

第10回 6月19日の講義内容

- § 4. 予備的分析
 - 作表
 - 独立性検定、相関

6/19/08

1

§ 4. 予備的分析

- 作表
 - 単純集計
 - クロス集計: ex.) 分割表
 - 平均や標準偏差など統計量を見るだけでなく、分布の形状にも注意
 - Bi-modalのケース
 - 歪んだ分布のケース

6/19/08

2

§ 4. 予備的分析(2)

- 独立性の検定: カイ2乗検定
 - 関係性の有無をチェック
 - 統計手法のみからは因果関係はわからない
- 相関係数
 - 相関の有無、程度をみる
 - 検定を行う場合には、帰無仮説に注意!!
 - $H_0: r=0$
 - 相関 = 二変数の線形関係の程度を表す尺度
 - 相関係数からは非線形関係の存在は分らない

6/19/08

3

クロス集計の例

(上野(2004)「日本企業の多角化経営と組織構造」『組織科学』vol.37(3))

		事業数の増減			合計
		減少	維持	増加	
既存主力 事業への 投資	縮小	6	4	3	13
	維持	26	44	18	88
	拡大	24	26	16	66
合計		56	74	37	167

6/19/08

4

クロス集計の例(2)

- 独立性の検定

$$c = \frac{\sum_{i=1}^k \sum_{j=1}^l (x_{ij} - e_{ij})^2}{e_{ij}}$$

$$c \xrightarrow{d} \chi^2((k-1) \times (l-1))$$

- 前表のケース
 - $c=2.98$
 - $\chi^2(4)$ の上位5%有意点は9.4877

➡ 帰無仮説(独立性)を棄却できない

6/19/08

5

相関係数の種類

- Pearsonの積率相関係数

- 一般に「相関係数」と呼んでいるもの

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) \left(\sum_{i=1}^n (y_i - \bar{y})^2\right)}}$$

- Non-parametricな関係性の指標

- グッドマン = クラスカルの順序連関係数
- Goodman-Kruskal Gamma
- Spearmanの順位相関係数
 - 順位変数の相関

次週説明

6/19/08

6

Spearmanの順位相関係数(1)

- 右の表のデータ
 - 標本の大きさ: 5
 - Xの平均: 3.2
 - Yの平均: 10.8
 - Xの2乗和
9+16+36+1+4=66
 - Yの2乗和
100+81+225+64+144=614
 - XとYの積和
30+36+90+8+24=188

	X	Y
	3	10
	4	9
	6	15
	1	8
	2	12

6/19/08

7

Spearmanの順位相関係数(2)

- まずPearsonの積率相関係数を求める。

$$\sum_{i=1}^5 (X_i - \bar{X})^2 = \sum_{i=1}^5 X_i^2 - 5\bar{X}^2 = 66 - 5 \times (3.2)^2 = 14.8$$

$$\sum_{i=1}^5 (Y_i - \bar{Y})^2 = \sum_{i=1}^5 Y_i^2 - 5\bar{Y}^2 = 614 - 5 \times (10.8)^2 = 30.8$$

$$\sum_{i=1}^5 (X_i - \bar{X})(Y_i - \bar{Y}) = \sum_{i=1}^5 X_i Y_i - 5\bar{X}\bar{Y} = 188 - 5 \times 3.2 \times 10.8 = 15.2$$

$$r = \frac{\sum_{i=1}^5 (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^5 (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^5 (Y_i - \bar{Y})^2}} = \frac{15.2}{\sqrt{14.8} \sqrt{30.8}} = 0.71$$

(Pearsonの相関係数)

6/19/08

8

Spearmanの順位相関係数(3)

- 下表のデータX, Yを、小さい順に順位をつけ変換し、W, Zとする

X	Y	W	Z
3	10	3	3
4	9	4	2
6	15	5	5
1	8	1	1
2	12	2	4

6/19/08

9

Spearmanの順位相関係数(4)

- 順位データW, Zの相関係数を求める。

$$\sum_{i=1}^5 (W_i - \bar{W})^2 = \sum_{i=1}^5 W_i^2 - 5\bar{W}^2 = 55 - 5 \times 3^2 = 10$$

$$\sum_{i=1}^5 (Z_i - \bar{Z})^2 = \sum_{i=1}^5 Z_i^2 - 5\bar{Z}^2 = 55 - 5 \times 3^2 = 10$$

$$\sum_{i=1}^5 (W_i - \bar{W})(Z_i - \bar{Z}) = \sum_{i=1}^5 W_i Z_i - 5\bar{W}\bar{Z} = 51 - 5 \times 3 \times 3 = 6$$

$$r^* = \frac{\sum_{i=1}^5 (W_i - \bar{W})(Z_i - \bar{Z})}{\sqrt{\sum_{i=1}^5 (W_i - \bar{W})^2} \sqrt{\sum_{i=1}^5 (Z_i - \bar{Z})^2}} = \frac{6}{\sqrt{10} \sqrt{10}} = 0.6$$

Spearmanの順位相関係数

6/19/08

10