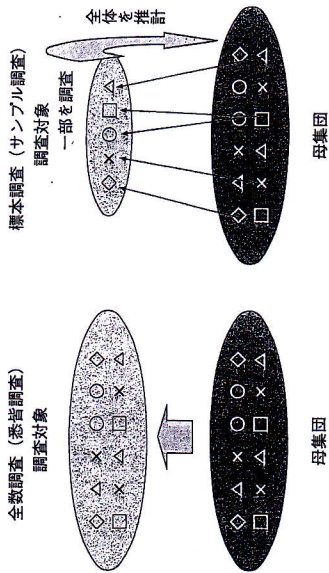


図1 全数調査と標本調査



標本抽出の無作為化

母集団から標本を取り出すためにある個体を選ぶとき、その母集団中のどの個体も選ばれる確率が等しく、かつそれぞれの個体が独立して選ばれる必要がある。そのような標本の取り出し方を、無作為抽出 randomization と呼び、母集団の状態を標本から偏りなく推定する上で重要となる。無作為抽出の方法論は多数あるが、次の4つが代表的な方法である。

■ 単純無作為抽出法

標本を抽出する際に、人為的要素（作意や主観など）の混入を防ぐために、カードまたはくじをひいたり、乱数表（無作為に並んだ数値の表）を用いたりして、個体番号を決める。またコンピュータで擬似乱数を発生させて標本を抽出する方法もある。

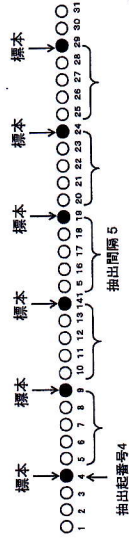
04 43 13 37 00	79 68 96 26 60	70 39 83 66 56	62 03 55 86 57	77 55 33 62 02
84 90 90 66 77	63 99 25 69 02	09 04 03 35 78	19 79 95 07 21	02 84 48 51 97
28 55 53 09 48	86 28 30 02 35	71 30 32 06 47	93 74 21 86 33	49 90 21 69 74
05 33 18 08 51	51 78 57 26 17	34 87 96 23 95	89 99 93 39 79	11 28 94 15 52
28 89 65 87 08	13 50 63 04 23	25 47 57 91 13	52 62 24 19 94	91 67 48 57 10
30 29 43 65 42	78 66 28 55 80	47 46 41 90 08	55 98 78 10 70	49 92 05 12 07
01 85 54 96 72	66 86 65 64 60	56 59 75 36 75	46 44 33 63 71	54 50 06 44 75
10 91 46 96 86	19 83 52 47 53	65 00 51 93 51	30 80 05 19 29	56 23 27 19 03
95 74 62 60 53	51 57 32 22 27	12 72 72 77 77	44 67 32 23 13	67 95 07 76 30
05 85 40 25 24	73 52 93 70 50	48 21 47 74 63	17 27 27 51 26	35 96 29 00 45

乱数表の見本

先頭の数字を目を閉じて選び、そこから順に数値を読み取っていく。例えば1~80の番号から任意の数値を10個取り出すには、始点の乱数から順に数値を読み、その範囲に入らなかった値があれば抜き出し、10個抽出されれば終了する。また、11~20のランダムな並びを作るには、その範囲の数値に相当するものを順に抜き取り書き出ししていく。同じ数値が選ばれたときは、2回目は兼用する。この抽出操作を全ての数値が揃うまで行うが、全ての数値が揃うのに、先の例よりも時間がかかる。

■ 系統抽出法

母集団の全個体に通し番号をつける。それ以降の個体（抽出開始番号）だけは乱数表などでランダムに選ぶ。それ以降の個体は、その数字から始めて一定間隔で順に抽出する。



抽出方法	メリットとデメリット
単純無作為抽出法 ● 乱数表を使って無作為抽出をする最も基本的な方法 系統抽出法（等間隔抽出法） ● 最初の1人を乱数表で抽出し、2人目以降は、母集団の構成員数を抽出しようとする標本数で割って算出した抽出間隔をもとに、名簿上で等間隔に抽出していく方法	● 少人数の抽出に適している ● 効率が悪い ● 標本数が多い場合に適している ● 名簿の番号に周期性がある（部署別など）場合は要注意
層化抽出法（層別抽出法） ● 母集団を属性などによって階層化し、その階層ごとに抽出する方法。階層ごとの人口規模に応じて標本数を各階層に割り当てる 2段階抽出法（多段階抽出法） ● 母集団からまず第一次抽出単位を抽出し、さらにその中から第二次抽出単位を抽出する方法。たとえばある県を母集団とする場合、まずその県の市町村から抽出し、抽出された市町村の中から再び標本を抽出する。第三次抽出単位以上を抽出する場合、多段階抽出	● 手続きを適切に行えば、予算を抑え、標本誤差を最小化できる ● 予算を抑えられる ● 標本誤差が大きくなる
層別2段階抽出法 ● はじめに抽選を人口規模や産業構造によって階層化しておき、各層から地域を抽出したあと、第二次抽出を行う方法	● 全国を対象とする調査に適している

表1 無作為抽出の方法