Formulas used:

1. Gaussian (normal) densite

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

2. Independence of i.i.d samples:

$$P(X_{1:n}|\theta) = \prod_{i=1}^{n} P(x_i|\theta)$$

e.g. $\Rightarrow P(2$ heads in a row$) = (0.5)(0.5)$

$= 0.25$

3. Product Log Rule:

$$\log\left(\prod_{i=1}^{n} a_i\right) = \sum_{i=1}^{n} \log a_i.$$

$$X_i \sim \mathcal{N}(\mu, \sigma^2)$$

where $\mu \in \mathbb{R}$ is the mean and $\sigma^2 > 0$ is the variance.

(i) Write down the likelihood function $P(x_{1:n} | \mu, \sigma^2)$

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

$$P(x_{1:n} | \mu, \sigma^2) = \prod_{i=1}^{n} P(x_i | \mu, \sigma^2)$$

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

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

$$\boxed{P(x_{1:n} \mid \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^{n} e^{\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}}$$

(iii) Derive $\log P(x_{1:n} \mid \mu, \sigma^2)$

$$= \log (2\pi\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^{n} e^{\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}$$

$$= \log (2\pi\sigma^2)^{-\frac{n}{2}} + \log \prod_{i=1}^{n} e^{\left(\frac{-(x_i-\mu)^2}{2\sigma^2}\right)}$$

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

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

$$\boxed{\log P(x_{1:n} \mid \mu, \sigma^2) = -\frac{n}{2} \log (2\pi\sigma^2) - \left(\frac{1}{2\sigma^2}\right) \sum_{i=1}^{n} (x_i-\mu)^2}$$

(iii) Which terms of $\log P(X_{1:n} | \mu, \sigma^2)$ depend on μ (the mean of the data)

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

Why do we care?

We might want do differentiation w.r.t μ!