3. Assume X has N points $x_1, x_2, \ldots, x_n$ that are i.i.d from a Gaussian distribution. Then, $\mathcal{E}[X] = \mu$ & $\mathcal{E}[(x-\mu)^2] = \sigma^2$.

The sample mean $\mu_{ML} = \frac{1}{N}\sum_{n=1}^{N} x_n$ and sample variance $\sigma^2_{ML} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \mu_{ML})^2$ derived from the maximum likelihood estimate. However, $\mathcal{E}[\sigma^2_{ML}] \neq \sigma^2$

$$\mathcal{E}[\sigma^2_{ML}] = \mathcal{E}\left[\frac{1}{N}\sum_{n=1}^{N}(x_n - \mu_{ML})^2\right] = \frac{1}{N}\mathcal{E}\left[\sum_{n=1}^{N}(x_n^2 - 2x_n\mu_{ML} + \mu_{ML}^2)\right]$$

$$= \frac{1}{N}\mathcal{E}\left[\sum_{n=1}^{N} x_n^2 - \sum_{n=1}^{N} 2x_n\mu_{ML} + \sum_{n=1}^{N}\mu_{ML}^2\right]$$

Since $\sum_{n=1}^{N} x_n = N\mu_{ML}$ and $\sum_{n=1}^{N}\mu_{ML}^2 = N\mu_{ML}^2$,

$$\mathcal{E}[\sigma^2_{ML}] = \frac{1}{N}\mathcal{E}\left[\sum_{n=1}^{N} x_n^2 - 2N\mu_{ML}^2 + N\mu_{ML}^2\right] = \frac{1}{N}\mathcal{E}\left[\sum_{n=1}^{N} x_n^2 - N\mu_{ML}^2\right]$$

$$= \frac{1}{N}\mathcal{E}\left[\sum_{n=1}^{N} x_n^2\right] - \mathcal{E}[\mu_{ML}^2]$$

$$= \underline{\underline{\mathcal{E}[x_n^2] - \mathcal{E}[\mu_{ML}^2]}}$$

When expanded, $\sigma^2 = \mathcal{E}[x_n^2] - (\mathcal{E}[x_n])^2$ and $\sigma_{ML}^2 = \mathcal{E}[\mu_{ML}^2] - (\mathcal{E}[\mu_{ML}])^2$

where $\mathcal{E}[x_n] = \mathcal{E}[\mu_{ML}] = \mu$

Therefore, $\mathcal{E}[\sigma^2_{ML}] = (\sigma^2 + \mu^2) - (\sigma^2_{ML} + \mu^2)$ with rearrangement

$$\mathcal{E}[\sigma^2_{ML}] = \sigma^2 - \sigma^2_{ML} = \sigma^2 - \mathrm{Var}(\mu_{ML})$$

$$= \sigma^2 - \mathrm{Var}\left(\frac{1}{N}\sum_{n=1}^{N} x_n\right) = \sigma^2 - \frac{1}{N^2}\mathrm{Var}\left(\sum_{n=1}^{N} x_n\right)$$

$$= \sigma^2 - \frac{1}{N^2}N\sigma^2 = \sigma^2 - \frac{1}{N}\sigma^2 = \frac{N-1}{N}\sigma^2 = \mathcal{E}[\sigma^2_{ML}]$$

As shown, $\mathcal{E}[\sigma^2_{ML}] \neq \sigma^2$ and actually underestimates the true variance by a factor of $\frac{N-1}{N}$. Therefore the MLE of the variance of a Gaussian is biased