In the name of God

Department of Physics
Shahid Beheshti University

ADVANCED COURSE ON COMPUTATIONAL PHYSICS

Exercise Set 5

(Due Date: 1403/08/20)

1. von Neuman algorithm to generate random data with arbitrary PDF: Using this method do:

A: Suppose that the PDF reads as: $p(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ for $x \in [-5, +5]$

B: For $p(x) = \sin(x^2/100) + \frac{1}{\cos(x^3/100)} + x^{-3}$ and $x \in [1, 5]$.

The size of your generated data be $N = 10000$.

2. According to Pearson correlation coefficient, compute the degree of correlation between 0.2.txt and 0.5.txt as well as with themselves.

3. Compute $C(\tau) = \langle x(t+\tau)x(t)\rangle$ for 0.2.txt and 0.5.txt and 0.8.txt data sets. Interpret your results.

4. Non-linear correlation. There are many methods to compute non-linear correlation coefficient. According to Wang, Qiang, Yi Shen, and Jian Qiu Zhang. "A nonlinear correlation measure for multivariable data set." Physica D: Nonlinear Phenomena 200.3-4 (2005): 287-295, and use the Eqs. (1), (2) and (3) of mentioned paper, compute the mutual information between all pairs of 0.2.txt, 0.5.txt and 0.8.txt.

5. Linear and non-linear correlation coefficients. Pearson's coefficient is a familiar measure to quantify the linear-correlation, while for assessing non-linear relation and even to determine the degree of correlation in the presence of outliers the Spearman's correlation coefficient is used. For all available pairs of 0.2.txt, 0.5.txt and 0.8.txt data sets, compute Spearman's and Pearson's correlation coefficient compare your results. Where:

$$\rho_p \equiv \frac{\langle [x - \langle x\rangle][y - \langle y\rangle]\rangle}{\sigma_x \sigma_y}$$

$$\rho_s \equiv 1 - 6\frac{\sum_i d_i^2}{N(N^2-1)}$$

and $d_i \equiv [Rank(x_i) - Rank(y_i)]$ and $Rank$ means the order of value of variable in a set. Suppose that for $\{x\} : \{20, 100, 30, 50, 160, 10\}$. Then the $Rank(x) : \{5, 2, 4, 3, 1, 6\}$.

Good luck, Movahed