

## 利 用 上 の 注 意

- 1 項目名に（人口〇〇当たり）とあるもので特に断り書きのない場合は、都道府県編では総務庁統計局公表の当該年 10 月 1 日現在推計人口（国勢調査年については国勢調査人口）を使用し、市町村編では千葉県企画部統計課公表の当該年 10 月 1 日現在常住人口（国勢調査年については国勢調査人口）を使用している。
- 2 時点の「〇〇年」は暦年（1月～12月）、「〇〇年度」は会計年度（4月～翌年3月）を示している。
- 3 表中データ値は、原則として単位未満を四捨五入してある。このため、各データの合計が総数と一致しない場合がある。
- 4 記号の用法等は、次のとおりである。
  - 「\*」 推計値、または、暫定値
  - 「0」 データが「0」、または、単位未満のもの
  - 「-」 データが無い、または、「指標計算式」の分母データが「0」のもの
- 5 都道府県編は、データ値の高い順に都道府県を配列している。また、同順位の場合は、都道府県コード順に配列している。

市町村編は、平成10年4月1日現在市町村コード順に配列している。よって、データの時点にかかわらず旧「印西町」については「印西市」で表示している。

- 6 平均値、標準偏差及び偏差値について
  - (1) データ全体のばらつき具合等の特徴（〇〇から〇〇の範囲に広がっている等）や個別データの特徴（平均よりどのくらい上なのか下なのか等）を数値で表したものが平均値、標準偏差及び偏差値である。
  - (2) 算出方法  
 (データ全体を代表する値)
 
$$\text{平均値 } (\bar{X}) = \frac{\text{個別データの総和}}{\text{データ総数}}$$
 (データ全体のばらつき具合を表す値)

$$\text{標準偏差 } (\sigma) = \sqrt{\frac{(\text{個別データ値} - \text{平均値})^2 \text{の総和}}{\text{データ総数}}}$$

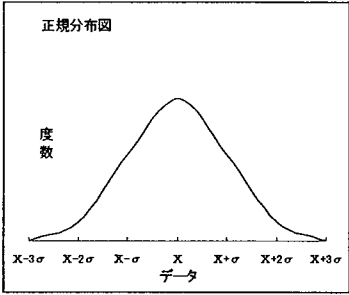
(平均値を「50」と表した場合、データ全体のばらつき具合から個別データの全体から見た位置を表す値)

$$\text{偏差値} = \frac{10 (\text{個別データ値} - \text{平均値})}{\text{標準偏差}} + 50$$

### (3) 標準偏差

データ集団をいくつかの級に分けた級ごとのデータ数を「度数」とよび、データと度数を対応させた図を「度数分布図」とよぶ。

このデータ集団が右図のように平均値を中心に左右対称でつり鐘型となる分布状態を「正規分布」とよび、この図を「正規分布図」とよぶ。



正規分布であるデータ集団と平均値、標準偏差との間には、次の関係が成立している。

- ①  $\bar{X} \pm \sigma$  の範囲に全データの 68.3% がはいつている
- ②  $\bar{X} \pm 2\sigma$  の範囲に全データの 95.5% がはいつている
- ③  $\bar{X} \pm 3\sigma$  の範囲に全データの 99.7% がはいつている

よって、 $\sigma$  が大きければ広がり幅は広くなり、 $\sigma$  が小さければ広がり幅は狭くなる。このことから、標準偏差  $\sigma$  によって、データ集団がばらつきの大きい集団なのか、ばらつきの小さい集団なのかがわかる。

#### (4) 偏差値

上記のように、データ集団のばらつき具合は様々であるため、実数データからだけでは、あるデータが集団の中でどの位置にいるのかわかりにくい。

このため、平均値を一律「50」とし、ばらつき（つまり標準偏差）を標準化して個々データの値を算出したものが「偏差値」である。

〈例〉

下記のようなデータ集団A、Bがあり、平均値は  $\bar{X}_A = \bar{X}_B = 70$  と等しい。

A : 60, 65, 65, 70, 70, 70, 70, 75, 75, 80

B : 40, 50, 60, 65, 70, 70, 75, 80, 90, 100

ここで、それぞれの集団の標準偏差を算出すると、

$\sigma_A \approx 5.5$ 、 $\sigma_B \approx 16.9$

となる。

このことから、集団A、Bは平均値は同じであってもばらつき具合に大きな違いがあり、集団Aはばらつきが小さく、集団Bはばらつきが大きいので、このままではデータ比較がしにくい。

そこで、集団Aにおけるデータ「75」と集団Bにおけるデータ「80」の偏差値を算出すると、

偏差値(Aの75)  $\approx 59.1$ 、偏差値(Bの80)  $\approx 55.9$

となる。

このことから、実数で見ると「75」より「80」の方が高いが、集団全体の中での位置づけからみると、集団Aにおける「75」の方が集団Bにおける「80」より高いことがわかる。

都道府県編では、平均値として各都道府県データ値の単純平均値を用いている。(よって、全国値と平均値は一致しないため、千葉県データ値が全国値を上回る場合でも偏差値が50を下回る場合がある。)

また、数値のみから事の善し悪しをいちがいに判断することは適当ではないため、ここでは一律にデータ値の高いものが偏差値が高くなるように算出している。

なお、市町村編では、平均値及び標準偏差のみ算出している。

#### 7 本書についての問い合わせ先

千葉県企画部統計課企画指導班 (TEL 043-223-2217)