Durham University

EXAMINATION PAPER

Examination Session:	Year:	Exam Code:
May/June	2020	MATH1617-WE01

Title:
Statistics I

Time (for guidance only):	2 hours	
Additional Material provided:		
Materials Permitted:		
Calculators Permitted:	Yes	Models Permitted: There is no restriction on the model of calculator which may be used.

Instructions to Candidates:	Credit will be given for your answers to all questions. All questions carry the same marks. Please start each question on a new page. Please write your CIS username at the top of each page. Show your working and explain your reasoning.

Revision:	

Q1 Let $X_1, \ldots, X_n$ be a random sample of size n from a population with unknown mean ϑ and known variance $\sigma^2 > 0$. Let $\overline{X}_n$ denote the sample mean. Throughout your answer, explicitly state any assumptions that you make.

1.1 Show that $(\overline{X}_n)^2$ is a biased estimator for Θ^2 by evaluating its bias.
[Hint: $\mathbb{V}\text{ar}\left(Y \mid \vartheta\right) = \mathbb{E}\left(Y^2 \mid \vartheta\right) - \left(\mathbb{E}\left(Y \mid \vartheta\right)\right)^2$.]

1.2 Let $\sigma = 1$, $n = 2$, $X_1 = 1$, and $X_2 = 3$. Provide an unbiased estimate for Θ^2.

Q2 Assume that $X \mid \vartheta$ has probability density function $f(x \mid \vartheta) = \vartheta e^{-\vartheta x}$. Suppose that we wish to test the hypotheses

$$H_0 \colon \vartheta \geq 1, \tag{1}$$
$$H_1 \colon \vartheta < 1. \tag{2}$$

Consider the test that rejects H_0 if $X \geq c$.

2.1 Show that $\mathbb{P}\left(X \geq c \mid \vartheta\right)$ is a decreasing function of ϑ.

2.2 For what value of c does the test have size α_0 (where $0 < \alpha_0 < 1$)?

2.3 State the definition of a p-value, and compute the p-value of the test when we observe $X = 2$.

2.4 Again when $X = 2$, would you reject H_0 at the 0.025 level?

Q3 An operating system for a personal computer has been studied extensively, and it is known that the standard deviation of the response time following a particular command is $\sigma = 7$ milliseconds.

3.1 A new version of the operating system is installed, and we wish to estimate the mean response time Θ for the new system to ensure that a 95% confidence interval for Θ has a length of at most 5 milliseconds. What sample size would you recommend? Explicitly state any assumptions that you make.

3.2 A random sample of 10 response times $X_1, \ldots, X_{10}$ has a sample mean of 52 milliseconds. Construct a 95% frequentist prediction interval for the next response time X_{11}, and provide a statistical interpretation of this interval. Explicitly state any additional assumptions that you make.

3.3 If $[a, b]$ is the interval that you obtained in part **3.2**, explain why

$$\mathbb{P}\left(a \leq X_{11} \leq b \mid x_1 \ldots x_{10}\right) \tag{3}$$

may not be equal to 0.95.

3.4 How do we call an interval $[c, d]$ such that $\mathbb{P}\left(c \leq X_{11} \leq d \mid x_1 \ldots x_{10}\right) = 0.95$?

Q4 Suppose that $X_1, \ldots, X_n$ form a random sample from the following distribution with parameter ϑ, where $\vartheta \in [0,1]$ is unknown:

$$\mathbb{P}\left(X_i = x_i \mid \vartheta\right) = \begin{cases} \frac{3}{4}\vartheta & \text{if } x_i = 0 \\ \frac{1}{4}\vartheta & \text{if } x_i = 1 \\ 1-\vartheta & \text{if } x_i = 2 \end{cases} \tag{4}$$

4.1 Consider the estimator $\hat{T} \coloneqq a + b\sum_{i=1}^n X_i$. Find values for a and b such that $\hat{T}$ is an unbiased estimator for Θ.

4.2 Find the Bayes estimator for Θ under squared error loss, assuming that Θ follows a uniform distribution on $[0,1]$.

You may find it useful to recall that, for any $\alpha > 0$ and $\beta > 0$, we have that $Z \sim \text{Beta}(\alpha, \beta)$ has probability density function $f(z) \overset{(z)}{\propto} z^{\alpha-1}(1-z)^{\beta-1}$ and expectation $\mathbb{E}\left(Z\right) = \frac{\alpha}{\alpha+\beta}$.

4.3 What is the maximal possible bias of the Bayes estimator for Θ that you derived in part **4.2**?

4.4 Calculate the estimate $\hat{t}$ and also the Bayes estimate when $n = 5$ and $X_1 = X_2 = X_4 = X_5 = 0$ and $X_3 = 1$, and with the values of a and b that you derived in part **4.1**. Which of the two estimates, $\hat{t}$ or Bayes, would you rather use? Explain in one sentence.

Q5 Suppose that we model bus waiting times $X_1, X_2, \ldots, X_n$ as $X_i \mid \vartheta \sim \text{U}(0, \vartheta)$ (the uniform distribution on $[0, \vartheta]$) and where the X_i are i.i.d. conditional on Θ. Assume that Θ has the following density function, called the Pareto density, with hyperparameters $\alpha_0 > 0$ and $\beta_0 > 0$:

$$f(\vartheta) \coloneqq \begin{cases} \alpha_0 \beta_0^{\alpha_0} \vartheta^{-\alpha_0 - 1} & \text{if } \vartheta \geq \beta_0, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

We observe a random sample of size n: $X_1 = x_1, \ldots, X_n = x_n$.

5.1 What is the maximum likelihood estimate of Θ?

5.2 Show that the posterior density is again a Pareto density, but with different parameters α_n and β_n. Determine α_n and β_n as a function of the data.

5.3 Let now $\alpha_0 = 1$ and $\beta_0 = 5$. A random sample of size 2 is observed with $x_1 = 2$ and $x_2 = 10$. If the value of Θ is to be estimated by using the squared error loss function, what is the Bayes estimate for Θ? Derive its numerical value in this case. Compare with the numerical value of the maximum likelihood estimate and briefly discuss in one sentence which of these two estimates you prefer, and why.

5.4 Again with $\alpha_0 = 1$, $\beta_0 = 5$, $n = 2$, $x_1 = 2$, and $x_2 = 10$, derive the probability density function of the posterior predictive distribution, as a function of x_3, up to a normalisation constant.