Date: 1403/08/24

In the name of God

Department of Physics
Shahid Beheshti University

ADVANCED COURSE ON COMPUTATIONAL PHYSICS

First midterm exam

(Time allowed: 3 hours)

Theoretical part:

1. Widely-used commands in terminal:
A: What is the command to connect a cluster? (assume that the valid IP is 192.168.220.100) (2 points)
B: What is the command to copy a file from cluster to our local computer? (assume that the valid IP is 192.168.220.100) (2 points)
C: What is the command to make a script as an executable file? (1 points)

2. Binomial PDF:
A: Prove that the variance of a binomial probability is

$$\sigma^2 = Np(1-p)$$

(5 points)
B Considering the result proved in Part A, now prove that for small $P(x)$, the error of the probability density function becomes

$$\sigma_m = \frac{1}{Nh},$$

where, h is the bin width. (5 points)

3. The concept of Probability Density Function (PDF):
Suppose that

$$\bar{\rho}(x) \equiv \left\langle \frac{1}{h} K\left(\frac{x-y}{h}\right) \right\rangle_y = \int_{-\infty}^{\infty} \frac{1}{h} K\left(\frac{x-y}{h}\right) \rho(y) dy$$

Where K is a kernel such that $K(A) = 1$ for $|A| \leq \frac{1}{2}$, and $K(A) = 0$ otherwise. Also we define

$$\text{Bias}[\bar{\rho}(x)] \equiv \rho(x) - \bar{\rho}(x)$$

A: What is the result of the following limit:

$$\lim_{h \to 0} \text{Bias}[\bar{\rho}(x)] =?$$

explain your result. (5 points)
B: Using the changing of variable as $s = \frac{x-y}{h}$ and by Tylor expansion of $\rho(x+sh)$ up to $\mathcal{O}(h^3)$, compute $\bar{\rho}(x)$. Finally, determine $\lim_{h\to 0} \text{Bias}(\bar{\rho}(x))$. (5 points)

C: Let's define L as

$$L \equiv Var(\bar{\rho}(x)) + (\text{Bias}[\bar{\rho}(x)])^2$$

Where $Var(\bar{\rho}(x))$ and $\text{Bias}[\bar{\rho}(x)]$ which have been obtained in part B of question 2 and and part B, respectively. Find the best value of h by minimizing the L. (5 points)

CONTINUED

4. The concept of Correlation:
A: Explain different categories of correlation. (5 points)
B: We have recorded two sets of data, $\{x\}$ and $\{y\}$ in an experimental set up. We imagine that the statistical relation of two mentioned series is given by $y \triangleq ax + b$. Now we define an estimator for error as $\epsilon \equiv \langle (y - (ax + b))^2 \rangle$. By minimizing the ϵ, try to find the statistical meaning of $CoV \equiv \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle$. (5 points)

5. Moments and Cumulants: For a Gaussian distribution of random variable, x, with variance equates to σ^2, we use the mapping $x \to y = x - \langle x \rangle$. Show that $\mathcal{K}_1$ and $\mathcal{K}_3$ for y are zero. What is the second cumulant? (5 points)

6. For the non-linear Langevin equation as:

$$\dot{v} = \gamma(t, v) + \eta(t)$$

clarify the physical meaning of each term in mentioned equation. What are the typical properties of η? (10 points)

Computational part:

7. Fitting algorithm. According to the data set provided for you, there are 10 pairs, each pair $(\bar{x}_i, \bar{y}_i)$ is subject to asymmetric errors (see Figure 1). The error-bars for each point as represented by $\bar{x}_i{}^{\sigma_x^+}_{\sigma_x^-}$ and $\bar{y}_i{}^{\sigma_y^+}_{\sigma_y^-}$, are $\sigma_x^- = 0.5$, $\sigma_x^+ = 5$, $\sigma_y^- = 2$, $\sigma_y^+ = 5$.
A: To map our fitting problem with the asymmetric error-bars to a simple case such that the abundance of data points for x around corresponding mean value, namely $\bar{x}_i$ and similarly for $\bar{y}_i$ are constant on different boundary, do following tasks:
1) For each pair, $(\bar{x}_i, \bar{y}_i)$, generate some data points randomly for the range of $x_j^{(i)} \in [\bar{x}_i - \sigma_x^-, \bar{x}_i + \sigma_x^+]$ for $j = 1, ..., 20$ with constant PDF. (10 points)
2) Do the same task as the part 1 just for $\bar{y}_i$. (10 points)
3) Now, for each i, collect the pairs as $(x_j^{(i)}, y_j^{(i)})$. where $i = 1, ..., 10$ and $j = 1, ..., 20$. Finally you have new data set with (x_k, y_k), where $k = 1, ..., 10 \times 20$ and write them in a text file. (5 points)

B: Now by this new data set in your hand, and considering a linear line as theoretical prediction, $Y_{theo.} = ax + b$ with two free parameters, compute the a_{best} and b_{best}. (15 points)

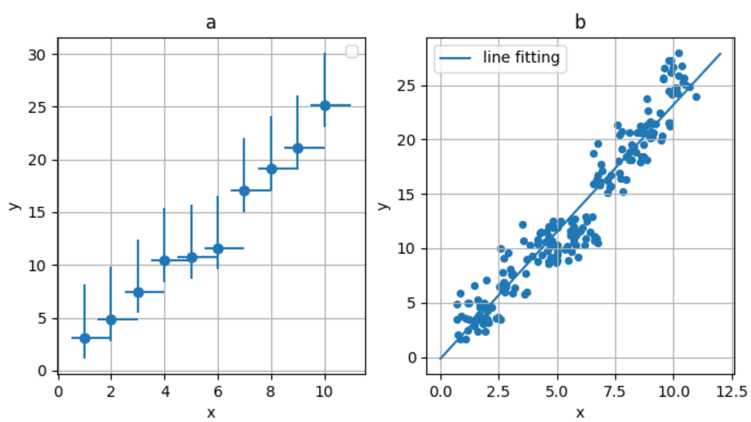

Figure 1: The asymmetric data sets around each filled circle symbols.

Good luck, Movahed

بسم الله الرحمن الرحيم

Answer Key for First midterm

Numerical Analysis 1403. and Advanced computational Physics

① Commands

5 Points

Ⓐ ssh user@192.168.220.100

Ⓑ scp user@192.168.220.100:/address/.

It means here

Ⓒ chmod u+x filename.sh

o+x

g+x

② Binomial PDf

Ⓐ

$$P(n) = \binom{N}{n} p^n (1-p)^{N-n}$$

$$\langle n \rangle = \sum_{n=0}^{N} n \binom{N}{n} p^n (1-p)^{N-n}$$

$$= Np$$

$$\sigma^2 = \langle (n - \langle n \rangle)^2 \rangle = \sum_{n=0}^{N} (n - \langle n \rangle)^2 \binom{N}{n} p^n (1-p)^{N-n}$$

$$\sigma^2 = p^2 N(N-1) + Np - p^2 N^2$$

5 points

$$= \cancel{p^2 N^2} - p^2 N + Np - \cancel{p^2 N^2}$$

$$= Np - Np^2 = Np(1-p)$$

(B)

$$P(x) = \lim_{N\to\infty} \frac{n(x)}{N}$$

$$\frac{dP(x)}{dx} = p(x) \; : \; \text{PDF} \quad \text{or} \quad \frac{\Delta P(x)}{\Delta x} = p(x)$$

$$\sigma^2_{p(x)} = \left(\frac{\sigma_{\Delta P(x)}}{\Delta x}\right)^2 = \frac{NP(1-P)}{N^2(\Delta x)^2}$$

$$\sigma_m^2 = \frac{\sigma^2_{p(x)}}{\langle n(x)\rangle} = \frac{\cancel{NP(x)}(1-P(x))}{N^2\Delta x^2 \cancel{NP(x)}} = \frac{(1-P(x))}{N^2 \Delta x^2}$$

5-Points

$$\boxed{\sigma_m = \frac{1}{N\Delta x}} \quad \text{or} \quad \boxed{\sigma_m = \frac{1}{Nh}}$$

(3) (A)

$$\bar{\rho}(x) = \left\langle \frac{1}{h} K\left(\frac{x-y}{h}\right)\right\rangle_y = \int dy \; \frac{K\left(\frac{x-y}{h}\right)}{h} \rho(y)$$

$$\text{Bias}[\bar{\rho}(x)] \equiv \rho(x) - \bar{\rho}(x)$$

$$\lim_{h\to 0} \text{Bias} = \rho(x) - \lim_{h\to 0} \int dy \left(\frac{K\left(\frac{x-y}{h}\right)}{h}\right) \rho$$

$$\lim_{h\to 0} \frac{K\left(\frac{x-y}{h}\right)}{h} = \delta_D(x-y)$$ Looks like Dirac Delta function

$$\lim_{h\to 0} \text{Bias}(\bar{\rho}(x)) = \rho(x) - \int dy\ \delta_D(x-y)\,\rho(y)$$

$$= \rho(x) - \rho(x) = 0$$

In this case the Kernel include only those values equates to $\underline{x}$ and therefore our Bias corresponds to $\rho(x)$ itself.

5 points

(B)

$$S \equiv \frac{x-y}{h} \rightarrow sh = x-y \Rightarrow y = x - sh$$

$$\boxed{dy = -h\,ds}$$

$$\rho(y) = \rho(x - sh) = \rho(x) - sh \left.\frac{d\rho}{dx}\right|_{h=0} + \frac{(sh)^2}{2!}\frac{d^2\rho}{dx^2} + \underbrace{O(sh^3)}_{\text{or } O(h^3)}$$

Therefore

$$\bar{\rho}(x) = \left\langle \frac{1}{h} K\left(\frac{x-y}{h}\right)\right\rangle_y = \int dy\ \frac{K\left(\frac{x-y}{h}\right)}{h} \times \left[\rho(x) - sh\left.\frac{d\rho}{dx}\right|_{h=0} + \frac{(sh)^2}{2!}\left.\frac{d^2\rho}{dx^2}\right|_{h=0} + O(h^3)\right]$$

$$\bar{\rho}(x) = \int (-h\,ds)\, \frac{K(s)}{h}\, \rho(x) + \int (-h\,ds)\, \frac{K(s)}{h}\, (-sh)\, \rho'$$

$$+ \int (-h\,ds)\, \frac{K(s)}{h}\, \frac{(-sh)^2}{2!}\, \rho'' + \mathcal{O}(h^3)$$

$$= -\rho(x) + h\rho' \int ds\, \underbrace{s}_{\text{odd}}\, \underbrace{K(s)}_{\text{even}} - \frac{h^2 \rho''}{2!} \underbrace{\int ds\, s^2\, K(s)}_{\int_{-1/2}^{+1/2} ds\, s^2 = \left.\frac{s^3}{3}\right|_{-1/2}^{+1/2} = \frac{1}{12}} + \mathcal{O}(h^3)$$

$$\bar{\rho}(x) = -\left[\rho(x) + \frac{h^2 \rho''}{2!}\, \frac{1}{12}\right]$$

$$\lim_{h\to 0} \text{Bias} = \rho(x) - (-\rho(x)) = 2\rho(x)$$

5-Points

(c) $$L \equiv \text{Var}(\bar{\rho}(x)) + (\text{Bias}[\bar{\rho}(x)])^2$$

$$= \sigma_m^2 + (\text{Bias}[\bar{\rho}(x)])^2$$

$$L = \frac{1}{N^2 h^2} + \left(\rho(x) - \left(-\rho(x) - \frac{h^2 \rho''}{24}\right)\right)^2$$

$$0 = \left.\frac{dL}{dh}\right|_{h=h_{Best}} = \checkmark$$

5-Points

4) A:

☆ Value $\begin{cases} \to \text{Coefficient} \\ \text{Function} \end{cases}$

☆ Type $\begin{cases} \to \text{Linear} \\ \text{Non-linear} \end{cases}$

☆ Relation $\begin{cases} \to \text{Auto-Correlation} \\ \to \text{Cross-Correlation} \end{cases}$

☆ Statistics $\begin{cases} \to \text{Two-Point} \\ \to \text{n-Point} \end{cases}$

☆ Kind $\begin{cases} \text{Weighted} \\ \text{Un-Weighted} \end{cases}$

5-Points

B) $\epsilon = \langle (y-(ax+b))^2 \rangle = \langle y^2 + a^2x^2 + b^2 + abx - 2axy - 2yb \rangle$

$$\begin{cases} \dfrac{\partial \epsilon}{\partial a} = 0 \to 2a\langle x^2 \rangle + 2b\langle x \rangle - 2\langle xy \rangle = 0 \\ \dfrac{\partial \epsilon}{\partial b} = 0 \to 2b + 2a\langle x \rangle - 2\langle y \rangle = 0 \end{cases}$$

$$\boxed{b = \langle y \rangle - a\langle x \rangle}$$

$$a = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2} \triangleq \text{Cov}$$

5-Points

5)

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\langle x\rangle)^2}{2\sigma^2}}$$

$$y = x - \langle x\rangle \rightarrow p(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}}$$

$$K_1 = M_1 = \langle y\rangle = \langle x - \langle x\rangle\rangle = \langle x\rangle - \langle x\rangle = 0$$

$$K_3 = M_3 - 3M_1M_2 + 2M_1^3$$
$$= M_3 = \langle yyy\rangle = \int dy \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} y^3 = 0$$

$$K_2 = M_2 - M_1^2 = M_2 - K_1^2 = M_2 = \sigma^2$$

5-points

6) Non-Linear Langevin Eq.

$$\frac{dv}{dt} = \underbrace{\gamma(t,v)}_{\text{Deterministic Part}} + \underbrace{\eta(t)}_{\text{Stochastic Part}}$$

$$\langle \eta(t) \rangle = 0$$

$$P(\eta(t)) = \mathcal{N}(0, \sigma_\eta)$$

$$\langle \eta(t)\, \eta(t') \rangle = \sigma_\eta^2\, \delta_D(t - t')$$

10 – Points

Computational part

A Use Random Generator to generate Data for each Point with Center $\bar{x}_i$ and $\bar{y}_i$

$$x_j^{(i)} = a + R(b - a) = (\bar{x}_i - \sigma_x^-) + R\left[\sigma_x^+ - \sigma_x^-\right]$$

$$y_j^{(i)} = (\bar{y}_i - \sigma_y^-) + R'\left[\sigma_y^+ - \sigma_y^-\right]$$

R and R' are Random Number with Flat Probability Distribution

B

$$Y_{theo} = ax + b$$

$$\chi^2 = \sum_{K=1}^{200} \left[y_K - (a x_K + b) \right]^2$$

$$\frac{\partial \chi^2}{\partial a} = 0 \quad \text{and} \quad \frac{\partial \chi^2}{\partial b} = 0$$

$$\boxed{a = \frac{\sum_{K=1}^{N} (x_K - \bar{x}) y_K}{\sum_{K=1}^{N} (x_K - \bar{x})^2}} \qquad \bar{x} = \frac{1}{N} \sum_{K=1}^{N} x_K$$

$$\boxed{b = \frac{1}{N} \sum_{K=1}^{N} y_K - a \frac{1}{N} \sum_{K=1}^{N} x_K}$$

$$a \simeq 2.37$$

$$b \simeq 2.47 \times 10^{-2}$$