EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2014

MODULE 3 : Basic statistical methods

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.	A textile manufacturer produces long rolls of wide cotton cloth for making bed linen. The spun cotton thread he uses has occasional irregularities which produce flaws in the woven fabric. As part of a quality control exercise, a random sample of 180 ten-metre lengths of the cloth is examined carefully and the number of flaws in each length is noted. The results of this exercise are shown in the table below.

Number of flaws per ten-metre length, x	Frequency, f
0	30
1	58
2	49
3	29
4	7
5	7
>5	0

(i)	Calculate the sample mean and sample variance, showing all working clearly. Comment on the sample mean and variance values in relation to the Poisson distribution.

(6)

(ii)	Perform a χ^2 goodness-of-fit test at the 5% significance level to investigate the null hypothesis that the number of flaws in ten-metre lengths of cloth has a Poisson distribution. Show all your working and report your conclusions.

(14)

2.   (i)   A pharmaceutical company is considering introducing new, easier to open, packaging for a drug used in the treatment of arthritis. The company seeks the views of two groups of patients. One group consists of those who have been using the existing packaging for a long time and the other group consists of new users. The preferences are shown below.

	Prefer new packaging	Prefer existing packaging
Long-term users	35	32
Recent users	25	8

Investigate, at the 5% significance level, whether there is evidence of a difference between the two groups of patients in their packaging preferences and briefly state your conclusions.

(9)

(ii)   After a six month trial period, the group of 67 long-term users is asked again which packaging they prefer. The initial preferences for these patients together with their preferences after the trial period are shown below.

		Preference after trial period	
		New packaging	Existing packaging
Initial preference	New packaging	28	7
	Existing packaging	14	18

State what statistical test you would use to examine whether the preferences for the new packaging for these long-term users have changed after the six month trial period. Briefly outline the circumstances when this test should be used, indicating how they apply here. Carry out the test, at the 5% significance level, stating your null and alternative hypotheses clearly and reporting your conclusions.

(11)

3. A sample of 15 people in various industries was taken in 2008, and the people were asked to rate the immediate business prospects for their companies on each of several factors. The ratings from each person were combined into an overall 'business optimism' score for that person, where higher scores indicate greater optimism. These people were contacted again in 2012 and asked to do the same thing. The scores are shown below.

Person	2008 score	2012 score
1	66.1	66.9
2	64.5	60.7
3	67.1	64.4
4	65.6	62.6
5	64.7	65.3
6	68.3	65.5
7	66.2	64.3
8	64.1	65.2
9	59.8	61.4
10	62.5	59.9
11	58.3	56.3
12	61.9	57.8
13	68.5	64.2
14	66.5	66.4
15	65.6	63.2

(i)   Perform a Wilcoxon signed-rank test on these data using a 5% significance level. State your conclusions clearly.

(10)

(ii) Now carry out a sign test using a 5% significance level and state your conclusions clearly.

(7)

(iii) Compare the conclusions reached in the two tests above, explaining any similarities or differences.

(3)

4. (i) Explain the meaning of the following terms which are used in hypothesis testing.

    (a) Type I error.

    (b) Significance level.

    (c) Type II error.

    (6)

(ii) A car manufacturer states that the fuel consumption of its newest model in typical urban driving conditions is Normally distributed with a mean of 35 miles per gallon and a standard deviation of 1.8 miles per gallon. A consumer group tests a sample of 9 cars under urban driving conditions and aims to challenge the manufacturer's statement if the mean fuel consumption from this sample is 34 miles per gallon or less.

    (a) What level of significance is the consumer group using in its test?

    (7)

    (b) If the mean fuel consumption in urban driving conditions for the new model is actually 32.5 miles per gallon, what is the probability that the consumer group challenges the manufacturer's statement? (You may assume that the standard deviation remains unchanged.)

    (5)

    (c) What is the name of the quantity you have calculated in part (b)?

    (2)

5

BLANK PAGE

BLANK PAGE

BLANK PAGE