

目次

はじめに iii

1	データ分析事始め	1
1.1	なぜ平均を求めるか	1
1.2	カテゴリカルデータ	2
1.3	順序尺度のデータ	3
1.4	統計データの樹葉図	5
1.5	データの代表値	7
1.6	範囲と四分位偏差	7
1.7	平均偏差と標準偏差(分散)	8
1.8	変動係数	9
1.9	相関比	10
1.10	回帰直線	13
1.11	相関係数	15
2	確率と確率変数の概念	18
2.1	不確実な現象と統計的現象	18
2.2	標本空間と確率空間	19
2.3	事象の確率	20
2.4	相互排反性と加法定理	21
2.5	事象の独立性と乗法定理	22
2.6	ベイズの定理	23
2.7	スロット・マシン	25
2.8	確率変数と確率分布	26
2.9	独立なふるい	27
2.10	確率変数の独立性	28

3 確率分布のいろいろ	30	5.4 2×2 分割表	57
3.1 硬貨投げの実験	30	5.5 学歴と恋愛志向の強弱	59
3.2 サイコロ投げの実験	31	5.6 インチキサイコロを見破る法	60
3.3 パスカルの三角形と 2 項係数	32	5.7 傾向分析	61
3.4 ベルヌーイ試行と 2 項分布	33	5.8 年齢と虫歯の数	63
3.5 2 項分布のグラフ	34	5.9 統計的仮説検定の論理	64
3.6 2 項分布から正規分布へ	35	6 情報量規準とは何か	66
3.7 標準正規分布	36	6.1 統計的方法とモデル	66
3.8 正規分布	38	6.2 理論モデルと経験モデル	67
3.9 初婚夫婦の年齢差	39	6.3 モデルの当てはまりの“悪さ”	68
3.10 カイ 2 乗分布	39	6.4 対数尤度と最大対数尤度	70
3.11 t -分布と F -分布	40	6.5 正規分布の最尤推定	71
4 推定の考え方	42	6.6 カルバック・ライブラーの情報量	72
4.1 不偏推定	42	6.7 どちらが近いか	73
4.2 みそ汁の味見と基石の数当て	43	6.8 正規分布に対する KLI 情報量	75
4.3 最尤推定	45	6.9 硬貨投げのモデル	76
4.4 硬貨投げ実験の出現比率	46	6.10 情報量規準 AIC	77
4.5 視聴率調査	47	7 情報量規準によるモデル選択	80
4.6 比率に対する区間推定	48	7.1 出生性比のモデル選択	80
4.7 身長に関する推論	50	7.2 薬の治癒効果	81
4.8 一般の母集団に対する平均の推定	50	7.3 年齢と恋愛志向の関係	85
4.9 平均の差に対する区間推定	51	7.4 作物の収穫量の比較	86
4.10 ストッキングの寿命	52	7.5 平均が異なるモデル	89
5 検定の考え方	54	7.6 肥料の効果の比較	90
5.1 試してみる	54	7.7 分散の値も異なるかもしれない場合	92
5.2 理論と実際のずれ	55	7.8 分散だけが異なるモデル	93
5.3 20 歳代後半の女性の未婚率	56		

8	分散分析	97
8.1	実験計画法	97
8.2	測定値のばらつきの分解	98
8.3	分散比と分散分析	99
8.4	肥料と地域と収穫量	100
8.5	1 要因モデルの AIC 分散分析	103
8.6	2 要因モデルの AIC 分散分析	107
9	回帰分析	111
9.1	回帰モデル	111
9.2	モデルの当てはまりをみる決定係数	114
9.3	残差平方和だけで適合度が計れるか	115
9.4	推定値の安定性をみる t -値	118
10	重回帰モデル	121
10.1	回帰分析の適用上の問題点	121
10.2	重回帰モデル	122
10.3	重回帰モデルの t -値	124
10.4	情報量規準によるモデル選択	127
10.5	職員数と収入保険料	130
10.6	最低気温と緯度・標高の関係	130
11	時系列分析	135
11.1	移動平均と指数平滑化	135
11.2	自己回帰モデル	138
11.3	2 次の自己回帰モデル	141
11.4	一般の自己回帰モデル	143
11.5	回帰直線のモデル	145
11.6	2 次曲線による回帰モデル	147

11.7	多項式回帰モデルの情報量規準	148
------	----------------	-----

12	主成分分析	152
12.1	多元情報の縮約	152
12.2	美女のプロポーシオン	153
12.3	ラグランジュの未定係数法	156
12.4	2 次形式の最大・最小化問題	158
12.5	主成分分析	161
12.6	主成分分析の実際	163
12.7	百貨店の伸びの主成分分析	165

索引	168
----	-----