Linear Regression Derivation

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad \hat{y}_i = ax_i + b$$

$$= \sum \left[y_i - (ax_i + b)\right]^2$$

Minimize S.S.E. w.r.t unknowns a & b.

$$\therefore \frac{\partial(SSE)}{\partial a} = 0 \quad \text{and} \quad \frac{\partial(SSE)}{\partial b} = 0$$

$$\frac{\partial}{\partial a}(SSE) = 0 = \sum 2\left[y_i - (ax_i + b)\right]\cdot(-x_i)$$

$$\therefore \sum\left(ax_i^2 + bx_i - x_i y_i\right) = 0$$

$$\therefore \sum\left(ax_i^2 + bx_i\right) = \sum x_i y_i$$

$$a\sum x_i^2 + b\sum x_i = \sum x_i y_i \qquad \text{— } (1)$$

$$\frac{\partial(SSE)}{\partial b} = 0 = \sum 2\left(y_i - (ax_i + b)\right)\cdot(-1)$$

$$\therefore \sum(-y_i + ax_i + b) = 0$$

$$\therefore \sum ax_i + \sum b = \sum y_i$$

$$\sum \frac{ax_i}{N} + \frac{b\cdot N}{N} = \frac{\sum y_i}{N} \qquad \text{....} \ (2)$$

$$\therefore a\bar{x} + b = \bar{y}$$

$$\therefore \boxed{b = \bar{y} - a\bar{x}}$$

From (1) $$a\frac{\sum x_i^2}{N} + b\frac{\sum x_i}{N} = \frac{\sum x_i y_i}{N}.$$

$$a\overline{x_i^2} + b\bar{x} = \overline{xy}$$

$$a\overline{x^2} + \bar{x}(\bar{y} - a\bar{x}) = \overline{xy}$$

$$a\overline{x^2} - a\bar{x}^2 + \bar{x}\bar{y} = \overline{xy}$$

$$a(\overline{x^2} - \bar{x}^2) = \overline{xy} - \bar{x}\bar{y}$$

$$\therefore \quad a = \frac{\overline{xy} - \bar{x}\bar{y}}{(\overline{x^2} - \bar{x}^2)} \quad \text{~~~} \quad (A).$$

$$b = \bar{y} - a\bar{x}$$

$$= \bar{y} - \bar{x}\left[\frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}\right]$$

$$= \frac{\bar{y}(\overline{x^2} - \bar{x}^2) - \bar{x}(\overline{xy} - \bar{x}\bar{y})}{(\overline{x^2} - \bar{x}^2)}$$

$$= \frac{\bar{y}\overline{x^2} - \cancel{\bar{y}\bar{x}^2} - \bar{x}\,\overline{xy} + \cancel{\bar{x}^2\bar{y}}}{(\overline{x^2} - \bar{x}^2)}$$

$$b = \frac{\bar{y}\overline{x^2} - \bar{x}\cdot\overline{xy}}{\overline{x^2} - \bar{x}^2} \quad \text{~~~} \quad (B).$$