Estimated Least-Squares Regression Coefficients

define estimated regression line :

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$\hat{\beta}_0$: estimated intercept

$\hat{\beta}_1$: estimated slope

numerator term

$$\hat{\beta}_1 = \frac{\sum y_i x_i - \frac{(\sum y_i)(\sum x_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \qquad \begin{matrix} \} S_{XY} \\ \} S_{XX} \end{matrix}$$

denominator term

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

where $\bar{y} = \sum y_i / n$

$\bar{x} = \sum x_i / n$

ex: purity of O_2 concentration (y)
vs. hydrocarbon level (x)

TABLE 11.1 Oxygen and Hydrocarbon Levels

Observation Number	Hydrocarbon Level $x(\%)$	Purity $y(\%)$
1	0.99	90.01
2	1.02	89.05
3	1.15	91.43
4	1.29	93.74
5	1.46	96.73
6	1.36	94.45
7	0.87	87.59
8	1.23	91.77
9	1.55	99.42
10	1.40	93.65
11	1.19	93.54
12	1.15	92.52
13	0.98	90.56
14	1.01	89.54
15	1.11	89.85
16	1.20	90.39
17	1.26	93.25
18	1.32	93.41
19	1.43	94.98
20	0.95	87.33

$\sum x_i = 23.92$

$\sum y_i = 1843.21$

$n = 20$

$\bar{x} = \frac{23.92}{20} = \underline{1.1960}$

$\bar{y} = \frac{1843.21}{20} = \underline{92.1605}$

$\sum x_i^2 = 0.99^2 + 1.02^2 + 1.15^2 + \ldots$

$= \underline{29.2892}$

$\sum y_i^2 = 90.01^2 + 89.05^2 + \ldots$

$= \underline{170{,}044.5321}$

$\sum x_i y_i = 0.99 \cdot 90.01 + 1.02 \cdot 89.05 + \ldots$

$= \underline{2214.6566}$

numerator term of slope :

$$S_{xy} = \sum y_i x_i - \frac{\sum y_i \sum x_i}{n}$$

$$= 2214.6566 - \frac{1843.21 \cdot 23.92}{20}$$

$$S_{xy} = 10.18$$

denominator term of slope :

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$= 29.2892 - \frac{23.92^2}{20}$$

$$S_{xx} = 0.68088$$

least-squares estimate of slope:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{10.18}{0.6809} = 14.95$$

unit: $\frac{O_2 \text{ purity}}{\text{hydrocarbon level}}$ [%/%]

least-squares estimate of intercept:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$= 92.1605 - 14.95 \cdot 1.1960$$

$$\hat{\beta}_0 = 74.28$$

unit: O_2 purity [%]

estimated least-squares regression line:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 74.28 + 14.95x$$

(O_2 purity [%])

-- what do we do with it?

for example, @ $x = 1.00\%$ hydrocarbon level:

line predicts

$$\hat{y} = 74.28 + 14.95 \cdot 1.00$$

$$= 89.23\% \ O_2 \text{ purity}$$

cpi calculator: look up what \$400 in 1973 is worth today!