**CONCEPTUAL QUESTIONS:**

* What type of network works best with simple classification tasks? What type of network suits the complex classification tasks with images? Suppose we have data in sequence, in other words, each record is dependent on the previous one, which network can handle this problem?
* How does the LSTM work? There are two types of activation functions used in it, state them.

Given the normal distribution PDF as show below, determine the maximum likelihood for **Mean**.

A black and white math equation

Description automatically generated

Hints:

1. Apply ln function to L (Likelihood function)
2. Use the properties of the logs to make the calculations much easier!
3. The final form of your likelihood function should be like the following.

A math equation with black text

Description automatically generated

1. A red line with black dots

   Description automatically generatedApply derivative with respect to the parameter then find its maximum.