# LEAST SQUARES ESTIMATION

Derivation of the Estimates $b_0$ and $b_1$ for a Straight Line

*Reference: N. R. Draper and H. Smith, Applied Regression Analysis, 3rd Edition, Chapter 1*

**@S. S. Roy | Date: 27 July 2026**

## 1. The Model and the Quantity to Minimise

For a single predictor X and response Y, the straight-line model for n observations is:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \qquad i = 1, 2, \ldots, n$$

$\beta_0$ (intercept) and $\beta_1$ (slope) are the unknown parameters; $\varepsilon_i$ is the random error of the i-th observation. Least squares estimates $\beta_0$ and $\beta_1$ by the values $b_0$ and $b_1$ that make the sum of the squared errors as small as possible. That sum is the sum-of-squares function S:

$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i\right)^2$$

Geometrically, S is the total of the squared vertical distances between the data points and the line. The fitted line is the one that minimises this total.

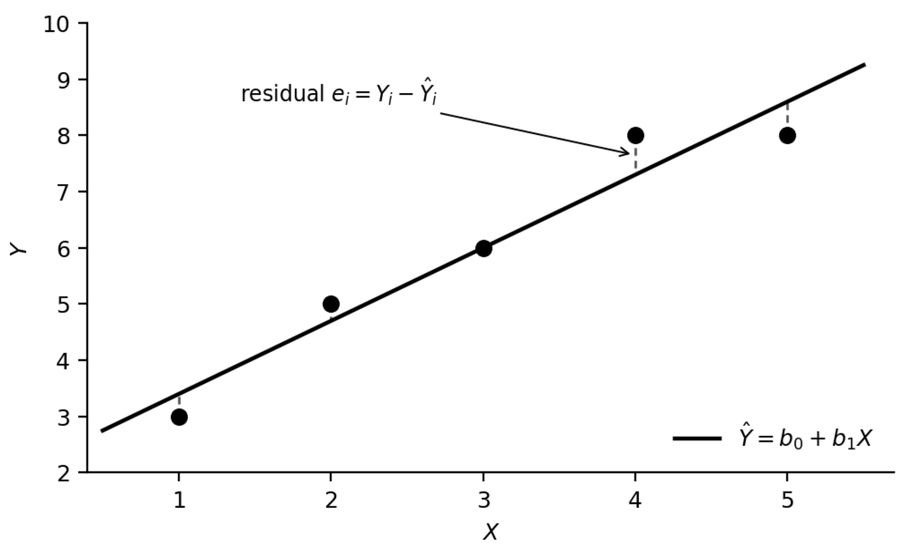


*Figure 1. Each dashed segment is a vertical deviation. Least squares minimises the sum of the squares of these deviations.*

## 2. Derivation of $b_0$ and $b_1$

S is a function of the two unknowns $\beta_0$ and $\beta_1$. A minimum occurs where both partial derivatives are zero.

### Step 1 — differentiate S with respect to each parameter

Differentiating with respect to $\beta_0$ (treating $\beta_1$ as constant), using the chain rule on the squared bracket:

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i\right)$$

Differentiating with respect to $\beta_1$, the inner derivative brings down a factor of $-X_i$:

$$\frac{\partial S}{\partial \beta_1} = -2\sum_{i=1}^{n} X_i\left(Y_i - \beta_0 - \beta_1 X_i\right)$$

### Step 2 — set the derivatives to zero at the solution

At the minimising values $b_0$ and $b_1$, both partial derivatives vanish:

$$\left.\frac{\partial S}{\partial \beta_0}\right|_{b_0,b_1} = 0, \qquad \left.\frac{\partial S}{\partial \beta_1}\right|_{b_0,b_1} = 0$$

Dividing each equation by −2 and substituting $(b_0, b_1)$ for $(\beta_0, \beta_1)$:

$$\sum\left(Y_i - b_0 - b_1 X_i\right) = 0, \qquad \sum X_i\left(Y_i - b_0 - b_1 X_i\right) = 0$$

### Step 3 — expand into the normal equations

Expanding the sums (and noting $\Sigma b_0 = n \cdot b_0$) gives the two **normal equations**:

$$n\, b_0 + b_1 \sum X_i = \sum Y_i$$
$$b_0 \sum X_i + b_1 \sum X_i^2 = \sum X_i Y_i$$

These are two linear equations in the two unknowns $b_0$ and $b_1$.

### Step 4 — solve for $b_0$ from the first equation

Dividing the first normal equation by n gives the intercept directly in terms of the slope and the two means ($\bar{X} = \Sigma X_i/n$, $\bar{Y} = \Sigma Y_i/n$):

$$b_0 = \frac{\sum Y_i - b_1 \sum X_i}{n} = \bar{Y} - b_1 \bar{X}$$

### Step 5 — substitute into the second equation and solve for $b_1$

Replacing $b_0$ in the second normal equation:

$$\left(\bar{Y} - b_1 \bar{X}\right)\sum X_i + b_1 \sum X_i^2 = \sum X_i Y_i$$

Collecting the $b_1$ terms on the left and simplifying $\Sigma X_i = n\bar{X}$, $\Sigma Y_i = n\bar{Y}$:

$$b_1\left(\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}\right) = \sum X_i Y_i - \frac{\left(\sum X_i\right)\left(\sum Y_i\right)}{n}$$

The bracket on the left is S_XX and the right side is S_XY. Dividing gives the slope:

$$\boxed{b_1 = \frac{\sum X_i Y_i - \left(\sum X_i\right)\left(\sum Y_i\right)/n}{\sum X_i^2 - \left(\sum X_i\right)^2/n} = \frac{S_{XY}}{S_{XX}}}$$

and back-substitution gives the intercept:

$$\boxed{b_0 = \bar{Y} - b_1 \bar{X}}$$

### Step 6 — confirm it is a minimum

The second derivatives are positive, so the stationary point is indeed a minimum, not a maximum or saddle point:

$$\frac{\partial^2 S}{\partial \beta_1^2} = 2\sum X_i^2 > 0, \qquad \frac{\partial^2 S}{\partial \beta_0^2} = 2n > 0$$

**Result.** The least-squares estimates are $b_1 = S\_XY / S\_XX$ and $b_0 = \bar{Y} - b_1\bar{X}$, where $S\_XY = \Sigma X_iY_i - (\Sigma X_i)(\Sigma Y_i)/n$ and $S\_XX = \Sigma X_i^2 - (\Sigma X_i)^2/n$. Because $b_0 = \bar{Y} - b_1\bar{X}$, the fitted line always passes through the point $(\bar{X}, \bar{Y})$.

## 3. Short Numerical Example

**Problem.** Fit a least-squares straight line $\hat{Y} = b_0 + b_1X$ to the five points below.

| X | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Y | 3 | 5 | 6 | 8 | 8 |

**Step 1 — the sums.**

$$n = 5, \quad \sum X = 15, \quad \sum Y = 30, \quad \sum XY = 103, \quad \sum X^2 = 55$$

**Step 2 — the means.**

$$\bar{X} = \frac{15}{5} = 3, \qquad \bar{Y} = \frac{30}{5} = 6$$

**Step 3 — corrected sums.**

$$S_{XX} = 55 - \frac{15^2}{5} = 10, \qquad S_{XY} = 103 - \frac{(15)(30)}{5} = 13$$

**Step 4 — slope.**

$$b_1 = \frac{S_{XY}}{S_{XX}} = \frac{13}{10} = 1.3$$

**Step 5 — intercept.**

$$b_0 = \bar{Y} - b_1\bar{X} = 6 - 1.3(3) = 2.1$$

**Fitted line:**

$$\hat{Y} = 2.1 + 1.3\,X$$

Check: at $X = \bar{X} = 3$, $\hat{Y} = 2.1 + 1.3(3) = 6.0 = \bar{Y}$, confirming the line passes through $(\bar{X}, \bar{Y})$. The positive slope 1.3 means Y increases by about 1.3 units for each unit increase in X.