MLE

Probability density function for a sample of n independent normal variable is given by equation

$$\gamma(\mu,\sigma^2) = \sum_{i=1}^{n} f(x_i|\mu,\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right)$$

Taking log on both sides with the assumption that $\log(x_1) > \log(x_2)$ when $x_1 > x_2$. So, we can shift our objective to maximise the $\log(\gamma(\mu,\sigma^2))$, i.e.

$$\log(\gamma(\mu,\sigma^2)) = \frac{-n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \Rightarrow \boxed{EQ\,2}$$

Negative log likelihood

We can ~~[illegible]~~ minimize the negative log likelihood, keeping same objective i.e.

$$NLL \to -\log(\gamma(\mu,\sigma^2)) = \frac{n}{2}\log(2\pi\sigma^2) + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \Rightarrow \boxed{EQ3}$$

To find the best parameter where NLL is minimum we ~~we~~ will need gradients. Thus, let us take partial first derivates

$$\frac{\partial}{\partial\mu}\left\{-\log(\gamma(\mu,\sigma^2))\right\} = 0 + \frac{-2}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)$$

$$\frac{\partial}{\partial\mu}\left\{-\log(\gamma(\mu,\sigma^2))\right\} = -\frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} \Rightarrow \boxed{EQ\,4}$$

Q. When will this first derivate be equal to 0?

A. when $\mu = \bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \Rightarrow \boxed{EQ\ 5}$

$$\frac{\partial}{\partial \sigma}\left\{-\log(\gamma(\mu,\sigma^2))\right) = \frac{+n}{\sigma} - \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \mu) \Rightarrow \boxed{EQ\ 6}$$

Q. When will this first derivate be equal to 0?

A. when $\sigma = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \Rightarrow \boxed{EQ\ 7}$

Simple Linear Regression

Let us assume data is in this form:

$x: [x_1, x_2, \ldots\ldots x_n]$ and $Y: [y_1, y_2, \ldots\ldots, y_n]$

The form of simple linear regression is given by linear equation $y = f(x) = b_0 + b_1 x$

where b_0 is intercept of equation line (i.e. when input is not available $\{x=0\}$ then what is the value of y)

b_1 is slope of the equation line which denotes the corrilation of y with x (i.e. if x is increased by a unit how will y respond to this change in x)

This line will take the best fit line when b_0 and b_1 are chosen in such a way that error or residual sum of square be minimum for the data set ~~a~~ x and y

$$RSS = \sum_{i=1}^{n} e_i^2 \quad \text{where } e_i = y_i - \hat{y}_i = y_i - b_0 - b_1 x_i$$

$$RSS = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

here for historical data where x and y are known x_i and y_i are seen as constants

b_0 and b_1 are variable which is ~~our~~ our task now to find

Thus RSS become $f(b_0, b_1)$

our optimization objective becomes to ~~maximum~~ minimize RSS

$$\min_{b_0, b_1} RSS \Rightarrow \min f(b_0, b_1) = \min_{b_0, b_1} \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) \Rightarrow \underline{EQ\ 1}$$

To solve for best b_0 and b_1, we need to get stationary point i.e. when both partial derivates are zero

$$\frac{\partial f}{\partial b_0} = 0 \quad \text{and} \quad \frac{\partial f}{\partial b_1} = 0$$

Let us first solve for partial derivate w.r.t b_0

$$\frac{\partial}{\partial b_0} = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_1)^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} \frac{\partial}{\partial b_0} (y_i - b_0 - b_1 x_1)^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} 2(y_i - b_0 - b_1 x_i)(-1) = 0$$

$$\Rightarrow n\bar{y} - nb_0 - b_1 n\bar{x} = 0$$

$$\Rightarrow \bar{y} = b_0 + b_1\bar{x}$$

$$\Rightarrow b_0 = \bar{y} - b_1\bar{x}$$

derivates are solved using rules

$$\frac{\partial Q^2}{\partial P} = 2Q \frac{dQ}{dP}$$

$$\sum_{i=1}^{n} y_i = n\bar{y}$$

$$\sum_{i=1}^{n} x_i = n\bar{x}$$

Now partial derivative w.r.t b_1

$$\frac{\partial}{\partial b_1} \sum_{i=1}^{n} (y_i - b_0 - b_1 x_1)^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} 2(y_i - b_0 - b_1 x_i)(-x_1) = 0$$

$$\Rightarrow \sum x_i y_i - b_0 n\bar{x} - b_1 \sum x_i^2 = 0$$

replacing b_0 using (a)

$$\sum x_i y_i - (\bar{y} - b_1\bar{x}) n\bar{x} - b_1 \sum x_i^2 = 0$$

$$\sum x_i y_i - \bar{y} n\bar{x} + b_1 n\bar{x} - b_1 \sum x_i^2 = 0$$

$$b_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \Rightarrow \underline{\underline{EQ\ 2}}$$

$$b_0 = \bar{y} - b_1\bar{x} \Rightarrow \underline{\underline{EQ\ 3}}$$