得る
テキストのExcelとJMPの出力結果と比較して
Rの使い方を確認する

JMPの層別散布図 p.203

●確率楕円
オプションメニュー
▼> [確率楕円]
> [0.50]

回帰分析
→ 回帰直線

相関分析
→ 確率楕円



表示4.1.6
男性の方が身長・体重共に
女性より大き
男女間で、楕円の形がほぼ同じ
→ 身長と体重の関係はほぼ同

Rに関するPDF

テキストの
該当ページ

層別散布図

p.203

●表示4.1.6 JMPによる散布図と層別散布図

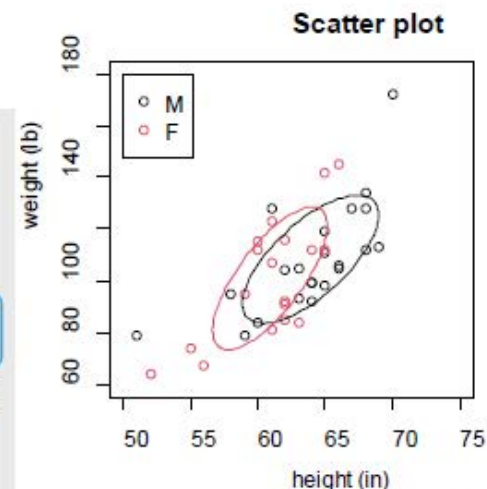
スクリプトファイル：Green1-4-1a.R

利用した関数：plot、subset、polygon、ellipse::ellipse

方法：確率楕円の表示

```
df_m <- subset(df, sex == "M",  
              select = c(height, weight))  
df_f <- subset(df, sex == "F",  
              select = c(height, weight))  
  
polygon(  
  ellipse(cov(df_m), # 楕円のデータ  
          centre = colMeans(df_m), # 楕円を描く中心  
          level = 0.50), # 確率楕円の信頼率  
  lty = 1, # 線の種類 1は実線  
  lwd = 1, # 線の太さデフォルト  
  border = 1) # 境界線の色
```

分散共分散行列



11

補足：Rを使った学習

●提供する2種類のPDFファイル

テキストを解説したPDF (Excel と JMP)

テキストの
該当ページ

自身でRを操作して出力結果を得る
テキストのExcelとJMPの出力結果と比較して
Rの使い方を確認する

テキストの
該当ページ

JMP [二変量の関係] p.31

●分散分析表

表示 1.1.7 解析 (2) 効果と分散分析表

要因	平方和	自由度	平均平方	F 比	p 値
水準間	3.10	3	1.033	7.584	0.0022
残差	2.18	16	0.136	1.000	
全体 (検算)	5.28	19	0.278		

表示 1.1.10 JMP のデータ表と [二変量の関係] の出力

要因	自由度	平方和	平均平方	F 値	p 値(Prob>F)
水準間	3	3.100000	1.03333	7.5841	0.0022
残差	16	2.180000	0.13625		
全体(修正済み)	19	5.280000			

全体：個々の観測値の2乗和
全体(修正済み)：個々の観測値と
総平均の差の2乗和

* : P < 0.05 を示す

Rに関するPDF

あてはめの要約

p.31

●表示1.1.10 JMPのデータ表と[二変量の関係]の出力

スクリプトファイル：Green2-1-1.R

利用した関数：lm、summary

lmの出力(一部)

```
## Residuals:
##   Min     1Q   Median     3Q      Max
## -0.600 -0.225  0.000  0.200  0.600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.3000    0.1651  62.396 < 2e-16 ***
## groupA2      0.5000    0.2335   2.142  0.04795 *
## groupA3      0.9000    0.2335   3.855  0.00140 **
## groupA4      1.0000    0.2335   4.284  0.00057 ***
##
## Residual standard error: 0.3691 on 16 degrees of freedom
## Multiple R-squared:  0.5871, Adjusted R-squared:  0.5097
## F-statistic: 7.584 on 3 and 16 DF,  p-value: 0.002242
```

Excelの分散分析表

JMPの分散分析表

Rの分散分析表

JMP「あてはめの要約」



- 参考文献

芳賀・永田（2009）データ解析に役立つ Excel 関数,
日科技連

- 作成

片瀬雅彦

- 監修

松本一彦、長谷文雄

- 作成時期

2018年8月21日

- 改訂

2019年3月15日、2024年7月12日