

MA2261 Linear Statistical Models - DLI, Year 2022-2023

Coursework 1

INSTRUCTIONS AND DEADLINE:

Please submit *electronically* one piece of written/typed work per person in a single PDF file by **Friday 24/Mar/2023 at 4pm UK time/23:59 China time**.

Please use this page as the cover page for your submission. Write below your student ID and sign it.

Student ID:

Signature:

MARKING CRITERIA:

- This problem sheet is worth 100 points. Scores for each main question are indicated at the beginning of each.
- Clearly justify and explain your answers. If you are using the R software for calculations, a printout of your answers without a full explanation of the formulas you are using and your reasoning will not score full marks.
- A true/false question answered without justification will get zero marks.
- Computational mistakes will be penalized more in coursework than in exam marking, since you have plenty of time and tools to check your calculations when doing the coursework.

Question 1 [15 marks]

Decide if the following statements are true or false. Justify your answers.

- i) A power set must be a σ -algebra, vice versa.
- ii) There is a 50% chance of rain on Earth and a 10% chance of rain on Mars. Therefore, there is a 45% chance that it will rain in neither place.
- iii) Let $f(x)$ be a density function of a continuous random variable. Then it must have $\int_{-\infty}^{\infty} x f(x) dx = 1$.
- iv) The central limit theorem ensures that the sampling distribution of the sample mean approaches normal as the sample size increases.
- v) The chi-square distribution is left-skewed and can take on values between $-\infty$ and ∞ .

Question 2 [15 marks]

A factory production line is manufacturing bolts using three machines, A, B and C. Of the total output, machine A is responsible for 25%, machine B for 35% and machine C for the rest. It is known from previous experience with the machines that 5% of the output from machine A is defective, 4% from machine B and 2% from machine C. A bolt is chosen at random from the production line and found to be defective. What is the probability that it came from

- i) machine A?
- ii) machine B?
- iii) machine C?

Question 3 [15 marks]

Among 10000 random digits, find the probability P that the digit 3 appears:

- i) 9998 times.
- ii) at least 975 and no more than 1025 times.
- iii) at most 950 times.

Question 4 [25 marks]

A continuous random variable X has the following density function:

$$f_X(x) = \begin{cases} (C - \alpha)x^{1-\alpha}, & 0 \leq x \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

where α is an unknown parameter satisfying $\alpha < 2$ and $\alpha \neq 1$.

- i) Determine the value of the constant C .
- ii) Find $E[X]$ and $Var(X)$.
- iii) For independent observations x_1, \dots, x_n of X , write down the likelihood function, where the aim is to find an estimate for the parameter α .
- iv) Find the log-likelihood function.
- v) Find the maximum likelihood estimate for α in terms of $x_i, i = 1, \dots, n$.

Question 5 [20 marks]

Assume that the weight of a product follows a normal distribution with unknown mean and variance. Now we randomly selected 16 pieces from a batch of products. The sample average weight is 820g, and the sample standard deviation is 60g. Decide whether to reject $H_0 : \mu = 800\text{g}$ when the significance level is $\alpha = 0.05$.

Question 6 [10 marks]

Suppose that X_1, \dots, X_{10} are random variables, each an observation from the same distribution $X \sim N(\mu, 4)$, where μ is unknown. Suppose that the following values x_1, x_2, \dots, x_{10} are recorded:

12.4, 3.5, 9.8, 11.9, 16.5, 9.6, 10.9, 11.1, 2.9, 5.8

- i) Find a 95% confidence interval for μ .
- ii) Find a 99% confidence interval for μ .