

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
2	1.1	図1.1 米国のマテリアルズゲノム・イニシアティブ
5	1.1	図1.2 材料設計手法の変遷
6	1.1	表1.1 現場の研究者・技術者がケモ・マテリアル・インフォマティクスに期待していること
7	1.1	図1.3 ケモ・マテリアル・インフォマティクスを構成する3つの分野
7	1.1	図1.4 ニコチン分子の化学式、化学構造、SMILES表記
8	1.1	表1.2 利用者視点から見た R言語とPython言語の比較
8	1.1	図1.5 母集団と標本の関係
9	1.1	表1.3 母平均の検定の手順
10	1.2.2	図1.6 Rプログラミングでよく利用されるグラフィックス
12	1.2.3	表1.4 3種類の離散型の確率分布の確率関数、期待値、分散
13	1.2.5	図1.7 7組のデータについての相関分析の結果
14	1.2.6	図1.8 主成分分析・多次元尺度構成法・自己組織化マップの分析結果
15	1.2.7	図1.9 分割法(k-平均法)と階層法(凝集法)の分析結果
16	1.2.8	図1.10 線形重回帰分析の例
16	1.2.9	表1.5 機械学習による分類問題で利用される主な手法
17	1.2.10	図1.11 化学構造データを利用するケモ・マテリアル・インフォマティクスの解析の流れ
19	1.3.2	図1.12 Rのインストールを開始する画面
19	1.3.2	図1.13 RStudioのインストールを開始する画面
20	1.3.2	図1.14 RStudio上での作業フォルダの選択
20	1.3.2	図1.15 RStudioの画面構成
21	1.3.2	図1.16 RStudio上で新規にプログラムの作成を開始する操作
21	1.3.2	図1.17 RStudio上でプログラムを実行させる操作
22	1.3.2	図1.18 RStudio上で作成したプログラムのスクリプトを保存する操作

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
24	2.1	図2.1 RStudioの実行画面(1)
25	2.1	図2.2 RStudioの実行画面(2)
28	2.2	表2.1 read.delim関数の引数およびその意味
28	2.2	表2.2 df1 のデータ
29	2.2	表2.2 df12のデータ
29	2.2	表2.3 dfcsv1のデータ
31	2.3.1.1	表2.5 typeof関数を実行した時の戻り値とその意味
32	2.3.1.1	表2.6 ベクトルの作成で使われる関数と実行内容
33	2.3.1.3	図2.3 order関数とsort関数の説明図
33	2.3.1.4	表2.7 集合の演算で使われる関数と実行内容
35	2.3.2.1	表2.8 行列作成時によく使われる関数と実行内容
37	2.3.2.2	表2.9 行列にラベルをつける時によく使われる関数と実行内容
38	2.3.2.3	表2.10 行列演算でよく使われる関数と実行内容
44	2.3.2.6	図2.4 プログラム2_3_2_6で出力されるヒストグラム
49	2.3.2.8	図2.5 正規化テーブル
52	2.3.2.9	表2.11 行列shortdataの内容
53	2.3.2.9	表2.12 行列longdataの内容
53	2.3.2.9	表2.13 行列shortdata2の内容
53	2.3.2.9	表2.14 行列longdata2の内容
56	2.3.3.2	表2.15 行列dataNewの内容
60	2.3.5	表2.16 Apply系関数の書式と演算内容
61	2.3.5.1	図2.6 プログラムRapply01b.Rの演算内容
64	2.3.5.2	図2.7 プログラムRapply02.Rの計算の流れ
70	2.4	表2.17 データの可視化によく利用される高水準作図関数
71	2.4	表2.18 データの可視化によく利用される低水準作図関数
71	2.4	表2.18 データの可視化によく利用される低水準作図関数
71	2.4.1	表2.19 plot関数の引数と機能
72	2.4.1	表2.20 plot関数の引数typeの指定値とプロットの形式
73	2.4.1.1	図 2.8 プログラム2_4_1_1から出力される折れ線グラフ
74	2.4.1.1	図 2.9 プログラム2_4_1_1から出力される2本の折れ線グラフ
74	2.4.2	図2.10 箱ヒゲ図の見方
75	2.4.2.1	図2.11 プログラム2_4_2_1の出力結果の 箱ヒゲ図
76	2.4.2.2	表2.21 行列dataAllの内容
78	2.4.2.2	図2.12 プログラム2_4_2_2の出力結果の 箱ヒゲ図
79	2.4.3.1	図2.13 プログラム2_4_3_1の出力結果の ヒストグラム
80	2.4.3.2	表2.22 行列df1の内容
80	2.4.3.2	表2.23 行列df2の内容

第2章の図・表(2/2)

ページ	章節項	タイトル
81	2.4.3.2	図2.14 プログラム2_4_3_2のPDF出力結果の ヒストグラム
82	2.4.4.1	表2.24 行列alldataの内容
83	2.4.4.1	図2.15 プログラム2_4_4_1のPDF出力結果の 散布図
84	2.4.4.2	表2.25 行列alldataの内容
86	2.4.4.2	図2.16 プログラム2_4_4_2のPDF出力結果の 散布図
87	コラム	表2.26 特殊なデータ型(NULL, NA, NaN, Inf)の比較

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
94	3.2.1	図3.1 5回サイコロを投げた時の目の出方と各々の確率
95	3.2.1	図3.2 プログラム3_2_1Bの作図結果
97	3.3	図3.3 二項分布の確率密度関数と累積分布関数の比較
98	3.2.3	図3.4 二項分布の統計検定の手順
99	3.2.4	図3.5 二項分布の確率密度関数 $p(k)$ と95%区間推定に使う領域
103	3.2.6	図3.6 二項分布と正規分布の確率密度関数の比較
105	3.3.1	図3.7 超幾何分布の確率密度比較 (プログラム3_3_1の計算結果)
106	3.3.2	図3.8 超幾何分布の確率密度関数と累積分布関数の比較
106	3.3.3	表3.1 ケース1の分割表(左側:カウント、右側:確率値)
107	3.3.3	表3.2 ケース2の分割表 (左側:カウント、右側:確率値)
108	3.3.3	図3.9 Fisherの直接確率計算法の手順
108	3.3.3	表3.3 ケース1の分割表
108	3.3.3	表3.4 ケース2の分割表
111	3.4.1	図3.10 二項分布とポアソン分布の確率密度関数の比較
111	3.5	表3.5 3種類の分布とRプログラミングで利用する関数
111	3.5	表3.6 3種類の検定で用いたRプログラミングの関数

第4章の図・表

ページ	章節項	タイトル
114	4.2.1	図4.1 xの関数であるx, x2, x3, x4, x5のグラフ出力
115	4.2.2	図4.2 平均値を1, 2, 3としたときの正規分布の確率密度関数の描画
116	4.2.3	図4.3 標準偏差が変化したときの正規分布の確率密度関数の描画
116	4.2.4	図4.4 正規分布N(0, 1)の確率密度関数と累積分布関数の比較
117	4.2.5	図4.5 正規分布N(0, 1)およびN(0, 2)の分位数関数の描画
120	4.3.1	図4.6 自由度nのカイ二乗分布の確率密度関数
123	4.4.1	図4.7 t分布と標準正規分布の確率密度関数
125	4.5	図4.8 t検定のイメージと解析に利用するR関数
126	4.5	図4.9 1群のt検定の概要
127	4.5.1	表4.1 10種類の材料に対して触媒D添加とL添加による反応率の増減
128	4.5.1	表4.2 1群の検定に利用するデータ
129	4.6	表4.3 パラメトリック手法とノンパラメトリック手法の比較
130	4.6.1	表4.4 6種類の材料に対して2種類の処理による特性変化量
130	4.6.1	表4.5 クラスの分類結果
131	4.6.1	表4.6 順位和と累積確率の表
131	4.6.2	表4.7 6種類の材料に対して2種類の処理による特性変化量
131	4.6.2	表4.8 クラスの分類結果
132	4.6.2	表4.9 順位和と累積確率の表
137	4.6.4	図4.10 Wilcoxon順位和検定結果の箱ひげ図
137	4.7	表4.10 1xm 分割表の構成
139	4.7.1	表4.11 2x2 分割表の構成
139	4.7.2	表4.12 製品ごとのアレルギー発症者数のデータ
141	4.8	図4.11 Wilcoxonの符号順位和検定と正規分布を利用した検定の使い分け
141	4.8	図4.12 二項分布を利用したFisher直接確率検定と正規分布を利用した検定の使い分け
142	4.9	図4.13 4つの確率分布(正規分布、カイ二乗分布、t分布、F分布)の特徴と関係
142	4.9	図4.14 対応のある2群(組)の平均の差の検定手法の選択
143	4.9	図4.15 対応のない2群(組)の平均の差の検定手法の選択(*は本書に例題が無いことを示す)
143	4.9	図4.16 分割表を利用した1要因の独立性の検定手法・適合性の検定手法の選択(*は本書に例題が無いことを示す)
143	4.9	図4.17 分割表を利用した2要因の独立性の検定手法・適合性の検定手法の選択(*は本書に例題が無いことを示す)

第5章の図・表

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
146	5.1	図5.1 教師無し学習と教師あり学習
147	5.1	図5.2 2次元ベクトルの向きと大きさの比較
148	5.1	図5.3 変数XとYの変化の関係
149	5.2	表5.1 相関係数の値と相関関係の強さ
150	5.2.1	表5.2 DataBloodType.csvのデータ
153	5.2.1	図5.4 血液型割合に関する2次元散布図、3次元散布図、散布図行列
155	5.2.2	図5.5 血液型割合に関する散布図行列、直線回帰、ヒストグラム、相関係数の複合表示
154	5.2.2	表5.3 Datalong.csvのデータ
157	5.2.3	図5.6 物性間で数値が存在した頻度のカウント数および相関行列の値のヒストグラム
159	5.2.3	図5.7 ピアソン相関係数と区間推定の関係
164	5.3.1.1	図5.8 寄与率と累積寄与率の散布図
164	5.3.1.1	図5.9 第1主成分と第2主成分に関する主成分スコアと因子負荷量のプロット
165	5.3.1.1	表5.5 ワインに寄与する成分とコニャックに寄与する成分の分離結果
165	5.3.1.1	図5.10 第1主成分と第3主成分に関する主成分スコアと因子負荷量のプロット
166	5.3.1.1	図5.11 主成分スコア、寄与率、因子負荷量の関係
166	5.3.1.1	図5.12 主成分分析の計算の流れ
169	5.3.2.1	図5.13 計算結果の固有値の棒グラフ表示
170	5.3.2.1	図5.14 固有値への寄与相関の散布図
171	5.3.2.2	図5.15 固有値への寄与相関の散布図
172	5.3.3	図5.16 自己組織化マップの構造
172	5.3.3	図5.17 自己組織化マップのアルゴリズムの模式図
176	5.3.3.1	図5.18 自己組織化マップの結果表示
177	5.4	図5.19 階層法(最短距離法)と分割法(k平均法)の比較
178	5.4.1	図5.20 「最短距離法」の計算過程
178	5.4.1	図5.21 樹形図(デンドログラム)
179	5.4.1	図5.22 「重心(距離)法」の計算過程
179	5.4.1	図5.23 最短距離法と重心(距離)法の樹形図比較
181	5.4.1.1	表5.6 hclust関数の引数methodとその説明
182	5.4.1.2	図5.25 「最長距離法」による樹形図(2)
183	5.4.1.3	図5.26 カラーバー無しのヒートマップ描画
186	5.4.1.3	図5.27 カラーバーを追加したヒートマップ描画
187	5.4.2	図5.28 k平均法の模式図
190	5.4.2.1	図5.29 クラスタ数とGap関数の関係
191	5.4.2.1	図5.30 ギャップ関数の変化量の棒グラフ
192	5.4.2.1	図5.31 クラスタ分析結果の散布図
192	5.4.2.1	表5.7 出力データ(fdata)の内容
193	5.5	表5.8 主成分分析と多次元尺度法の比較

第6章の図・表

ページ	章節項	タイトル
196	6.1	図6.1 実測データを2種類の関数でフィッティングした結果の比較
200	6.2.1	図6.2 元データの総合得点と回帰モデルで計算される総合得点の関係
205	6.2.2	図6.3 100M走と走り幅跳びの成績の相関図
206	6.2.3	図6.4 選手と回帰モデルの残差の散布図
207	6.2.3	図6.5 残差の正規Q-Qプロットと正規分布に相当する直線
209	6.2.5	図6.6 目的変数と7種類の説明変数の説明
211	6.2.5	表6.1 単相関係数と重相関係数の比較
217	6.3.1	図6.8 潜在変数の数とMSEPの関係(1)
217	6.3.1	図6.9 潜在変数の数とMSEPの関係(2)
219	6.3.1	図6.10 潜在変数の数と決定係数 Q^2 との関係
219	6.3.1	図6.11 目的変数 y の実測値と予測値の関係(1: PLSの結果)
219	6.3.1	図6.12 leave-one-out法の考え方とその決定係数 Q^2 の求め方
220	6.3.1	図6.13 目的変数 y の実測値と予測値の関係(2: PLSの結果)
221	6.3.1	図6.14 leave-one-out法の後の考え方とその決定係数 R^2 の求め方
221	6.3.1	図6.15 PLSの回帰係数の棒グラフ
222	6.3.1	図6.16 目的変数 y の実測値と予測値の関係(3: 線形重回帰分析の結果)
222	6.3.1	表6.2 単相関係数、重回帰係数、PLSの回帰係数の比較
225	6.4	図6.17 リッジ回帰とラッソ回帰の比較
228	6.4.1	図6.18 正規化パラメータ λ と説明変数(x_1 から x_8 まで)の回帰係数の関係
228	6.4.1	図6.19 正規化パラメータ λ 、説明変数(x_1 から x_8 まで)の回帰係数、GCVスコアの関係
229	6.4.1	図6.20 目的変数の予測値と実測値の関係図
231	6.4.2	図6.21 正規化パラメータ λ と平均交差検証誤差 cvm 値の関係
231	6.4.2	図6.22 平均交差検証誤差 cvm 値の区間推定およびゼロでない説明変数の数の変化
232	6.4.2	図6.23 正規化パラメータ λ の値と回帰係数の関係
233	6.4.2	図6.24 目的変数の実測値と予測値の散布図
234	6.5	表6.3 各種手法の回帰係数の比較

第7章の図・表

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
237	7.1	図7.1 分類問題における「教師あり学習」の流れ
238	7.1	図7.2 caretパッケージに含まれる機械学習の機能一覧
239	7.1	表7.1 機械学習のカテゴリー名、この章で学ぶプログラム名、所属するアルゴリズムの個数
239	7.1	図7.3 パッケージ” caret” のマニュアルの内容
240	7.2	表7.2 アヤメのデータ(“irisAll.csv”の内容)
240	7.2	図7.4 アヤメの3種
244	7.2	図7.5 第1主成分得点と第2主成分得点を(x, y)座標とした散布図
245	7.2	図7.6 第1主成分得点と第3主成分得点を(x, y)座標とした散布図
246	7.3	図7.7 1次元のデータ配置とグループ分け
246	7.3	図7.8 1次元のデータ配置とマハラノビス距離
247	7.3	図7.9 2次元のデータ配置とマハラノビス距離
251	7.3.1	図7.10 分類評価のための統計量
253	7.3.1	図7.11 第1判別得点と第2判別得点の散布図
256	7.4	図7.12 k最近傍法の説明
259	7.4.1	図7.13 最近傍の数kと判別精度の関係図
261	7.5	図7.14 観測データと事後確率との関係
263	7.5.1	図7.15 3種のアヤメの事後確率の散布図
267	7.5.2	図7.16 アヤメの4種類の特徴量の確率密度関数
268	7.6	図7.17 決定木の図
270	7.6.1	図7.18 木の構造
271	7.7	図7.19 ニューラルネットワークモデルの構成
272	7.7	図7.20 シグモイド関数
274	7.7.1	図7.21 隠れ層のユニット数と識別精度(Accuracy)の関係
276	7.7.1	図7.22 構築されたニューラルネットワークモデルの構造
277	7.8	図7.23 サポートベクトルマシーン(○印はデータ, X1, X2は説明変数)
282	7.9	図7.24 「バギング」、「ランダムフォレスト」、「ブースティング」の比較
285	7.10	表7.3 正誤表が格納されたファイル名、内容、作成プログラム
287	7.10.1	図7.25 8種の分類手法比較のヒートマップ表示

第8章の図・表

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
292	8.2	図8.1 Nicotinic acidのモルファイル
293	8.2	図8.2 Nicotinic acidのSMILESと化学構造の関係
294	8.3.1.1	図8.3 RCDK01.Rを用いた化学構造の描画結果
294	8.3.1.2	表8.1 DataTableChemicalStructure.csvの先頭8行目までの内容
296	8.3.1.2	図8.4 RCDK02.Rを用いた化学構造の描画結果
296	8.3.2	図8.5 モルファイルからSMILIESへの変換
298	8.3.2	表8.2 モルファイル情報からSMILES形式のデータを作成したファイル
299	8.3.3	表8.3 SMILES形式の化学構造からの物性予測
302	8.3.4	表8.4 rcdkで利用できる「分子記述子」
305	8.3.5	表8.5 get.fingerprint関数に搭載されているフィンガープリント
307	8.3.5	図8.6 20種の化合物の樹形図
307	8.3.5	図8.7 Evonineの構造情報と生合成する生物の情報
308	8.4.1	図8.8 SDFDataTable(sdfset)の出力図
310	8.4.1	図8.9 20種の化合物の化学構造描画
313	8.4.2.1	表8.6 ファイルDataTableChemicalStructure.csv の内容
316	8.4.2.3	図8.10 プログラム8_4_2_3の処理概要
319	8.4.3	表8.7 行列blockMatrixの内容
320	8.4.3	表8.8 化学フィンガープリントの種類の比較
321	8.4.3	表8.9 行列fpsetBIの内容
321	8.4.3	表8.10 行列moldataの内容
322	8.4.3	図8.11 クラスタ分析結果の樹形図表示
324	8.4.4	表8.11 分子記述子計算用の関数と計算される分子の特性
325	8.4.4	表8.12 出力ファイル. ” MolPropData01.csv” の内容
328	8.4.5.2	図8.12 cluster.visualization関数の描画結果

第9章の図・表

<u>ページ</u>	<u>章節項</u>	<u>タイトル</u>
330	9.1	図9.1 ニューラルネットワーク（左）と深層学習モデル（右）
333	9.3	図9.2 入力が3ノード、中間層が4ノード、出力が1ノードから構成されるネットワークモデル
337	9.4	図9.3 学習曲線
339	9.4	図9.4 学習を100エポックまで進めた場合の学習曲線
340	9.4	図9.5 予測値v. s. 実測値の散布図
341	9.5	図9.6 画像のRed, Green, Blue, Brightnessへの分解
342	9.5	図9.7 プーリングプロセスの説明
344	9.5	図9.8 手書き数字の例
345	9.5	図9.9 手書き文字の誤予測された例
346	9.5	図9.10 グラフコンボリューション・ニューラルネットワーク

別サイトのコラムの図・2022/7/30作成

<u>資料page</u>	<u>関連する 章節項</u>	<u>タイトル</u>
6	4.1	図 Quantile-Quantile(QQ)プロット
11	5.1	図C.1 F分布の確率密度関数と累積分布関数
16	5.3	図C.2 ミンコフスキーのパワー距離
17	5.4	図C.3 新しい軸とデータとの位置関係