

## 相関係数の計算

2つの群のデータを

第1群  $\{x_i\}_{i=1}^n$ 、第2群  $\{y_i\}_{i=1}^n$

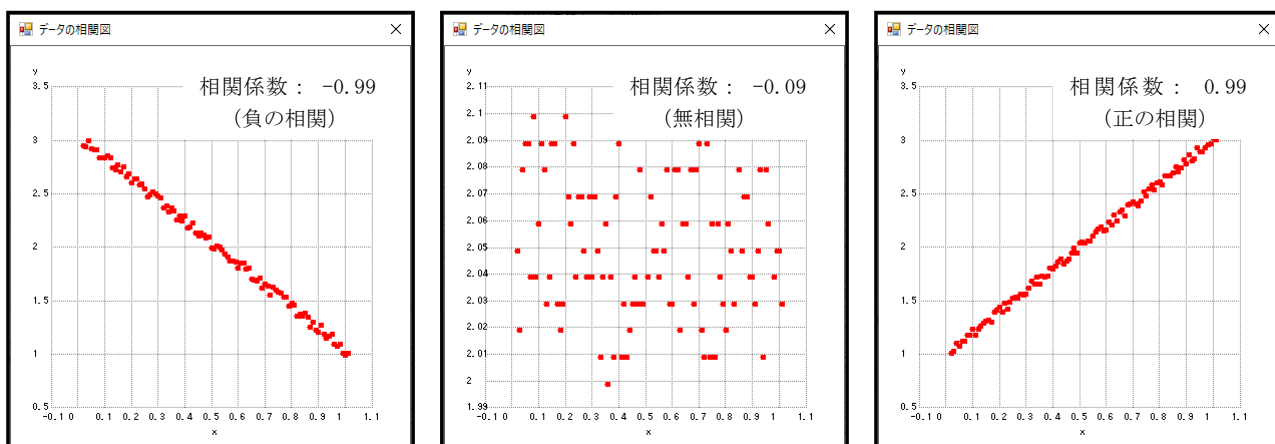
としたとき、相関係数  $r$  は以下で計算される。

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{として}$$
$$r = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}}$$

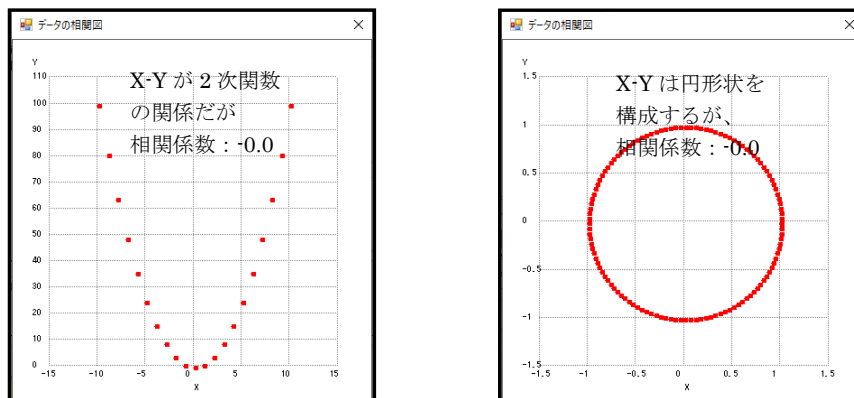
相関係数は $[-1.0 \sim 1.0]$ の数値で2つの間に直線的な関係があるかを示すもので、以下の通りである。

- $-1.0$  は負の相関が最も強いことを示す
- $0.0$  は全く無相関であることを示す
- $1.0$  は正の相関が最も強いことを示す

下図に 上記3パターンにおける例を示す。



相関係数は あくまでも 直線的な（線形な）関係を調べるもので、非線形な相関の表現には向かない。



また 統計量  $t = \frac{|r| \sqrt{n-2}}{\sqrt{1-r^2}}$  は 自由度  $n-2$  の  $t$  分布に従う。

この  $t$  を 用いて母集団の相関係数が 0 かどうかを検定できる。この場合、

帰無仮説 : 母相関係数は 0 つまり無相関である。

対立仮説 : 母相関係数は 0 でない。

$t$  が自由度  $n-2$  の  $t$  分布の棄却域に入るとき ( $p$  値が有意水準以下)、対立仮説を採用する。