UNIVERSITY OF CAMBRIDGE

ECT1
ECONOMICS TRIPOS PART 1

Tuesday 7 June 2022 9:00am – 12:00pm

Paper 3

QUANTITATIVE METHODS IN ECONOMICS

This paper is divided into four sections.
Section A answer **four** compulsory questions.
Section B answer **one** question out of two questions.
Section C answer **four** compulsory questions.
Section D answer **one** out of two questions.

Answers from Section A and B (Mathematics) and from Section C and D (Statistics) must be completed on separate pages with the letter of the Section written at the top of the page.

Section A and C will each carry 30% of the marks for this paper. Section B and D each carry 20% of the marks.

Each question within each section carries equal weight.

Write your **Blind Grade Number** (not your name) on your answers.

Candidates are asked to note that there may be a reduction in marks for scripts with illegible handwriting.

If you identify an error in this paper, please alert the **Invigilator**, who will notify the **Examiner**. A **general** announcement will be made if the error is validated.

STATIONERY REQUIREMENTS
20 Page booklet x 2
Rough work pads
Tags

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAMINATION
Calculator – students are permitted to bring an approved calculator.
New Cambridge Elementary Statistical Tables

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator.

SECTION A - Answer **ALL FOUR** questions from this Section.

1. Consider the function

$$f(x) = \begin{cases} x \ln 2 + 1 & \text{for } x < 0 \\ 2^x & \text{for } x \geq 0 \end{cases}$$

(a) Find all x, where $f(x)$ is continuous.

(b) Find all x, where $f(x)$ is differentiable, and report the corresponding derivative.

2. Consider the subset S of $\mathbb{R}$ that consists of the union of the intervals $[2^{-2i-1}, 2^{-2i}]$ over all non-negative integers $i = 0, 1, 2, ...$

(a) Argue that S is bounded, but not compact.

(b) Let $g(x)$ be 1 if $x \in S$ and zero otherwise. Find the value of $\int_0^1 g(x)dx$.

3. Let $h(x, y) = -x^2 + xy - y^2$

(a) Show that this function is strictly concave.

(b) Find the maximum of $h(x, y)$ subject to the constraint $x^2 + y^2 = 2$.

4. Let $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

(a) What geometric linear transformation does A represent?

(b) Find A^{2022}.

SECTION B - Answer **ONE** question from this Section.

5. A person called Flat Stanley lives in a square

$$S = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$$

on a two-dimensional Euclidean space $\mathbb{R}^2$. His business is transporting goods produced at location $(0, 0)$ to the shore of a lake. The lake is represented by

$$L = S \cap \{(x, y) : xy - x/2 \geq a\},$$

where $0 < a < 1/2$. The cost of transporting goods from $(0, 0)$ to a point (x, y) is $C(x, y) = x + y - xy$.

(a) Show that the cost function is neither concave, nor convex.

(b) Suppose that Stanley would like to choose a point (x^*, y^*) on the shore where $C(x, y)$ is minimized. Write down the Lagrangian for Stanley's optimization problem. Derive the corresponding first order conditions and solve them, thus finding the candidate solution point (x^*, y^*).

(c) Note that on the shore $y = 1/2 + a/x$. Substitute this into $x + y - xy$ to obtain the cost of delivery on the shore as a function of x only. Using the function that you obtained, verify that the candidate solution point (x^*, y^*) found in (b) indeed delivers the minimum.

(d) Suppose that initially $a = 1/8$. However, due to climate change, a decreases by 0.01. Using the Envelope Theorem, approximate the corresponding change in optimal cost.

6. Consider the differential equation describing growth of capital in the economy:

$$I(t) = \frac{d}{dt}K(t) + \delta K(t),$$

where $I(t)$ is investment at time t, $K(t)$ is the capital stock at time t, and δ is a positive depreciation rate.

(a) Suppose that $\delta = 0.1$ and t is measured in years. How long will it take for the capital stock to decrease by 50% in the absence of investment?

(b) Assume now that the depreciation rate is itself a function of time, so that $\delta = \delta(t) = 0.1e^{-t}$. Further, let $K(0) = 100$. Find an explicit expression for $K(t)$ in the absence of investment.

(c) Under the assumptions of (b), find the second order Taylor approximation to $K(t)$ at $t = 0$. Using the Taylor theorem, find an upper bound on the error of this approximation at $t = 1$.

(d) Suppose now that δ is fixed at 0.1, $K(0) = 100$, and $I(t) = t$. Find the value of the capital stock at $t = 1$.

SECTION C - Answer **ALL FOUR** questions from this Section.

7. (a) For a birthday, I got a weighted coin. I was told that the probability of heads, $P(heads)$, was either 0.2 or 0.8 but I cannot recall which one. To test if the probability is 0.8 I toss the coin 5 times and reject this hypothesis if I see 0, 1 or 2 heads. What is the size and power of this test?

   (b) Over a 52 week period I record each week my household consumption of electricity (in kWh), $X_i, \quad i = 1, \ldots, 52$ and obtain the following data

$$\sum_{i=1}^{52} X_i = 6224.2 \qquad \text{and} \qquad \sum_{i=1}^{52} X_i^2 = 857719.7$$

   Stating clearly any assumptions you make, calculate (to 2dp) a 95% confidence interval for my weekly consumption of electricity.

8. The half-normal random variable X has density function $f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2}$ for $0 \leq x < \infty$.

   (a) Show that for the half-normal variable $E(X) = \sqrt{2/\pi}$.

   (b) The standard normal random variable Z has density function $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ and has mean 0 and variance 1. For the standard normal variable show that $E(Z^2)$ is also 1 and hence show that for the half-normal variable

$$Var(X) = 1 - 2/\pi$$

9. The level of output in Oceania Y_t is given by

$$\ln(Y_t) = gt + \varepsilon_t \quad t = 1, 2, \ldots$$

where $g > 0$ is a fixed number, $t = 1, 2, \ldots$ measures years, and $\varepsilon_t \sim N(0, \sigma^2)$ is an i.i.d. random variable.

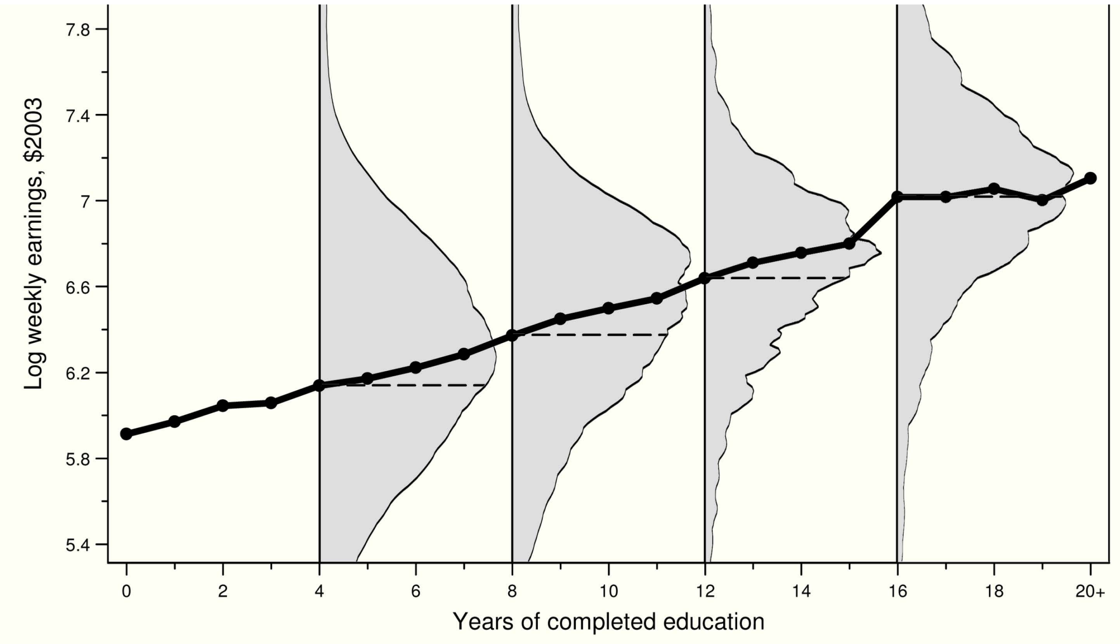
   (a) One measure of growth rate of the level of output is $\ln(Y_{t+1}/Y_t)$. Show that the expected value of this measure is g. What is the variance of this measure?

A statistician has data on the level of output in Oceania over the period $t = 1, \ldots, 21$ and wishes to estimate g. Two estimators are proposed

$$\hat{g} = \frac{\ln(Y_{21}) - \ln(Y_1)}{20} \qquad \text{and} \qquad \tilde{g} = \frac{\sum_{t=1}^{t=21} t \ln(Y_t)}{\sum_{t=1}^{t=21} t^2}$$

   (b) Calculate the bias and mean squared error of each estimator. Which would you prefer to use? You may use without proof that $\sum_{t=1}^{t=21} t^2 = 3311$.

10. The following figure plots the sample conditional expectation function (CEF) of log weekly earnings ($\log W_i$) given education (Ed_i) for a sample, $i = 1, ..., N$, of middle-aged white men from the 1999 US Census. Years of education is a categorical variable taking values from 0-20. The distribution of earnings is also plotted for values: 4, 8, 12, and 16 years of education.



An analyst chooses to represent the relationship between log weekly earnings and education using the following OLS regression model

$$\log W_i = \alpha + \beta Ed_i + u_i.$$

(a) From the figure, there is some evidence that the conditional distribution of u_i given Ed_i is not invariant to years of education. In what sense does this compromise any inference conducted on the OLS estimator $\hat{\beta}$?

(b) We wish to investigate potential differences in the returns to education for males and females. The following equations present separate linear regression equations for males (m) and females (f).

$$\begin{aligned} \log W_i &= \alpha_m + \beta_m Ed_i + u_i \text{ for males} \\ \log W_i &= \alpha_f + \beta_f Ed_i + u_i \text{ for females.} \end{aligned}$$

(i) Using these equations, derive a single linear regression model for males and females.

(ii) Explain how you would use the solution to (i) to test the hypothesis that the returns to education for males and females are equal.

SECTION D - Answer **ONE** question from this Section.

11 A Bernoulli random variable takes on the discrete values 0 and 1 and encodes a variety of binary outcome situations (for example the outcomes "NO" and "YES" can be coded as 0 and 1, respectively). We describe the probability mass function of a Bernoulli random variable as follows

$$\Pr(X = x) = \begin{cases} 1-p & \text{for } x = 0 \\ p & \text{for } x = 1 \end{cases},$$

where Pr denotes probability and p is a parameter with $0 \leq p \leq 1$. For a Bernoulli random variable, $E(X) = p$ and $Var(X) = p(1-p)$.

(a) Given an i.i.d. sample $X_i, \quad i = 1, \ldots, N$, an estimator of p is $\hat{p} = \frac{1}{N}\sum_{i=1}^{N} X_i$. Show that $\hat{p}$ is unbiased and calculate its mean squared error.

At a game reserve the big treat is to see the lions drinking at the water hole in the evening. Whether the lions drink or not can be modelled as a Bernoulli random variable so $\Pr(drink) = p$ and $\Pr(no\ drink) = 1 - p$, where p may depend on the weather. The management record the temperature, T, and humidity, H, at lunchtime over a two week period and also whether the lions drank that evening and obtain the following results.

Day	T	H	Lions Drink
1	hot	high	no
2	hot	high	no
3	hot	high	yes
4	mild	high	yes
5	cool	normal	yes
6	cool	normal	no
7	cool	normal	yes
8	mild	high	no
9	cool	normal	yes
10	mild	normal	yes
11	mild	normal	yes
12	mild	high	yes
13	hot	normal	yes
14	mild	high	no

At lunch on day 15 the temperature is cool and the humidity is high.
To make a recommendation the management initially decide to use only the information on humidity and whether the lions drink. They will recommend guests go to the water hole if they estimate the probability that lions will drink exceeds their estimate that they won't.

(b) Use the data in the Table and the estimator discussed in part (a) to estimate the following probabilities:
(i) $\Pr(drink)$,
(ii) $\Pr(no\ drink)$,
(iii) $\Pr(H = high|drink)$,
(iv) $\Pr(H = high|no\ drink)$
and hence use Bayes' Theorem to calculate whether $\Pr(drink|H = high)$ is larger or smaller than $\Pr(no\ drink|H = high)$. What will the management recommend?

To improve prediction they decide to use both temperature and humidity data, so now want to compare $\Pr(drink|T = cool, H = high)$ and $\Pr(no\ drink|T = cool, H = high)$. A guest points out that there are zero observations in the cool-high category so we cannot estimate a non-zero probability within this category. The guest suggests as an alternative to estimate

$$\Pr(T = cool, H = high|A) = \Pr(T = cool|A)\Pr(H = high|A),$$

where A may take two values: $A = drink$ and $A = no\ drink$.

(c) What assumption does this suggestion make?

(d) Use this assumption to determine whether $\Pr(drink|T = cool, H = high)$ or $\Pr(no\ drink|T = cool, H = high)$ is larger. With this extra information what will the management recommend?

.

12. A researcher has a sample of 50 individuals with similar education, but differing amounts of training. She hypothesizes that hourly earnings, $Earn$, may be related to hours of training, $Train$, according to the relationship:

$$Earn_i = \beta_1 + \beta_2 Train_i + u_i, \ \ i = 1, ..., 50.$$

(a) An analyst is prepared to test the null hypothesis $H_0 : \beta_2 = 0$ against the alternative hypothesis $H_1 : \beta_2 \neq 0$ at the 5%, 1% and 0.1% levels. What should she report if:
(i) $\hat{\beta}_2 = 0.3$ and $s.e.\hat{\beta}_2 = 0.12$;
(ii) $\hat{\beta}_2 = 0.55$ and $s.e.\hat{\beta}_2 = 0.12$;
(iii) $\hat{\beta}_2 = 0.1$ and $s.e.\hat{\beta}_2 = 0.12$.
Here $\hat{\beta}_2$ denotes the OLS estimator of β_2, and $s.e.\hat{\beta}_2$ denotes the standard error of $\hat{\beta}_2$.

(b) (i) Explain why in this instance it would have been preferable to perform one-sided tests instead of two-sided tests.
(ii) Perform them and state whether the use of one-sided tests makes any difference.

(c) An estimate $\hat{\beta}$ and a hypothetical value β are said to be incompatible if

$$\frac{\hat{\beta} - \beta}{s.e.\hat{\beta}} > t_{crit} \text{ or } \frac{\hat{\beta} - \beta}{s.e.\hat{\beta}} < -t_{crit}$$

such that the null hypothesis is rejected against a two-sided alternative. Here t_{crit} denotes a critical value from the t distribution. Given these inequalities, find an expression for the confidence interval for β, namely the set of hypothetical values β that are compatible with the estimate $\hat{\beta}$.

(d) In the example above, for $\hat{\beta}_2 = 0.55$ and $s.e.\hat{\beta}_2 = 0.12$, calculate a 95% confidence interval. What do you conclude from this calculation?

.

END OF PAPER