Indian Institute of Technology Bombay

EP219 Data analysis and interpretation
Assignment 2
Dated : 6 - 10 - 2016
Deadline: 14-10-2016 at 10 am

— — — — — — — — — — — — — — — —

Please rotate roles for group members for this assignment.
Try to write your own functions to calculate the mean and standard deviation and to find best fit lines.

Problem 1 Consider two correlated Gaussian random variables x and y with 0 mean with covariance matrix given by,

$$C = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}.$$

a) Create a python code to make the contours of the probability density function of x and y.

b) Generate random pairs of real numbers x and y that have this distribution.

c) Make a 2D histogram of the the generated pairs (x,y) using a heatmap or colormap to represent the height of the histogram.

d) For each pair construct the variable $z \equiv \frac{1}{50}\left(6x^2 + 4xy + 9y^2\right)$.

e) Plot a histogram of z. Find the mean value of z and the standard deviation. Is this consistent with the expected distribution for z?

(See the midsem solutions if you need some help)

Problem 2 Consider the weight (in kg) of a thousand one-year-old baby pandas in the file "pandas.txt". Assuming that the fluctuations in weight are purely statistical, estimate the mean weight and the error on the mean weight. What is the size of the typical fluctuations of the weight of each baby panda about this mean?

Problem 3 Consider the readings for the length of an Aluminum rod (measured in mm) as a function of the temperature in Celsius (varying from 0-10 degrees C) in the file "linearexpansion.txt".

a) Plot the data on a clearly labeled graph with temperature on the x-axis and length on the y-axis. This is called a scatter plot.

b) Assuming uncorrelated identical statistical errors for each reading, minimize the least squares distance to find the best fit line that fits the length vs temperature. What is the best estimate for the length of the rod at 0 degrees Celsius? What is the coefficient of linear expansion in mm/Kelvin?

c) Show the best fit line on top of the graph that you plotted earlier. If we trusted the linear extrapolation up to 15 degrees Celsius, what would you expect the length of the rod to be? Estimate the error would you expect on this length.