

代表値の種類

~代表値の選択の仕方~

●代表値とは？

- 統計データを求める際、度数分布の中心的傾向を数値で要約的に示す特性値。
- 代表値とは一組のデータに含まれる情報を1つの数値で代表するもので、**計算的代表値**と**位置的代表値**の2種類がある。

計数的代表値-1

- 計数的代表値(数学的代表値)は、**算術平均、幾何平均、調和平均、平方平均**などに細分される。

算術平均(arithmetic mean)

- 算術平均(相加平均)は、**データの総和をデータの個数で割った値**で最も多様される代表値。
- 単に平均という場合には、算術平均が一般的である。

幾何平均(geometric mean)

- 個々のデータの相乗積のN乗根幾何平均(相乗平均)という。
- 幾何平均によって、比率のような相対量の平均を算出する。

調和平均(harmonic mean)

- 各データの逆数の算術平均の逆数を、調和平均という。
- 逆数の形になっているデータを平均するときを使う。

平方平均(Root Mean Square)

- 大きなデータを特に強く影響させる方法。

位置的代表値

- 位置的代表値(具体的代表値)は、中央値、最頻値、四分位数、などに細分される。

これらは、度数分布中の特定位置を占める一つまたは少数のデータを全データの代表値とみなすもので、データ中の極端な値の影響を受けることが小さいという長所がある。

中央値(Median,中位数)

- 測定値を小さい順に並べたとき、ちょうど真ん中にくる値である。分布の両端に大きな値や小さな値があっても影響されない。
- 有効ケース数を n 、各ケースの測定値を X_i ($i = 1, 2, \dots, n$) とすると、以下の式で定義される。

最頻値(Mode)

- 一組のデータの中で最も多く存在する値、最も頻りに現れる値のこと。
- 又、並み数、典型数ともいう。

最頻値(Mode)

- 一組のデータの中で最も多く存在する値、最も頻りに現れる値のこと。
- 又、並み数、典型数ともいう。
- 最多度数データなので、現実的・具体的な代表値として理解しやすく、実在する値。
- しばしば、商品の流行を代表するので、流行値ということもある。

分位数

- 分位数とは、中位数と同じ発想から、四分位数、五分位数、十分位数、百分位数など各%点でわけたもの。
- 四分位数とは、昇順に配列された全データ列の、 $1/4$ 、 $2/4$ 、 $3/4$ に位置するデータの値をそれぞれ、第1四分位数 Q_1 、第2四分位数 Q_2 、第3四分位数 Q_3 と呼びます。

まとめ

- 一般的に代表値といっても数多くの種類が存在し、ケースや求める値によって使い方も異なってくる。
- また、それぞれ代表値には特徴があり、よく理解する必要がある。