が最小になるように直線 $y=ax+b$ ……………①　を求めることにするのです。前図より，ずれ d_i は，

$$d_i = y_i - (ax_i - b)$$

ですから，データの大きさが n であったとしますと，

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \{ y_i - (ax_i + b) \}^2$$

この $\sum_{i=1}^{n} d_i^2$ が最小になるように a，b を定めるのです。（このようにして，a，b の値を求める方法を最小二乗法といいます。）

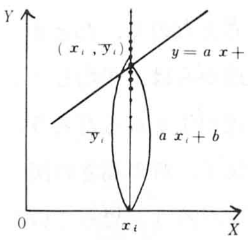
ここでは，結果だけを書いておくことにします。

$$a = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$$

$$b = \frac{1}{n}(\Sigma y - a\Sigma x), \quad (b = \overline{y} - a\overline{x}) \cdots\cdots ②$$

このようにして得られた a，b を係数とする直線 $y=ax+b$ を，**y の x への回帰直線**といい，a を**回帰係数**といいます。
(注)

この y の x への回帰直線は，x の各値に対する y の中心的傾向を示すものです。すなわち，同じ x の値 x_i に対する y の値が，$y_1, y_2, \cdots y_K$ とあったとしますと，これらの点 $(x_i, y_1), (x_i, y_2) \cdots (x_i, y_K)$ は，左図のように縦の直線 $x=x_i$ 上に並んでいるわけです。



このとき，$y_1, y_2, \cdots\cdots y_K$ の平均値を $\overline{y}_i$ としますと，直線 $y=ax+b$ は，点 $(x_i, \overline{y}_i)$ の近くを通るというのです。

なお，②の式を①に代入しますと，次の式が得られます。

$$y - \overline{y} = a(x - \overline{x})$$

これは，y の x への回帰直線は，点 $(\overline{x}, \overline{y})$ を通ることを示しています。

(例27) 次ページの（表12）から，y の x への回帰直線を求めよ。

(解)　p 130より

$\Sigma x = 870$,　$\Sigma y = 810$