# 第５章 統計の見方・考え方・使い方（読みとる）

まとめた統計について比率や代表値などを計算して特徴をはっきりさせたり，他との比較をして，どんな問題点があるかなどを読みとる。

## １．統計分析の手法

統計は利用されなければ価値がありません。

統計を利用するには，人口がいくら，農家数がいくらというように，そのまま使うこともありますが，多くの場合，数理的計算によって加工・分析を行い，いろいろな活用をはかります。

統計は前に説明したように，集団現象を数量的に表したものであり，その集団をどのようにみるかによって，いろいろな加工・分析の方法が考えられます。一つの集団について，その特徴をつかむために，これまで説明してきたように，いろいろな分類をして，表を作ったり，グラフを描いたりすることも，加工・分析の一方法であるといってよいでしょう。

そのほか，次のような加工・分析の方法がありますので，これから説明していきましょう。

**度数分布の分析**
**統計比例数の分析**

なお，この加工・分析の方法とは別に，最近，広く行われるようになった標本調査の基礎についても，最後に説明することにします。

## ２．度数分布の分析

### (1) 度数分布表と度数分布図

表１は，Ａ中学校２年生の男子と女子の体重を調べたものです。この表について，次のことを考えてみましょう。

**問１**　番号13番の男子の体重はいくらですか。

**問２**　番号32番の女子は，この学校の２年生女子のなかで体重は重い方から数えて何番目でしょうか。

**問３**　番号30番の男子は，この学校の２年生男子のなかで体重は重い方でしょうか。

**問４**　男子と女子を全体として比較した時，体重はどちらが重いでしょうか。

**表１　Ａ中学校２年生の体重**

| 男子 番号 | 体重 (kg) | 番号 | 体重 (kg) | 女子 番号 | 体重 (kg) | 番号 | 体重 (kg) |
|---|---|---|---|---|---|---|---|
| 1 | 40.0 | 22 | 49.5 | 1 | 58.2 | 22 | 39.0 |
| 2 | 47.5 | 23 | 56.2 | 2 | 48.5 | 23 | 40.5 |
| 3 | 52.2 | 24 | 51.5 | 3 | 41.0 | 24 | 50.0 |
| 4 | 51.4 | 25 | 41.0 | 4 | 43.0 | 25 | 37.3 |
| 5 | 44.5 | 26 | 42.8 | 5 | 46.6 | 26 | 43.6 |
| 6 | 44.6 | 27 | 55.6 | 6 | 42.5 | 27 | 47.4 |
| 7 | 56.2 | 28 | 41.5 | 7 | 44.5 | 28 | 50.5 |
| 8 | 55.0 | 29 | 36.2 | 8 | 44.2 | 29 | 49.8 |
| 9 | 44.4 | 30 | 53.4 | 9 | 52.0 | 30 | 52.0 |
| 10 | 48.8 | 31 | 42.0 | 10 | 49.4 | 31 | 47.2 |
| 11 | 49.3 | 32 | 43.0 | 11 | 45.8 | 32 | 56.5 |
| 12 | 44.6 | 33 | 45.6 | 12 | 46.0 | 33 | 44.5 |
| 13 | 48.0 | 34 | 46.0 | 13 | 46.6 | 34 | 45.0 |
| 14 | 40.5 | 35 | 60.8 | 14 | 47.5 | 35 | 51.4 |
| 15 | 43.8 | 36 | 46.2 | 15 | 59.2 | 36 | 52.6 |
| 16 | 42.4 | 37 | 38.5 | 16 | 45.4 | 37 | 46.5 |
| 17 | 61.0 | 38 | 39.4 | 17 | 41.0 | 38 | 47.0 |
| 18 | 37.6 | 39 | 40.6 | 18 | 41.5 | 39 | 46.8 |
| 19 | 44.2 | 40 | 54.0 | 19 | 42.6 | 40 | 42.2 |
| 20 | 37.0 | 41 | 52.3 | 20 | 40.5 | | |
| 21 | 64.0 | | | 21 | 44.8 | | |

問１の答えは48.0kg，問２の答えは３番目だということは，すぐにわかったと思いますが，問３・４はわかりにくかったのではないでしょうか。

このように，表１では，個々の生徒の体重はよくわかりますが，ある生徒が，全体の中で，重い方か，軽い方かとか，男子と女子とでは，全体としてどんな違いがあるか，というようなことは，すぐにはわかりません。

そこで，このようなことを知りたいときには，まず，目的に合うように，18から20ページで学習したことをもとにして，この資料を整理してみましょう。

次の表２は，Ａ中学校２年生男子の体重を**階級**ごとに分類したものです。

**表２　Ａ中学校２年生男子の体重階級別人数**

| 体　重（kg）以上 ～ 未満 | 人　数 |
|---|---|
| 35.0 ～ 38.0 | 3 |
| 38.0 ～ 41.0 | 5 |
| 41.0 ～ 44.0 | 7 |
| 44.0 ～ 47.0 | 8 |
| 47.0 ～ 50.0 | 5 |
| 50.0 ～ 53.0 | 4 |
| 53.0 ～ 56.0 | 4 |
| 56.0 ～ 59.0 | 2 |
| 59.0 ～ 62.0 | 2 |
| 62.0 ～ 65.0 | 1 |
| 計 | 41 |

この表の各階級に入る資料の個数を**度数**といい，このような表を**度数分布表**といいます。

続けて変わるような資料を分類するときには，階級の数を10程度にした方がよいでしょう。

この表２を利用すると，前にあげた問３について考えやすくなりましたね。

次に，度数分布の様子をグラフにすると，その状態が一層明らかになります。

図１は，表２をグラフにしたものです。このグラフは，階級の幅を底辺とし，度数を高さとする長方形を順々に並べて度数の分布を表しています。このようなグラフを**ヒストグラム**（柱状グラフ）といいます。

また，このヒストグラムの一つ一つの長方形の上の辺の中点を結んでいくと，図２のようなグラフができます。このようなグラフを**度数分布折れ線グラフ**（度数分布多角形）といいます。

**図１**

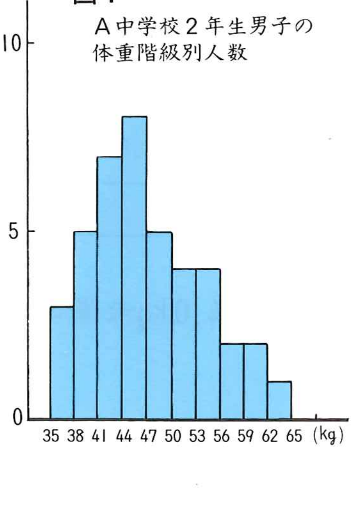
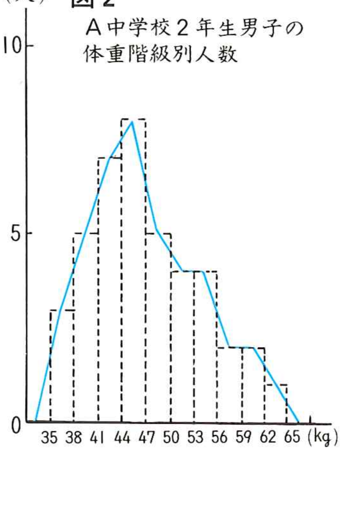



**図２**



### ヒストグラム・度数分布折れ線グラフ（度数分布多角形）を書くときの注意

ヒストグラムは，棒グラフによく似ていますが，長方形と長方形の間をあけずに書く必要があり，また，その両側は度数がなくても，底辺の幅一つ分だけをあけておくようにします。

度数分布折れ線グラフも，両側の度数0の階級の中点まで線を結んでおきます。

表3は，B中学校2年生男子の体重の度数分布表です。

表2と表3の二つの度数分布表から，二つの中学校の男子の体重を比較してみましょう。

ところが，二つの中学校の男子の生徒数が，異なっているために，分布の違いをはっきりとはつかむことができないことに気づくでしょう。

このような場合には，各階級の度数の全体に対する割合（その階級の**相対度数**という。）を求めていきます。

$$\text{ある階級の相対度数} = \frac{\text{その階級の度数}}{\text{全体の度数}}$$

**表3　B中学校2年生男子の体重階級別人数**

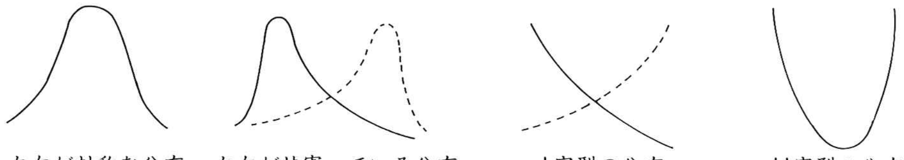
| 体　重（kg）<br>以上　未満 | 人　数 |
|---|---|
| 29.0 ～ 31.0 | 2 |
| 31.0 ～ 35.0 | 5 |
| 35.0 ～ 38.0 | 6 |
| 38.0 ～ 41.0 | 10 |
| 41.0 ～ 44.0 | 15 |
| 44.0 ～ 47.0 | 15 |
| 47.0 ～ 50.0 | 15 |
| 50.0 ～ 53.0 | 5 |
| 53.0 ～ 56.0 | 11 |
| 56.0 ～ 59.0 | 8 |
| 59.0 ～ 62.0 | 2 |
| 計 | 94 |

例えば，表3のB中学校の階級41.0kg以上44.0kg未満の相対度数は$\frac{15}{94} \fallingdotseq 0.16$となります。

度数分布折れ線グラフの折れ線にそってえがいた，なめらかな曲線を度数分布曲線といいます。

度数分布曲線の形から集団の状態をだいたい知ることができます。

この世の中のいろいろなできごとを度数分布曲線に表すと，次のようないくつかの形になってあらわれてきます。



左右が対称な分布（自然のできごとによくあらわれます）

左右が片寄っている分布（社会や経済のできごとによくあらわれます）

J字型の分布（社会や経済のできごとによくあらわれます）

U字型の分布（年齢別の死亡率などにあらわれます）

表4は，A・B中学校2年生男子の体重の相対度数分布表です。

このように，全体の度数が異なる資料を比較するときは，相対度数分布表を作って比較すると，分布の違いがつかみやすくなります。

**表4　　A・B中学校2年生男子の体重階級別人数**

| 体重（kg） | A中学校 | | B中学校 | |
|---|---|---|---|---|
| | 度数（人） | 相対度数 | 度数（人） | 相対度数 |
| 以上　未満 | | | | |
| 29.0 ～ 32.0 | 0 | 0.00 | 2 | 0.02 |
| 32.0 ～ 35.0 | 0 | 0.00 | 5 | 0.05 |
| 35.0 ～ 38.0 | 3 | 0.07 | 6 | 0.06 |
| 38.0 ～ 41.0 | 5 | 0.12 | 10 | 0.11 |
| 41.0 ～ 44.0 | 7 | 0.17 | 15 | 0.16 |
| 44.0 ～ 47.0 | 8 | 0.20 | 15 | 0.16 |
| 47.0 ～ 50.0 | 5 | 0.12 | 15 | 0.16 |
| 50.0 ～ 53.0 | 4 | 0.10 | 5 | 0.05 |
| 53.0 ～ 56.0 | 4 | 0.10 | 11 | 0.12 |
| 56.0 ～ 59.0 | 2 | 0.05 | 8 | 0.09 |
| 59.0 ～ 62.0 | 2 | 0.05 | 2 | 0.02 |
| 62.0 ～ 65.0 | 1 | 0.02 | 0 | 0.00 |
| 計 | 41 | 1.00 | 94 | 1.00 |

## ① 平　均

資料の一つ一つの数値がわかっている場合，平均は，それらの数値を加えて，その合計を個数で割れば計算できます。

$$平均=\frac{資料の数値の合計}{資料の個数}$$

**表7　　C中学校3年生男子の身長**

| 番号 | 身長(cm) | 番号 | 身長(cm) | 番号 | 身長(cm) |
|---|---|---|---|---|---|
| 1 | 167.2 | 11 | 163.8 | 21 | 166.5 |
| 2 | 151.7 | 12 | 164.4 | 22 | 173.2 |
| 3 | 162.2 | 13 | 160.9 | 23 | 164.0 |
| 4 | 165.7 | 14 | 152.3 | 24 | 169.1 |
| 5 | 169.9 | 15 | 165.3 | 25 | 154.8 |
| 6 | 156.5 | 16 | 164.9 | 26 | 167.4 |
| 7 | 158.4 | 17 | 166.3 | | |
| 8 | 162.8 | 18 | 174.8 | | |
| 9 | 156.4 | 19 | 157.9 | | |
| 10 | 168.7 | 20 | 147.4 | | |

この方法で，表7の資料から，3年生男子の身長の平均を計算すると，

$$\frac{身長の合計}{人数}=\frac{4232.5}{26}\fallingdotseq 162.8\ (\mathrm{cm})$$

となります。

**表8　　C中学校3年生男子の身長階級別人数**

| 身　長　(cm)<br>以上　　未満 | 度数(人) |
|---|---|
| 145.0 ～ 150.0 | 1 |
| 150.0 ～ 155.0 | 3 |
| 155.0 ～ 160.0 | 4 |
| 160.0 ～ 165.0 | 7 |
| 165.0 ～ 170.0 | 9 |
| 170.0 ～ 175.0 | 2 |
| 計 | 26 |

表8は表7を度数分布表にしたものです。

資料が度数分布表で与えられている場合は，階級の中央値（階級値）の欄を作ってそれぞれの階級ごとに「階級値×度数」を計算し，その合計を資料の個数で割れば平均が求められます。

$$平均=\frac{（階級値×階級の度数）の合計}{資　料　の　個　数}$$

**表9　度数分布表からの平均計算表**

| 身　長　(cm) | 度数(人) | 階級値(cm) | 階級値×度数 |
|---|---|---|---|
| 以上　未満 | | | |
| 145.0 ～ 150.0 | 1 | 147.5 | 147.5 |
| 150.0 ～ 155.0 | 3 | 152.5 | 457.5 |
| 155.0 ～ 160.0 | 4 | 157.5 | 630.0 |
| 160.0 ～ 165.0 | 7 | 162.5 | 1137.5 |
| 165.0 ～ 170.0 | 9 | 167.5 | 1507.5 |
| 170.0 ～ 175.0 | 2 | 172.5 | 345.0 |
| 計 | 26 | | 4225.0 |

表9は，表8に，「階級値」と「階級値×度数」の欄を作って計算したものです。

この表から計算すると，

$$平均=4225.0\div26=162.5\ (\mathrm{cm})$$

となります。

この計算方法でやっても，真の平均 $\left(\dfrac{4232.5}{26}\fallingdotseq162.8\right)$ との誤差はほんのわずかであるということがわかります。

ところが，この方法では階級値や度数が大きいと計算に時間がかかるので，もっと簡単な方法として，**仮の平均**を使う方法が考えられています。

仮の平均を使う例として，5人の身長が，148cm，152cm，153cm，156cm，163cmのときについて考えてみましょう。

（5人の身長の平均は

$$\frac{148+152+153+156+163}{5}=\frac{772}{5}=154.4\ (\mathrm{cm})$$

）

このとき，真ん中と思われる数値153cmを仮の平均として，それぞれの身長とこの仮の平均153cmとの差を計算すると

－5cm（148－153＝－5），－1cm，0cm，＋3cm，＋10cm

となります。

これらの差の平均は

$$\frac{(-5)+(-1)+0+(+3)+(+10)}{5}$$

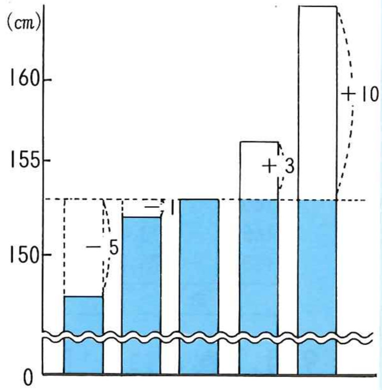
$$=\frac{+7}{5}=+1.4\ (\text{cm})$$

となり，身長の平均は

（仮の平均）＋（差の平均）＝（平均）

153cm＋（＋1.4cm）＝154.4cm

として求められます。




このような考え方を度数分布表に利用したのが，仮の平均を使う簡便法で，表10のように，「仮の平均との差」と「仮の平均との差×度数」の欄を作って計算します。

表10の計算の手順を説明しますと，階級値167.5cmを仮の平均とし，各階級値と仮の平均との差（167.5cmより大きいときは正の数，小さいときは負の数で表す。）を求め，それぞれの階級ごとに「仮の平均との差×度数」を計算し，その合計を度数で割ったものを仮の平均に加えると，

**表10　仮の平均を用いた平均計算表**

| 身長（cm） | 度数 | 階級値（人） | 仮の平均との差 | （仮の平均との差）×（度数） |
|---|---|---|---|---|
| 以上　未満 | | | | |
| 145.0～150.0 | 1 | 147.5 | −20.0 | −20.0 |
| 150.0～155.0 | 3 | 152.5 | −15.0 | −45.0 |
| 155.0～160.0 | 4 | 157.5 | −10.0 | −40.0 |
| 160.0～165.0 | 7 | 162.5 | −5.0 | −35.0 |
| 165.0～170.0 | 9 | 167.5 | 0.0 | 0.0 |
| 170.0～175.0 | 2 | 172.5 | 5.0 | 10.0 |
| 計 | 26 | | | −130.0 |

$$\text{平均}=167.5+\frac{-130.0}{26}=167.5+(-5)=162.5\ (\text{cm})$$

となり，これは，49ページで求めた平均と同じになります。

平均は代表値として非常によく使用されますが，平均を扱うときに注意をしなければならないことが二つあります。

○　一つは，平均は，一つ一つの数値の大きさをすべて同じ大きさにならした場合の値ですが，そのような値が必ずしも実際にあるとは限りません。

> 例えば，平成７年国勢調査による島根県の一般世帯一世帯あたりの家族数の平均は3.08人ですが，このように端数のついた人数の家はありません。

○　二つは，数値のなかに，極端に大きい値，または極端に小さい値があるとき，平均は，それによって影響されることが強いという欠点があります。

> 例えば，10人のうち，9人までがひと月1,000円の小遣いであるのに，ただ一人だけ20,000円の小遣いをもらっている人がまじっていたら，小遣いの平均は，この一人の金額の大きさの影響を受けて，2,900円になります。しかし，この金額は他の９人の小遣い1,000円とは大きくはなれており，小遣いの代表値としては適当ではありません。

このように，平均には，いくつかの欠点もありますが，集団が大きく，分布が対称のときには，安定した性質を示すことから代表値として広く利用されています。

### ②　メジアン（中央値）

集団の一つ一つの資料を大きさの順に並べたとき，真ん中にくる資料の数値を**メジアン**といいます。

メジアンは，数値のなかに極端に大きい値，または，極端に小さい値があるとき，平均に比べて影響をうけることが少ないので，平均よ

りも中心の傾向を表すのに適しているといわれています。

| |
|---|
| 168.5, 159.6, 160.4, 163.9, 164.0, 166.8, 169.4, 162.6, 158.2, 169.3, 170.4 |

上の数値は、11人の生徒の身長を測った結果です。これを身長の高い順（低い順でもよい）に並べかえると，メジアンは簡単に求めることができます。

生徒数が11人で奇数ですから，身長の高い方から数えて6番目の生徒の値，166.8cmがメジアンになります。

下の数値のように，生徒数が12人で偶数のときには，6番目と7番目との身長の平均を計算して，メジアンとします。

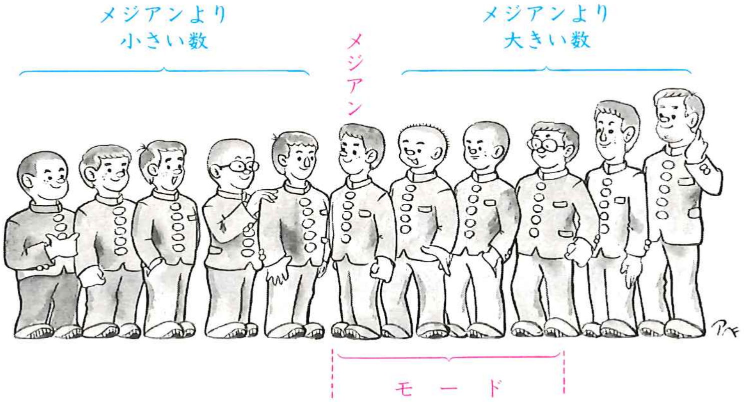
つまり，164.0cmと166.8cmとの平均165.4cmがメジアンです。

| |
|---|
| 158.2, 159.6, 160.4, 162.6, 163.9, 164.0, 166.8, 167.5, 168.5, 169.3, 169.4, 170.4 |

### ③　モード（最頻値）

資料の数値のなかで，最も多く出てくる数値のことで，度数分布表でいえば，度数の最大である階級の数値，または，階級値で，度数分布曲線では一番高いところの位置に相当する数値ということになります。

例えば，48ページの表8では，最大度数の階級は165.0cm以上～170.0cm未満で，その階級値167.5cmがモードになります。



## (3) 散布度

表11は，３年Ａ組の数学と英語のテストの成績を，度数分布表にしたものです。

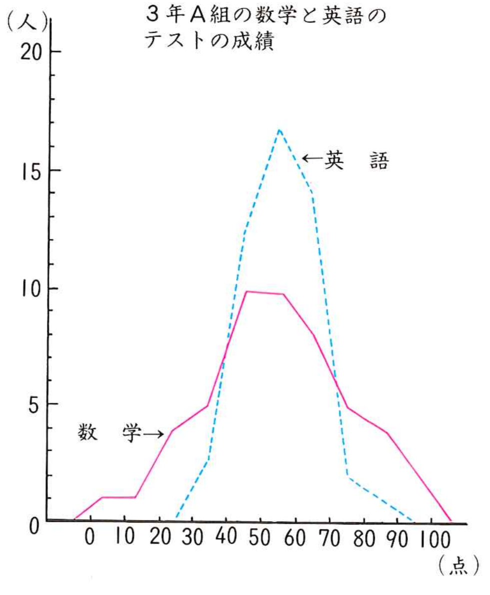
表11　３年Ａ組の数学と英語のテストの成績

| テストの点数 | 数学のテストの成績の度数（人） | 英語のテストの成績の度数（人） |
|---|---|---|
| 91～100 | 2 | 0 |
| 81～ 90 | 4 | 1 |
| 71～ 80 | 5 | 2 |
| 61～ 70 | 8 | 14 |
| 51～ 60 | 10 | 17 |
| 41～ 50 | 10 | 13 |
| 31～ 40 | 5 | 3 |
| 21～ 30 | 4 | 0 |
| 11～ 20 | 1 | 0 |
| 0～ 10 | 1 | 0 |
| 計 | 50 | 50 |

数学と英語のそれぞれの平均を計算すると，数学は54.4点，英語は55.4点で，ほとんど差がありません。

しかし，この度数分布表を度数分布折れ線グラフにして重ね合わせてみると，次の図５のようになります。

図５




これらの表やグラフをみると数学では，30点以下や91点以上の人がいますが，英語にはそのような人はいないことがわかります。

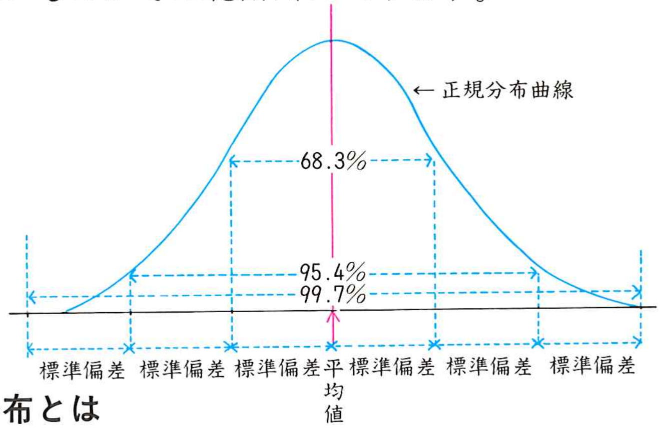
このように，二つの集団の平均が等しくても，同じような分布になるとは限りません。

だから，集団の特徴を表すときは，代表値だけでは不十分で，資料の一つ一つの値が分布の中心から離れている程度，つまり，ちらばりの程度をみることも大切です。

### ○標準偏差の性質

標準偏差は，平均と組み合わせて，統計分析を行う上で重要な役割を持っていて，標本調査を行ったり，調査を分析したりする場合に，なくてはならない大切な数値です。

例えば，集団が正規分布をしているときは，平均の両側に標準偏差の幅をとると，その範囲内には，集団の個数の68.3%(約$\frac{2}{3}$)が含まれ，同様に，標準偏差の2倍の幅をとると95.4%，3倍の幅をとれば99.7%となり，ほとんど100%に近いものがその範囲内に入ります。




### ○正規分布とは

ある事柄の集まりを，数量の大きさによっていくつかのグループに分け，ヒストグラムや度数分布折れ線グラフに表すことは，43ページで勉強しましたね。

このようなグラフを書いてみると，事柄の性質によって，いろいろな形の分布があることがわかります。その中で，左右が対称で，なめらかな美しい山形をした分布を正規分布といいます。

自然界の事柄には，このような分布になるものがたくさんあります。

## 3．比率の分析

ある部分が全体の中でどれだけの割合を占めているかとか，二つの相異った出来事や事柄について，その大きさや変化を比較するとき，統計比例数あるいは比率という数値をよく用います。

普通に用いられる比率は，構成比率・発生比率・対立比率・指標比率などです。

### (1) 構成比率

右の表は，松江市と浜田市の産業別就業者数とその百分率を表したものです。この表での百分率は，それぞれの産業で働いている人数を総数で割って100をかければ得られることはよく知っていますね。これによると，松江市で最も働く人の多い産業は，サービス業で，全体の31.2%を占め，農業は4.8%であるといったことがよくわかります。

**産業別就業者数**

平成７年10月１日

| 産業（大分類） | 松江市 実数 | 松江市 百分率 | 浜田市 実数 | 浜田市 百分率 |
|---|---|---|---|---|
| | 人 | % | 人 | % |
| 総数 | 74,130 | 100.0 | 25,211 | 100.0 |
| 農業 | 3,527 | 4.8 | 1,072 | 4.3 |
| 林業 | 72 | 0.1 | 48 | 0.2 |
| 漁業 | 424 | 0.6 | 730 | 2.9 |
| 鉱業 | 41 | 0.1 | 22 | 0.1 |
| 建設業 | 7,434 | 10.0 | 2,942 | 11.7 |
| 製造業 | 7,136 | 9.6 | 4,220 | 16.7 |
| 電気・ガス・熱供給・水道業 | 698 | 0.9 | 247 | 1.0 |
| 運輸・通信業 | 3,271 | 4.4 | 1,670 | 6.6 |
| 卸売・小売業，飲食店 | 19,804 | 26.7 | 6,125 | 24.3 |
| 金融・保険業 | 3,514 | 4.7 | 6,666 | 2.6 |
| 不動産業 | 579 | 0.8 | 80 | 0.3 |
| サービス業 | 23,152 | 31.2 | 6,378 | 25.3 |
| 公務（他に分類されないもの） | 4,267 | 5.8 | 974 | 3.9 |
| 分類不能の産業 | 211 | 0.3 | 37 | 0.1 |

資料出所：総務庁統計局「平成７年国勢調査報告」

このように，部分の大きさの，全体に対する割合を示したものを，**構成比率**といい，一般には，全体を100とした百分率（パーセント，%）で表されます。

### (2) 発生比率

右の表は，二つの中学校の生徒数と，ある日の欠席者数を示したものです。

欠　席　者　数

| 学 校 名 | 生 徒 数 | 欠席者数 |
|---|---|---|
| A中学校 | 1,553 人 | 21 人 |
| B中学校 | 238 | 4 |

この表から，生徒数100人につき何人の欠席者がいるかを調べるには，どうしたらよいでしょうか。欠席者数を生徒数で割って100をかけるといいですね。

A中学校 $\dfrac{21}{1,553}\times 100 \fallingdotseq 1.35$(人)　　B中学校 $\dfrac{4}{238}\times 100 \fallingdotseq 1.68$(人)

A中学校は生徒数100人につき1.35人の欠席者，B中学校は生徒数100人につき1.68人の欠席者がいることになります。

このようにすると，生徒数の違った学校についても，どちらが欠席者が多いか少ないかの比較ができます。

ある集団を100とみなして，その中で，ある事柄がどのような割合で発生しているかを表す数値を，**発生比率**といいます。

出生率・死亡率などもこれに入ります。

### (3) 対立比率

右の表で，A中学校とB中学校とでは，1人当たりの校庭の面積は，どちらが広いでしょうか。

生徒数と校庭の面積

| 学 校 名 | 生 徒 数 | 校庭の面積 |
|---|---|---|
| A中学校 | 1,553 人 | 9,000 ㎡ |
| B中学校 | 238 | 6,000 |

それを調べるには，校庭の面積を生徒数で割ればわかりますね。

A中学校 $\dfrac{9,000}{1,553} \fallingdotseq 5.80$(㎡)　　B中学校 $\dfrac{6,000}{238} \fallingdotseq 25.2$(㎡)

A中学校は1人当たり5.80㎡，B中学校は1人当たり25.2㎡となり，B中学校の方が1人当たりの面積が広いことがわかります。

このように，ある事柄１単位の大きさに対して，もう一方の事柄の平均の大きさを表したものを，**対立比率**（あるいは**密度**）といいます。人口を面積（㎢）で割って計算した１㎢当たりの人口，すなわち，人口密度は，人口の疎密の程度を比べるのによく用いられています。

### (4) 指標比率

**ある中学校の１年生男子の平均体重の動き**

| 年　　月 | 平均体重(kg) | 指数（４月＝100） |
|---|---|---|
| 平成７年４月 | 36.8 | 100 |
| 5 | 37.9 | 103 |
| 6 | 38.1 | 104 |
| 7 | 38.4 | 105 |
| 8 | … | … |
| 9 | 39.3 | 107 |
| 10 | 39.5 | 107 |
| 11 | 40.3 | 110 |
| 12 | 40.9 | 111 |
| ８年１月 | 41.6 | 113 |
| 2 | 42.3 | 115 |
| 3 | 42.7 | 116 |

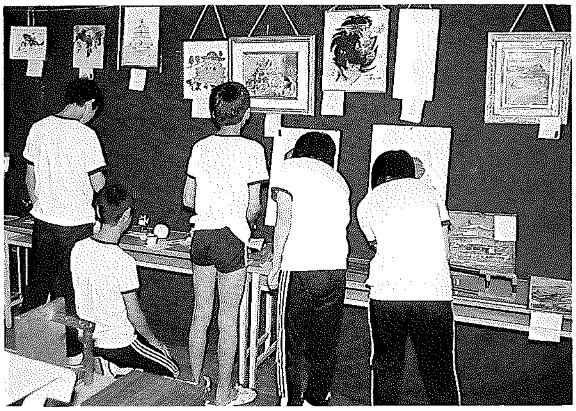
左の表は，ある中学校の１年生男子36人の体重を，毎月初めに測定して平均し，書き並べたものですが，このように，同じ種類の統計を，時間の順序に並べたものを**時系列統計**といっています。

そこで，各月の平均体重を，４月の平均体重36.8kgで割って，これに100をかけたものをそれぞれ記入していくと，４月の平均体重を100として，他の月の体重がいくらになっているかということが表されます。

このような数値を，**指標比率**（あるいは**指数・指標比例数**）といい，各月の平均体重が４月に比べて，どれだけ増えたかということがよくわかります。例えば，平成８年３月は，７年４月に比べて16％増えたということです。

このように，指数は，多くの場合，ある数値を100(これを基準という。)として，時間的な動き，あるいは場所的な比較をわかりやすくするために使われます。物価指数や生産指数などがこの例です。

## 4．標本調査の基礎


力作ぞろいの文化祭

今，仮に文化祭について，学校全体の生徒の意見を早急に調べたいとしたら，みなさんはどうしますか。すべての生徒に意見を聞くと一番確かな結果が得られますが，その時間がなかったら，代表者の意見を調べる方法がありますね。その代表者の意見でもって全体の考え方を推定するのです。しかし，ここで大切なことは，代表者を3年生だけとか，文化委員だけとかの特定の人に限っては，すべての生徒の意見を推定することは不可能です。できるだけ無作為に，その代表者を選定しなければなりませんね。そして，またその代表者も多い方が，より確かな意見と考えられるでしょう。

ある学級の体重の平均を出したり，学校全体で，ある意見に賛成の人の割合を出したりするように，集団が小さいときは全数調査が可能ですが，全国民にアンケートをとったり，すべての乾電池の点灯時間の平均を調べたりすることは，難しいことです。乾電池の場合，すべてを調べたら売るのがなくなってしまいますね。

そこで，資料の一部について調査し，それによって全体を推定する方法があります。その方法を**標本調査**といい，調べようとする全体を**母集団**，その中から取り出した一つ一つの資料を**標本**といいます。母集団からいくつかの標本を取り出し，標本の平均を出し，そして母集団の平均を推定するのです。

標本の取り出し方の代表的なものに，**無作為抽出法**があります。

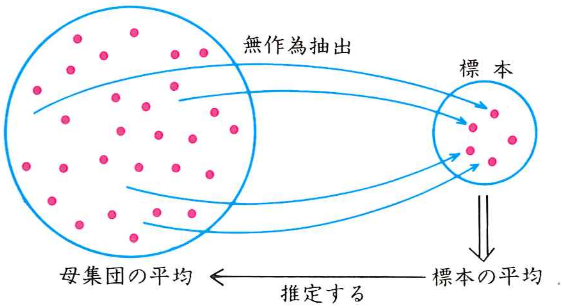
無作為抽出法は，標本を取り出すとき，くじ引きのように，無作為に選ぶ方法です。

我が国では，最近，このような方法で，標本調査が広く行われるようになっています。

例えば，右の表は，松江市に住んでいる世帯の1か月平均の費目別支出額を表したものです。

これは，松江市の約5万4千世帯の中から，96世帯を無作為に取り出し，その世帯に，毎日“家計簿”を記入してもらい，その結果を取りまとめたものです。

また，標本調査にはこれらの無作為抽出法のほかに，母集団から何らかの考え(意図・目的)で標本を選び出す**有意抽出法**もあります。




**松江市の1世帯当たり**
**1か月間の消費支出**
平成7年

| 費 目 | 支出金額 |
|---|---|
| 消費支出 | 308,072 円 |
| 食料 | 78,207 |
| 住居 | 15,359 |
| 光熱・水道 | 21,913 |
| 家具・家事用品 | 10,359 |
| 被服及び履物 | 19,210 |
| 保健医療 | 6,983 |
| 交通・通信 | 30,168 |
| 教育 | 12,487 |
| 教養娯楽 | 31,082 |
| その他の消費支出 | 82,303 |

資料出所：総務庁統計局「家計調査年報」


一部分で全体を知る