Dugga 2024-05-23

TBMT19 / TBMT37

Write your Dugga-ID on all pages and your answers in Swedish or English. There are in total 5 questions, each worth 3 points. You need at least 10 points with 2 points per question, or 12 points in total to pass. Good luck! /William

1 Model formulation and model parts

You have the following model:

$$\frac{dA}{dt} = -v1 + v2 + v7 \qquad v1 = k1 \cdot A \cdot u1 \qquad v2 = k2 \cdot B$$

$$\frac{dB}{dt} = v1 - v2 - v3 + v4 \qquad v3 = B \cdot Vmax \cdot \frac{u2}{kM + u2} \qquad v4 = k4 \cdot C$$

$$\frac{dC}{dt} = v3 - v4 - v5 + v6 \qquad v5 = k5 \cdot C \qquad v6 = k6 \cdot D$$

$$\frac{dD}{dt} = v5 - v6 - v7 \qquad v7 = k7 \cdot D$$

$$[k1, k2, Vmax, kM, k4, k5, k6, k7] = [1, 1, 2, 2, 4, 1, 5, 3]$$

$$[A(0), B(0), C(0), D(0)] = [1, 0, 0, 1]$$

$$[u1, u2] = [10, 2]$$

(a) Give an interaction graph which represents the model. (1 point)

(b) List the model parameters and the model inputs. (1 point)

(c) Update the reaction rate v3 to not be *saturated* with respect to u2. (1 point)

2 Model simulation

(a) Use the Euler forward method to simulate the following model with step-length $\Delta t = 0.1$. What are the values in t = 0.1, t = 0.2, and t=0.3? (1 point)

$$d/dt(A) = -v1 \qquad v1 = k1 \cdot A$$

$$k1 = 2, \quad A(0) = 10$$

(b) If we simulate the model for a long time, what values will A go towards? (1 point)

(c) Why do we need to numerically simulate systems biology models? (1 point)

3 Parameter estimation

(a) You have a model simulation ($\hat{y}$) of four different time-points for a given parameter set (θ):

$$\hat{y}(\theta) = [11, 24, 31, 29]$$

You also have corresponding experimental data (y) with measurement uncertainties (SEM):

$$y = [10, 28, 30, 30], \qquad SEM = [2, 4, 2, 1]$$

What is the cost for the parameter set θ? (1 points)

(b) Why do we square the residuals? (1 point)

(c) Are *local* optimization algorithms guaranteed to find the optimal solution? (1 point)

4 Statistical tests

Assume that we again have the same model simulation and data as in question 3 above.

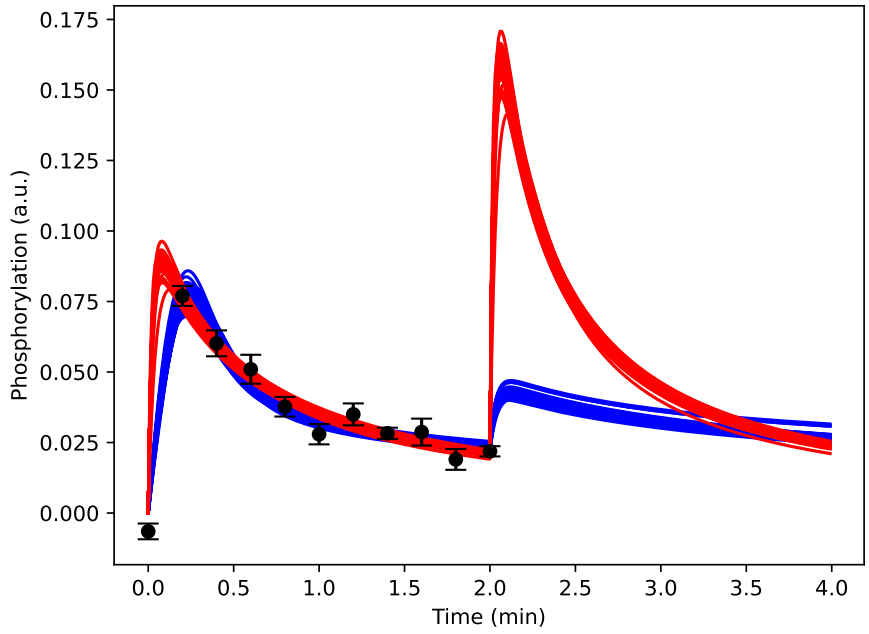
(a) If the χ^2-limit is 2.4, should you reject the model simulation given the parameter set θ? What is the conclusion of the test? (1 point)
If you did not calculate the cost in 3(a), assume that the cost is 3.2.

(b) What is the null hypothesis of the χ^2-test, and what do you conclude if you *reject* the model simulation with the χ^2-test? (1 point)

(c) Based on your conclusion from 4(a), should you reject the *model* if the parameter set θ is not the optimal parameter set? (1 point)

5 Predictions and experimental design

Assume that you have the following model predictions from two different models: a *red* model and a *blue* model. The predictions correspond to the phosphorylation of a protein over time.



(a) If your experimental partner wanted to measure the response at t = 3.5, what would you tell them? What could be the outcome(s) of the experiment? (1 point)

(b) What would you say, if you would try to convince them to measure in one other time-point? Also give your preferred time-point. (1 point)

(c) If one of the models is able to predict the new data correctly, does that mean that the model is the correct one? (1 point)

Answers: Dugga 2024-05-23(revised: February 28, 2025)

1 Model formulation and model parts

You have the following model:

$$\frac{dA}{dt} = -v1 + v2 + v7 \qquad v1 = k1 \cdot A \cdot u1 \qquad v2 = k2 \cdot B$$

$$\frac{dB}{dt} = v1 - v2 - v3 + v4 \qquad v3 = B \cdot Vmax \cdot \frac{u2}{kM + u2} \qquad v4 = k4 \cdot C$$

$$\frac{dC}{dt} = v3 - v4 - v5 + v6 \qquad v5 = k5 \cdot C \qquad v6 = k6 \cdot D$$

$$\frac{dD}{dt} = v5 - v6 - v7 \qquad v7 = k7 \cdot D$$

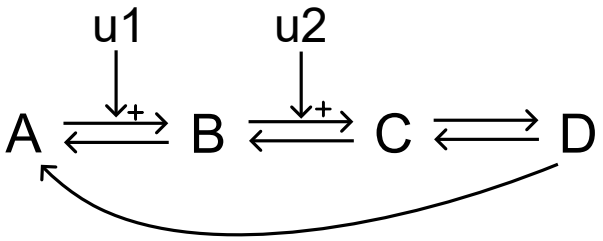
$$[k1, k2, Vmax, kM, k4, k5, k6, k7] = [1, 1, 2, 2, 4, 1, 5, 3]$$

$$[A(0), B(0), C(0), D(0)] = [1, 0, 0, 1]$$

$$[u1, u2] = [10, 2]$$

(a) Give an interaction graph which represents the model. (1 point)

Answer:
The interaction graph should look something like this:



(b) List the model parameters and the model inputs. (1 point)

Answer:
The model parameters are: k1, k2, Vmax, kM, k4, k5, k6, k7.

The initial conditions are also often considered to be parameters: A(0), B(0), C(0), and D(0).

The model inputs are: u1 and u2.

(c) Update the reaction rate v3 to not be *saturated* with respect to u2. (1 point)

Answer:
The reaction $B \rightarrow C$ corresponds to the reaction rate v3.

To make v3 no longer be saturated with respect to $u2$ we can update the reaction rate to:

$$v3 = k3 \cdot B \cdot u2$$

Value for the new introduced parameter (made up): $k3 = 3$.

2 Model simulation

(a) Use the Euler forward method to simulate the following model with step-length $\Delta t = 0.1$. What are the values in t = 0.1, t = 0.2, and t = 0.3? (1 point)

$$d/dt(A) = -v1 \qquad v1 = k1 \cdot A$$

$$k1 = 2, \quad A(0) = 10$$

Answer:
In the Euler forward method, we start with the initial values of the states and calculate the values of the ODEs at the initial time-point. We then take a small step in the direction of the ODEs, and use the new values to calculate the ODEs at the next time-point. This is repeated until the time-frame asked for by the user has been simulated.

The Euler forward method can be written as:

$$x(t + \Delta t) = x(t) + d/dt(x(t)) \cdot \Delta t$$

For this model, to calculate the values of A in t = 0.1, t = 0.2, and t = 0.3 we start by

calculating the values for t = 0.1 using the following steps:

$$v1(0) = k1 \cdot A(0) = 2 \cdot 10 = 20$$

$$d/dt(A(0)) = -v1(0) = -20$$

$$A(0.1) = A(0) + d/dt(A(0)) \cdot \Delta t = 10 - 20 \cdot 0.1 = 8$$

Then we calculate the values for t = 0.2 using the following steps:

$$v1(0.1) = k1 \cdot A(0.1) = 2 \cdot 8 = 16$$

$$d/dt(A(0.1)) = -v1(0.1) = -16$$

$$A(0.2) = A(0.1) + d/dt(A(0.1)) \cdot \Delta t = 8 - 16 \cdot 0.1 = 6.4$$

Finally, we calculate the values for t = 0.3 using the following steps:

$$v1(0.2) = k1 \cdot A(0.2) = 2 \cdot 6.4 = 12.8$$

$$d/dt(A(0.2)) = -v1(0.2) = -12.8$$

$$A(0.3) = A(0.2) + d/dt(A(0.2)) \cdot \Delta t = 6.4 - 12.8 \cdot 0.1 = 5.12$$

So, the values of A in t = 0.1, t = 0.2, and t = 0.3 are:

$$A(0.1) = 8, \quad A(0.2) = 6.4 \quad A(0.3) = 5.12$$

(b) If we simulate the model for a long time, what values will A go towards? (1 point)

Answer:
Since A is decreasing with a constant rate, it will go towards zero as time goes to infinity. This is because the rate of change of A is proportional to A, and the rate is negative. Therefore, A will go towards zero as time goes to infinity.

(c) Why do we need to numerically simulate systems biology models? (1 point)

Answer:
In systems biology, we often have complex models with many states and reactions, and it is often not feasible to solve these models analytically. Therefore, we need to use numerical simulation to solve the models and obtain the model simulations.

3 Parameter estimation

(a) You have a model simulation ($\hat{y}$) of four different time-points for a given parameter set (θ):

$$\hat{y}(\theta) = [11, 24, 31, 29]$$

You also have corresponding experimental data (y) with measurement uncertainties (SEM):

$$y = [10, 28, 30, 30], \qquad SEM = [2, 4, 2, 1]$$

What is the cost for the parameter set θ? (1 points)

Answer:
The cost for the parameter set θ is calculated as follows:

$$\text{residuals} = y - \hat{y} = [10, 28, 30, 30] - [11, 24, 31, 29] = [-1, 4, -1, 1]$$

$$v(\theta) = \sum_{t=1}^{N} \frac{(y(t) - \hat{y}(t, \theta))^2}{SEM(t)^2} = \frac{(-1)^2}{2^2} + \frac{4^2}{4^2} + \frac{(-1)^2}{2^2} + \frac{1^2}{1^2}$$

$$= \frac{1}{4} + \frac{16}{16} + \frac{1}{4} + \frac{1}{1} = \frac{1}{4} + 1 + \frac{1}{4} + 1 = 2.5$$

(b) Why do we square the residuals? (1 point)

Answer:
To make sure that the residuals are positive (and to penalize large residuals more than small residuals).

(c) Are *local* optimization algorithms guaranteed to find the optimal solution? (1 point)

Answer:
No, neither local nor global optimization algorithms are guaranteed to find the optimal solution. Local optimization algorithms are only guaranteed to find *local* minimum not the global minimum.

4 Statistical tests

Assume that we again have the same model simulation and data as in question 3 above.

(a) If the χ^2-limit is 2.4, should you reject the model simulation given the parameter set θ? What is the conclusion of the test? (1 point)
If you did not calculate the cost in 3(a), assume that the cost is 3.2.

Answer:
The χ^2-limit is 2.4, and the cost is 2.5 (or 3.2). Since the cost is larger than the χ^2-limit, we should reject the model simulation given the parameter set θ. The conclusion of the test is that we reject the model given the parameter set θ.

(b) What is the null hypothesis of the χ^2-test, and what do you conclude if you *reject* the model simulation with the χ^2-test? (1 point)

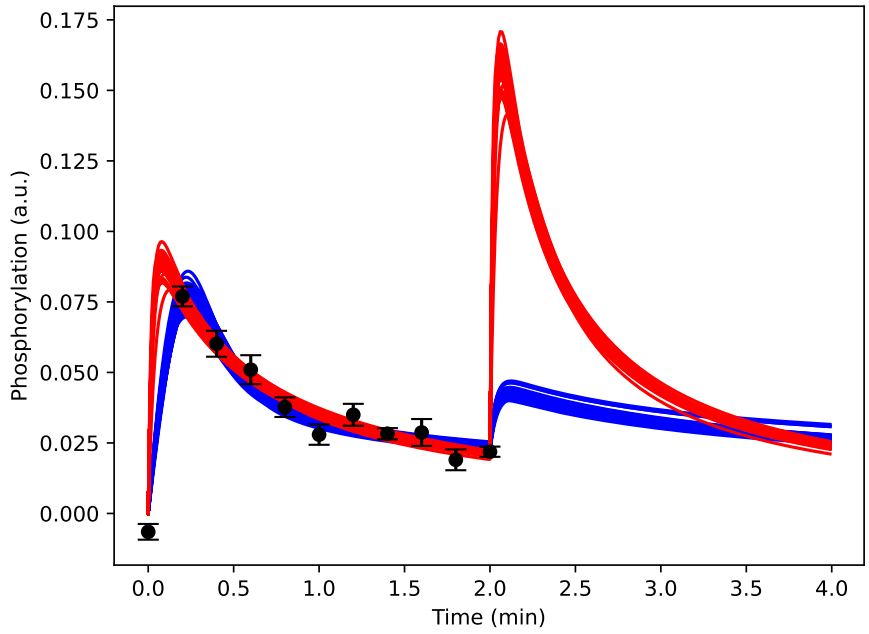
Answer:
The null hypothesis of the χ^2-test is that the model simulation agrees with the data given the parameter set θ. If we reject the model simulation with the χ^2-test, we conclude that the model simulation does not agree with the data given the parameter set θ.

(c) Based on your conclusion from 4(a), should you reject the *model* if the parameter set θ is <u>not</u> the optimal parameter set? (1 point)

Answer:
No. There could be other parameter sets that fit the data better, and we should not reject the *model* based on one parameter set. We should instead try to find the parameter set that fits the data the best. If we also reject the model simulation with the best parameter set, we should reject the model.

5 Predictions and experimental design

Assume that you have the following model predictions from two different models: a *red* model and a *blue* model. The predictions correspond to the phosphorylation of a protein over time.



(a) If your experimental partner wanted to measure the response at t = 3.5, what would you tell them? What could be the outcome(s) of the experiment? (1 point)

Answer:
I would say that this is not a good idea. Since the models overlap in this point, we will not be able to distinguish between the models. The outcome of the experiment could be that we are not able to reject any of the models. However, we might still be able to reject both models if neither agree with the data. However, there are many other time-points where the models differ more, and we would be able to reject at least one of the models.

(b) What would you say, if you would try to convince them to measure in one other time-point? Also give your preferred time-point. (1 point)

Answer:
It would be preferable to measure around 2.25, since this is where the models differ the most and the uncertainty is low. If the experimental data has a reasonable uncertainty, we will be able to reject at least one of the models.

(c) If one of the models is able to predict the new data correctly, does that mean that the model is the correct one? (1 point)

Answer:
No, it does not. The model could still be wrong. We need to continue validating the model by making additional predictions and experiments.