

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り(確認テスト・相互採点・リフレクションの記入)

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

【内容目標】調査方法の違いや注意点を理解しよう

□全数調査と標本調査 ← 標本調査については中学のおさらいです

統計的な調査には、対象全体からデータを集めて調べる **全数調査** という方法がある。たとえば、5年に1度行われる国勢調査は全数調査である。全数調査は、多くの費用と労力を必要とするので簡単には行えない。これに対して、調査の対象全体からその一部を抜き出して調べる **標本調査** という方法がある。

調査の対象全体が多数のときは、標本調査を行って、少ない費用と労力で全体の傾向をつかむことが有効な方法として採用されている。たとえば、テレビ番組の視聴率を調べる方法は標本調査である。

出口調査で選挙結果が分かるからつまらない。  
開票率5%で当確なんておかしいと  
数学者・秋山仁と話したら。  
「それが統計学ですよ」  
「まだ開票率5%なの？」  
「あなたね、味噌汁作って味見するのに  
丼鉢でグーッと飲む？」  
「・・・小皿ですよ」  
「それが5%よ」

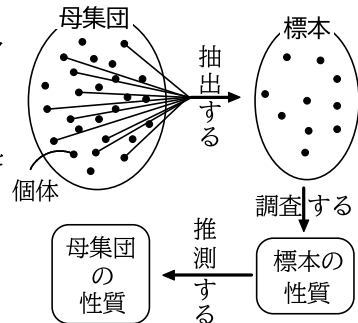
左の話は立川志の輔さんの落語にある話です。  
味噌汁全飲みが全数検査ということになり、鍋の味噌汁全体に均等に味噌が行き渡っていると仮定して、鍋一杯の全量に対してほんの少しだけ小皿にとって味見して、鍋の味噌汁の味噌の濃度を考えるのが標本調査となります。ただし、「ちゃんと混ぜてるのか」とか「どれぐらいのちょっとならよいのか」という課題があります。そのような事をこの後考えていきます。

標本調査の場合、調査の対象全体を **母集団** という。母集団に属する個々の対象を **個体** といい、個体の総数を **母集団の大きさ** という。

また、調査のため母集団から抜き出された個体の集合を **標本** といい、母集団から標本を抜き出すことを **抽出** という。標本に属する個体の総数を **標本の大きさ** という。

標本調査の目的は、抽出された標本から母集団のもつ性質を正しく推測することにある。そのためには、標本が偏りなく公平に抽出されることが必要である。

母集団の各個体を等しい確率で抽出する方法を **無作為抽出** といい、無作為抽出によって選ばれた標本を **無作為標本** という。



□無作為抽出の方法

無作為抽出では、**乱数さい** や **乱数表** などが使われる。

乱数さいは、正二十面体のさいころで、0から9までの数字が2回ずつ刻まれている。たとえば、青白赤の3個の乱数さいを投げて、それぞれに出た目の数を、3桁の数の百、十、一の各位の数とすることで、0から999までの数の1つを無作為抽出することができる。



	10			20							
	07	11	51	83	35	12	35	37	59	13	64
	77	75	52	21	20	82	26	52	29	06	47
	44	25	57	16	26	73	48	47	15	19	32
	86	79	39	57	33	99	69	31	32	42	36
	82	66	49	35	24	65	79	11	37	74	06
	62	66	89	51	70	44	20	34	68	35	55
	98	06	44	45	67	76	92	33	82	72	49
	71	28	23	55	52	67	72	00	05	18	07
	67	72	64	67	06	84	71	08	74	47	76
10	12	73	22	89	84	99	87	32	57	09	40
	42	25	31	79	54	98	70	23	91	40	41

世界最初の乱数表はティベット (L.H.C Tippett) が1927年頃に作ったもの。  
乱数表はどこから始めても、どのような規則に従って選んでも0から9までの数字がほぼ等しい規則で出現する。

また、非常に大きい母集団から標本を抽出するときは、コンピュータを利用すると便利である。右の写真は、コンピュータを利用して、1から100000までの10万個の数から40個の数字を無作為に抽出したものである。

	A	B	C	D	E	F	G	H
1	70120	92561	14289	11126	45989	8359	96016	84718
2	51038	57182	93874	76599	77620	60457	61400	24480
3	56201	57559	82828	89808	3698	60134	11024	74483
4	70638	14696	89537	571	21772	81843	51397	3732
5	38653	82897	62425	54345	86044	65549	78386	17034

コンピュータが発生するのは真の意味の乱数ではなく疑似乱数と呼ばれる非常に大きな周期をもつ周期列である。

コラム

標本の抽出方法

統計的な推測は、標本調査にもとづく推定や仮説検定によって行われます。母集団のもつ性質を標本を通じて推測するわけですから、標本の抽出を適切に行うことは非常に重要になります。標本を抽出する方法にもいくつか種類があります。

■層化無作為抽出法

母集団にいくつかの層（たとえば、性別、年代別、職業別など）が含まれる場合に、各層からデータを偏りなく得るために層ごとに無作為抽出する方法を、層化無作為抽出法といいます。層によるばらつきを小さくして観測の精度を上げるなどのねらいがあります。

■クラスター抽出法

母集団を地域など複数の部分集団（クラスター）に分割し、いくつかの部分集団を抽出して、その集団に対しては全数調査を行う方法を、クラスター抽出法といいます。あらかじめ部分集団ごとの名簿があれば時間と費用を軽減することができます。

標本調査を行う場合は、どのような抽出方法が適しているかも検討することが大事です。

【練習】 標本の抽出方法について、どのような方法があるか調べてみよう。

また、それらの抽出方法について、よい点や問題点をまとめてみよう。

□復元抽出と非復元抽出

母集団から標本を抽出するのに、

独立な確率変数（独立な試行や反復試行のときをイメージしよう）

毎回もとにもどしながら個体を1個ずつ抽出することを **復元抽出** という。

これに対して、個体をもとにもどさないで標本を抽出することを **非復元抽出** という。

標本調査のときは非復元抽出であることが多い

互いに独立ではない（条件付き確率のようなもの）

例19) 1から100までの番号札100枚の中から、大きさ5の標本を抽出するときの標本の総数を求める。ただし、抜き出した順序を区別する。

復元抽出では、 $100^5$  通りの標本ができる。

もとに戻して抽出するので積の法則により  
 $100 \times 100 \times 100 \times 100 \times 100 = 100^5$

非復元抽出では、 ${}_{100}P_5$  通りの標本ができる。

もとに戻さないで、異なる100個の番号から異なる5個を取って並べる順列なので、その総数は  ${}_{100}P_5$

ちなみに  $100^5 = (10^2)^5 = 10^{10} = 10 \times 10^9$

${}_{100}P_5 = 100 \cdot 99 \cdot 98 \cdot 97 \cdot 96 = 9034502400 \approx 9 \times 10^9$