

2001.5.9 現代日本論演習 I (田中重人)

第5回「係数」

1. 自由度 (degree of freedom)
2. クロス表分析のふたつの系列
3. 2×2 クロス表の性質
4. 係数 (phi coefficient)

【自由度】

2×2 クロス表では、周辺度数が所与なら、1つのセル度数が決まればほかも決まる

	1	2	合計
1	a	$g - a$	g
2	$i - a$	$h - i + a$	h
合計	i	j	N

3 × 3 クロス表：セル度数が 4 つ決まれば...

	1	2	3	合計
1				f
2				g
3				h
合計	i	j	m	N

$k \times l$ クロス表の自由度 (degree of freedom)

$$\text{d.f.} = (k - 1)(l - 1)$$

【クロス表分析の2つの系列】

「%の差」系 (期待度数との差)
= 連関係数

オッズ比系 (乗法モデル)
= 対数線形分析、ロジット分析

この授業で取り上げるのは前者だけ

【2×2 クロス表の性質】

以下、つぎの記号法を使う

	1	2	合計
1	a	c	g
2	b	d	h
合計	i	j	N

(1) 行%は1列についてだけ比較すればよい:

$$\frac{a}{g} - \frac{b}{h} = \frac{d}{h} - \frac{c}{g}$$

(2) 行%の差がゼロなら列%の差もゼロ

(3) $g=i$ なら行%の差と列%の差は同じ:

$$\frac{a}{g} - \frac{b}{h} = \frac{a}{i} - \frac{c}{j}$$

(例 1) 行%の差 = 8 %

60%	40%	100%
52%	48%	100%

(例 2) 行・列とも%に差なし

52	48	100
52.0%	48.0%	100.0%
66.7%	66.7%	

26	24	50
52.0%	48.0%	100.0%
33.3%	33.3%	

78	72	150
52.0%	48.0%	100.0%

(例 3) 行・列とも 10%の差

70	30	100
70.0%	30.0%	100.0%
70.0%	60.0%	

30	20	50
60.0%	40.0%	100.0%
30.0%	40.0%	

100	50	150
52.0%	48.0%	100.0%

【 係数 】

2×2 クロス表の「連関」の尺度

$$\phi = \frac{ad - bc}{\sqrt{ghij}}$$

この係数の意味は？

(分子だけ取り出して考えてみよう)

【SPSS での 係数の計算】

「クロス集計表」の「統計」で「ファイとクラマーの V」をチェック

【 係数の別定義 】

Pearson 積率相関係数 (教科書 72 頁) の特殊ケース

(四分点相関係数 4-fold point correlation)

の絶対値 $|r|$ を使うこともある (教科書 110 頁)

【%の差と 係数】

の絶対値 = 行・列の比率差の幾何平均：

$$|\phi| = \sqrt{\underbrace{\left| \frac{a}{g} - \frac{b}{h} \right|}_{\text{行 \% の差}} \underbrace{\left| \frac{a}{i} - \frac{c}{j} \right|}_{\text{列 \% の差}}} = \sqrt{\left| \frac{c}{g} - \frac{d}{h} \right| \left| \frac{b}{i} - \frac{d}{j} \right|}$$

周辺度数がつりあっていれば ($g=i$)、行%の差と列%の差は等しい。