

第10回 相関係数と回帰分析(実習)

- ・ 分析のために量的なデータを準備する。
- ・ 相関係数を計算する。
- ・ 回帰式を求める。

1. 量的データを準備する

ここでは、クロス表分析で用いたのと同じデータを、量的データに見立てて、分析を試みます。所得、学歴、年齢を量的データとして、以下のように準備します。

(1) 所得：各カテゴリーの中間の値をとります。

q52 仕事収入	新変数 income2
1 0～100万	50
2 100～300万	200
3 300～500万	400
4 500～700万	600
5 700～900万	800
6 900～1100万	1000
7 1100～1300万	1200
8 1300万～	1400
9 DK.NA	(欠損値)

SPSS 上では次のように新変数 income2 を用意します。

①最終行につきの入力する。

```
compute income2=q52.
recode income2 (1=50) (2=200) (3=400) (4=600) (5=800) (6=1000)
              (7=1200) (8=1400).
missing values income2 (9).
freq income2.
```

② compute 以下の部分を実行し、変数変換がうまくいっているかどうか確認します。

実行の仕方は、最終行においた場合、カーソルを compute の位置に置いて、「実行」→「最後まで」とするか、compute 行から、freq までを選択（輝度反転させ）て、「実行」→「選択部分」とする。（ただし、立ち上げてからまだ一度も実行していない場合には、「実行」→「すべて」とする）。

(2) 学歴：教育年数をしめす新変数 edctn を用意します。

q43 学歴	新変数 edctn
1 中学卒	9
2 高校卒	12
3 短大・高専卒	14
4 大学卒	16

①最終行につきのように入力する。

```
compute edctn=q43.  
recode edctn (1=9) (2=12) (3=14) (4=16).  
freq edctn.
```

② compute 以下を実行し、変数変換がうまくいっているかどうか確認する。

(3) 年齢 age を、5 歳刻みの素データの中間年をとった量的変数として用意する。

q42b 年齢	age
1 20 ~ 24	22
2 25 ~ 29	27
3 30 ~ 34	32
4 35 ~ 39	37
5 40 ~ 44	42
6 45 ~ 49	47
7 50 ~ 54	52
8 55 ~ 59	57
9 60 ~ 64	62
10 65 ~ 69	67
11 70 ~	70

※本来、70 歳以上はいないはずなのだが、なぜかサンプルに紛れ込んでいるので、70 歳とみなす。

①最終行につきのように入力する。

```
compute age=q42b.  
recode age(1=22) (2=27) (3=32) (4=37) (5=42) (6=47)  
          (7=52) (8=57) (9=62) (10=67) (11=70).  
missing values age(99).  
freq age.
```

②この部分を実行して、変数変換がうまくいっているかどうかを確認する。

2. 相関係数を計算する

相関係数を求める spss のコマンドは、corr.

corr 変数 1 with 変数 2. 変数 1 と変数 2 の相関係数を求める。

corr 変数 1 変数 2 with 変数 3. 変数 1 と変数 3、変数 2 と変数 3 の相関係数を求める。

corr 変数 1 with 変数 2 変数 3. 変数 1 と変数 2、変数 1 と変数 3 の相関係数を求める。

corr 変数 1 変数 2 変数 3. 変数 1 と変数 2、変数 1 と変数 3、変数 2 と変数 3 の相関係数を求める (すべての組み合わせについての相関係数のマトリックスが出力される)。

【1】年齢と所得、学歴と所得の相関係数をそれぞれ求めよ。

①ここでは、つぎのように入力する。

```
corr age edctn with income2.
```

②カーソルをこの行において「実行」→「現在の位置」

③結果表を読みましょう。相関係数がどこに出力されているか、分かりますね。

④結果表を、MS ワードにコピーし、「【1】年齢と所得、学歴と所得の相関係数」という見出しをつけておいてください。

【2】性別 q42a と所得 income2 との相関係数を求めよ。

性別は厳密には、量的変数ではありません。しかし、2つの値しかない2値変数(バイナリー変数)ですので、ダミー変数として、量的分析の対象とすることがあります。ダミー変数は、コード(0, 1)の変数のことです。

性別 q42a をもとに新変数 gender (男性 0 女性 1) をつくり、所得 income2 との相関係数を求め、【1】の結果と比較してみましょう。

(新変数 gender は、これまでの変数変換の例を参考に自分でつくってみましょう)。

結果は、同じワード文書に貼り込み、【2】性別(ダミー)と所得の相関係数という見出しをつけてください。

【1】【2】ができたなら、「相関係数」という文書名をつけて、自分のフォルダに保存したうえで、V-Campus CHORUSの「社会学データ実習」→「レポート課題」→「相関係数」をひらき、ファイルを転送してください。

3. 回帰式を求める

【例題】所得(単位:万円)を従属変数として、年齢(単位:歳)、学歴(単位:年)を

それぞれ独立変数とする 2 つの単回帰式を求めなさい。

$$\text{INCOME2} = a + b\text{AGE}$$

$$\text{INCOME2} = a + b\text{EDCTN}$$

SPSS のコマンドは、regression

regression /dependent 従属変数名 /method=enter 独立変数名.

所得を年齢に回帰させるには、

regression /dependent income2 /method=enter age.

所得を学歴に回帰させるには、

regression /dependent income2 /method=enter edctn.

結果の表から定数と B の値を読み取ると、

$$\text{INCOME2} = 251.9 + 3.78\text{AGE}$$

この式の解釈：年齢が 1 歳上がるごとに所得が約 3.8 万円増加する。

$$\text{INCOME2} = -212.6 + 46.7\text{EDCTN}$$

この式の解釈：教育年数が 1 年上がるごとに、所得が 46.7 万円増加する。

【3】年齢 10 歳刻みの age10 を使って、年齢から所得を予測する回帰式を求め、その結果を解釈しなさい。

結果は、「係数」の表を貼り付け、例題を参考に、回帰式を完成させ、解釈を 1 行で述べること。また、「【3】年齢 age10 から所得を予測する回帰式」という見出しをつけること。

【4】学歴 q43 は、1 = 中学卒 2 = 高校卒 3 = 短大・高専卒 4 = 大学卒というコードが与えられている。この変数を使って、所得を予測する回帰式を求め、その結果を解釈しなさい。

結果は、「係数」の表を貼り付け、例題を参考に、回帰式を完成させ、解釈を 1 行で述べること。また、「【3】学歴 q43 から所得を予測する回帰式」という見出しをつけること。

※このふたつの課題について、分からなければ、助教の先生や T.A. に質問してかまわない。

【3】【4】ができたなら、「回帰式」という文書名をつけて、自分のフォルダに保存したうえで、V-Campus CHORUS の「社会学データ実習」→「レポート課題」→「回帰式」をひらき、ファイルを転送してください。