Polished Writing for Week 4

I am very glad to explain the concept of Maximum Likelihood Estimators to you based on my understanding. I hope this could help you recall this part of knowledge.

Suppose we are given a set of data, which is a set of observations sampled from the population, and we design a model for the data with some parameter beta. Now we hope to estimate the parameter beta based on the given observations. Firstly, consider the meaning of likelihood, it is indeed a measure of how likely to observe the given data based on the value of the parameter. Mathematically, it is a function of the model parameter and we usually express it as a cumulative product of the probability density over all the observations. And the Maximum Likelihood Estimator for the model parameter is the one that maximizes the likelihood function.

For example, consider the simple linear regression model as you know, it has one parameter beta_0 for the intercept and one parameter beta_1 for the slope. Given a set of inputs and outputs, the maximum likelihood estimators for the two parameters are the ones that are most likely to produce the observed data. Equivalently, the MLEs of beta_0 and beta_1 maximize the likelihood function of the outputs corresponding to the inputs under the linear model. In this case, the likelihood depends on how large each error term is, under the assumption that the error terms follow a normal distribution identically and independently.

In general, MLE is commonly used for statistical inference. It aims to find the optimal estimate of the parameter that is most likely to produce the observations. Does my explanation help you recall this concept? If there is anything unclear, please feel free to ask.