

## 第10講 回帰分析

田中重人 (東北大学文学部教授)

[テーマ] 回帰分析の基礎

### 1 変数間の関連を確認するための加工

年齢や学歴のような変数は、こまかくわかれていたり、順序に統一性がなかったりするので、そのままでは使いくい。

適当なカテゴリーに統合して、クロス表を見る：

- 年齢 → 10 歳刻みなど
- 学歴 → 初等・中等・高等の3区分 (前回資料 参照)

間隔尺度あるいは順序尺度としてあつかい、相関係数を見る：

- 年齢はそのまま比率尺度として使える
- 学歴は、「教育年数」(その学歴取得に必要な標準的年限) に変換することが多い (前期第4講資料 参照)

SPSS のシンタックスはつぎのようになる：

```
recode q1_2a
( 20 thru 29 = 20 )
( 30 thru 39 = 30 )
( 40 thru 49 = 40 )
( 50 thru 59 = 50 )
( 60 thru 70 = 60 )
into age10.
```

```
recode q6_1
(1 thru 2 = 1) (3 thru 5 = 2) (6 thru 7 = 3)
(12 = 1) (13 = 2) (14 thru 17 = 3)
into edu3.
```

```
recode q6_1
(1 = 6) (2 = 8) (3 thru 5 = 11) (6 = 14) (7 = 17)
(12 = 9) (13 = 12) (14 = 14) (15 = 16) (16 = 18)
into eduyear.
```

### 2 回帰分析のモデルとパラメータ

#### 2.1 独立変数がひとつだけのモデル (単回帰分析)

従属変数として Q39g を、独立変数として Q1\_2a を投入して回帰分析を実行 (前回資料参照)。

切片 (A) と回帰係数 (B) を確認。

## 2.2 最小2乗法

回帰分析では、最小2乗法 (least square method) で係数を求める。これは、適当な直線  $A + BX$  によって  $Y$  の値を近似する方法であり、 $Y$  と  $A + BX$  とのずれの大きさを評価するために、差の2乗和をとる。この2乗和  $\sum(Y - A - BX)^2$  が最小になるように  $A$  と  $B$  の組み合わせを求める。

**回帰係数  $B$  の意味:**  $X$  が1つ増えたとき  $Y$  がどれだけ増えるか

教科書 78-81 頁参照

## 2.3 独立変数が複数の場合 (重回帰分析)

学歴を「教育年数」(上記参照)に変換したものを、独立変数に追加。この場合、回帰係数 ( $B$ ) が独立変数の数だけあることになる。

$$Q39g = \text{切片} + B_1X_1 + B_2X_2 \quad (1)$$

やはり最小2乗法で係数を求めるので、 $\sum(Y - A - B_1X_1 - B_2X_2)^2$  が最小になるように  $A$  と  $B_1$  と  $B_2$  の組み合わせを求める。

- 独立変数がひとつの場合と何が異なるか?
- 「コントロール」することの意味
  - 媒介効果
  - 擬似相関 (教科書 pp. 86-91)
- 分散分析表から独立変数の影響力の大きさを読む

## 3 期末レポート

**期限:** 2/3 (金)

**提出先:** Google Classroom

**内容:** 相関係数、対応のある分析、多変量解析について、それぞれ適当な分析をして結果を解釈する。すべての分析について、推定または検定結果をつける。データは何を使ってもよいが、SSM データ以外のものを使うときはデータについての説明をつけること。

**備考:** レポート提出後に、SSM データのコピーをすべて消去すること。

## 文献

吉川徹・轟亮 (1996) 「学校教育と戦後日本の社会意識の民主化」『教育社会学研究』 58:87-101.