

## 21 個の例題

「エクセル統計 実用多変量解析編」で扱う例題です。

### 01 重回帰分析

次のデータは 25 人の健康な男性についてある検査をした結果です。  
A, B は運動負荷をかけたあとのある種の測定値, C1, C2, C3, C4 は安静時の測定値です。  
測定値 A を予測する回帰式を測定値 B と測定値 C1, C2, C3, C4 から求めなさい。  
さらに, この回帰式が予測に役立つかどうか検定しなさい。

	A	B	C1	C2	C3	C4
	20	19	2	8.7	17.9	6.4
	23	24	3	7.3	17.7	5.1
	⋮	⋮	⋮	⋮	⋮	⋮
	10	11	5	6.2	16.5	6.3
	43	79	6	5.8	17.4	8.3

### 02 重回帰分析

中学 3 年生女子 1000 人の「身長」, 「体重」, 「座高」を測定して集計したところ, 次のデータを得た。このデータから, 身長を予測する回帰式を求めなさい。この回帰式が予測に役立つかどうか検定しなさい。

◎ 平均と標準偏差		平均	標準偏差
身長		156.7	5.8
体重		50.9	6.5
座高		84.8	2.1

  

◎ 相関行列		身長	体重	座高
身長		1		
体重	0.61		1	
座高	0.74	0.34		1

### 03 変数選択-重回帰分析

例題 01 のデータを用いて, 目的変数の予測に役立つ説明変数を選択しなさい。

### 04 主成分分析

20 歳代女性の健康診断データから, 「身長」, 「体重」について, 主成分分析しなさい。

	身長	体重
	157.7	51.9
	163.8	47.7
	⋮	⋮
	148.5	47.1
	155.1	44.9

### 05 主成分分析

次のデータは 10 人の「数学」, 「物理」, 「化学」の試験の点数です。試験結果の良い順に順位を付けます。物理の平均点が低いのでそのまま合計すると不公平になるという意見が出ました。点数を補正して合計点を出すことにします。主成分分析を用いて考察してみましょう。

No.	数学	物理	化学	素点合計
1	98	28	90	217
2	70	43	82	197
⋮	⋮	⋮	⋮	⋮
9	81	20	82	192
10	47	45	82	184
平均点	71.9	43.6	80.3	

### 06 因子分析

主婦 10 人に日常生活について次のアンケートを取りました。

- Q.1 体重をチェックしていますか?  
1. 全然していない。 2. あまりしていない。 3. 時々する。  
4. している。 5. 毎日する。
- Q.2 買い物は歩いて行きますか?  
1. 全く歩かない。 2. あまり歩かない。 3. 時々歩く。  
4. 歩く。 5. 必ず歩く。
- Q.3 寝る 2 時間前は食べ物を口にしませんか?  
1. 全く食べない。 2. あまり食べない。 3. 時々食べる。  
4. 食べる。 5. 必ず食べる。

このデータについて因子分析を行いなさい。

	体 重	徒 歩	食 べ 物
4	3	2	
3	3	4	
⋮	⋮	⋮	
3	4	1	
2	3	5	

### 07 因子分析

中学生 1500 人に国語, 数学, 英語, 理科, 社会の試験を行った結果の相関行列です。これについて因子分析を行いなさい。

	国語	数学	英語	理科	社会
国語	1	0.28	0.45	0.15	0.75
数学	0.28	1	0.1	0.64	0.37
英語	0.45	0.1	1	0.48	0.38
理科	0.15	0.64	0.48	1	0.25
社会	0.75	0.37	0.38	0.25	1

### 08 判別分析: 線形判別関数

健康人グループ A 群と患者グループ B 群に対して, 血中のある物質 X1 と X2 を測定したところ, 次のデータを得た。

このデータから健康人と患者とを判別する線形判別関数を求めなさい。

判別においてどの変数が大きくかわっているか検定しなさい。

K さんは X1 = 100, X2 = 5.5 でした。K さんはどちらのグループに属すると考えられるでしょうか。

	X1	X2
A	85	5.5
	78	4.2
	⋮	⋮
	75	5.8
	88	4.5
B	135	8.2
	110	6.4
	⋮	⋮
	130	7.5
	148	9
	118	6.2

### 09 判別分析: 2 次判別関数

2 グループ A 群, B 群に対する肝機能検査の AST (GOT), ALT (GPT),  $\gamma$ -GTP の測定データを得た。2 群の判別関数を求めなさい。

O さんは AST = 25, ALT = 18,  $\gamma$ -GTP = 23 でした。O さんはどちらの群に属すると考えられるでしょうか。

	AST	ALT	$\gamma$ -GTP
A	24	48	32
	15	15	13
	⋮	⋮	⋮
	26	43	32
	23	28	26
B	16	9	10
	21	14	7
	⋮	⋮	⋮
	19	9	15
	23	10	20

### 10 判別分析: 線形判別関数 (変数選択)

白血病患者の 3 グループ A 群, B 群, C 群に対する細胞表面マーカー CD2, CD3, CD4, CD5, CD7, CD8 の測定データより, 3 群を判別することができるか検討しなさい。有用な変数を選択する変数選択法で判別関数を求めなさい。

	CD2	CD3	CD4	CD5	CD7	CD8
A	1	0	65	4	1	0
	2	1	0	2	75	1
	⋮	⋮	⋮	⋮	⋮	⋮
	4	1	49	1	1	0
	2	3	5	3	2	2
B	7	1	0	0	76	1
	0	0	14	1	4	0
	⋮	⋮	⋮	⋮	⋮	⋮
	13	1	2	16	22	1
	3	0	1	1	90	1
C	16	15	8	28	15	20
	2	1	1	1	2	1
	⋮	⋮	⋮	⋮	⋮	⋮
	1	1	0	1	2	1
	6	6	5	7	8	3

### 11 判別分析: 正準判別分析

例題 10 について正準判別分析法で 3 群を判別しなさい。

### 12 正準相関分析

糖尿病と高脂血症との関係を調べる目的で次の血中物質を測定しました。糖尿病に関しては空腹時血糖値, ヘモグロビン A1c (HbA1c), 高脂血症については総コレステロール (TC), 中性脂肪 (TG), HDL コレステロールとしました。健康人と, 高血糖と思われる人, 合わせて 13 人のデータです。糖尿病と高脂血症との関係を分析しなさい。

糖尿病		高脂血症			
血糖値	HbA1c	TC	TG	HDL	
75	3.7	199	75	95	
78	4.2	137	30	76	
⋮	⋮	⋮	⋮	⋮	
137	5.9	282	155	86	
148	9	255	180	73	

### 13 正準相関分析

学生 80 人の期末試験の結果について, 前期 (X) の 3 科目, X1, X2, X3 と後期 (Y) の 4 科目, Y1, Y2, Y3, Y4 を集計したところ, 次の相関行列を得た。

前期と後期の成績の関係をこのデータから分析しなさい。

	X1	X2	X3	Y1	Y2	Y3	Y4
X1	1	0.932	0.878	0.416	0.067	0.242	0.296
X2	0.932	1	0.804	0.294	0.005	0.198	0.213
X3	0.878	0.804	1	0.51	0.08	0.233	0.141
Y1	0.416	0.294	0.51	1	0.81	0.778	0.445
Y2	0.067	0.005	0.08	0.81	1	0.805	0.671
Y3	0.242	0.198	0.233	0.778	0.805	1	0.691
Y4	0.296	0.213	0.141	0.445	0.671	0.691	1

### 14 クラスタ分析: 量的データの例

右のデータは平成 15 年の関東地方の都県別の人口千対の出産率と死亡率を表す人口動態です。都県の分類を試みなさい。

	出生率	死亡率
茨城	9	8.3
栃木	9.1	8.5
群馬	9.2	8.5
埼玉	9.1	6.4
千葉	8.9	6.8
東京	8.2	7.3
神奈川	9.4	6.4

### 15 クラスタ分析: 質的データの例

次の表は 30 歳代女性 10 人の余暇についての回答表です。

1. ガーデニング 2. ドライブ 3. 読書 4. スポーツ  
5. 音楽鑑賞 6. 映画鑑賞 7. グルメ食べ歩き 8. パソコン

表中の数値 1 は選択した項目を意味しています。(複数選択可能)

余暇の分類を行いなさい。次に, 回答者について分類しなさい。

	ガーデニング	ドライブ	読書	スポーツ	音楽鑑賞	映画鑑賞	グルメ	パソコン
1	1	0	0	0	1	0	1	0
2	1	1	0	1	0	0	0	0
3	1	0	1	0	1	1	1	0
4	0	1	0	0	0	0	1	1
5	1	0	1	0	1	0	0	0
6	1	0	0	0	1	0	0	1
7	0	1	1	1	0	0	0	1
8	0	0	1	0	1	1	0	0
9	0	1	0	1	0	0	0	1
10	1	1	0	1	0	0	1	1

### 16 数量化Ⅰ類

治療薬の効果について薬別, 性別, 年齢群別に集計したところ, 次のデータを得た。薬別, 性別, 年齢群によって治療薬の効果は推定しなさい。

効果	治療薬	性別	年齢群
26	A	女	壮年者
-0.32179	B	女	高齢者
⋮	⋮	⋮	⋮
27	B	女	壮年者
34	B	男	高齢者

### 17 数量化Ⅱ類

兄弟の有無を性格によって分析する目的で, 大学生に次のアンケートを取りました。

- Q.1 あなたは好奇心旺盛ですか。 1. はい 2. いいえ  
Q.2 あなたは計画的に物事を処理しますか。 1. はい 2. いいえ  
Q.3 あなたは自制心が強いですか。 1. はい 2. いいえ  
Q.4 あなたは協調性がありますか。 1. はい 2. いいえ  
Q.5 あなたは兄弟がいますか。 1. いる 2. なし

次のデータより, 数量化Ⅱ類による分析を行いなさい。

次に, Q.5 に答えないで, アンケートの答えを順に (2, 2, 1, 1) とした K さんに兄弟がいるかどうか判別しなさい。

外的基準は兄弟の有無です。第 1 列に外的基準を配置します。

兄 弟	好奇心	計画的	自制心	協調性
1	1	1	1	1
1	2	2	1	2
⋮	⋮	⋮	⋮	⋮
2	2	2	2	2
2	1	1	2	2

### 18 数量化Ⅱ類

治療薬の効果をも 3 段階に分類し, 薬別, 性別, 年齢群別に集計したところ, 次のデータを得た。数量化Ⅱ類によって, 治療薬の効果は分析しなさい。

効果	治療薬	性別	年齢群
1	A	女	壮年者
2	B	女	高齢者
⋮	⋮	⋮	⋮
2	B	女	壮年者
3	B	男	高齢者

### 19 数量化Ⅲ類

次の表は 30 歳代女性 10 人の余暇についての回答表です。

1. ガーデニング 2. ドライブ 3. 読書  
4. スポーツ 5. 音楽鑑賞 6. 映画鑑賞  
7. グルメ食べ歩き 8. パソコン

表中の数値 1 は選択した項目を意味しています。(複数選択可能)

このデータを数量化Ⅲ類で分析しなさい。

	ガーデニング	ドライブ	読書	スポーツ	音楽鑑賞	映画鑑賞	グルメ	パソコン
1	1	0	0	0	1	0	1	0
2	1	1	0	1	0	0	0	0
3	1	0	1	0	1	1	1	0
4	0	1	0	0	0	0	1	1
5	1	0	1	0	1	0	0	0
6	1	0	0	0	1	0	0	1
7	0	1	1	1	0	0	0	1
8	0	0	1	0	1	1	0	0
9	0	1	0	1	0	0	0	1
10	1	1	0	1	0	0	1	1

### 20 数量化Ⅳ類: 質的データの例

ある疾患の治療薬 A, B, C の有効性を 10 人の医師 d1, d2, ..., d10 がチェック (複数選択可) をして, 次のデータを得た。治療薬の評価の類似度を解析しなさい。また医師の評価の類似度についても解析しなさい。数値 1 は有効を意味しています。

	A	B	C
d1	1	0	1
d2	1	1	0
⋮	⋮	⋮	⋮
d9	0	1	1
d10	1	0	0

### 21 数量化Ⅳ類: 類似度行列の例