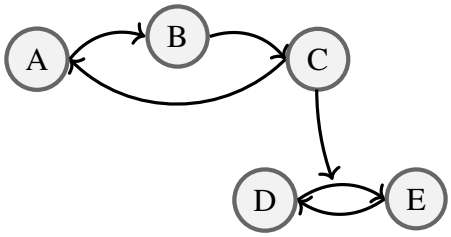
Dugga A 2023

TBMT37 / TBMT19

Please, write your Dugga-ID on all pages and your answers in Swedish or English. You need at least 2 points per assignment or 12 points in total to pass. Good luck! /Elin

1 Model formulation and model parts

Consider the following interaction graph.



(a) Write down the ordinary differential equations that corresponds to the interaction graph. Assume that we have measured the concentration of E. Make any other necessary assumptions and include in the answer. Introduce parameters with values of your choice. Make sure your suggested model is complete. (2 points)

(b) List all model states. What characterizes model states? (1 point)

2 Model simulation

What is a model simulation, how is a model simulation computed, and why do we simulate models in systems biology? (3 points)

3 Parameter estimation

(a) Give example of a cost function and explain all parts of the equation (2 points)

(b) What is the difference between local and global optimization algorithms? (1 point)

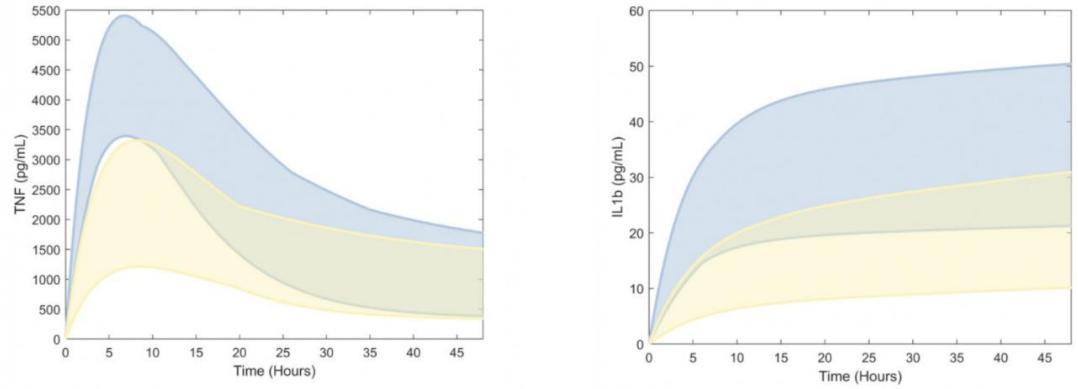
4 Statistical tests

(a) Formulate a null hypothesis that can be used to reject a model! (1 point)

(b) What do you conclude when you do not reject the null hypothesis in a chi2-test? (1 point)

(c) Give example of a situation when you would use cross validation! (1 point)

5 Predictions and experimental design



You have two models, blue model and yellow model, both models are in agreement with training data from the inflammatory responses in monocytes (not shown). Here you see predictions with uncertainty from the two models. The predicted concentration of tumor necrosis factor (TNF) is shown to the left and the predicted concentration of interleukin 1 beta (IL1b) is shown to the right. Suggest a measurement were you know you will draw at least one conclusion when you get the data. Also, give an example of a measurement were you do not know if you will draw any conclusion when you get the data. Motivate your answers. (3 points)

Answers: Dugga A 2023

1

(a) Identify model states: $x1 = [A], x2 = [B], x3 = [C], x4 = [D], x5 = [E]$

Identify reaction rates:

$$v1 = k1 \cdot x1$$
$$v2 = k2 \cdot x2$$
$$v3 = k3 \cdot x3$$
$$v4 = k4 \cdot x3 \cdot x4$$
$$v5 = k5 \cdot x5$$

Assumptions: mass-action kinetics

Formulate ODEs:

$$d/dt(x1) = -v1 + v3$$
$$d/dt(x2) = v1 - v2$$
$$d/dt(x3) = v2 - v3$$
$$d/dt(x4) = -v4 + v5$$
$$d/dt(x5) = v4 - v5$$

What is measured?
The measurement equation, $\hat{y} = x5$ shows that we can measure the concentration of E.

$$\hat{y} = x5$$

Parameters and their values:
$k1 = 3$, $k2 = 1$, $k3 = 2$, $k4 = 4$, $k5 = 1$
$x1(0) = 20$, $x2(0) = 0$, $x3(0) = 0$, $x4(0) = 5$, $x5(0) = 0$
All parameter values are made up.

(b) The model states are: $x1 = [A], x2 = [B], x3 = [C], x4 = [D], x5 = [E]$
Model states are being derived with respect to time and are therefore useually changing with time.

2

A model simulation is the numerical solution to the ordinary differential equations. The solution is based on small time steps in the direction of the flow/gradient. Euler Forward is one method that can be used to compute the solution. We use numerical simulation since most models are complex and thus the ordinary differential equations does not have analytical solutions.

3

(a) A cost function can look like this

$v(p) = \sum \frac{(y(t)-\hat{y}(t,p))^2}{SEM(t)^2}$,

where the sum is over all measured time points. The input to a cost function is the values of the parameters, p, and the output to a costfunction is the agreement between model simulations and data, $v(p)$. The cost function computes the residuals $(y(t) - \hat{y}(t,p))$, i.e. the distance between data $y(t)$ and model simulation $\hat{y}(t,p)$, weighted by the uncertainty of data, SEM(t). The terms are squared so that negative terms become positive and thus all terms have a positive contribution to the calculated cost.

(b) A local optimization algorithm searches only downhill in the cost landscape and thus can only find local optima, while a global optimization algorithm searches both uphill and downhill for the global optimum.

4

(a) E.g. H0: The residuals (difference between model simulation and data) are small compared to the uncertainty in data, or H0: There is no difference between the model simulation and the data.

(b) You conclude nothing when you do not reject the null hypothesis in a statistical test.

(c) When you want to test if the model is too complex/have too many parameters compared to the amount of data and you think that the model might be overfitted to data due to this complexity.

5

A measurement of the concentration of TNF at 5h will give at least one of the following conclusions: reject blue model, reject yellow model, or reject both models. The reason for this is that the core prediction / predictions with uncertainty for the concentration of TNF are not overlapping at that time. If we instead for example would measure the concentration of IL1b at 40 hours, there is no guarantee that we will be able to conclude anything since the predictions are overlapping. If the corresponding data would be, say 25 pg/ml, none of the models would be rejected.