

■平均値

分布するデータのすべての値を足し合わせて、データ数(回答数)で割った値。なお、平均値は極端に高い数値、または低い数値などのはずれ値がある場合、適切な統計量とはいえなくなるので注意が必要である。たとえば、年収が200万円の人9人、1000万円の人1人だと、平均は280万円になるが、多くの人は平均値を下回っている。

■中央値と最頻値

中央値は、分布するデータを大ききの順に並べて中央に位置する値。データ数が偶数の場合は、中央値2つの平均を取る。最頻値は、分布するデータのなかで最も頻繁に現れる値をいう。

■歪度と尖度

- ①歪度：データの分布が平均値を中心に左右対称になっっているかどうかを示す指標のことをいう。データが低い方に偏っている場合を「正」、左右対称を「0」、高い方に偏っている場合を「負」と表す。
- ②尖度：データの分布の中心部が突出しているかどうかを示す指標のこと。中心部のデータが突出している場合を「正」、低い場合を「負」と表す。

表2-3 中央値と最頻値

No	年収
1	1000万
2	900万
3	800万
4	800万
5	700万
6	700万
7	500万
8	500万
9	500万
10	300万
11	200万



■度数と度数分布(単純集計)  
各段間の単純集計で選択肢ごとの度数の分布を表の形式に整理する。度数分布を度数と相対度数(パーセント)で表示し把握しやすくする(度数分布表、単純集計表)。

表2-1 単純集計表の例(年代の回答)

No	カテゴリ	件数	割合 (%)
1	20 歳代	157	10.0
2	30 歳代	222	14.2
3	40 歳代	254	16.2
4	50 歳代	284	18.8
5	60 歳代	288	18.4
6	70 歳代	343	21.9
7	無回答	7	0.4
サンプル数 (%ベース)		1565	100.0

■クロス集計

対象者の属性項目と意識や態度を示す項目との関係を見る必要から、それぞれの段間を掛けあわせて集計すること。また、仮説の検証との関係で段間間のクロス集計を行うこともある。なおクロス集計表において、値の大小を比較するだけでは、変数間の関連を意味しているとは限らず、後述する統計的検定で検証する必要がある。

表2-2 クロス集計表の例(ボランティア活動の経験)

上段：度数 下段：%	年齢	1年間のボランティア活動経験		
		合計	ある	ない
合計		1565	481	1001
		100.0	30.7	64.0
20 歳代		157	20	135
		100.0	12.7	86.0
30 歳代		222	53	163
		100.0	23.9	73.4
40 歳代		254	74	175
		100.0	29.1	68.9
50 歳代		284	84	191
		100.0	32.0	65.0
60 歳代		288	117	156
		100.0	40.6	54.2
70 歳代		343	122	176
		100.0	35.6	51.3
無回答		7	1	5
		100.0	14.3	71.4

■t検定

2つの母集団の平均値が等しいかどうかの検討には平均の差の検定を用いる。この平均の差を検定する際に有効な検定法である。母集団の分散が知られていない場合に用いる。

■ $\chi^2$ (カイ二乗)検定

クロス表などにおいて、母集団において2変数の間に何の関係もない場合に期待される各セルの値(期待値)と実際に調査等で得た値(観測値)がどの程度一致するかを測定して関連の有無を検定する方法。

■相関係数

2つの変数の間の関連の度合を数値で測定する方法。ピアソンの積率相関係数に代表される。これは、2つの変数を標準偏差を用いて変数間を標準化するため、測定単位(たとえばcmとcmなど)の影響を受けない。この係数は、1に近いほど関連が強く、0に近づくにしたがい関連が低くなる。また、相関の方向性を正の相関(一方が増加すると、もう一方も増加する)と負の相関(一方が増加すると一方が減少する)で表す。

相関係数の値が大きいか小さいかは絶対的なものではなく、変数間の相関関係の強弱を示し、あくまで相対的なものである。

■回帰分析

2つの変数の測定値の間で、一方の変数の変動を他方の変数の変動によってどの程度正確に予測することが可能かという考えに基づいて、2変数間の関連を分析する統計的方法。

■多変量解析

①重回帰分析：変数が2個以上の場合の回帰分析。1つの被説明変数yを予測するために複数の説明変数 $x_1, \dots, x_p$ の1次結合を求める方法。

②因子分析：一般的な行動様式に関するデータを多項目にわたって調査した際、それらの項目のいくつかに共通した要因が作用していると考えられる場合、それぞれのデータが示す数多くの変数を、より少数の共通因子に輪約し、解釈しようとする方法。