

# EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY



## HIGHER CERTIFICATE IN STATISTICS, 2009

### 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  ${}^nC_r$ .*

This examination paper consists of 3 printed pages **each printed on one side only**.

This front cover is page 1.

Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. (i) Explain what kinds of information can be obtained better from an open-ended question in a questionnaire, rather than from a question in closed form. (5)
- (ii) Give three reasons for asking a question in closed form rather than in open-ended form in order to obtain information. (3)
- (iii) Teenagers are to be asked to complete a questionnaire which includes a short section on their use of mobile phones. Topics of interest are: who pays for their calls, how often they use a mobile phone, how long their calls usually last, their reasons for making calls, the approximate proportion of their calls that are received as opposed to sent, and on what percentage of occasions that they use the phone they use it to send a text message (as opposed to making a call). Draft one question on each of these six topics, with three questions in closed form and three in open-ended form, saying which questions are of which form. (12)
2. (a) An experiment was performed in which 228 students were asked to produce what they believed to be a sequence of 25 random digits. The frequencies with which each digit was chosen in the experiment were as follows.

<i>Digit</i>	0	1	2	3	4	5	6	7	8	9
<i>Frequency</i>	375	637	676	709	607	522	534	573	515	552

- (i) Comment briefly on whether the group of students as a whole appears to have chosen digits at random. (You are not expected to perform any tests.) (5)
- (ii) Calculate the sample mean and variance of the digits generated by the students. (5)
- (iii) For a population of 5700 ( $= 228 \times 25$ ) digits with each digit occurring equally often, the mean is 4.5 and the variance is 8.25. Compare the sample mean and sample variance calculated in part (ii) with these values and comment. (3)
- (iv) State another property that you would look for in deciding whether a sequence of digits had been randomly generated. (2)
- (b) A simple random sample is to be taken from a population consisting of 350 elements. Explain how to allocate random numbers to enable this to be done. (5)

3. (a) The information below is derived from the National Diet and Nutrition Survey of British Adults, 2000–2001. It relates to a random sample of adults aged 19–64 in private households in mainland Britain.

<b>Age (years)</b>	<b>Number of men</b>	<b>Number of women</b>
19–24	61	78
25–34	160	211
35–54	393	486
55–64	<u>152</u>	<u>183</u>
	<u>766</u>	<u>958</u>
<b>Region</b>		
Scotland and the North	248	326
Central, South West and Wales	274	350
London and the South East	244	282
<b>Current smoker</b>	236	313

Write a report comparing the representation of men and women within the various categories in this achieved sample. Your report should include diagrams and summary statistics. You should hand in any calculations needed to produce the diagrams and summary statistics, making clear that these calculations are not part of the report.

(12)

- (b) A diary has been devised in which respondents are to record information about their diet over a period of one week. This diary contains, for each day, columns in which the respondent can enter the times that drink or food was consumed, the items consumed and the approximate quantities consumed. State four advantages and four disadvantages of asking respondents to complete a diary of this nature.

(8)

4. A survey is to be undertaken to investigate the shopping habits of people living on a housing estate at the edge of a town of moderate size.

- (i) Suggest a suitable sampling scheme.

(4)

- (ii) Design a set of eight to ten questions as part of a questionnaire to obtain information about shopping for food. State whether the questions are for self-completion or are to be used in an interview.

(11)

- (iii) Outline the kinds of tabulations you would plan to obtain from the survey. You may assume that appropriate background information about the respondents is available and that questions other than those you have written in (ii) are in the questionnaire.

(5)