

STATISTICS

SUBJECT 4041

PAPER 1

GENERAL COMMENTS

There was an improvement in the general standard. The candidates found 75% of the questions to be more straight-forward and a higher percentage of candidates obtained good marks. In Section A, many candidates omitted question 4 and failed to obtain all marks in Question 5. On the whole, candidates seemed to have made more attempt to learn the basic work for this section. In Section B, question 7 was the least popular question and many candidates who attempted the question scored poorly.

COMMENTS ON INDIVIDUAL QUESTIONS

QUESTION 1

This question was answered well by most candidates. However, in (a)(ii), some candidates gave y =discrete. They could not think that discrete is a category in quantitative.

- Answers:**
- | | | |
|-----|-------|---|
| (a) | (i) | Variable |
| | (ii) | Quantitative |
| | (iii) | Continuous |
| (b) | (i) | the number of children in each student's family |
| | (ii) | the different hour colours |

QUESTION 2

There were many correct answers especially to part (i). Candidates could not give correctly the disadvantages of using the pictograms. Some candidates only stated that they are inaccurate. The statement just like that would be wrong. Accuracy has to be related to fractional images and not the whole image. Part (b) the graphs in the question paper were supposed to have been labeled A and B. The candidates anyway stated the graph they wanted in different description of its location.

QUESTION 3

This was a very popular question for almost all candidates except for one centre. Very few candidates confused the variance with the standard deviation. Some candidates did not follow instructions on the degree of accuracy.

Answers: $\pi = 7,55$ $S = \sqrt{3,7473} = 1,94$

QUESTION 4

This was a very popular question only 7% of the candidates correctly answered part (a) (ii). No one was able to give the correct answers to parts (a) (ii) and (b).

Answers: (a) (i) $\frac{60-47}{10} = \frac{y_1-50}{8} \Rightarrow y_1 = 60,4\%$

Test 2: $\frac{60-56}{4} = \frac{y_2-50}{8} \Rightarrow y_2 = 58\%$

(ii) First test was better.

(b) There was a higher range of marks for the first test (i.e.) the results were more spread out.

QUESTION 5

This question was well answered by few candidates. Candidates failed to use the interpolation method. They failed to notice that time is continuous. Candidates gave various answers when they calculated the median time. They could not locate the position of the median (i.e.) the 18th value. Some gave the L.C.B. as 14 which was wrong.

Answers: (i) L.C.B. = 13,5

(ii) Median $= 13\frac{1}{2} + 2 \cdot \frac{5}{15}$
 $= 13\frac{1}{2} + \frac{10}{15} = 14,2$

QUESTION 6

Most candidates scored all marks in this question. They defined Price Index well and they were able to use the weights to find the weighted aggregate Price Index. Various reasons were given to part (d) and this was well done.

Answers: (a) (ii) $a = 126$ (iii) $b = \$973544 = 974\ 000$
 $C = \$747\ 967 \approx 748\ 000.$

(b) Weighted aggregate Price Index = 122,83

(c) Expenditure = \$102 million

QUESTION 7

This was the least popular question in the whole paper. The few candidates who attempted this question poorly scored candidates find probability questions difficult. Candidates failed to state whether given events were mutually exclusive or independent.

- Answer:**
- (a) (i) A and B not mutually exclusive because $P(A) + P(B) > 1$
 - (ii) 0,4
 - (iii) 1
 - (b) C and D not independent
 - $0,25 \times 0,4 = 0,1$ but the given one is 0,08
 - (c) (i) $\frac{4}{7}$ (ii) $\frac{3}{7}$ (iii) 0,165
 - (iv) 0,099 (v) 0 (vi) 0,934

QUESTION 8

This question was popular. Candidates scored good marks. However, some candidates gave the total for bananas and oranges yet this was not required on part (ii) of (a). Also, on part (b) candidates went ahead to find the total cost of all the fruits. Again, this was not required. There is need for candidates to answer only what is asked so that time is not wasted. The pie chart was well drawn.

- Answers:**
- (a) (i) 600 fruits
 - (ii) Bananas - 250
 - Oranges - 200
 - (b) (i) Apples - \$4 million
 - Bananas - \$5 million
 - Oranges - \$6 million

QUESTION 9

This was the most popular question answered by 99% of the candidates and all scoring very good marks. However, there were very few candidates who could not explain part (a) (iii) well. They failed to state that some information is lost when data is grouped.

The cumulative frequency curve was well done. Some candidates lost a mark when they did not calculate the mean to the nearest whole number.

- Answers:**
- (a) (i) $40 < x \leq 60$
 - (ii) $49,8 = 50$

(b)

Mark	$x \leq 20$	$x \leq 40$	$x \leq 60$	$x \leq 80$	$x \leq 100$
Cumulative Frequency	23	65	134	180	200

- (c) (i) median = 50
(ii) Q1 = 33
Q3 = 68
1QR = 35
(iii) 10

QUESTION 10

Many candidates attempted this question but the majority did not score very good marks. This question was rather too long as it needed many descriptive posts. Candidates ended up confusing the sampling methods. In part (d) (iii), some candidates used the correct method to come up with a sample from each stratum but then left the figures in a mixed fraction form of which there cannot be – e.g. $2\frac{1}{2}$ people.

- Answers:** (a) (i) Simple random sampling
(ii) Student B correct
(iii) $n = \frac{800}{50} = 16$

- (b) (i) Stratified Sampling

(ii)

Form	1	2	3	4	L6	U6
Sample	10	12	11	9	5	3

- (e) Quota sampling

QUESTION 11

This is another very popular question where candidates scored good marks. The graph was well done. Candidates managed the required scale. Parts (a) and (b)(i)–(iii) were well answered. In part (iv) and (v) some candidates lost some marks when they could not predict the value of y and giving the reason why y may be unreliable.

- Answers:** (a) (i) B
(ii) C
(iii) E
(b) (i) M $(\bar{x}, \bar{y}) = (6;8)$
L $(\bar{x}, \bar{y}) = (3;6)$
U $(\bar{x}, \bar{y}) = (9;10)$
(ii) $m = \frac{2}{3}$
C = 4
(iv) when $x = 21, y = 18$