

目次

まえがき	i
本書の構成	ii
第 1 部 統計学の基礎と検定の考え方	1
1 章 t 検定 (数式なし)	2
1-1 期待値 (平均値)	2
1-2 分散	3
1-3 有意差	4
1-4 サンプルサイズ	6
1-5 t 値	7
1-6 t 検定	8
1-7 p 値	9
1-8 p 値と検定	10
1-9 検定は偉いのか?	11
2 章 t 検定 (数式あり)	13
2-1 期待値と平均値	13
2-2 分散の数式	16
2-3 不偏分散	18
2-4 標準偏差	21
2-5 標準誤差	21
2-6 t 値の計算例	22
3 章 R の簡単な使い方	24
3-1 R とは	24
3-2 R のインストール	24
3-3 計算の実行方法と F5 ボタン	25
3-4 R による簡単な計算	26

3-5	「変数に格納する」という言葉の意味と使い時.....	28
3-6	コメントの使い方.....	30
3-7	Rによるt検定.....	30
3-8	t検定を一瞬で行う方法.....	34
第2部	統計モデル基礎：正規線形モデル.....	36
1章	分散分析(ANOVA).....	37
1-1	分散分析の使い時.....	37
1-2	分散分析によるデータのモデル化.....	37
1-3	分散分析による検定1：誤差と効果の比をとる.....	38
1-4	分散分析による検定2：誤差と効果の大きさを測る.....	40
1-5	計算例1：データ.....	41
1-6	計算例2：「効果の分散の大きさ」を計算する.....	43
1-7	計算例3：「誤差の分散の大きさ」を計算する.....	44
1-8	用語：自由度、偏差平方和、残差平方和、平均平方.....	45
1-9	計算例4：p値の意味.....	46
1-10	計算例5：「R」を用いてp値を求める.....	47
1-11	Rによる分散分析1：データの準備.....	48
1-12	Rによる分散分析2：データフレームの使い方.....	49
1-13	Rによる分散分析3：モデルの作成と検定.....	50
2章	統計モデルから見た分散分析.....	53
2-1	統計モデルの意味とその必要性.....	53
2-2	統計モデルと誤差.....	54
2-3	分散分析とモデル選択.....	55
2-4	数値例.....	57
2-5	単純なモデル・複雑なモデル.....	60
2-6	用語1：帰無仮説・対立仮説.....	61

2-7	補足：検定の非対称性.....	62
2-8	用語 2：応答変数・説明変数.....	63
2-9	数式でみる分散分析モデル.....	64
2-10	R でみる分散分析モデル.....	65
3 章	回帰分析.....	68
3-1	回帰分析とは.....	68
3-2	回帰モデルの考え方.....	68
3-3	線形回帰モデルの仮定.....	69
3-4	パラメタの推定方法—最小二乗法.....	71
3-5	R による最小二乗法 1：解析の流れ.....	73
3-6	R による最小二乗法 2：データの準備とプロット.....	73
3-7	R による最小二乗法 3：予測関数の作成.....	74
3-8	R による最小二乗法 4：残差平方和の計算.....	76
3-9	R による最小二乗法 5：残差平方和計算関数の作成.....	78
3-10	R による最小二乗法 6：最適化関数 optim.....	79
3-11	R の lm 関数による回帰分析.....	80
3-12	回帰分析の検定 1：検定の流れ.....	82
3-13	回帰分析の検定 2：誤差と効果に分ける方法.....	82
3-14	回帰分析の検定 3：モデル選択による方法.....	84
3-15	回帰分析の検定 4：R で実行.....	86
3-16	用語： R^2 （決定係数）.....	87
3-17	summary 関数.....	88
4 章	正規線形モデル.....	91
4-1	カテゴリ変数と連続変数.....	91
4-2	R による解析 1：解析の流れ.....	92
4-3	R による解析 2：データの読み込み.....	92

4-4	Rによる解析3: データの図示と pairs 関数.....	94
4-5	Rによる解析4: lm 関数によるモデリング.....	95
4-6	Rによる解析5: 検定.....	96
4-7	Rによる解析6: summary 関数によるモデルの要約.....	97
4-8	Rによる解析7: 係数の抽出.....	99
4-9	Rによる解析8: モデルの係数を用いた予測.....	100
4-10	Rによる解析9: predict 関数による予測.....	101
4-11	Rによる解析10: 予測用データセットの作成.....	102
4-12	Rによる解析11: モデルの図示.....	103
第3部	正規線形モデルによるデータ解析.....	107
1章	Type II ANOVA とモデル選択.....	108
1-1	解析の準備.....	108
1-2	モデル選択と「フツアの」分散分析の違い.....	109
1-3	「フツアの」分散分析の仕組み.....	111
1-4	Type II ANOVA.....	113
1-5	Type II ANOVA のメリット.....	114
1-6	Rによる解析1: パッケージのダウンロード.....	114
1-7	Rによる解析2: ライブラリの読み込み.....	116
1-8	Rによる解析3: Anova 関数の実行.....	116
2章	Type II ANOVA の応用.....	118
2-1	売り上げに効いている説明変数はどれ?.....	118
2-2	解析の順序.....	119
2-3	解析手順1: 目的の設定.....	119
2-4	解析手順2: データの読み込みと図示.....	120
2-5	よくない解析の例.....	123
2-6	変数を一つずつ検定すべきでない理由.....	126

2-7	解析手順 3: 統計モデルの作成.....	128
2-8	解析手順 4: 検定 (Type II ANOVA)	129
2-9	解析手順 5: モデル選択.....	129
2-10	解析手順 6: モデルの解釈.....	131
2-11	解析手順 7: 意思決定.....	132
2-12	補足: モデルに入らなかった変数について	133
2-13	補足: モデル選択の前と後の比較.....	134
2-14	統計モデルを知ることの意味.....	135
第 4 部	確率と統計データ	137
1 章	確率変数	138
1-1	手に入れたデータ・手に入れていないデータ	138
1-2	手に入れていないデータを扱う意義.....	138
1-3	データを扱う学問になぜ「確率」が必要か.....	140
1-4	データと確率変数	140
1-5	確率変数と確率分布.....	141
2 章	データが得られるプロセス.....	144
2-1	母集団と標本とサンプリングの関係.....	144
2-2	母集団からのサンプリングがイメージしにくい例	144
2-3	無限母集団と確率分布と確率変数の関係.....	145
2-4	時系列データとデータ生成過程 (DGP)	146
2-5	データが得られるプロセス.....	148
2-6	補足: サンプリングと偏り	148
3 章	データを解析するプロセス.....	150
3-1	手に入れていないデータを扱う方法.....	150
3-2	データから母集団分布を推定する方法.....	150
3-3	データを比較する方法.....	151

3-4	確率分布の比較と検定.....	151
3-5	確率分布の使い分け.....	152
第5部	確率分布と統計モデル.....	154
1章	確率密度関数.....	155
1-1	確率分布.....	155
1-2	パラメトリックとノンパラメトリック.....	156
1-3	確率密度関数から確率分布を求める.....	157
1-4	確率密度関数とパラメトリック法.....	158
1-5	仮定を置くということの意味.....	158
1-6	どのような確率密度関数を使うべきか.....	159
2章	正規分布.....	160
2-1	誤差理論の基礎 1：誤差とは何か.....	160
2-2	誤差理論の基礎 2：誤差のでき方.....	161
2-3	誤差理論の基礎 3：誤差の特徴.....	161
2-4	正規分布の使い時.....	162
2-5	正規分布の確率密度関数.....	163
2-6	連続変数と正規分布.....	163
2-7	確率密度関数から確率を求める方法.....	164
2-8	面積と積分と確率の関係.....	164
2-9	正規分布の確率密度関数：常に0以上.....	165
2-10	期待値と分散と正規分布の関係.....	167
2-11	正規分布の確率密度関数：面積 1.....	167
2-12	正規分布を推定する.....	168
2-13	正規分布の記法.....	169
2-14	Rによる正規分布 1：確率密度の計算 <code>dnorm</code>	169
2-15	Rによる正規分布 2：ある値以下になる確率 <code>pnorm</code>	173

2-16	Rによる正規分布 3: ある確率になる基準値 $qnorm$	175
2-17	Rによる正規分布 4: 乱数の発生 $morm$	176
2-18	補足: 正規分布の推定と局外パラメタ	179
3 章	統計モデルと確率分布	180
3-1	統計モデルの定義	180
3-2	統計モデルと誤差	181
3-3	統計モデルと母集団分布とデータの比較の関係	182
3-4	統計モデルによる予測	182
3-5	最小二乗法による期待値の推定	183
3-6	正規線形モデルの定義	185
4 章	パラメトリックブートストラップ検定	186
4-1	復習: 検定と p 値	186
4-2	パラメトリックブートストラップ検定	187
4-3	復習: 帰無仮説と対立仮説	187
4-4	Rによる計算の流れ	188
4-5	Rの文法: for ループ (繰り返し処理)	189
4-6	Rの文法: for ループによる計算結果の格納処理	190
4-7	Rの文法: 条件を満たす値の数え方	192
4-8	Rによる PB 検定 1: 統計モデルの作成	194
4-9	Rによる PB 検定 2: 標本から F 比を計算	195
4-10	Rによる PB 検定 3: シミュレーションデータ作成	197
4-11	Rによる PB 検定 4: F 比を多数計算する	198
4-12	Rによる PB 検定 5: p 値の算出	201
4-13	補足: 独立で同一の確率分布 (i.i.d)	202
5 章	正規分布から派生した確率分布	204
5-1	中心極限定理	204

5-2	中心極限定理の応用.....	206
5-3	標本分布	207
5-4	期待値の標準偏差としての標準誤差.....	208
5-5	t 分布	209
5-6	t 分布と信頼区間	211
5-7	χ^2 分布.....	213
5-8	F 分布	214
第 6 部	一般化線形モデル	216
1 章	一般化線形モデルの長所	217
1-1	一般化線形モデルとは.....	217
1-2	一般化線形モデルの使い時.....	217
2 章	尤度と最尤法.....	218
2-1	モデルとパラメタ	218
2-2	補足：同時確率.....	219
2-3	尤度	219
2-4	対数尤度	220
2-5	最尤法	222
2-6	R による最尤法.....	223
2-7	補足：最小二乗法と最尤法.....	224
3 章	一般化線形モデルの推定	227
3-1	ポアソン分布	227
3-2	線形予測子	228
3-3	リンク関数	229
3-4	リンク関数とモデルの解釈.....	230
3-5	補足：一般化線形モデルから見た正規線形モデル	231
3-6	R でみるポアソン分布	231

3-7	Rによるポアソン回帰1: 計算の流れ	233
3-8	Rによるポアソン回帰2: データの準備	234
3-9	Rによるポアソン回帰3: 対数尤度の計算	234
3-10	Rによるポアソン回帰4: 関数の作成	236
3-11	Rによるポアソン回帰5: パラメタ推定	236
3-12	Rの glm 関数によるポアソン回帰	237
3-13	線形予測子から直接求められた予測結果	238
3-14	応答変数の予測値	239
3-15	線形予測子の係数の解釈	239
3-16	GLMによりグラフにひかれた線の意味	240
4章	Deviance と尤度比検定	243
4-1	正規線形モデルの検定方法の復習	243
4-2	誤差の計算: 正規分布を仮定できるとき	244
4-3	誤差の計算: 正規分布を仮定できない時	244
4-4	一般化された残差平方和としての Deviance	245
4-5	Rによる Deviance 1: データの準備	246
4-6	Rによる Deviance 2: カンペキ対数尤度の計算	247
4-7	Rによる Deviance 3: GLM の対数尤度の計算	247
4-8	Rによる Deviance 4: Deviance の計算	248
4-9	Deviance 残差	249
4-10	Deviance の差の検定	249
4-11	Deviance の差の検定における Type II ANOVA	251
4-12	尤度比と Deviance の関係	252
4-13	最小二乗法と最尤法の関係	253
4-14	補足: summary 関数	254
4-15	補足: 一般化線形モデルによる解析の手順	257

第7部	一般化線形モデルによるデータ解析.....	258
1章	ロジスティック回帰.....	259
1-1	ロジスティック回帰.....	259
1-2	二項分布.....	259
1-3	Rでみる二項分布.....	260
1-4	ロジスティック関数とロジット関数.....	262
1-5	オッズと対数オッズ.....	263
1-6	Rによるロジスティック回帰1：データの図示.....	264
1-7	Rによるロジスティック回帰2：推定と検定.....	268
1-8	成功確率の予測値.....	269
1-9	オッズと係数の関係.....	271
1-10	ロジスティック回帰の図示.....	272
1-11	補足：モデルの予測結果を意思決定に使用する.....	274
2章	交互作用.....	277
2-1	交互作用とは何か.....	277
2-2	データの例.....	277
2-3	交互作用をRで表す.....	279
2-4	交互作用を入れたモデルの係数.....	280
2-5	用語：交互作用と主効果.....	281
2-6	Type III ANOVAによる検定.....	281
2-7	交互作用を入れたモデルの図示.....	283
3章	分割表に対するGLM（対数線形モデル）.....	285
3-1	分割表の検定.....	285
3-2	数値例：ナマズと地震の関係について.....	285
3-3	χ^2 検定の考え方1：期待度数を求める.....	287
3-4	χ^2 検定の考え方2：統計量を求める.....	288

3-5	χ^2 検定 : R による計算	289
3-6	対数線形モデル : 考え方	290
3-7	R による対数線形モデル 1 : データの読み込み	291
3-8	R による対数線形モデル 2 : モデルの推定	292
3-9	R による対数線形モデル 3 : 検定	293
3-10	複雑な分割表への対数線形モデルの適用	293
4 章	ゼロ切断モデル	297
4-1	ゼロ切断モデルとは	297
4-2	条件付き確率分布	297
4-3	ゼロ切断分布	298
4-4	R によるゼロ切断モデル 1 : データの作成	299
4-5	R によるゼロ切断モデル 2 : モデルの推定	301
4-6	R によるゼロ切断モデル 3 : 検定	303
4-7	R によるゼロ切断モデル 4 : 図示	303
4-8	一般化線形モデルの発展系	305
第 8 部	情報理論と統計学	307
1 章	AIC の考え方	308
1-1	モデル選択と AIC	308
1-2	予測のずれと確率分布	308
1-3	検定と AIC	309
2 章	AIC を用いたデータ解析	310
2-1	R による AIC 1 : AIC の計算	310
2-2	R による AIC 2 : step 関数による自動モデル選択	311
2-3	R による AIC 3 : MuMin による総当たり法の実行	316
2-4	「ぜひ入れたい変数」があるときの変数選択の仕方	319
3 章	AIC と相対エントロピー	320

3-1	確率と対数と情報量.....	320
3-2	コイン投げの例.....	321
3-3	確率分布の違いと情報量.....	322
3-4	相対エントロピーと分布間擬距離.....	323
3-5	相対エントロピーの最小化.....	323
3-6	数式の解釈.....	324
3-7	対数尤度のバイアスと AIC.....	326
3-8	予測と情報量.....	327
逆引き R 関数.....		331
索引.....		333