CS 6140 Assignment 2

④ $y = f_w(x) = w_0 + w_1 x = mx + b$ where $w_0 = b$, $w_1 = m$

$\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$ ← dataset

find w_0 and w_1

$w_0 = b =$ y-intercept
$w_1 = m =$ slope

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

$$\boxed{\begin{aligned} w_0 &= \overline{y} - w_1 \overline{x} \\ w_1 &= \frac{\overline{x}\,\overline{y} - \overline{xy}}{(\overline{x})^2 - \overline{(x^2)}} \end{aligned}}$$

Solution for Min SSE

$$SSE = \sum_{i=1}^{n} y_i^2 - 2y_i m x_i - 2y_i b + m^2 x_i^2 + 2m x_i b + b^2$$

$$SSE = \sum_{i=1}^{n} y_i^2 - 2m \sum_{i=1}^{n} x_i y_i - 2b \sum_{i=1}^{n} y_i + m^2 \sum_{i=1}^{n} x_i^2 + 2mb \sum_{i=1}^{n} x_i + nb^2$$

$$\frac{\sum_{i=1}^{n} y_i^2}{n} = \overline{y^2}$$ ← mean of y^2 values.

$$\sum_{i=1}^{n} y_i^2 = n\overline{(y^2)}$$

$$\to SSE = n\overline{(y^2)} - 2mn\overline{(xy)} - 2bn\overline{y} + m^2 n\overline{(x^2)} + 2mbn\overline{x} + nb^2$$

Need to minimize SSE for m and b variables: (using partial derivatives)

$$\frac{\partial SSE}{\partial m} = 0 = -\cancel{2}\cancel{n}\overline{(xy)} + \cancel{2}m\cancel{n}\overline{(x^2)} + \cancel{2}b\cancel{n}\overline{x} = -\overline{(xy)} + m\overline{x^2} + b\overline{x}$$

$$\frac{\partial SSE}{\partial b} = 0 = -\cancel{2}\cancel{n}\overline{y} + \cancel{2}m\cancel{n}\overline{x} + \cancel{2}\cancel{n}b = -\overline{y} + m\overline{x} + b$$

$$\boxed{w_0 = b = \overline{y} - m\overline{x}} \star$$

$$0 = m\overline{(x^2)} - \overline{(xy)} + b\overline{x}$$

$$0 = m\overline{(x^2)} - \overline{(xy)} + (\overline{y} - m\overline{x})\overline{x}$$

$$0 = m\overline{(x^2)} - \overline{(xy)} + \overline{x}\,\overline{y} - m(\overline{x})^2 \Rightarrow m(\overline{x})^2 - m\overline{(x^2)} = \overline{x}\,\overline{y} - \overline{(xy)}$$

$$\boxed{w_1 = m = \frac{\overline{x}\,\overline{y} - \overline{xy}}{(\overline{x})^2 - \overline{x^2}}} \star$$

(8) N samples: $\{x_1, x_2, \ldots, x_N\}$

Solution continues onto next page, but best $\mu = \bar{x}$

☆5

$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

μ → mean, σ → std. deviation

Likelyhood → $L(\mu, \sigma | x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$

For all data:

Maximize

$$L(\mu, \sigma | x_1, x_2 \ldots x_n) = \prod_{i=0}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x_i-\mu)^2/2\sigma^2}$$

$$\frac{\partial L}{\partial \mu} = 0 \qquad \frac{\partial L}{\partial \sigma} = 0$$

Solve both for maximum estimates for μ & σ.

$$\ln(L()) = \ln\left(\prod_{i=0}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{(x_i-\mu)^2/2\sigma^2}\right)$$

Log likelyhood

$$= \sum_{i=0}^{n} \ln\left(\frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x_i-\mu)^2/2\sigma^2}\right)$$

$$= \sum_{i=0}^{n} \left(\ln\left((2\pi\sigma^2)^{-1/2}\right) + \ln\left(e^{-(x_i-\mu)^2/2\sigma^2}\right)\right)$$

$$= \sum_{i=0}^{n} \left(-\frac{1}{2}\ln(2\pi) - \ln(\sigma) - \frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$= -\frac{n}{2}\ln(2\pi) - n\ln(\sigma) - \sum_{i=0}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}$$

chain rule: $\left(\frac{(x_i-\mu)^2}{2\sigma^2}\right)' = \frac{2(x_i-\mu)}{2\sigma^2}$

$$\frac{\partial L}{\partial \mu} \ln(L()) = \sum_{i=0}^{n} \frac{x_i-\mu}{\sigma^2} = \frac{1}{\sigma^2}(n\bar{x} - n\mu) = \frac{n}{\sigma^2}(\bar{x} - \mu)$$

$\bar{x}$ ← average (x) value

$$\frac{\partial L}{\partial \sigma} \ln(L()) = -\frac{n}{\sigma} + \sum_{i=0}^{n} \frac{(x_i-\mu)^2}{\sigma^3} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2$$

Continued on next page →

(8) cont...

$$\frac{\partial L}{\partial \mu} \ln\left(L(\mu, \sigma \mid x_1, x_2 \dots x_n)\right) = \frac{n}{\sigma^2}(\bar{x} - \mu)$$

$$\frac{\partial L}{\partial \sigma} \ln\left(L(\mu, \sigma \mid x_1, x_2 \dots x_n)\right) = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\rightarrow \frac{n}{\sigma^2}(\bar{x} - \mu) = 0$$

$$\star \; \boxed{\mu = \bar{x}}$$ Mean of measurements.

$$\rightarrow -\frac{n}{\cancel{\sigma}} + \frac{1}{\sigma^{\cancel{3}2}} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$n = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\sigma^2 n = \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}}, \quad \mu = \bar{x}$$

Just for fun

$$\boxed{\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}}$$ Std deviation of measurements

∴ Mean of the data $(\bar{x})$ is where center of normal distribution 'μ' should go.

Std. Deviation should determine σ, or how wide the normal distribution should be.