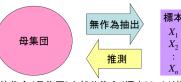
世論調査の数学 (統計的推測入門)

能本大学工学部 数理工学科 岩佐 学

標本調査 - 母集団と標本 -



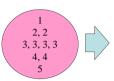
- ◆全体集合(母集団)を部分集合(標本)により推測する。
- ◆母集団分布:母集団に何がどれくらいあるか(多様性)
- ◆標本調査(統計的推測)の対象は母集団分布となる。

内閣支持率: 母集団は国民全体、支持者: 不支持者の比率 視聴率調査:ある地域の住民全体、どの番組がどれくらいの 人に視聴されているか?

無作為抽出

街頭インタ ビューは無 作為か?

- ◆標本抽出は無作為に行うことが必要
- ◆無作為とは、母集団に属する個体を等確率で選択すること を意味する(特定の個体が選ばれやすくなってはならない)



無作為に一つ抽出した個体の 出現値Xの確率分布は

ĺ	х	1	2	3	4	5	計
	確率	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	1

母集団分布(多様性) = 標本Xの確率分布(ランダム性)

無作為抽出により母集団の情報が標本に忠実にコピーされる

毎回の抽出の結果に関連性がない → 独立

P(X = i, Y = j) = P(X = i) P(Y = j)

抽出の反復の独立性

復元抽出は独立であるが、非復元抽出は独立でない。 独立性が成り立てば、確率計算が簡単になる。

母集団: 1 2 2

1 2 計 復元抽出 1/9 2/9 1/3 2 2/9 4/9 2/3 計 1/3 2/3

非復元抽出

-									
	XY	1	2	計					
	1	0	2/6	1/3					
	2	2/6	2/6	2/3					
	計	1/3	2/3	1					

煙★· V V

標本の持つ情報量

復元抽出

サイズ50の標本

サイズ500の標本

個体数 1億の 母集団

個体数

1000**の**

母集団



サイズ50の標本

サイズ500の標本

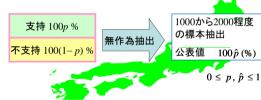
割合でなく個数

どちらが 情報量が

多い?

復元抽出であれば母集団の大きさと標本の情報量は無関係

世論調査(内閣支持率



- ◆標本における支持率 p̂ を発表 (真の支持率 p との誤差を認識する必要性)
- ◆調査主体による支持率の差(調査法,抽出法の差)

1万人の調査で±1%

視聴率調查

- ◆実施主体:ビデオリサーチ社 (http://www.videor.co.jp/top.htm)
- ◆母集団は各地方(関東、関西、名古屋、北部九州、長崎 など27地域)
- ◆世帯単位の標本抽出(600 or 200)
- ◆世帯視聴率と個人視聴率
- ◆機械による調査(2種類)とアンケートによる調査
- ◆調査世帯のローテーション(2年 or 3年) 毎週 600/24 = 25 世帯ずつ調査世帯が入れ替わる

サンプリング

- ◆地域毎に100~600の世帯を無作為抽出(系統抽出法)
- ◆関東地方 15万世帯 から 600世帯 を抽出とすると
- ① 15万世帯に1から150,000までの番号をつける。
- (2) 150.000 \div 600 = 25.000
- ③ 1 から 25,000 まで整数 x を一つ無作為に選ぶ。
- ④ x + 25.000 n (n=0.....599) により 600 の数値を決定。

乱数は一つだけ。簡便。

ナンバリングにより地域から万遍なく抽出可能。

非復元抽出

標準誤差 $\hat{p} = 0.2$ のとき2、

理論と経験

統計的推測

母集団

標本

全体集合 理論

部分集合 経験

現象

実験. 観測 データ

モデル 共通の理解

個人の体験,意見

統計的推測理論

= 経験を理論へ昇華させるために必須の理論