

## 第11回 7月6日の講義内容

### □ § 4. 予備的分析

- 作表
- 独立性検定、相関

7/6/05

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

7/6/05

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**

7/6/05

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<br>事業への<br>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 |

7/6/05

4

## クロス集計の例(2)

### □ 独立性の検定

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

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

#### ■ 前表のケース

□  $c_0 = 2.98$

□  $\chi^2(4)$  の上位5%有意点は9.4877



帰無仮説(独立性)を棄却できない

7/6/05

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

7/6/05

6

## グッドマン＝クラスカルの順序連関係数

### □ 順序のあるカテゴリー変数の関係性の指標

対となったデータ $(x_1, y_1)$ ,  $(x_2, y_2)$ を比較すると、下表の分類のどこかに入る。

あらゆるデータの対に対して、これを行い

$G = (\#A - \#B) / (\#A + \#B)$  を求める。

|             | $y_1 < y_2$ | $y_1 = y_2$ | $y_1 > y_2$ |
|-------------|-------------|-------------|-------------|
| $x_1 < x_2$ | A           | C           | B           |
| $x_1 = x_2$ | C           | C           | C           |
| $x_1 > x_2$ | B           | C           | A           |

7/6/05

7

## グッドマン＝クラスカルの順序連関係数(2)

### □ クロス表の例

標本数  $n = 167$       ペアの数 13,861

#A = 2556

#B = 2516

$G = (2556 - 2516) / (2556 + 2516) = 0.0078$

連関ナシ

7/6/05

8

【課題】サンプリング 解説

- 7/6(水)昼現在、13名提出
  - うち、3名は誤りあり
  - 有効答案10
  - 次ページ以降にまとめ
- 母集団分布は右表のとおり

|     |     | 回答者数  | 比率    |
|-----|-----|-------|-------|
| 視聴率 |     | 1,182 | 39.4% |
| 携帯  | 不所持 | 108   | 3.6%  |
|     | D社  | 1,379 | 46.0% |
|     | A社  | 1,149 | 38.3% |
|     | V社  | 362   | 12.1% |

7/6/05

9

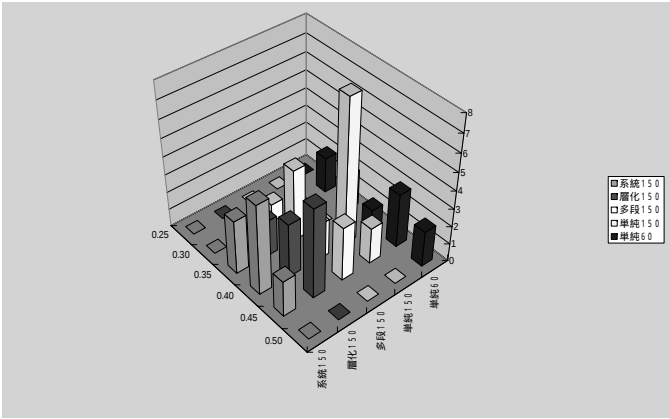
標本視聴率（母集団視聴率=39.4%）

| 視聴率  | 系統150 | 層化150 | 多段150 | 単純150 | 単純60 |
|------|-------|-------|-------|-------|------|
| 0.25 | 0     | 0     | 0     | 0     | 0    |
| 0.30 | 0     | 0     | 1     | 0     | 2    |
| 0.35 | 3     | 2     | 4     | 0     | 2    |
| 0.40 | 5     | 3     | 2     | 8     | 1    |
| 0.45 | 2     | 5     | 3     | 2     | 3    |
| 0.50 | 0     | 0     | 0     | 0     | 2    |

7/6/05

10

分布



7/6/05

11