

例題 18： [SampleData] の [DB 型 打撃成績.SFD6] は、プロ野球における 13 の打撃指標を 54 名の選手についてまとめたものである。これから、13 の指標の共通性を因子分析で調べよ。また各選手が各共通因子をどの程度持っているかを、因子スコアの形で表せ。



選手名	打率	得点	安打	二塁打	三塁打	本塁打	打点	盗塁	犠打	犠飛	四球	三振	併殺打
坂本 (巨)	0.3455	91	160	27	3	22	73	13	1	6	77	62	6
鈴木 (広)	0.3333	72	151	26	7	27	93	16	3	3	50	77	10
筒香 (D)	0.3185	85	144	27	4	40	102	0	0	2	86	102	5
菊池 (広)	0.3172	92	178	22	2	13	55	13	23	3	40	101	3
山田 (ヤ)	0.3145	102	145	25	3	38	100	30	0	4	94	93	15
福留 (神)	0.3089	51	135	24	3	11	56	0	0	5	59	75	6
坂口 (ヤ)	0.3003	73	152	14	5	0	39	6	5	5	60	65	5
倉本 (D)	0.2998	38	155	19	2	1	38	2	5	3	21	95	13
新井 (広)	0.2997	65	134	22	2	18	98	0	0	4	54	98	12
...	...	n=54	以下省略

データファイルを読み込み、[統計] メニューの [多変量解析] の [因子分析] を実行します。このデータでは、13 変数全てを**対象変数**に指定します。推定因子数は、**参照** ボタンで 2~5 と表示されますが 3 つとして、あとは標準のバリマックス回転法を利用して、**実行** ボタンを押します。すると出力結果は次のようになります。

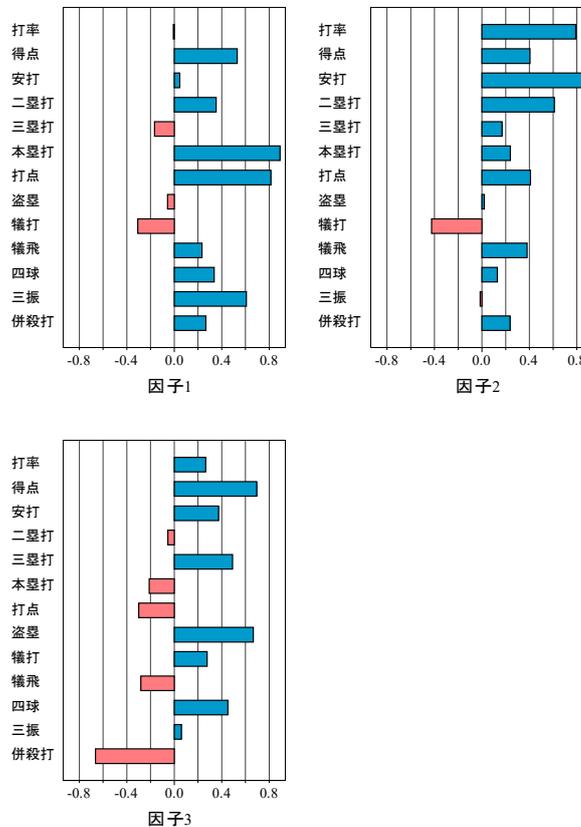
計算出力例

《 因子分析 》

[第1頁:] 有効データ数=54

《 因子負荷行列と寄与率 》 (バリマックス法)

変数名	因子1	因子2	因子3	共通性
打率	-0.007	0.790	0.265	0.695
得点	0.530	0.406	0.695	0.928
安打	0.047	0.849	0.373	0.862
二塁打	0.354	0.610	-0.054	0.500
三塁打	-0.167	0.169	0.491	0.298
本塁打	0.892	0.240	-0.211	0.899
打点	0.814	0.409	-0.298	0.919
盗塁	-0.057	0.020	0.666	0.447
犠打	-0.307	-0.424	0.277	0.351
犠飛	0.234	0.381	-0.282	0.280
四球	0.336	0.130	0.451	0.334
三振	0.608	-0.014	0.061	0.373
併殺打	0.267	0.237	-0.664	0.569
固有値	2.601	2.534	2.320	
寄与率	0.200	0.195	0.178	
累積寄与率	0.200	0.395	0.573	



因子負荷行列では、抽出された 3 因子が、もとの打撃パラメータのどれと関連が強いかが示されてい

ます。因子1は、本塁打、打点、得点および三振と高い正の相関があり、かつ犠打と逆相関などの特性から、**長打力**を表す因子と解釈できます。因子2は、打率と安打、2塁打、打点、犠飛との正の相関、犠打との逆相関から、**巧打**と解釈されます。また因子3は、得点、3塁打と盗塁と正の相関、併殺打と逆相関することから、**俊足度**を表す因子と考えられます。

なお、共通性の推定値は、その値が低いと分析からの除外を考慮すべきですが、この例では、犠飛打が一番低いものの、因子2との相関を無視できないため、全ての変数をそのまま残すことになります。

この結果に基づき、因子得点を StatFlex 形式で求め、それを因子の大きさ順並べ替えると次ぎようになります。

	因子1	因子2	因子3
1 山田(ヤ)	2.218	0.515	1.037
2 筒香(D)	1.932	0.546	0.258
3 メヒア(西)	1.795	-0.227	-0.741
4 レアード	1.565	-0.014	-0.906
5 パシテ	1.413	-0.065	-1.019
6 松田(心)	1.140	-0.013	0.020
7 中田(日)	1.105	0.250	-1.251
8 丸(広)	1.094	0.832	1.284
9 ロベス(D)	1.084	-0.165	-1.061
10 テスバイ	1.083	0.644	-0.439
11 ウーラ	1.010	-0.343	-0.304
12 柳田(心)	0.770	0.085	0.931
13 ゴメス(神)	0.727	-0.451	-0.768
14 平田(中)	0.656	-1.030	-0.221
15 鈴木(広)	0.642	1.224	-0.184
16 新井(広)	0.630	0.585	-0.720
17 ビシエ	0.611	-0.523	-0.475
18 田中(広)	0.531	-0.532	2.631
19 坂本(巨)	0.434	1.384	0.992
20 ギャレット	0.399	-1.254	-0.940

	因子1	因子2	因子3
1 角中(口)	-0.787	1.870	0.456
2 浅村(西)	0.026	1.561	-0.641
3 内川(ソ)	-0.046	1.536	-1.597
4 坂本(巨)	0.434	1.384	0.992
5 鈴木(広)	0.642	1.224	-0.184
6 菊池(広)	-0.144	1.119	1.625
7 倉本(D)	-1.907	1.033	-0.997
8 糸井(才)	0.015	0.967	0.443
9 大島(中)	-1.262	0.953	1.679
10 村田(巨)	-0.226	0.936	-1.378
11 秋山(西)	0.114	0.928	1.574
12 丸(広)	1.094	0.832	1.284
13 テスバイ	1.083	0.644	-0.439
14 長野(巨)	-0.987	0.630	0.029
15 新井(広)	0.630	0.585	-0.720
16 福留(神)	-0.729	0.569	-0.578
17 筒香(D)	1.932	0.546	0.258
18 坂口(ヤ)	-1.076	0.544	1.017
19 鈴木(口)	-0.737	0.517	-0.264
20 山田(ヤ)	2.218	0.515	1.037

	因子1	因子2	因子3
1 田中(広)	0.531	-0.532	2.631
2 大島(中)	-1.262	0.953	1.679
3 菊池(広)	-0.144	1.119	1.625
4 秋山(西)	0.114	0.928	1.574
5 西川(日)	-0.594	0.070	1.537
6 中島(直)	-0.381	-2.061	1.442
7 丸(広)	1.094	0.832	1.284
8 山田(ヤ)	2.218	0.515	1.037
9 坂口(ヤ)	-1.076	0.544	1.017
10 坂本(巨)	0.434	1.384	0.992
11 金子(侑)	-0.822	-0.911	0.971
12 桑原(D)	0.110	-0.660	0.955
13 柳田(心)	0.770	0.085	0.931
14 今宮(心)	0.203	-1.294	0.591
15 西野(才)	-1.196	-0.293	0.561
16 梶谷(D)	0.384	-1.469	0.553
17 角中(口)	-0.787	1.870	0.456
18 糸井(才)	0.015	0.967	0.443
19 中村(具)	-0.499	-0.292	0.411
20 筒香(D)	1.932	0.546	0.258

いずれの因子も、因子スコアの平均値は0.0、標準偏差はおよそ1.0となるように調整されているため、スコアが2.0を超える値は他の選手よりもその因子得点が著しく高いこととなります。部分的にしか表示されていませんが、因子1は、山田(ヤ)と筒香(D)が高く、メヒア(西)がそれに続き、因子2は角中(口)が高く、浅村(西)・内川(ソ)で高め、因子3は田中(広)が際だっており、大島(中)、菊池(広)がそれに続くことがわかります。

Rによる分析結果(1)

```
> dat = read.csv("DB型 打撃成績.CSV", header=T)
> res = factanal(data=dat, factors=3, ~打率+得点+安打+二塁打+三塁打+本塁打+打点+盗塁+犠打+
+ 犠飛+四球+三振+併殺打)
> res$loadings
```

Loadings:

	Factor1	Factor2	Factor3
打率		0.790	0.265
得点	0.530	0.406	0.695
安打		0.849	0.373
二塁打	0.354	0.610	
三塁打	-0.167	0.169	0.491
本塁打	0.892	0.240	-0.211
打点	0.814	0.409	-0.298
盗塁			0.666
犠打	-0.307	-0.424	0.277
犠飛	0.234	0.381	-0.282
四球	0.336	0.130	0.451
三振	0.608		
併殺打	0.267	0.237	-0.664

	Factor1	Factor2	Factor3
SS loadings	2.601	2.534	2.320
Proportion Var	0.200	0.195	0.178
Cumulative Var	0.200	0.395	0.573

例題 19： [SampleData] の [DB 型 因子分析.SFD6] は、基準範囲設定調査で得られた健常男性 273 例の 11 項目の主要臨床検査結果である。これから、各臨床検査値の間にどのような共通性があるかを調べよ。



データファイルを読み込み、[統計] メニューの [多変量解析] の [因子分析] を実行します。このデータでは、全 11 検査を**対象変数**として指定します。推定因子数を 3 とし、因子回転法はバリマックス法を指定して、**実行** ボタンを押すと、次の出力結果が得られます。



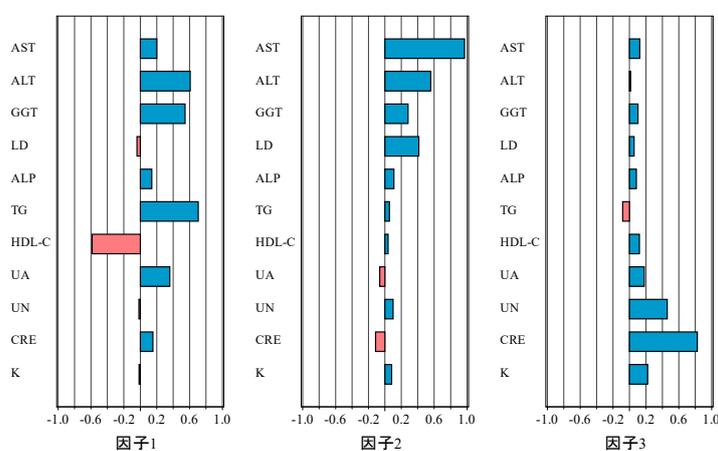
計算出力例

《(因子分析)》

[第1頁:] 有効データ数=273

《(因子負荷行列と寄与率) (バリマックス法)》

変数名	因子1	因子2	因子3	共通性
AST	0.201	0.969	0.124	0.995
ALT	0.608	0.559	0.014	0.682
GGT	0.545	0.281	0.104	0.386
LD	-0.039	0.413	0.057	0.175
ALP	0.141	0.111	0.083	0.039
TG	0.706	0.054	-0.083	0.509
HDL-C	-0.586	0.039	0.120	0.359
UA	0.359	-0.063	0.179	0.165
UN	-0.016	0.101	0.460	0.222
CRE	0.153	-0.113	0.826	0.718
K	-0.016	0.083	0.221	0.056
固有値	1.723	1.552	1.032	
寄与率	0.157	0.141	0.094	
累積寄与率	0.157	0.298	0.392	



この結果から、**因子 1** は ALT、GGT、TG、UA と強い正の相関、HDL-C と負の相関があることから、**栄養過多**を表す因子、**因子 2** は、AST、ALT、LD との相関から**肝細胞障害**を表す因子、**因子 3** は、UN、CRE との相関から**腎機能**を表す因子と解釈できます。

一方、共通性の推定値から、ALP と K のそれは非常に低く、この分析にほとんど関係しておらず省略可能であることが分かります。



Rによる分析結果 (2)

```
> dat = read.csv("DB型 因子分析.csv", header=T)
> res = factanal(data=dat, factors=3, ~log(AST)+log(ALT)+log(GGT)+LD+ALP+
+           log(TG)+HDL.C+UA+UN+CRE+K, rotation="varimax")
> res$loadings
```

Loadings:

	Factor1	Factor2	Factor3
log(AST)	0.201	0.969	0.124
log(ALT)	0.608	0.559	
log(GGT)	0.545	0.281	0.104
LD		0.413	
ALP	0.141	0.111	
log(TG)	0.706		
HDL.C	-0.586		0.120
UA	0.359		0.179
UN		0.101	0.460
CRE	0.153	-0.113	0.826
K			0.221

	Factor1	Factor2	Factor3
SS loadings	1.723	1.552	1.032
Proportion Var	0.157	0.141	0.094
Cumulative Var	0.157	0.298	0.392