STATISTICAL METHODS IN EXPERIMENTAL PHYSICS: FINAL EXAM

A. PASTORE

Instructions

The open-book exam needs to be completed by Friday the 31st of March at 2.30pm. This exam will count for 80% of the final mark.

- For each submission put either your student number or your name on top of your copy. Anonymous submission will count as 0.
- Respect the deadlines.
- For submission I do not accept Matlab codes (alone), I accept only printed files or pdf files you can send to me. The pdf file can be obtained from your Matlab code directly. Please add notes so I can follow your reasoning.
- Follow announcement on VLE.
- There is a box for submission in the main corridor for manual submission or send me a scan by email: alessandro.pastore@york.ac.uk. If you send me your copy in electronic format, please check that your name appears on the attached file and not only in the email. Please avoid submitting multiple files, try to merge them (especially if you send multiple .pdf files) in a unique file (when possible).
- The threshold material is indicated with (*) and it represents standard bookwork. This part of the exam <u>can</u> be done also without Matlab.
- Each section (question) counts as 25% of total mark. By answering threshold material only (assuming it *all* correct) it is sufficient to pass the exam (with minimum mark).
- To solve the 4 question completely, it should take you at least one full day of work, please consider this as an estimate so consider carefully how to organise your workload to avoid running out of time.
- If you do not understand the question, you have doubts, problem please do not hesitate to contact me by email. I will present in my office starting the 27th till the 31st so if you have doubts feel free to pass.

1. Probability distributions (25% total mark)

Exercise 1.1. Show that the probabilities p assigned by a Poisson $P(y;\lambda)$ probability distribution satisfy the requirement $0 \leq p \leq 1$ for all values of y and $\sum_{y=0}^{\infty} P(y;\lambda) = 1$

Exercise 1.2 [*]. Suppose that a team of quality-control inspectors visit a given production site. This event may happen Y_i=0,1,2,3,… time in a single day, but on average there is only $\bar{Y} = 1$ visit per day. Assume Y follows a Poisson distribution P(Y_i;$\bar{Y}$).

a) What is the probability of having no visits in 1 day?
b) What is the probability of having 1 visit in 1 day?
c) What is the probability of having 2 visit in 1 day?
d) What is the probability of having at least 1 visit in 1 day?

Exercise 1.3 [*]. The probability density function of x is given by

$$f(x) = \begin{cases} a + bx^2 & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

If the mean of $f(x)$ is $\mu = \frac{3}{5}$ find a, b [Hint: remember the PDF is normalised!]

2. Maximum Likelihood Estimator (MLE) (25% total mark)

Exercise 2.1 [*]. Find the MLE for the probability p of the Bernulli distribution

$$f(x_i; p) = p^{x_i}(1-p)^{1-x_i}$$

where $0 \leq x_i \leq 1$ for every i.

Exercise 2.2 [*]. Find the MLE for the mean μ of the log-normal distribution distribution

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}}$$

defined only for $x > 0$.

Exercise 2.3. The attached data set *exercise23.cvs* $\{t_i\}$ (in seconds) has been recorded with a maximum acceptance time T, which is unknown to us. The data is hypothesised to originate from an exponential distribution

a) Assume infinite acceptance. The distribution reads $P(t, \tau) = \frac{1}{\tau}e^{-t/\tau}$. Using MLE estimate τ and its error σ_τ.

b) Now assume there is an acceptance time T. Evaluate T from the data.

c) Define a procedure to find from the data the estimated value of τ (use for T the values you estimated before). $P(t, \tau) = \frac{1}{\tau}e^{-t/\tau}\frac{1}{1-e^{-T/\tau}}$

d) Give the value of τ as well as its error.

3. Error propagation (25% total mark)

Exercise 3.1 [*]. Suppose that our theoretical model connects the variable x that is measured by our apparatus to the quantity y as

$$y(x) = \frac{1}{\sqrt{4 - x^2}}$$

defined for $-2 < x < 2$. Suppose that the error on x is constant and given by σ_x. Determine the expression of the error for σ_y. Can you find the value of x which minimises such an error?

Exercise 3.2. The know resistor $R = 0.470k\Omega$ has an error of $\varepsilon_R = 5\%$. R and two resistor R_1, R_2 are connected in two circuits as shown in Fig.1. We measure the the currents $I_{R1} = 12.2mA$, $I_{RR1} = 25.8mA$, $I_{R2} = 72.4mA$ and $I_{RR2} = 76.3mA$. The ammeter has random errors of $\sigma_I = 0.5mA$ <u>and</u> systematic of $\varepsilon = 5\%$. Under the assumption that the intrinsic resistance of the ammeter is negligible, the values of the resistors can be calculated using Ohm's law. We have

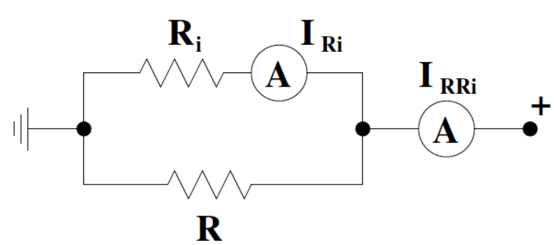

FIGURE 1. Circuit used to measure resistor R_i $i = 1, 2$.

$$R_i I_{Ri} = \frac{1}{\frac{1}{R_i} + \frac{1}{R}} I_{RRi} \qquad i = 1, 2$$

a) Derive expressions for R_1 and R_2

b) Give the 5×5 covariance matrix (error matrix) for the five observables $R, I_{R1}, I_{RR1}, I_{R2}, I_{RR2}$

c) Calculate the covariance matrix for the two derived resistances R_1, R_2

d) Based on your error analysis suggest a way to improve the experiment to improve the determination of R_1, R_2.

4. Non-linear Fit (25% total mark)

The attached file *exercise4.cvs* describe an exponential equilibration process for a physical observable

$$f(x) = a(1 - e^{-bx})$$

Each data point has a fixed uncertainty in y of $\sigma = 0.12$. The uncertainty on x is negligible.

a) Fit the data set $\{x, y\}$ to the theoretical function f(x) to find estimates of $\hat{a}$ and $\hat{b}$ as well the confidence interval of the parameters (90% for example, but please specify your choice)

b) Make a 3D plot of $\chi^2(a, b)$, determine a, b and evaluate and discuss the goodness-of-fit.

c) Calculate the covariance matrix V_{ab} for the parameters $\{a, b\}$.

d) Calculate the correlation matrix. Are a, b, correlated?

e) Without doing calculations, how do the parameters $\{a, b\}$ change if define a new error $\sigma = 0.012$?