BINDURA UNIVERSITY OF SCIENCE EDUCATION

DIPLOMA IN SCIENCE EDUCATION

MT006: Statistics 11/

DM006: Inferential statistics

Time: 2 hours

Candidates may attempt ALL questions in Section A and at most TWO questions in Section B. Each question should start on a fresh page.

SECTION A (40 marks)

Candidates may attempt ALL questions being careful to number them A1 to A4.

A1. Define the following terms

(a) Inferential statistics [2]
(b) Estimator [2]
(c) Type 1 error [2]
(d) Type 11 error [2]
(e) Unbiased estimator [2]

A2. (a)Explain three factors that affect the power of a test. [6]

(b) State two estimation methods.[2]

(c) Explain three factors that affect the interval width. [3]

A3. State three advantages and two disadvantages of sampling.[5]

A4. A random sample of 25 children with the mean IQ = 100, and the standard deviation 15 was selected by an IQ expect from the general population. It is known that IQ is approximately normally distributed.

(a) What is the probability of getting a sample mean of 108 or higher? [3]
(b) What is the probability of getting a sample mean of 92 or lower? [3]
(c) How high would the sample mean have to be for you to say that the probability of getting a mean that high (or higher) was 0.05 (or 5%)? [4]
(d) How low would the sample mean have to be for you to say that the probability of getting a mean that low (or lower) was 0.05 (or 5%)? [4]

SECTION B(60 marks)

Candidates may attempt TWO questions being careful to number them B5 to B7

B5.(a) When a production machine is properly calibrated, it requires an average of 25 seconds per unit produced, with a standard deviation of 3 seconds. For a simple random sample of $n = 36$ units, the sample mean is found to be 26.2 seconds per unit. When the machine is properly calibrated, what is the probability that the mean for a simple random sample of this size will be at least 26.2 seconds? [5]

(b) For many years, the mean gas mileage on a long trip was known to be 26.5 miles per gallon for a certain type of automobile. Each automobile in a simple random sample of those with the new design is driven 100 miles and the following gas mileages are recorded:

26.2 26.5 27.2 27.3 27.1 27.6 25.9 27.5 26.8.

(i) When a newly designed engine is incorporated into the automobile, is there is any evidence that the mean gas mileage with the new design is different from 26.5 miles per gallon at 0.05 significance level. [10]

(ii) Why would a box plot be an appropriate graphical display for the data used in this hypothesis test? [3]

(iii) Construct the box plot, and comment on whether there is any reason to think the t statistic is not appropriate.[3]

(iv) Considering the results of the hypothesis test, decide which of the Type I or Type II errors is possible,and describe this error. [3]

(v) Decide whether H_o would have been rejected or would not have been rejected with each of the following significance levels: $\alpha = 0.01$ and $\alpha = 0.10$.[6]

B6.(a) The campaign manager for a political candidate claims that 55% of registered voters favor the candidate over her strongest opponent. Assuming that this claim is true, what is the probability that in a simple random sample of 300 voters, at least 60% would favor the candidate over her strongest opponent? [6]

b)Before introducing a new rule, the librarian ofa university library decided to find out how members might react to this rule.

(i) Explain why the librarian decided to take a random sample of members rather thanask all the members. [4]

(ii) Suggest a suitable sampling frame.[2]

(iii) Identify the sampling units.[2]

c)A sample of 11 circuits from a large normal population has a mean resistance of 2.20 ohms. We know from past testing that the population standard deviation is 0.35 ohms. Determine a 95% confidence interval for the true mean resistance of the population. [8]

d) It is known that 3 % of frozen pies delivered in a canteen are broken. What is the probabilitythat, on a morning when 500 pies are delivered, 5 % or more are broken? [8]

B7. An old film is treated with a chemical in order to improve the contrast. Preliminary test on 9samples drawn from a segment of the film produced the following results.

Sample	A	B	C	D	E	F	G	H	I
x	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
y	49	60	66	62	72	64	89	90	96

The quantity x is a measure of the amount of chemical applied, and y is the contrast index, which takes values between 0 (no contrast) and 100 (maximum contrast).

(a) Plot a scatter plot diagram to illustrate the data. [5]

(b) It was subsequently discovered that on of the samples of the film was damaged and produced an incorrect result. State which sample you think was. [2]

In all subsequent calculations this incorrect sample is ignored. The remaining data can be summarized as follows:

$$\sum x = 23{,}5 \quad \sum x^2 = 83.75 \quad \sum y = 584 \quad \sum y^2 = 44\,622 \quad \sum xy = 1883 \quad n = 8$$

(c) Calculate the product moment correlation coefficient. [8]

(d) State, with a reason, whether it is sensible to conclude from your answer to part (c) that x and y are linearly related. [2]

(e) The line of regression of y on x has equation $y = a + bx$. Calculate the value of a and b, each correct to three significant figures. [9]

(f) Use your line of regression to estimate what the contrast index corresponding to the damaged piece of film would have been if the piece had been undamaged. [2]

(g) State, with a reason; whether it would be sensible to use your regression equation to estimate the contrast index when the quantity of chemical applied to the film is zero. [2]

END OF PAPER