

## 第9講 多変量解析入門

田中重人 (東北大学文学部准教授)

[テーマ] 多変量解析の種類と、因子分析の基本的な考えかた

### 1 前回課題について

「差」の変数をつくったときの「平均」「標準誤差」の意味。

対応のあるデータの場合、平均値の差の信頼区間を求める際の数式の標準誤差 (standard error) を、相関係数を用いて調整する。この点が、通常の (対応のない) 平均値の場合と異なる。

→ 相関図を描いて考えてみよう

### 2 対応のある分析について: 結果の書きかた

#### 2.1 個々の結果を表示する十分なスペースがある場合

クロス表 (または相関図) をいちいち示すのが基本 (別紙参照)。各セルには、度数と全体での%を書く。統計量などは表の下に書く。必要な統計量は分析法によって違うので注意。

- 対応のある t 検定 → 相関係数、平均値の差、有意水準 (対応のある検定であることを明記)
- 符号検定 →  $x > y$  ケースと  $x < y$  ケースの比率、有意水準

#### 2.2 スペースがあまりない場合

対応のある t 検定であれば、各変数の平均と SD の表をのせる。表の下に、人数、相関係数、平均値の差、有意水準 (対応のある検定であることを明記) を書く。

符号検定であれば、 $x > y$ ,  $x = y$ ,  $x < y$  各ケースの比率の表をのせる。表の下に、有意水準 (符号検定であることを明記) を書く。

#### 2.3 多数の変数間の関連を示す場合

ハッセ図 (Hasse diagram) が使える。平均値などの高い順に変数を並べ、有意な差がある変数どうしを線でむすぶ。具体例は別紙参照。

### 3 多変量解析とは

3つ以上の変数をつかう分析法を「多変量解析」(multivariate analysis) という。次の2種類に分けられる(大野, 1998, p.48-56)。

- 類似関係型: 似た変数同士をまとめたり、潜在因子を取り出したりするもの。因子分析, クラスター分析など
- 因果関係型: 原因と結果の関係を追究するもの。回帰分析, 分散分析, 一般線型モデルなど

この授業では前者をあつかう。

### 4 課題

相関係数行列を出力して、「似ている」変数のグループを探す。

### 文献

大野高裕 (1998) 『多変量解析入門』 同友館.

三土修平 (1997) 『初歩からの多変量統計』 日本評論社.

表1 自分にとって大切なこと

高い地位を得ること(x)	家族の信頼・尊敬を得ること (y)				合計
	1	2	3	4	
1. そう思う	13 (5.4)	1 (0.4)	0 (0.0)	1 (0.4)	15 (6.3)
2. どちらかといえぼさう思う	35 (14.6)	12 (5.0)	2 (0.8)	0 (0.0)	49 (20.5)
3. どちらかといえぼさう思わない	79 (33.1)	37 (15.5)	9 (3.8)	0 (0.0)	125 (52.3)
4. そう思わない	32 (13.4)	15 (6.3)	3 (1.3)	0 (0.0)	50 (20.9)
合計	159 (66.5)	65 (27.2)	14 (5.9)	1 (0.4)	239 (100.0)

度数 (全体%) を示す。

平均値の差=1.48 ( $x=2.88, y=1.40$ ),  $p<0.05$  (対応のある t 検定による)。  $r=0.073$ 。

$x>y$  ケース84.1%,  $x<y$  ケース1.7%,  $p<0.05$  (符号検定)。

対応のある t 検定の場合

符号検定の場合

表2 自分にとって大切なこと

	平均	SD
高い地位を得ること	2.88	0.81
家族の信頼・尊敬を得ること	1.40	0.62

平均値の差=1.48,  $p < 0.05$  (対応のある  $t$  検定による)。  $r = 0.073$ 。  $N = 239$ 。

表3 自分にとって大切なこと

	$N$	(%)
$x > y$	201	(84.1)
$x = y$	34	(13.6)
$x < y$	4	(1.7)
合計	239	(100.0)

$x$ : 高い地位を得ること,  $y$ : 家族の信頼・尊敬を得ること。  
 $p < 0.05$  (符号検定)。