

第5回 クロス表のエラボレーション

エラボレーション（精密化）：2変数間の関連を調べるクロス表に、第3の変数を絡めることにより、分析結果をさらに精密にすることができる。

例1（ザイゼル『数字で語る』p.170-171）

表 1-1 XX 食品の年齢別使用状況

	40歳未満	40歳以上
使用している	28	20
使用していない	72	80
	100	100
	(1,224)	(952)

XX 食品の使用には、年齢が関係している。追加要因として、性別も関係しているかもしれない（仮説）。

表 1-2 性別・年齢別の XX 食品の使用状況

	男		女	
	40歳未満	40歳以上	40歳未満	40歳以上
使用している	36	23	20	17
計の実数	619	480	605	472

性別も関係しているが、年齢も関係している。

XX 食品の使用は、女性よりも男性に多く、40歳以上よりも40歳未満に多い。

例2 偽りの相関を暴く例（ザイゼル、p.186-187）。

偽りの相関（擬似相関）：2つの変数のあいだに関連があるように見えるが、真の要因が第3の要因にあるような関連のこと。

表 2-1 配偶関係別キャンデーを食べる割合

	未婚	既婚
いつもキャンデーを食べる%	75	63
実数	999	2,010

表 2-2 年齢別キャンデーを食べる割合

	25歳未満	25歳以上
いつもキャンデーを食べる%	80	58
実数	1,302	1,707

表 2-3 年齢別・配偶関係別キャンデーを食べる割合

	未婚		既婚	
	25 歳未満	25 歳以上	25 歳未満	25 歳以上
いつもキャンデーを食べる%	79	60	81	58
実数	799	200	503	1,507

配偶関係がキャンデーの嗜好に及ぼす効果は「偽り」の相関。真の要因は、年齢であった。

例 3 東京総合社会調査のデータで、年齢と学歴が所得に及ぼす効果はどちらも真の関連であるのかどうか？

仮説 1 どの年齢階層においても、学歴が高くなるほど所得が多くなる傾向がある。

仮説 2 どの学歴階層においても、年齢が高くなるほど所得が多くなる傾向がある。

この 2 つの仮説を 2 つの 3 重クロス表によって検証する。

3 重クロス表のコマンド

`cross 独立変数 by 従属変数 by 統制変数 /cells=count row /statistics chisq phi.`

	仮説 1 の場合	仮説 2 の場合
独立変数	学歴 q43	年齢
従属変数	所得 income	所得
統制変数	年齢 age10	学歴

仮説 1 の場合のコマンド

`cross q43 by income by age10 /cells=count row /statistics chisq phi.`

出力結果

Q43 学歴とINCOMEとAGE10のクロス表

AGE10		INCOME					合計
		1.00 -300	2.00 300-700	3.00 700-1100	4.00 1100-		
1.00 20-29	Q43 学歴	1 中学卒	度数	1	1	0	2
			Q43 学歴の%	50.0%	50.0%	.0%	100.0%
	2 高校卒	度数	16	5	0	21	
		Q43 学歴の%	76.2%	23.8%	.0%	100.0%	
	3 短大卒	度数	21	8	0	29	
		Q43 学歴の%	72.4%	27.6%	.0%	100.0%	
	4 大学卒	度数	25	18	4	47	
		Q43 学歴の%	53.2%	38.3%	8.5%	100.0%	
	合計		度数	63	32	4	99
			Q43 学歴の%	63.6%	32.3%	4.0%	100.0%
2.00 30-39	Q43 学歴	1 中学卒	度数	6	2	0	8
			Q43 学歴の%	75.0%	25.0%	.0%	100.0%
	2 高校卒	度数	21	18	0	39	
		Q43 学歴の%	53.8%	46.2%	.0%	100.0%	
	3 短大卒	度数	20	12	0	32	
		Q43 学歴の%	62.5%	37.5%	.0%	100.0%	
	4 大学卒	度数	22	40	17	84	
		Q43 学歴の%	26.2%	47.6%	20.2%	100.0%	
	合計		度数	69	72	17	163
			Q43 学歴の%	42.3%	44.2%	10.4%	100.0%
3.00 40-49	Q43 学歴	1 中学卒	度数	1	1	0	2
			Q43 学歴の%	50.0%	50.0%	.0%	100.0%
	2 高校卒	度数	26	17	7	50	
		Q43 学歴の%	52.0%	34.0%	14.0%	100.0%	
	3 短大卒	度数	14	14	4	32	
		Q43 学歴の%	43.8%	43.8%	12.5%	100.0%	
	4 大学卒	度数	20	19	21	69	
		Q43 学歴の%	29.0%	27.5%	30.4%	100.0%	
	合計		度数	61	51	32	153
			Q43 学歴の%	39.9%	33.3%	20.9%	100.0%
4.00 50-59	Q43 学歴	1 中学卒	度数	10	9	2	21
			Q43 学歴の%	47.6%	42.9%	9.5%	100.0%
	2 高校卒	度数	40	22	25	91	
		Q43 学歴の%	44.0%	24.2%	27.5%	100.0%	
	3 短大卒	度数	8	9	4	21	
		Q43 学歴の%	38.1%	42.9%	19.0%	100.0%	
	4 大学卒	度数	5	13	22	54	
		Q43 学歴の%	9.3%	24.1%	40.7%	100.0%	
	合計		度数	63	53	53	187
			Q43 学歴の%	33.7%	28.3%	28.3%	100.0%
5.00 60-69	Q43 学歴	1 中学卒	度数	20	7	0	27
			Q43 学歴の%	74.1%	25.9%	.0%	100.0%
	2 高校卒	度数	46	22	10	78	
		Q43 学歴の%	59.0%	28.2%	12.8%	100.0%	
	3 短大卒	度数	7	4	1	12	
		Q43 学歴の%	58.3%	33.3%	8.3%	100.0%	
	4 大学卒	度数	14	11	7	42	
		Q43 学歴の%	33.3%	26.2%	16.7%	100.0%	
	合計		度数	87	44	18	159
			Q43 学歴の%	54.7%	27.7%	11.3%	100.0%

5つのクロス表が並んでいる。

カイ2乗検定

AGE10		値	自由度	漸近有意確率 (両側)
1.00 20-29	Pearson のカイ2乗	7.546 ¹	6	.273
	連続修正			
	尤度比	9.087	6	.169
	線型と線型による連関	4.425	1	.035
	有効なケースの数	99		
2.00 30-39	Pearson のカイ2乗	33.423 ²	9	.000
	連続修正			
	尤度比	41.983	9	.000
	線型と線型による連関	23.488	1	.000
	有効なケースの数	163		
3.00 40-49	Pearson のカイ2乗	22.570 ³	9	.007
	連続修正			
	尤度比	26.304	9	.002
	線型と線型による連関	15.937	1	.000
	有効なケースの数	153		
4.00 50-59	Pearson のカイ2乗	45.522 ⁴	9	.000
	連続修正			
	尤度比	49.236	9	.000
	線型と線型による連関	32.169	1	.000
	有効なケースの数	187		
5.00 60-69	Pearson のカイ2乗	38.080 ⁵	9	.000
	連続修正			
	尤度比	40.145	9	.000
	線型と線型による連関	25.784	1	.000
	有効なケースの数	159		

5つのクロス表の検定結果がまとめて表になっている。

1. 6セル(50.0%)は期待度数が5未満です。最小期待度数は.08です。
2. 9セル(56.3%)は期待度数が5未満です。最小期待度数は.25です。
3. 7セル(43.8%)は期待度数が5未満です。最小期待度数は.12です。
4. 2セル(12.5%)は期待度数が5未満です。最小期待度数は2.02です。
5. 8セル(50.0%)は期待度数が5未満です。最小期待度数は.75です。

対称性による類似度

AGE10		値	漸近標準誤差 ¹	近似 T 値 ²	近似有意確率
1.00 20-29	名義と名義	ファイ	.276		.273
		Cramer の V	.195		.273
	有効なケースの数		99		
2.00 30-39	名義と名義	ファイ	.453		.000
		Cramer の V	.261		.000
	有効なケースの数		163		
3.00 40-49	名義と名義	ファイ	.384		.007
		Cramer の V	.222		.007
	有効なケースの数		153		
4.00 50-59	名義と名義	ファイ	.493		.000
		Cramer の V	.285		.000
	有効なケースの数		187		
5.00 60-69	名義と名義	ファイ	.489		.000
		Cramer の V	.283		.000
	有効なケースの数		159		

1. 帰無仮説を仮定しません。
2. 帰無仮説を仮定して漸近標準誤差を使用します。

5つのクロス表の Cramer V がまとめて表になっている。

20 歳代においては、学歴と所得とのあいだに有意な関連はない。
しかし、30 歳以上では、60 歳代まで、学歴と所得とのあいだに有意な関連があり、概して学歴が高くなるほど、所得が多くなる傾向が見られる。

【SPSS プログラムについての解説】

SPSS コマンド・シンタックスの一般的規則（文法）

- ①コマンド・シンタックスは原則として半角英数字で構成されています。英文字は大文字でも小文字でもかまいません。
- ②スペースは原則として半角で入力します。スペースはいくつ打ってもかまいません。
- ③コマンドは複数行にわたってかまいません。
- ④複数行にわたるときの改行は、スペースの打てる場所で行います。
- ⑤ 1 行の文字数は自由ですが、画面表示や印刷出力などとの関係を考えて、80 桁以内にとどめておいた方がよいでしょう。改行をうまく利用して、見やすいコマンド・シンタックスを構成することが、ミスを減らすコツです。
- ⑥コマンドの最後にはピリオドを打ちます。ピリオドによって一連のシンタックスの終了が示されるので、ピリオドがないと、つぎのコマンドとの区別がつかず、エラーとなります。

要は、英作文と同じです。単語が二行に渡らないように、文章を書く感覚で OK ということです。

data list コマンド：データの読み取り方を指示するコマンド

```
data list file 'C:\¥Documents and Settings¥Yasushi Matsumoto¥My Documents¥tgss2000授  
業用¥tgss2000.txt'
```

→データの所在を指示

```
fixed record=3 /1 chiten 2 sample 3-5 q1 6
```

→固定書式、3 レコード、1 行目、...

```
q2a 7 q2b 8 q2c 9 q2d 10 q2e 11 q2f 12 q3 13 q4 14
```

```
q5 15 q6 16 q7 17 q8 18 q9 19 q10 20 q11 21 q12 22
```

```
q13a 23 q13b 24 q13c 25 q13d 26 q13e 27 q13f 28
```

```
q14a 29-30 q14b 31-32 q14c 33-34 q14d 35-36
```

```
q15a 37-38 q15b 39-40 q15c 41-42 q15d 43-44
```

```
q16a 45-46 q16b 47-48 q16c 49-50 q16d 51-52
```

```
q17 53-54
```

```
q18a 55-56 q18b 57-58 q18c 59-60 q18d 61-62
```

```
q19a1 63 q19a2 64 q19a3 65 q19a4 66 q19a5 67
```

```
q19a6 68 q19a7 69 q19a8 70 q19a9 71
```

```
q19b1 72 q19b2 73 q19b3 74 q19b4 75 q19b5 76
```

```
q19b6 77 q19b7 78 q19b8 79 q19b9 80
```

```
/2 q19c1 2 q19c2 3 q19c3 4 q19c4 5 q19c5 6
```

→ 2 行目

```
q19c6 7 q19c7 8 q19c8 9 q19c9 10
```

```
q19d1 11 q19d2 12 q19d3 13 q19d4 14 q19d5 15
```

社会学データ実習

q19d6 16 q19d7 17 q19d8 18 q19d9 19
q19e1 20 q19e2 21 q19e3 22 q19e4 23 q19e5 24
q19e6 25 q19e7 26 q19e8 27 q19e9 28
q19f1 29 q19f2 30 q19f3 31 q19f4 32 q19f5 33
q19f6 34 q19f7 35 q19f8 36 q19f9 37
q20a 38 q20b 39 q20c 40 q20d 41 q20e 42
q20f 43 q20g 44 q20h 45 q20i 46
q21a 47 q21b 48 q21c 49
q22a 50 q22b 51 q22c 52 q22d 53 q22e 54 q22f 55
q23 56
q24a 57 q24b 58 q24c 59 q24d 60 q24e 61
q24f 62 q24g 63 q24h 64 q24i 65
q25 66 q26 67 q27 68 q28 69 q29 70-71
q30 72 q31 73 q32 74-76 q33 77 q34 78 q34sq 79-80
/3 q35 2 q36 3 q36sq 4-5 q37 6 q38 7 q39 8 q40 9-10
q41 11 q42a 12 q42b 13-14 q43 15 q44 16 q45 17
q46 18-20 q47 21 q48 22 q48s 23-25 q49 26
q50_1 27 q50_2 28 q50_3 29 q50_4 30 q50_5 31
q50_6 32 q50_7 33 q50_8 34 q50_9 35 q50_10 36
q50_11 37 q50_12 38 q50_13 39 q50_14 40
q51 41-42 q52 43 q53 44 q54 45 q55 46-47
q56 48 q57 49 q58 50 q59 51 q60 52 q60s 53-55 q61 56.

→ 3 行目

→ここでデータが終わる。