Zuse Institute Berlin (ZIB)
Freie Universität Berlin
Dr. Vikram Sunkara, Dr. Max von Kleist

4. Homework
Numerics for Bioinformaticians
WS 2016/17

Deadline: November 23, 10:00 (**before** the lecture)

The homework should be worked out in groups of two or three students. Each solution sheet must contain the names and student numbers ('Matrikulationnummer') of all group members and the exercise group (Wednesday/Friday). Solutions to homework 1&2 must be handed in in paper form, either hand written or printed out if generated electronically. Please staple all sheets. Programming tasks (homework 4) must be submitted to BioInfNumerik@hotmail.com by email. Before sending it, please 'zip' it.

Homework 1 (2 points)
Given

$$X_t = X_0 + \mathcal{P}(\alpha t) - \mathcal{P}\left(\int_0^t \beta X_s ds\right),$$

for some time $t > 0$, show that,

$$\frac{d\mathbb{E}[X_t]}{dt} = \alpha - \beta \cdot \mathbb{E}[X_t].$$

Hint 1) For a fixed time, the Poisson process is distributed according to a Poisson Distribution.
Hint 2) Use the Fundamental Theorem of Calculus.

Homework 2 (4 points)
The Boomerang model. Let us consider the Boomerang process over $\Omega = \mathbb{N}_0^2$.

$$R_1 : \emptyset \xrightarrow{1.0} A, \qquad R_2 : \emptyset \xrightarrow{1.0} B, \qquad R_3 : A + B \xrightarrow{0.01} \emptyset.$$

with stoichiometric matrix S and propensity function vector $r = \{r_i\}$

	r_1	r_2	r_3
A	1	0	-1
B	0	1	-1

$$\begin{aligned} r_1 &= 1 \\ r_2 &= 1 \\ r_3 &= 0.01 \times A \times B. \end{aligned}$$

Let the starting population be $A_0 = 10$ and $B_0 = 10$.
Let N be the number of realisations. For each $N \in \{10, 100, 1000, 10000\}$, compute N realisations up to time point $t_{final} = 10$. Then for each N take the sample mean,

$$\bar{x}^{(N)} := \sum_{i=1}^{N} \frac{x^{(i)}}{N}.$$

Let us denote $\mathbb{E}[A_{10}]$ to be the expectation of just the species at time 10. Given $\mathbb{E}[A_{10}] = 10$, plot on a log–log scale,

$$\left|\bar{x}_A^{(N)} - \mathbb{E}[A_{10}]\right| \text{ versus } N.$$

That is, on a log–log scale, plot the expected sample error of the species A against the true expectation.

Homework 3 (4 points)
Expect the Unexpected. Consider the Slögl model.

$$R_1 : 2x_1 \xrightarrow{k_1} 3x_1, \qquad R_2 : 3x_1 \xrightarrow{k_2} 2x_1 + \emptyset, \tag{1}$$

$$R_3 : \emptyset \xrightarrow{k_3} x_1, \qquad R_4 : x_1 \xrightarrow{k_4} \emptyset. \tag{2}$$

with propensities (*stochastic* reaction rates): $r_1 \ldots r_4$.

$$r_1 = k_1 \cdot x_1(x_1 - 1) \tag{3}$$

$$r_2 = k_2 \cdot x_1(x_1 - 1)(x_1 - 2) \tag{4}$$

$$r_3 = k_3 \tag{5}$$

$$r_4 = k_4 \cdot x_1 \tag{6}$$

with parameter values are $0.15, 0.0015, 20.0$ and 3.5 for $k_1, \ldots, k_4$ respectively. The initial value for x_1 is $x_1(t_0) = 40$. The stoichiometric matrix S is given by:

	r_1	r_2	r_3	r_4
x_1	1	-1	1	-1

Construct 200 realisations of the Slögl model leading up to $t_{final} = 5$. At t_{final}, plot a histogram of species x_1. Also mark the sample expectation of x_1 on the same figure. In this case, is the sample expectation a good statistic to describe the distribution of x_1?

Caution!! *Make sure all reactions leading out of the state space are zero. Furthermore make sure that you are storing the right state when you cross t_{final}.*

BONUS 2 Marks *If you have a fast solver, try running ten thousand realisations and plot the probability distribution (normalized histogram) of x_1 at t_{final}.*