



# EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

## HIGHER CERTIFICATE IN STATISTICS, 2015

### MODULE 1 : Data collection and interpretation

**Time allowed: One and a half hours**

*Candidates should answer **THREE** questions.*

*Each question carries 20 marks.*

*The number of marks allotted for each part-question is shown in brackets.*

*Graph paper and Official tables are provided.*

*Candidates may use calculators in accordance with the regulations published in the Society's "Guide to Examinations" (document Ex1).*

*The notation  $\log$  denotes logarithm to base  $e$ .*

*Logarithms to any other base are explicitly identified, e.g.  $\log_{10}$ .*

*Note also that  $\binom{n}{r}$  is the same as  ${}^n C_r$ .*

This examination paper consists of 8 printed pages.

This front cover is page 1.

Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. In 2011, the European Commission published a report on transport in the (then) 27 European Union countries. The table below, taken from the report, summarises the answers given to the question 'What is the main mode of transport you use for your daily activities?'. Answers were obtained from about 1000 people in each of 24 countries, and about 500 in each of Cyprus, Luxembourg and Malta. In the last column, 'DK/NA' denotes 'don't know or not applicable'.

The total numbers of responses under the headings 'AGE', 'EDUCATION', 'URBANISATION' and 'OCCUPATION' are slightly lower than the overall total of responses. You may assume that this arises because some responses were incomplete.

- (i) Write a report summarising the most important aspects of the data. You should incorporate appropriate diagrams into your report. (16)
- (ii) Explain briefly how you decided which information to focus on. (4)

	Total number of respondents	% Car	% Public transport	% Walking	% Cycling	% Motorbike	% Other	% No daily/regular mobility	%DK/NA
EU27	25570	52.9	21.8	12.6	7.4	2.1	1.4	1.6	0.2
SEX									
Male	12363	58.9	18.1	8.7	7.4	3.7	1.7	1.3	0.2
Female	13207	47.3	25.3	16.2	7.4	0.6	1.1	1.9	0.2
AGE									
15–24	3488	32.9	41	10.8	8.1	5.1	1.5	0.4	0.1
25–39	5764	60.8	18.7	10	6.6	2.5	1	0.1	0.2
40–54	7310	63.5	15.4	9.8	7.6	2.1	1.1	0.4	0.2
55+	8812	47	21.5	17.2	7.6	0.6	1.8	4	0.3
EDUCATION (end of)									
Until 15 years of age	4216	43.4	21.6	19.2	7.3	1.8	2.4	4.1	0.2
16–20	11080	57.2	18.9	12.1	7.3	2	1.2	1.2	0.1
20+	7154	62.4	17.9	9.2	7.1	1.6	0.9	0.6	0.3
Still in education	2443	26.6	46.5	10.9	9.2	4.7	1.4	0.3	0.3
URBANISATION									
Metropolitan	4679	43.1	37.1	9.7	5.6	2.6	0.8	0.9	0.3
Urban	11196	47.7	22.9	16.1	7.8	2.2	1.5	1.7	0.2
Rural	9602	63.7	13.1	9.7	8	1.7	1.6	1.8	0.3
OCCUPATION									
Self-employed	2384	70.6	10.5	6.6	5.8	2.3	3.8	0.2	0.2
Employee	8841	65.6	16.2	8	6.8	2.3	0.9	0.2	0.1
Manual worker	2217	56.8	19.5	11.9	6.2	3.9	1	0.7	0
Not working	12054	39.4	28.6	17.2	8.4	1.6	1.4	3.1	0.3

2. 89 university students all taking the same course were monitored for their use of online lecture materials. The students were classified according to the mode in which they learned best: Visual, Auditory, Read/Write, Kinaesthetic, Multimodal.

The table below shows, for various categories of student, the mean number of online lecture pages viewed and the mean examination scores. The corresponding standard deviations are also given.

Learner type	Online Lecture Pages Viewed									Examination Scores					
	Numbers			All		Male, M		Female, F		All		Male, M		Female, F	
	All	M	F	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Visual	3	1	2	4247	2143	6469	0	3135	1329	78.1	3.1	78.7	0.0	77.8	0.4
Auditory	9	5	4	2815	750	2780	698	2849	856	85.2	6.9	<i>a</i>	<i>b</i>	82.3	1.0
Read/Write	19	5	14	2941	1408	3180	1778	2854	1317	84.9	7.4	87.2	6.5	84.1	7.1
Kinaesthetic	18	10	8	2320	696	2336	778	2299	631	83.1	10.6	85.7	1.3	79.9	0.6
Multimodal	40	16	24	2492	993	2899	1064	2221	862	86.4	6.3	88.4	6.0	85.0	6.4
All	89	37	52	<i>c</i>	1113	2865	1217	2487	1024	85.0	7.5	87.1	8.3	83.5	6.8

- (i) For the 5 males who are auditory learners, the sum of their examination marks was 436; the sum of the squares of their marks was 38092. Calculate the values of entries *a* and *b* in the table. (3)
- (ii) Calculate the value of the entry *c* in the table. (2)
- (iii) Explain why two of the standard deviations in the table are shown as zero. (1)
- (iv) A lecturer in the university department is interested in whether or not there are systematic differences between
- males and females in examination scores,
  - learner types in the number of online pages viewed.
- Without carrying out any formal statistical tests**, discuss the strength of the evidence for such differences. (8)
- (v) The lecturer notes that the learner types with the highest mean numbers of pages viewed are those with the lowest mean examination scores. He therefore claims that online lectures are an ineffective way for students to learn. Discuss the strengths and weaknesses of this claim in the light of the information in the table. (You may wish to include sketch diagram(s) in your answer.) (6)

3. Quarries are potentially dangerous working environments. Typical dangers include trips and falls, incorrect use of equipment, and accidents involving vehicles.

The workers in a quarry, excluding managers and supervisors, are to be surveyed on their views about the health and safety of their working practices. The questionnaire below is a preliminary design.

<i>Please tick the appropriate box to show your level of agreement with each of the following statements.</i>	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1 Some rules and procedures do not need to be followed to get the job done safely					
2 Management rarely act quickly over our health and safety concerns					
3 I have often seen people taking risks here					
4 Managers and Supervisors would consider people who always keep strictly to the health and safety rules to be over cautious					
5 I see people taking risks here that I would not take myself					
6 There are good communications on this site about issues which affect me					
7 I can trust most people who I work with to work safely					
8 My manager doesn't take the safety committee or safety representatives seriously					
9 Supervisors seldom check that I am working safely					
10 Supervisors often ignore people who are not working to the health and safety rules and procedures					
11 Most of the workforce do not pay much attention to health and safety					
12 My health and safety training over emphasised the health and safety risks of my job					
13 People here do not work well in teams					

- (i) Comment on the design of the questionnaire. Identify strong and weak points in its overall structure. Indicate which questions you consider to be good and which bad. In the latter case give improved versions where appropriate. (10)
- (ii) Assume now that your modified version of the questionnaire is to be used. Write a suitable brief introduction to the questionnaire. (6)
- (iii) Past practice has been for workers to post their completed questionnaires, anonymously, in boxes provided on site. The completion rate has typically been about 50%. Explain why this completion rate might be a cause for concern. What steps could be taken to improve it? (4)

4. For each of the following sampling methods,
- explain briefly how the method operates,
  - describe a realistic situation in which the method would be used,
  - identify the advantages and disadvantages associated with the method.
- (i) Simple random sampling. (4)
- (ii) Systematic sampling. (4)
- (iii) Quota sampling. (4)
- (iv) Cluster sampling. (4)
- (v) Stratified random sampling. (4)

BLANK PAGE

BLANK PAGE

BLANK PAGE