**What is a likelihood function? Also add a formula and explain what it means.**

Likelihoods are the Y-axis values for any fixed data points where the distribution can be moved.

For example, we can consider a simple distribution of the weight of mice (plural of mouse). At first, we take a mouse and weigh it, which is 34 gram for example.

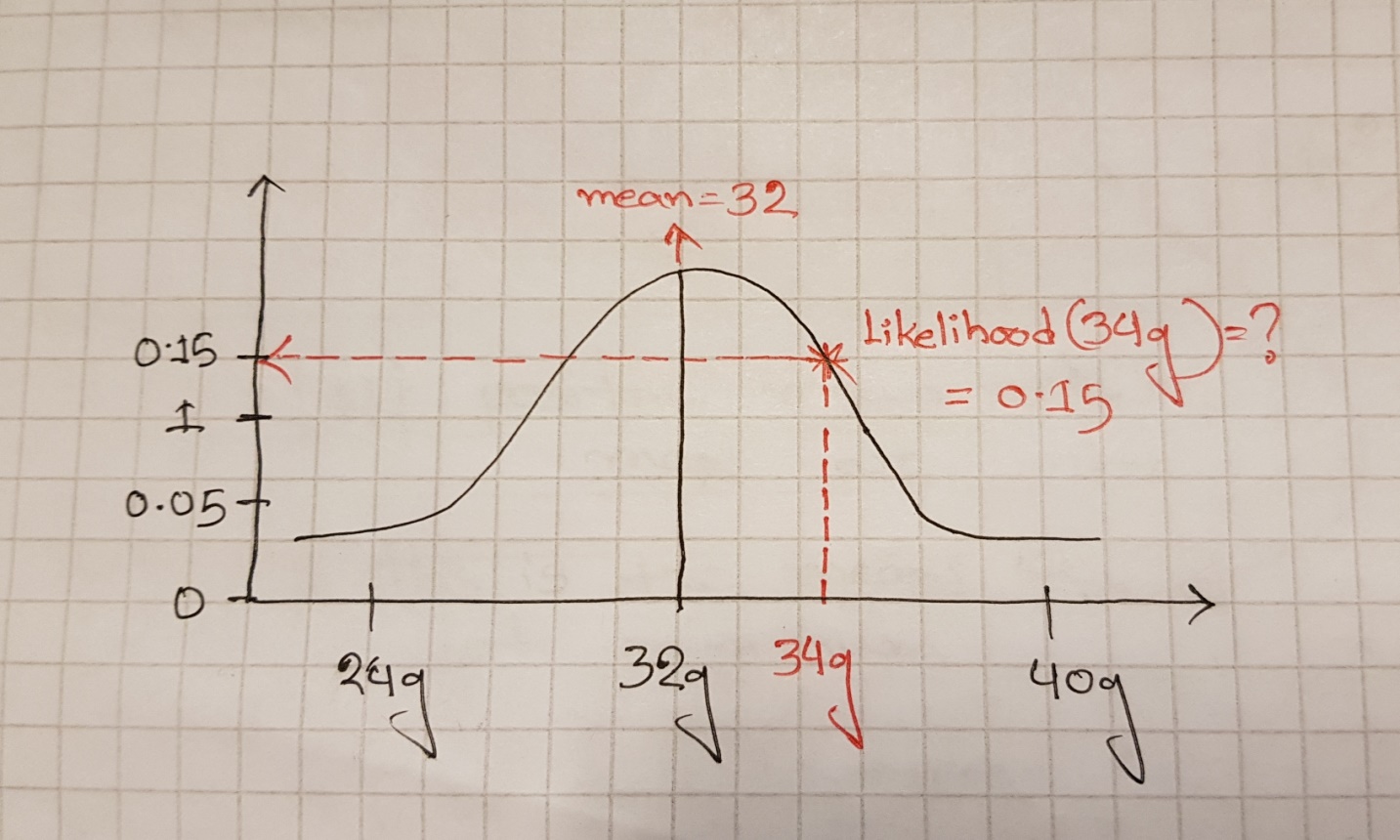
What is the likelihood of a mouse being 34 gram?

To calculate this, we have to locate the exact point of 34 gram and the Y-axis value of that point will be the likelihood. Here, let’s consider mean is 34 gram and Standard Deviation = 2.5.

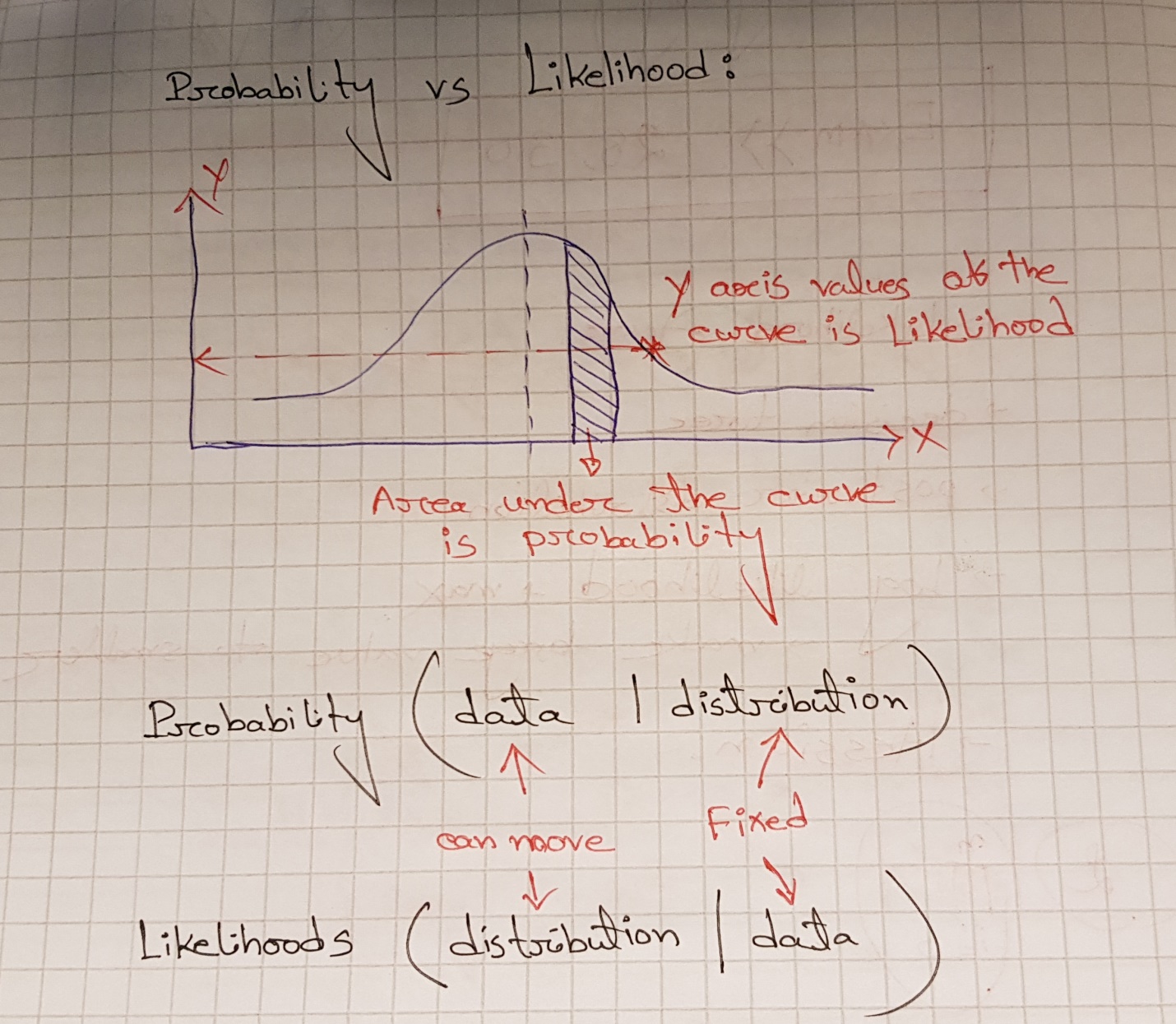
We express it,

L (mean = 32 and Standard Deviation = 2.5 | mouse weighs 34 gram)

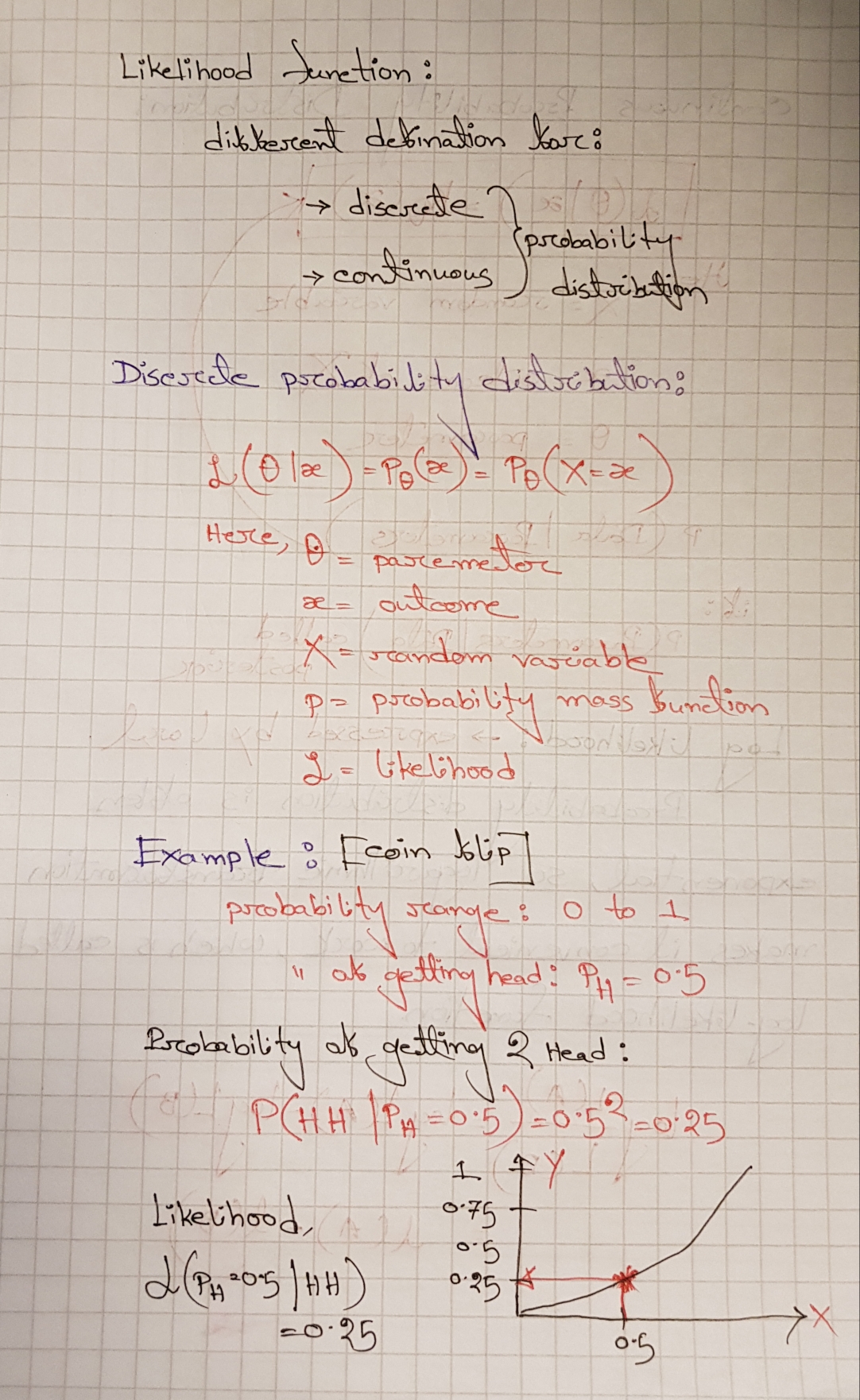
= Y-axis value of the curve = 0.15

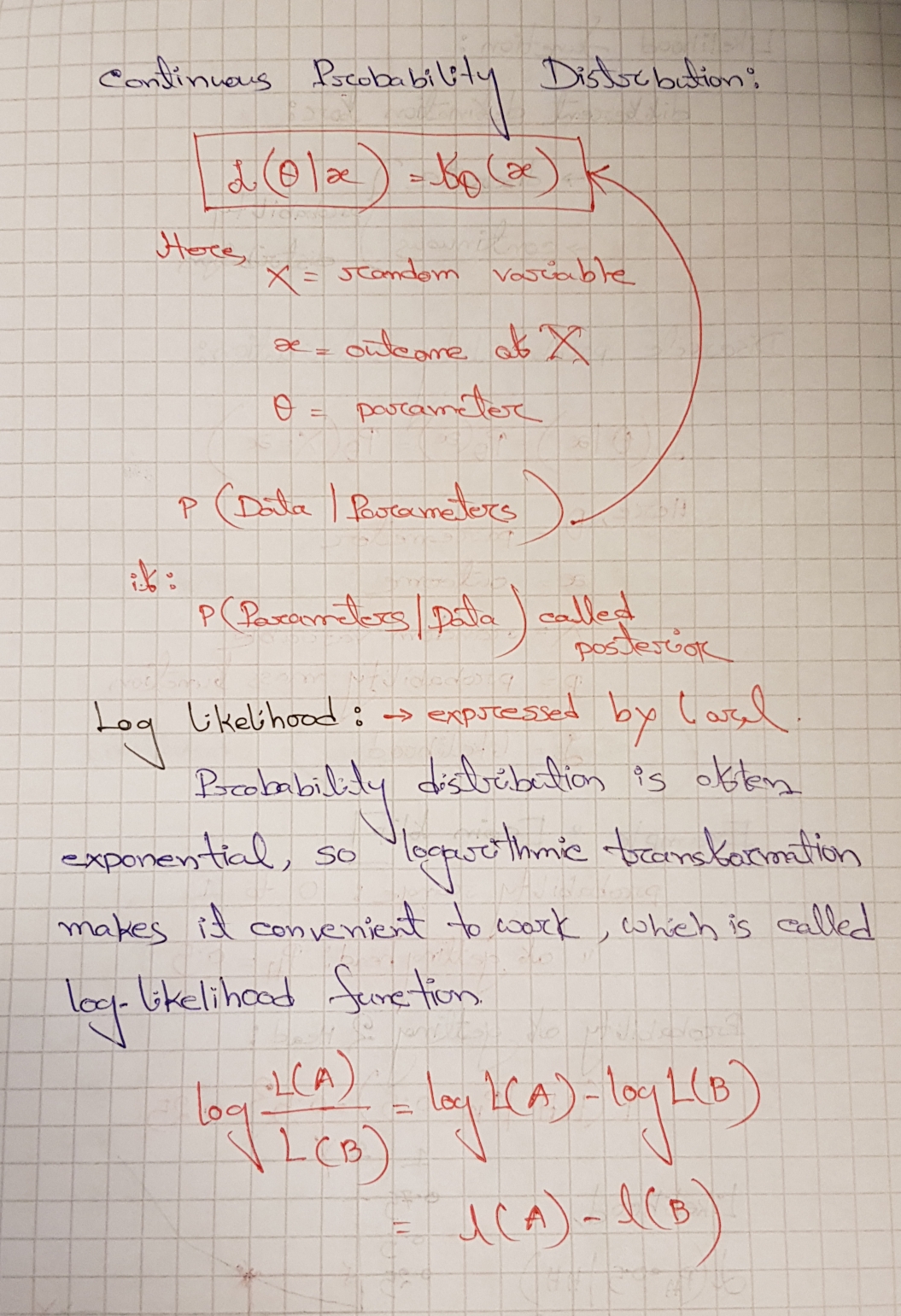


Probability vs Likelihood:



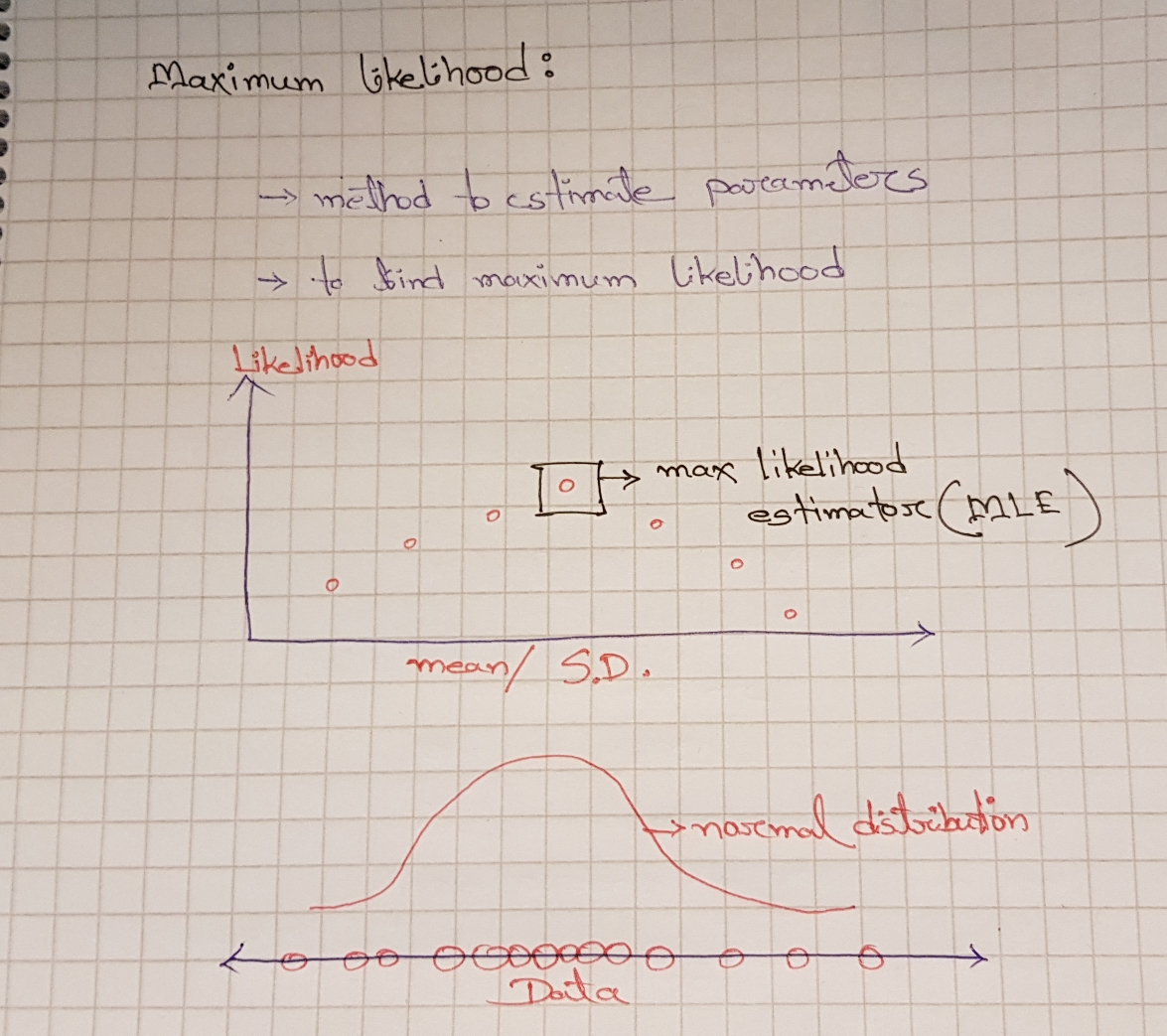
Likelihood Function:





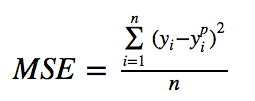
**Maximum likelihood estimation (MLE):**

In statistics, **maximum likelihood estimation** (**MLE**) is a method of [estimating](https://en.wikipedia.org/wiki/Estimator) the [parameters](https://en.wikipedia.org/wiki/Statistical_parameter) of a [statistical model](https://en.wikipedia.org/wiki/Statistical_model) given observations, by finding the parameter values that maximize the [likelihood](https://en.wikipedia.org/wiki/Likelihood) of making the observations given the parameters. We can find out the maximum likelihood of the mean or the standard deviation.

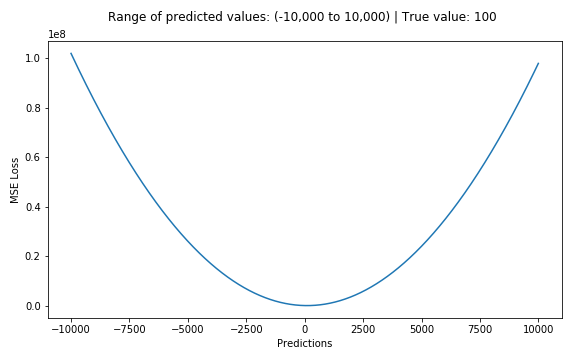


Write out MSE loss for linear regression. Could we also use this loss for classification?

[Mean Square Error (MSE)](https://medium.freecodecamp.org/machine-learning-mean-squared-error-regression-line-c7dde9a26b93) is the most commonly used regression loss function. MSE is the sum of squared distances between our target variable and predicted values.

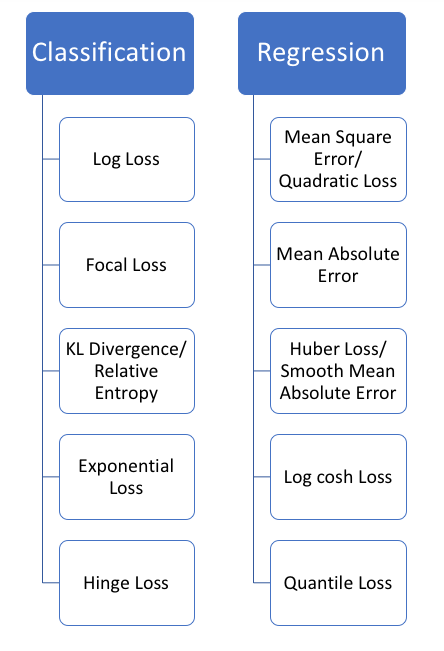


Below is a plot of an MSE function where the true target value is 100, and the predicted values range between -10,000 to 10,000. The MSE loss (Y-axis) reaches its minimum value at prediction (X-axis) = 100. The range is 0 to ∞.



Plot of MSE Loss (Y-axis) vs. Predictions (X-axis)

We cannot use [Mean Square Error (MSE)](https://medium.freecodecamp.org/machine-learning-mean-squared-error-regression-line-c7dde9a26b93) for classification. The following figure (source: [https://heartbeat.fritz.ai](https://heartbeat.fritz.ai/)) shows which loss functions should be used in case of Regression and classification.



Write out the Maximum likelihood Estimation (MLE) for linear regression. How is this related to the MSE loss for linear regression derived in the last point? Derive the relation between them.

MLE method can be used to find the best model parameters of a linear regression model. But when calculating parameters values for those statistical distribution models, we knew what kind of distributions was it and the relevant PDF function.