**SysBio Questionnaire for AS#4**

Name: Yunyoung Choi ID: 22100748

**NOTE: Answers for questions shown below should be written in English only. Points may be deducted if the sentence is incomplete, grammar is seriously wrong, or it is difficult to understand due to incorrect use of words and terminology.**

1. The difference between the measured data and the model result is called loss. Loss function is a function that calculates the loss. Loss function is used in parameter estimation to find the model parameters that minimizes loss. A popular loss function is the mean squared error (MSE). Describe the advantages (at least two) of MSE as a loss function.

| The advantage of MSE, firstly, is that the sum of the errors is unlikely to be offset because it is calculated by square first. If the error is the opposite sign, such as 3 and -3, the simple sum becomes 0, but if it is obtained by MSE, it becomes 18, and the error is clearly displayed. Second, as the error increases, the square value increases accordingly, so we can intuitively see how much the error increases. Also, It is easy to observe that the number decreases when the error value is minimized. |
| --- |

1. In this assignment, we used a simple algorithm that randomly assigns parameter values N times and finds the best set among the N parameter sets. Describe the disadvantages (at least two) of this algorithm.

| When finding values one by one at random, first of all, there is a disadvantage that it takes a lot of time. It will take a long time because the previous step does not affect the next step at all, because it is simply random and not logical reasoning. Second, there seems to be a disadvantage that the accuracy decreases when N is small. In fact, I simulated when N=10 in AS#4, but the value did not come out accurately as expected. I think the accuracy will vary depending on the range of parameters, and if the parameter is in a larger range than this assignment, I think it will be almost impossible to predict when N=10. |
| --- |

1. Let's assume that the number of unknown parameters of the model used in the previous assignment is 5 (a, b, c, d, n). We already know that even small changes in the parameter values can make a very big difference in the model results. If the algorithm we used can randomly adjust the 5 parameters in the range [0,10] in increments of 0.001, 1) **how many combinations of parameters can be generated in total?** 2) **If each trial takes 0.01 second on your laptop, how many hours will it take to check all the combinations?** Answer the two questions and explain the calculation logic and process in detail.

| 1. 0.000~9.999 :10000 iterations (per one parameter)   Number of cases that [a,b,c,d,n] can be generated?  : 10000⁵ = 10²º  So 10²º combinations of parameters can be generated in total.   1. To check all the parameters, it need (one parameter set checking time)\*(combinations of parameters) time.   So it will be (0.01s) X (10²º) = 10¹⁸s.  Convert 10¹⁸ seconds into hours: 10¹⁸ / 3600 = 277777777777777.8 hours. |
| --- |

1. As the number of unknown parameters increases in your model, finding the best parameter set becomes extremely difficult. Are there any ideas that allow the computer to perform parameter estimation faster than the algorithm used for this task? It's okay if your idea is wrong, so feel free to share your own thoughts.

| If we know the x and y coordinates of the part with the largest error, we will know how much the error is in that part, and we need to move parameters in the way of reducing errors. If we continue step by step in this way, we will be able to access the desired value faster and more accurately than randomly assigning values. |
| --- |

1. When you finish writing computer code (you know that it works), the first thing to do is to make sure that the code is error-free. **The fact that the code produces results well doesn't guarantee that the results are correct.** The most important part of this assignment is to make sure that the loss function accurately calculates the error. 1) How would you check this? How can you objectively test the accuracy of the loss function in your code? 2) Before submitting this assignment, did you do that?

| 1. I think it would be accurate to print the parameter values and then calculate by myself. If you match two to three with computers, it will be possible to check whether it did well or not. 2. There are quite a lot of calculations, so I calculated only one value with the smallest MSE by hand. There is a difference(0.5) between the computer-calculated and the hand-calculated value, which I think is a sufficient error because the person rounded from the second decimal place when calculating. |
| --- |

1. Please write down what you learned from this assignment or what was most memorable to you (Any answer will earn points).

| It was interesting to draw various graphs using one value list. I can express graphs so diversely even with the same value, and it was also amazing that all of such tools were already made. |
| --- |

1. Do you have any questions or things you don't understand on this topic? (Any question you ask will earn points)

| Using the MSE loss function, we can know how much error there is. By the way, I think it can estimate parameters more effectively if it knows where the error occurred(most significant or big value). Is there any way to find the part where the error occurred the most and adjust it? |
| --- |

1. If you have your team member or outside member who helped you understand, please write the name and what you are grateful for.

| Systematic trio team |
| --- |