CHARITÉ / HUMBOLDT UNIVERSITÄT ZU BERLIN
INSTITUTE FOR THEORETICAL BIOLOGY

Prof. Hanspeter Herzel
Dr. Grigory Bordyugov
Institute for Theoretical Biology
D-10115 Berlin

h.herzel@biologie.hu-berlin.de
Grigory.Bordyugov@hu-berlin.de
Tel: +49 30 2903 9101 / 9121
Fax: +49 30 2903 8801

MODULE IV - BIOINFORMATICS: ASSIGNMENT 3

Please return your solution in hard copy until Monday May 9th, 15:00.

1. Statistics

Download the data file `data.txt` from `http://bit.ly/1bloBZY`. The file consists of two columns of numbers, in the following the values in the first column will be referred to as x_i and the values in the second column – as y_i.

1. Perform a t-test on x_i and y_i and check the hypothesis that x_i and y_i are drawn from normal distributions with the same mean.
2. Calculate the mean and the standard deviation of x_i and y_i.
3. Calculate the 2×2 covariance matrix between x_i and y_i and interpret its elements.
4. Calculate the correlation coefficient between x_i and y_i. Does it correspond to any of the values in the covariance matrix?

2. Exponential decay

The isotope ^{35}S decays exponentially with a half-time of about 87 days.

1. Sketch the time course of the ^{35}S decay.
2. Formulate a differential equation for the concentration of ^{35}S .
3. Calculate the rate constant λ of the decay.
4. Provide a general formula relating the rate constant λ and the half-life time $t_{1/2}$.
5. After how many months will only 0.1% of the isotope remain?

3. Integrals

Calculate the following integrals:

$$\int_0^1 x\,\mathrm{d}x, \qquad \int_{-3}^{3} x^3\,\mathrm{d}x, \qquad \int_0^{\infty} \mathrm{e}^{-ax}\,\mathrm{d}x, \qquad \int_0^{2\pi} (a\sin\phi + b\cos\phi)\,\mathrm{d}\phi, \qquad \int_1^{\mathrm{e}} \frac{\mathrm{d}x}{x}, \qquad \int_0^{2\pi} \sin^2 x\,\mathrm{d}x.$$

4. Trigonometry

1. Using trigonometric identities, prove the following formula:

$$(1+\sin(\omega t))\,(1+\cos(\omega t)) = 1+\sin(\omega t)+\cos(\omega t)+\frac{1}{2}\sin(2\omega t)\,.$$

2. Express A and ϕ in terms of a and b such that the following identity holds:

$$a\sin(\omega t)+b\cos(\omega t) = A\cos(\omega t+\phi)\,.$$

5. Oscillations

The abundance $x(t)$ of protein A and abundance $y(t)$ of protein B is approximated by harmonic oscillations:

$$x(t) = 1.25\cos(\omega t)\,, \qquad y(t) = \cos(\omega t)+\sin(\omega t)\,,$$

where $\omega = \frac{2\pi}{24}$ and time t is measured in hours.

1. Determine the amplitudes of both oscillations, their periods and the phase difference between them. The result of exercise 5.2 may be helpful here.
2. Sketch graphs of $x(t)$ and $y(t)$. Pay attention to the proper amplitudes, periods and phases of the oscillations.
3. Is it protein A that peaks before protein B or vice versa?
4. Another protein C oscillates according to

$$z(t) = \cos(\omega t)+\cos(2\omega t)\,.$$

Sketch the graph of $z(t)$. What is the period of oscillations of protein C? How would you determine the peak phase of protein C?

6. Feedback and oscillations

Give three examples of feedback-induced oscillations in everyday life and explain how the delay of the feedback influences the period of the oscillations.

7. Binary classification

Please explain the following concepts in a couple of short sentences

1. True positive rate and false negative rate
2. Sensitivity and specificity of a binary classifier
3. ROC-curve and area under curve in the context of ROC-curves