

7/12講義後スライド変更

## 第12回 7月12日の講義内容

- § 4. 予備的分析(補足)
  - グッドマン=クラスカルの順序連関係数(再)
- 7月5日提出課題について
- [補足]層別抽出したときの母集団平均の推定について

7/12/06

1

## 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

7/12/06

2

## 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

7/12/06

3

## #Aの具体的な求め方

- セル $(i, j)$ を固定し、そのセルより右下にあるセルの度数を $N_{ij}^+$ とする。
- $n_{11}$ に対応する $N_{11}^+$ は $n_{22} + n_{23} + n_{32} + n_{33}$
- $n_{21}$ に対応する $N_{21}^+$ は $n_{32} + n_{33}$

	減少	不変	増加
減少	$n_{11}$	$n_{12}$	$n_{13}$
不変	$n_{21}$	$n_{22}$	$n_{23}$
増加	$n_{31}$	$n_{32}$	$n_{33}$

7/12/06

4

## #Bの具体的な求め方

- セル $(i, j)$ を固定し、そのセルより左下にあるセルの度数を $N_{ij}^-$ とする。
- $n_{12}$ に対応する $N_{12}^-$ は $n_{21} + n_{31}$
- $n_{13}$ に対応する $N_{13}^-$ は $n_{21} + n_{31} + n_{22} + n_{32}$

	減少	不変	増加
減少	$n_{11}$	$n_{12}$	$n_{13}$
不変	$n_{21}$	$n_{22}$	$n_{23}$
増加	$n_{31}$	$n_{32}$	$n_{33}$

7/12/06

5

## クロス表の例

- 標本数  $n=167$       ペアの数 13,861
- $\#A=2556$   
 $=6 \times (44+18+26+16)+4 \times (18+16)+25 \times (26+16)+44 \times 16$
- $\#B=2516$   
 $=4 \times (26+24)+3 \times (26+24+44+26)+44 \times 24+18 \times (24+26)$
- $G=(2556-2516)/(2556+2516)=0.0078$
- 連関ナシ

7/12/06

6

### 【課題】(7/5提出分) サンプリング 解説

- 今回の課題提出例
  - 有効答案8
  - 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		510	17.0%
携帯	不所持	150	5.0%
	D社	1575	52.5%
	A社	795	26.5%
	V社	480	16.0%

7/12/06

7

### 標本視聴率の分布 (母集団視聴率=17.0%)

7/12講義後スライド追加

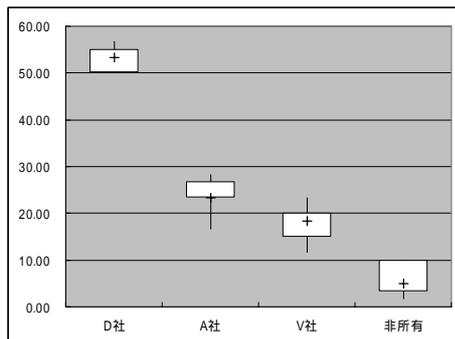
視聴率	単純60	単純150	系統150	二段150	層化150
6.5～9.5%	0	1	1	0	0
9.5～12.5%	0	0	0	0	1
12.5～15.5%	0	1	3	2	3
15.5～18.5%	1	3	1	3	2
18.5～21.5%	1	2	2	1	1
21.5～24.5%	3	0	0	0	0
24.5～27.5%	2	0	0	0	0

7/12/06

8

### 携帯の保有率の分布(単純60) 箱ヒゲ図(Box-Plot)

7/12講義後スライド追加

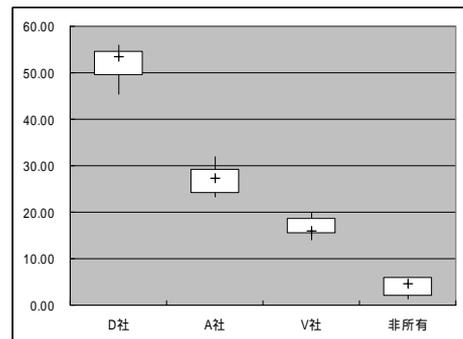


7/12/06

9

### 携帯の保有率の分布(単純150) 箱ヒゲ図

7/12講義後スライド追加



7/12/06

10

### 【補足】層別抽出した ときの母集団平均の推定について

7/12講義後スライド追加

- 層別抽出を行った場合、層ごとの事後抽出率(=抽出確率×回収率)が同一でないとき、抽出(後)標本の算術平均を母集団平均の推定量に用いてはいけません。
- このようなケースでは、抽出率調整を行って母集団平均の推定をする必要がある。

7/12/06

11

### 抽出率調整した推定量

7/12講義後スライド追加

- 抽出率調整の方法を例を示して説明する
- 仮定
  - 母集団は層Aと層Bの二層から構成される。
  - 母集団における層Aの個体数を $N_A$ 、層Bの個体数を $N_B$ とする。
  - 層Aの抽出割当数を $n_A$ 、層Bの抽出割当数を $n_B$ とする。
  - 層Aの回収数を $k_A$ 、層Bの回収数を $k_B$ とする。
  - 回収標本中の個体の回答を $x_i$ とする。

7/12/06

12

7/12講義後スライド追加

## 抽出率調整した推定量(2)

- 事後抽出率が層ごとに異なる場合、抽出後標本における層比率と母集団における層比率が異なるため、調整が必要となる。
- 抽出率調整後の母集団平均の推定量

$$\tilde{\mu} = \frac{1}{N_A + N_B} \times \left( \frac{N_A}{k_A} \sum_{i=1}^{k_A} x_i + \frac{N_B}{k_B} \sum_{j=1}^{k_B} x_j \right)$$

7/12/06

13

7/12講義後スライド追加

## 抽出率調整の有無の関係

- 抽出率調整を行わない推定量

$$\hat{\mu} = \frac{\sum_{i=1}^{k_A} x_i + \sum_{j=1}^{k_B} x_j}{k_A + k_B}$$

$$(c.f.) \tilde{\mu} = \frac{\sum_{i=1}^{k_A} x_i}{k_A + k_A \frac{N_B}{N_A}} + \frac{\sum_{j=1}^{k_B} x_j}{k_B \frac{N_A}{N_B} + k_B}$$

7/12/06

14

7/12講義後スライド追加

## 抽出率調整の有無の関係(2)

- 抽出率調整が不要な場合
  - 抽出後標本における層比率と母集団における層比率が同一のとき
  - 層ごとの事前抽出率が同一にコントロールしてあるときは、層ごとの回収率が同一のとき  
の二ケースのみ

$$\frac{k_A}{n_A} \times \frac{n_A}{N_A} = \frac{k_A}{N_A} = \frac{k_B}{N_B} = \frac{k_B}{n_B} \times \frac{n_B}{N_B}$$

$$\Rightarrow \begin{cases} k_A = k_B \frac{N_A}{N_B} \\ k_B = k_A \frac{N_B}{N_A} \end{cases}$$

回収率      事前抽出率

7/12/06

15