Analysing Enzymatic Kinetic Data

André Heck & Marthe Schut; University of Amsterdam

A. Information for lecturers

Unit description

Description: Students determine from experimental data the best fit of kinetic parameters v_{max} and K_m in the Michaelis-Menten model of enzyme kinetics and check whether this mode offers a good model-based description of the experimental data taken from literature. We guide the students' research steps and ask questions which will hopefully encourage deep thinking, trying out methods, and reflecting on intermediate results. At the end of this self-contained unit students have carried out a thorough analysis of enzyme kinetics and learned many mathematical methods and techniques that are of use in scientific practice.

Student and discipline level: This unit is concerned with data analysis in enzyme kinetics, introduced to first year bachelor students in biomedical sciences.

Prior knowledge: Expected student and teacher knowledge and skills are

- experience and some fluency in doing mathematical computation in the R environment;
- acquaintance with finite difference methods to compute numerical derivatives
- basic understanding of linear and nonlinear regression
- basic understanding of differential equations
- rather well-developed symbol sense and graph sense

It is assumed that students have prior practical experience in the R environment with

- plotting data
- manipulating and transforming data
- carrying out linear and nonlinear regression
- computing numerical derivatives via finite difference methods

Estimated duration: Expected time needed for students to carry out the assignment is 4 hours under the assumption that they have the required ICT expertise and mathematical abilities.

Learning objectives

At completion of the unit, students will be able to

- make the rather giant step from theory on enzyme kinetics and the Michaelis-Menten model to data analysis of real data from lab experiments (references to papers in the research literature are given in the student texts);
- compute numerically the rate of change of a chemical reaction from real data via a suitable finite difference method and know what the purpose of doing so is in the context of differential equations as models of a problem situation;
- read, interpret, and create plots of rate of change against reactant concentration and variations thereof with an eye toward parameter estimation;
- understand the purpose of data transformation in the context of regression analysis and carry out the required steps;

- carry out computations in the R environment as professionals and understand better what scientists cope with in practice;
- reflect on methods and techniques so that they can apply them in other contexts, too.

IBME character

The unit can be characterized as structured inquiry, meaning that students follow directions and hints in a structured teaching-learning path based on professional practice of parameter estimation, but conclusions are predominantly based on the investigation carried out by individual students or pairs of students.

The kind of student inquiry we have in mind is that the students personally experience that data analysis is in real practice not a routine task as suggested by many textbooks or a matter of entering the right command in a software environment, but normally involves quite a lot of decision making and exploration to come to satisfying answers. This also motivates the use of real data from research literature: students not only learn data analysis but also do data analysis and learn about the approaches biomedical scientists would apply. The techniques that students learn in the unit, such as linearization of the model to find good initial values for nonlinear parameter estimation and computational exploration of the sensitivity of parameters, are useful in other contexts as well.

Mathematical content

The main mathematical content, within the context of biomedical sciences, is

- visualization of data
- mathematical manipulation and transformation of data
- regression (linear and nonlinear)
- numerical differentiation
- parameter estimation
- mathematical modelling of enzyme kinetics

Technological pedagogical content knowledge

Data analysis is often taught in a procedural way with a focus on methods and techniques applied to artificial data. In real practice it is normally not a routine task and real data are not smooth, but the work involves quite a lot of decision making, handling data, and exploration to come to satisfying answers. Students lack in this way sufficient experiences of the real process of analysing real data.

Data analysis is not a matter of entering the right command or selecting the right menu item in a software environment. Students have to understand the underlying mathematics and the ICT tools to do the mathematics. For instrumentation theory it is known that this is not a trivial issue. In this unit for first year biomedical students, who lack much experience in using the R environment for mathematical work, students get directions in how to use R, and it is expected that this ICT use is discussed with them in lectures and tutorials (including live demonstrations).

Learning path

In the table below of student activities within the unit we typify them as part of the 7E learning cycle of inquiry.

Assignment	*Activity*	*E-emphasis*
1	giving meaning to kinetic variables	elicit
2	defining and understanding the Michaelis-Menten model as ODE	engage
3	estimating the initial concentration of the substrate	explore
4	transforming data to a linear model	engage
5	computing the reaction rate and constructing the Lineweaver-Burk plot	explore
6	estimating parameters via the Lineweaver-Burk plot	engage
7	doing a numerical sensitivity analysis of kinetic parameters	explore
8	using other transformations: the Eadie-Hofstee plot and the Hanes-Woolf plot	elaborate
9	doing nonlinear regression through the Michaelis–Menten formula	engage
10	doing nonlinear regression using the differential equation	elaborate

Grouped in ways suggested in the Pathways to Inquiry project (http://pti.lsu.edu), we distinguish the following inquiry abilities in the unit:

- *Testable questions:* answer questions through reflection and exploration. Students are frequently asked questions about the meaning and quality of outcomes. It is stressed that finding an answer to a question is not good enough, but that reflection on the outcome(s) is more important. By asking many small questions, students are encourage to ask themselves in future also questions without being prompted and think about things to be done ahead.
- *Scientific investigation*: carry out a regression analysis in a professional way, meaning that much attention is also paid to the quality of the mathematical methods and techniques used. When it comes to a method of numerical differentiation, students must select their method and verify that it works flawless or at least well enough for its purpose. A systemic approach is pursued to tackle the problem at hand in which one keeps an open eye for alternative methods and techniques.
- *Tools and techniques:* use appropriate tools and techniques to analyse and interpret data in similar way as practitioners do. For this reason, the popular, open source R environment is used. But also use of common techniques are promoted such as transforming data to a linear mode for which linear regression is applicable and using the results as initial values of parameters in an nonlinear regression method. Another example is the use of several methods of finite differences to compute a numerical derivative.
- *Using evidence:* create computer outcomes in the regression analysis process that can be used for interpreting the outcomes and for improving the quality of work. For example, students are explicitly invited to explore more than one data transformation and compare the results. The create plots to explore the quality of results. The results from linear regression after data transformation and nonlinear regression of raw data are compared to convince students that one must be careful with data transformations when it comes to parameter estimation. Graphs are convincing in this context as one sees immediately the better quality of nonlinear regression. This goes beyond what students traditionally learn in textbook and practical work.
- *Data and explanations:* think critically and logically to make the relationships between evidence and explanations. Many small questions are embedded in the assignments to let students think more deeply about the results that they obtained. For example in the task about numerical differentiation students are asked whether the expectation of a straight line is met for the computed numerical derivative and they are stimulated, if needed, to try to improve

the method of numerical differentiation. The idea behind this is that students learn that it is important to inspect not only the quality of the final result, but also to do this rigorously for intermediate results.

- *Alternative explanations:* recognize and analyse alternative explanations. This is mainly touched upon by looking at three methods of transforming the raw data into a linear model. Students are invited to explore the sensitivity of the kinetic parameters with respect to the linearization process.
- *Using mathematics skills:* use mathematics in all aspect of scientific inquiry. Students are expected to be acquainted with basis mathematical knowledge and skills, and be able to apply them. Main skills concern symbol sense (interpreting and manipulating algebraic expressions and relating them to other representational formats), graph sense (creating, reading and interpreting graphical representations), computational abilities with ICT (doing mathematics with R). Specific presumed mathematical knowledge concerns numerical differentiation, basics of differential equations (at least that they represent a relationship between a function and its derivatives), and basics of linear and nonlinear regression analysis (including the experience that linear regression is more routinely carried out than nonlinear regression).

Experiences

The unit has not yet been tried out in classroom.
(in the study year 2018/2019 no time was left so that a choice had to be made between this unit and a unit on mathematical modelling of immune response. the same was true in the study year 2019/2020; because of a course reform which provides more tutorial time, it is expected to be tried in the study year 2020/2021).

Student with special needs

No special resources have been included for students with special needs.

Assessment

Suggestion: let students do a similar analysis of published enzyme kinetic data.

Relevance of/to the real word

Analysis of enzyme kinetic data by the methods explored by the students in this unit is common practice in biomedical laboratory work. It helps students read textbooks and publications in biomedical journals.

B. Student learning activities

The unit consists of 10 assignments. The envisioned student engagement in each activity has already been listed in the description of the learning path in the section *Information for lecturers*. Suggestions of teacher-student interactions have been briefly discussed in that section as well.

Assignment 1: The meaning of kinetic variables

Learning objective: Understanding the meaning of the variables in the Michaelis-Menten formula $r = \frac{v_{max}s}{K_m+s}$.

Student activity: Answering questions that hopefully prompt them to look at the formula from several perspectives. Optionally this task can be used for group discussion in class.

Tool use: SOWISO environment (or Word/PDF/ePub export of theory page) and pen-and -paper

Assignment 2: Differential equations in chemical kinetics.

Learning objective: Formulating the Michaelis-Menten model in terms of a differential equation for the substrate concentration $\frac{ds}{dt} = -\frac{v_{max}s}{K_m+s}$ and the product concentration $\frac{dp}{dt} = \frac{v_{max}(s_0-p)}{K_m+s_0-p}$ where s_0 is the (unknown) initial substrate concentration.

Student activity: Formulating the initial value problem of enzyme kinetics according to the Michaelis-Menten model. Optionally this task can be used for group discussion in class.

Tool use: SOWISO environment (or Word/PDF/ePub export of theory page) and pen-and -paper

Assignment 3: Estimation of the initial substrate concentration.

Learning objective: Figuring out how to estimate the initial substrate concentration s_0 from a plot of the product concentration vs time. Student must realize that this estimate is unknown in the measured data taken from a published research paper, but that it should be equal to the maximum product concentration reached after enough time. It helps students realize that formulas have a meaning and can be used purposefully.

Student activity: Creating a plot of the product concentration vs time and use it to estimate the initial substrate concentration by reading the graph.

Tool use: R environment to create a data plot

Assignment 4: From nonlinear to linear regression: the Lineweaver-Burk plot.

Learning objective: Understanding how to transform measured data to a linear model that corresponds with the Lineweaver-Burk plot. For this purpose, students first derive the linear relationship and afterward linger upon the meaning of the variables. The Lineweaver-Burk plot is the standard method found in most textbooks. The purpose of the assignment is in essence that students understand that transformation to a linear problem is a way to make the problem more manageable. In this case, the linear problem will provide estimates for initial values of the ODE describing the enzyme kinetics.

Student activity: Deriving the linear relationship between $\frac{1}{r}$ and $\frac{1}{s}$ and then linger upon the meaning of the variables in this linear relationship when plotted against each other. Optionally this task can be used for group discussion in class.

Tool use: SOWISO environment (or Word/PDF/ePub export of theory page) and pen-and -paper.

Assignment 5: Numerical reaction rate and the Lineweaver-Burk plot.

Learning objective: Finding a way to compute the reaction rate as a numerical derivative of the product concentration and using it to create the Lineweaver-Burk plot and visually check that the scatter plot suggest to a straight line fit.

Student activity: Exploring methods (of finite difference, but this is only hinted at) to approximate the reaction rate numerically. This numerical quantity can then be used to create the Lineweaver-Burk plot that should be a scatter plot with point lying almost all on a straight line. Visual inspection of the quality of the numerical derivative on the basis of the Lineweaver-Burk plot suffices.

Tool use: R environment to determine the reaction rate by numerical differentiation of the data and to create the Lineweaver-Burk plot

Assignment 6: Parameter estimation via the Lineweaver-Burk plot.

Learning objective: Carrying out linear regression on transformed data, and learning to deal sensible with possible discrepancies between expected results and obtained results.

Student activity: Carry out the linear regression of the transformed data in the Lineweaver-Burk plot and determine on the basis hereof numerical estimates of the parameters v_{max} and K_m. In the answer model we compare the result with rougher estimate of these parameters from a plot of the reaction rate against the substrate concentration.

Tool use: R environment to carry out computational work and compare obtained results with expected results.

Assignment 7: Sensitivity of the estimation of kinetic parameters.

Learning objective: Finding out how sensitive the estimated values of the kinetic parameters are to changes in the estimated value of the initial substrate concentration. Students learn to be not satisfied with an obtained result, but also that a professional modeler inspect the quality of the obtained result.

Student activity: Using the R environment to find out how the values of the kinetic parameters change when one varies the initial substrate concentration s_0 by a small amount.

Tool use: R environment as a tool to carry out sensitivity analysis by doing various simulations.

Assignment 8: Eadie-Hofstee plot and Hanes-Woolf plot.

Learning objective: Exploring alternative transformations of data toward a linear model. Again this promotes exploration of the quality of a model.

Student activity: Exploring two alternative transformations of the Michaelis-Menten formula to a linear model, namely the Eady-Hofstee and the Hanes-Woolf plot, which are less often or only breifly mentioned in textbooks. Students must derive the linear relationships, plot them, and compare outcomes with each other, Special attention is asked to the sensitivity of the results in these methods on the estimated initial substrate concentration.

Tool use: R environment to carry out computational work and compare obtained results with results obtained through other methods.

Assignment 9: Nonlinear regression through the Michaelis-Menten formula.

Learning objective: Carrying out nonlinear regression though the Michaelis-Menten formula, using the computed reaction rate, the estimated substrate concentration, and the previously determined guesses of kinetic parameters. Students learn that nonlinear regression preceded by initial estimation of parameter values is a useful approach and can lead to better results than parameter estimation via linear regression of transformed data.

Student activity: Using the R environment to carry out the nonlinear regression through the Michaelis-Menten formula and determine on the basis hereof numerical estimates of the parameters v_{max} and K_m. Students checked graphically that these new parameter values lead to better

Tool use: R environment to carry out nonlinear regression fort he purpose of improving parameter estimates

Assignment 10: Nonlinear regression using the differential equation.

Learning objective: Fitting parameters in nonlinear differential equations to measured data by guess-and-improve methods. Students learn to use advanced methods as grey-box tools

Student activity: Using the `FME` package with its dependent R packages to fit parameters in nonlinear differential equations to measured data. In essence one repeatedly solves the ODE numerically for given parameter values and systematically improves the parameter values. Students compare outcomes of the `FME` package with those obtained with the Lineweaver-Burk plot to see that results have indeed improved and are not compatible with values found in the literature. We hope that this raises the students' appreciation and confidence in using advanced methods as black boxes in their data analysis. At least, it shows how professionals do this kind of work.

Tool use: R environment to apply state-of-the-art methods used by professional modellers

Suggestions for use

The unit can be shortened by leaving out assignments 8-10.

Lecturers can consider interactive discussions with their students in assignments 1-4 to engage them in the data analysis and discuss various aspects of mathematical inquiry.

Lecturers are advised to stimulate and make time for discussions amongst peers (for example, allowing students to compare and judge their approaches in assignments and the results obtained, to listen and react to each other's arguments, to come up with better computational ideas, etc.)