# PROBABILITY AND STATISTICS-2

# 9709

(March, June and November series 2020 – 2023 With marking scheme)

## SAMPLING AND ESTIMATION

## EXERCISE -2

**Manjula Balaji**

## 1) SP 2020 9709_6 Q1

Leaves from a certain type of tree have lengths that are distributed with standard deviation 3.2 cm. A random sample of 250 of these leaves is taken and the mean length of this sample is found to be 12.5 cm.

**(a)** Calculate a 99% confidence interval for the population mean length. [3]

**(b)** Write down the probability that the whole of a 99% confidence interval will lie below the population mean. [1]

## 2) SP 2020 9709_6 Q2

Describe briefly how to use random numbers to choose a sample of 10 students from a year-group of 276 students. [3]

## 3) March 2020 9709_62 Q2

Lengths of a certain species of lizard are known to be normally distributed with standard deviation 3.2 cm. A naturalist measures the lengths of a random sample of 100 lizards of this species and obtains an $\alpha$% confidence interval for the population mean. He finds that the total width of this interval is 1.25 cm.

Find $\alpha$. [5]

## 4) March 2021 9709_62 Q1

A construction company notes the time, $t$ days, that it takes to build each house of a certain design. The results for a random sample of 60 such houses are summarised as follows.

$$\Sigma t = 4820 \qquad \Sigma t^2 = 392\,050$$

**(a)** Calculate a 98% confidence interval for the population mean time. [6]

**(b)** Explain why it was necessary to use the Central Limit theorem in part **(a)**. [1]

## 5) March 2022 9709_62 Q1

The lengths, in millimetres, of a random sample of 12 rods made by a certain machine are as follows.

200 201 198 202 200 199 199 201 197 202 200 199

**(a)** Find unbiased estimates of the population mean and variance. [3]

**(b)** Give a statistical reason why these estimates may not be reliable. [1]

## 6) March 2022 9709_62 Q3

A random sample of 500 households in a certain town was chosen. Using this sample, a confidence interval for the proportion, $p$, of all households in that town that owned two or more cars was found to be $0.355 < p < 0.445$.

Find the confidence level of this confidence interval. Give your answer correct to the nearest integer. [5]

## 7) March 2023 9709_62 Q1

Anita carried out a survey of 140 randomly selected students at her college. She found that 49 of these students watched a TV programme called *Bunch*.

**(a)** Calculate an approximate 98% confidence interval for the proportion, $p$, of students at Anita's college who watch *Bunch*. [3]

Carlos says that the confidence interval found in **(a)** is not useful because it is too wide.

**(b)** Without calculation, explain briefly how Carlos can use the results of Anita's survey to find a narrower confidence interval for $p$. [1]

## 8) June 2020 9709_61 Q1

The lengths, $X$ centimetres, of a random sample of 7 leaves from a certain variety of tree are as follows.

5.2 4.8 5.5 6.1 4.8 3.9 4.4

**(a)** Calculate unbiased estimates of the population mean and variance of $X$. [3]

It is now given that the true value of the population variance of $X$ is 0.55, and that $X$ has a normal distribution.

**(b)** Find a 95% confidence interval for the population mean of $X$. [3]

## 9) June 2020 9709_62 Q4

The score on one spin of a 5-sided spinner is denoted by the random variable $X$ with probability distribution as shown in the table.

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0.1 | 0.2 | 0.4 | 0.2 | 0.1 |

**(a)** Show that $\text{Var}(X) = 1.2$. [2]

The spinner is spun 200 times. The score on each spin is noted and the mean, $\overline{X}$, of the 200 scores is found.

**(b)** Given that $P(\overline{X} > a) = 0.1$, find the value of $a$. [4]

## 10) June 2020 9709_63 Q1

A random sample of 100 values of a variable $X$ is taken. These values are summarised below.

$$n = 100 \qquad \Sigma x = 1556 \qquad \Sigma x^2 = 29\,004$$

Calculate unbiased estimates of the population mean and variance of $X$. [3]

## 11) June 2020 9709_63 Q5

Sunita has a six-sided die with faces marked 1, 2, 3, 4, 5, 6. The probability that the die shows a six on any throw is $p$. Sunita throws the die 500 times and finds that it shows a six 70 times.

**(a)** Calculate an approximate 99% confidence interval for $p$. [4]

**(b)** Sunita believes that the die is fair. Use your answer to part **(a)** to comment on her belief. [1]

**(c)** Sunita uses the result of her 500 throws to calculate an $\alpha$% confidence interval for $p$. This interval has width 0.04.

Find the value of $\alpha$. [5]

## 12) June 2021 9709_61 Q2

The time, in minutes, taken by students to complete a test has the distribution N(125, 36).

**(a)** Find the probability that the mean time taken to complete the test by a random sample of 40 students is less than 123 minutes. [3]

**(b)** Explain whether it was necessary to use the Central Limit theorem in the solution to part **(a)**. [1]

## 13) June 2021 9709_61 Q4

100 randomly chosen adults each throw a ball once. The length, $l$ metres, of each throw is recorded. The results are summarised below.

$$n = 100 \qquad \Sigma l = 3820 \qquad \Sigma l^2 = 182\,200$$

Calculate a 94% confidence interval for the population mean length of throws by adults. [6]

## 14) June 2021 9709_62 Q6

The heights, $h$ centimetres, of a random sample of 100 fully grown animals of a certain species were measured. The results are summarised below.

$$n = 100 \qquad \Sigma h = 7570 \qquad \Sigma h^2 = 588\,050$$

**(a)** Find unbiased estimates of the population mean and variance. [3]

**(b)** Calculate a 99% confidence interval for the mean height of animals of this species. [3]

Four random samples were taken and a 99% confidence interval for the population mean, $\mu$, was found from each sample.

**(c)** Find the probability that all four of these confidence intervals contain the true value of $\mu$. [2]

## 15) June 2021 9709_63 Q4

The masses, $m$ kilograms, of flour in a random sample of 90 sacks of flour are summarised as follows.

$$n = 90 \qquad \Sigma m = 4509 \qquad \Sigma m^2 = 225\,950$$

**(a)** Find unbiased estimates of the population mean and variance. [3]

**(b)** Calculate a 98% confidence interval for the population mean. [3]

**(c)** Explain why it was necessary to use the Central Limit theorem in answering part **(b)**. [1]

**(d)** Find the probability that the confidence interval found in part **(b)** is wholly above the true value of the population mean. [2]

## 16) June 2022 9709_61 Q1.

The diameters, $x$ millimetres, of a random sample of 200 discs made by a certain machine were recorded. The results are summarised below.

$$n = 200 \qquad \Sigma x = 2520 \qquad \Sigma x^2 = 31852$$

**(a)** Calculate a 95% confidence interval for the population mean diameter. [6]

**(b)** Jean chose 40 random samples and used each sample to calculate a 95% confidence interval for the population mean diameter.

How many of these 40 confidence intervals would be expected to include the true value of the population mean diameter? [1]

## 17) June 2022 9709_62 Q7

$X$ is a random variable with distribution Po(2.90). A random sample of 100 values of $X$ is taken.

Find the probability that the sample mean is less than 2.88. [5]

## 18) June 2022 9709_63 Q1

The number of characters in emails sent by a particular company is modelled by the distribution $N(1250, 480^2)$.

Find the probability that the mean number of characters in a random sample of 100 emails sent by the company is more than 1300. [3]

## 19) June 2022 9709_63 Q6

A random sample of 5 values of a variable $X$ is given below.

$$2 \qquad 3 \qquad 3 \qquad 5 \qquad a$$

**(a)** Find an expression, in terms of $a$, for the mean of these values. [1]

It is given that an unbiased estimate of the population variance of $X$, using these values, is 4. It is also given that $a$ is positive.

**(b)** Find and simplify a quadratic equation in terms of $a$ and hence find the value of $a$. [3]

## 20) June 2023 9709_61 Q4

A certain train journey takes place every day throughout the year. The time taken, in minutes, for the journey is normally distributed with variance 11.2.

**(a)** The mean time for a random sample of $n$ of these journeys was found. A 94% confidence interval for the population mean time was calculated and was found to have a width of 1.4076 minutes, correct to 4 decimal places.

Find the value of $n$. [3]

**(b)** A passenger noted the times for 50 randomly chosen journeys in January, February and March.

Give a reason why this sample is unsuitable for use in finding a confidence interval for the population mean time. [1]

**(c)** A researcher took 4 random samples and a 94% confidence interval for the population mean was found from each sample.

Find the probability that exactly 3 of these confidence intervals contain the true value of the population mean. [2]

## 21) June 2023 9709_61 Q6

A sample of 5 randomly selected values of a variable $X$ is as follows:

$$1 \quad 2 \quad 6 \quad 1 \quad a$$

where $a > 0$.

Given that an unbiased estimate of the variance of $X$ calculated from this sample is $\frac{11}{2}$, find the value of $a$. [3]

## 22) June 2023 9709_62 Q1

In a survey of 200 randomly chosen students from a certain college, 23% of the students said that they owned a car.

Calculate an approximate 93% confidence interval for the proportion of students from the college who own a car. [3]

## 23) June 2023 9709_63 Q2

A club has 264 members, numbered from 1 to 264. Donash wants to choose a random sample of members for a survey. In order to choose the members for the sample he uses his calculator to generate random digits. His first 20 random digits are as follows.

10612 11801 21473 22759

**(a)** The numbers of the first two members in the sample are 106 and 121.

Write down the numbers of the next two members in the sample. [2]

**(b)** To obtain the numbers for members after the 4th member, Donash starts with the second random digit, 0, and obtains the numbers 061 and 211.

Explain why this method will not produce a random sample. [1]

## 24) June 2023 9709_63 Q3

In a random sample of 100 students at Luciana's college, $x$ students said that they liked exams. Luciana used this result to find an approximate 90% confidence interval for the proportion, $p$, of all students at her college who liked exams. Her confidence interval had width 0.157 92.

**(a)** Find the two possible values of $x$. [4]

Suzma independently took another random sample and found another approximate 90% confidence interval for $p$.

**(b)** Find the probability that neither of the two confidence intervals contains the true value of $p$. [1]

## 25) October 2020 9709_61 Q2

In a survey, a random sample of 250 adults in Fromleigh were asked to fill in a questionnaire about their travel.

**(a)** It was found that 102 adults in the sample travel by bus. Find an approximate 90% confidence interval for the proportion of all the adults in Fromleigh who travel by bus. [3]

**(b)** The survey included a question about the amount, $x$ dollars, spent on travel per year. The results are summarised as follows.

$$n = 250 \qquad \Sigma x = 50\,460 \qquad \Sigma x^2 = 19\,854\,200$$

Find unbiased estimates of the population mean and variance of the amount spent per year on travel. [3]

A councillor wanted to select a random sample of houses in Fromleigh. He planned to select the first house on each of the 143 streets in Fromleigh.

**(c)** Explain why this would not provide a random sample. [1]

## 26) October 2020 9709_62 Q2

A six-sided die has faces marked 1, 2, 3, 4, 5, 6. When the die is thrown 300 times it shows a six on 56 throws.

**(a)** Calculate an approximate 96% confidence interval for the probability that the die shows a six on one throw. [3]

**(b)** Maroulla claims that the die is biased.

Use your answer to part **(a)** to comment on this claim. [1]

## 27) October 2021 9709_61 Q1

It is known that the height $H$, in metres, of trees of a certain kind has the distribution N(12.5, 10.24) A scientist takes a random sample of 25 trees of this kind and finds the sample mean, $\overline{H}$, of the heights

**(a)** State the distribution of $\overline{H}$, giving the values of any parameters. [2

**(b)** Find $P(12 < \overline{H} < 13)$. [3]

## 28) October 2021 9709_61 Q3

A random sample of 75 students at a large college was selected for a survey. 15 of these students said that they owned a car. From this result an approximate $\alpha\%$ confidence interval for the proportion of all students at the college who own a car was calculated. The width of this interval was found to be 0.162.

Calculate the value of $\alpha$ correct to 2 significant figures. [5]

## 29) October 2021 9709_61 Q6

The random variable $T$ denotes the time, in seconds, for 100 m races run by Tania. $T$ is normally distributed with mean $\mu$ and variance $\sigma^2$. A random sample of 40 races run by Tania gave the following results.

$$n = 40 \qquad \Sigma t = 560 \qquad \Sigma t^2 = 7850$$

**(a)** Calculate unbiased estimates of $\mu$ and $\sigma^2$. [3]

The random variable $S$ denotes the time, in seconds, for 100 m races run by Suki. $S$ has the independent distribution N(14.2, 0.3).

**(b)** Using your answers to part **(a)**, find the probability that, in a randomly chosen 100 m race, Suki's time will be at least 0.1 s more than Tania's time. [5]

## 30) October 2021 9709_62 Q1

The mass, in kilograms, of a block of cheese sold in a supermarket is denoted by the random variable $M$. The masses of a random sample of 40 blocks are summarised as follows.

$$n = 40 \qquad \Sigma m = 20.50 \qquad \Sigma m^2 = 10.7280$$

**(a)** Calculate unbiased estimates of the population mean and variance of $M$. [3]

**(b)** The price, \$$P$, of a block of cheese of mass $M$ kg is found using the formula $P = 11M + 0.50$.

Find estimates of the population mean and variance of $P$. [3]

## 31) October 2021 9709_62 Q2

Andy and Jessica are doing a survey about musical preferences. They plan to choose a representative sample of six students from the 256 students at their college.

**(a)** Andy suggests that they go to the music building during the lunch hour and choose six students at random from the students who are there.

Give a reason why this method is unsatisfactory. [1]

**(b)** Jessica decides to use another method. She numbers all the students in the college from 1 to 256. Then she uses her calculator and generates the following random numbers.

204393 162007 204028 587119 207395

From these numbers, she obtains six student numbers. The first three of her student numbers are 204, 162 and 7.

Continue Jessica's method to obtain the next three student numbers. [2]

## 32) October 2021 9709_62 Q3

The probability that a certain spinner lands on red on any spin is $p$. The spinner is spun 140 times and it lands on red 35 times.

**(a)** Find an approximate 96% confidence interval for $p$. [3]

From three further experiments, Jack finds a 90% confidence interval, a 95% confidence interval and a 99% confidence interval for $p$.

**(b)** Find the probability that exactly two of these confidence intervals contain the true value of $p$. [3]

## 33) October 2022 9709_61 Q1

The heights, in metres, of a random sample of 10 mature trees of a certain variety are given below.

5.9 6.5 6.7 5.9 6.9 6.0 6.4 6.2 5.8 5.8

Find unbiased estimates of the population mean and variance of the heights of all mature trees of this variety. [3]

## 34) October 2022 9709_61 Q5

A builders' merchant sells stones of different sizes.

**(a)** The masses of size $A$ stones have standard deviation 6 grams. The mean mass of a random sample of 200 size $A$ stones is 45 grams.

Find a 95% confidence interval for the population mean mass of size $A$ stones. [3]

**(b)** The masses of size $B$ stones have standard deviation 11 grams. Using a random sample of size 200, an $\alpha$% confidence interval for the population mean mass is found to have width 4 grams.

Find $\alpha$. [4]

## 35) October 2022 9709_62 Q1

Each of a random sample of 80 adults gave an estimate, $h$ metres, of the height of a particular building. The results were summarised as follows.

$$n = 80 \qquad \Sigma h = 2048 \qquad \Sigma h^2 = 52\,760$$

**(a)** Calculate unbiased estimates of the population mean and variance. [3]

**(b)** Using this sample, the upper boundary of an $\alpha\%$ confidence interval for the population mean is 26.0.

Find the value of $\alpha$. [4]

## 36) October 2022 9709_62 Q5

$X$ is a random variable with distribution B(10, 0.2). A random sample of 160 values of $X$ is taken.

**(a)** Find the approximate distribution of the sample mean, including the values of the parameters. [3]

**(b)** Hence find the probability that the sample mean is less than 1.8. [3]

## Marking Scheme

**1) .**

| | | |
|---|---|---|
| (a) | $z = 2.576$ | 1 |
| | $12.5 \pm z\frac{3.2}{\sqrt{250}}$ | 1 |
| | 12.0 to 13.0 (3 sf) | 1 |
| | | **3** |
| (b) | 0.005 or 0.5% | **1** |

**2) .**

| | |
|---|---|
| (Number students from 1 to 276)<br>Generate 3-digit random numbers | 1 |
| Ignore numbers > 276 | 1 |
| Ignore repeats | 1 |

**3) .**

| | |
|---|---|
| $2 \times z \times \frac{3.2}{10} = 1.25$ | **M1** |
| $z = 1.953$ | **A1** |
| $\phi$('*their* 1.953') (= 0.9746) | **M1** |
| $= 1 - 2(1 - $ '0.9746')<br>$= 0.9492$ | **M1** |
| $\alpha = 94.9$ or 95 | **A1** |

**4) .**

| | | |
|---|---|---|
| (a) | $\text{Est}(\mu) = \frac{4820}{60}$ or $\frac{241}{3}$ or 80.3 (3 sf) | **B1** |
| | $\text{Est}(\sigma^2) = \frac{60}{59}(\frac{392\,050}{60} - (\frac{4820}{60})^2)$ | **M1** |
| | $82.0904\left(\frac{14530}{177}\right)$ to 82.635 or SD = 9.0604 to 9.0904 (3sf) | **A1** |
| | $z = 2.326$ | **B1** |
| | $\frac{4820}{60} \pm z \times \sqrt{\frac{'82.0904'}{60}}$ | **M1** |
| | 77.6 to 83.1 (3 sf) | **A1** |
| (b) | Population distribution of times unknown | **B1** |

**5) .**

| | | |
|---|---|---|
| (a) | $\text{Est}(\mu) = \frac{1199}{6}$ or 199.833 or 200 or $\frac{2398}{12}$ [mm] | **B1** |
| | $\text{Est}(\sigma^2) = \frac{12}{11}\left(\frac{479226}{12} - \frac{'1199'^2}{6}\right)$ or $\frac{1}{11}\left('479226' - \frac{'2398'^2}{6}\right)$ | **M1** |
| | = 2.33 (3 sf) [$mm^2$] | **A1** |
| | | **3** |
| (b) | Small sample | **B1** |

**6) .**

| | |
|---|---|
| $\text{est}(p) = 0.4$ | **B1** |
| $'0.4' + z\sqrt{\frac{'0.4' \times (1 - '0.4')}{500}}$ [= 0.445] | **M1** |
| $z\left[= 0.045 \div \sqrt{\frac{'0.4' \times (1 - '0.4')}{500}}\right] = 2.054$ | **A1** |
| 0.98 – (1 – 0.98) | **M1** |
| 96% confidence | **A1** |

**7) .**

| | | |
|---|---|---|
| (a) | $\left[\frac{49}{140}=0.35\right]$ | |
| | $0.35 \pm z\sqrt{\frac{0.35(1-0.35)}{140}}$ | **M1** |
| | $z = 2.326$ | **B1** |
| | Confidence interval = 0.256 to 0.444 (3 sf) | **A1** |
| | | **3** |
| (b) | Find a smaller **percentage** confidence interval/ lower **level** of confidence | **B1** |

**8) .**

| | | |
|---|---|---|
| (a) | $\frac{\Sigma x}{7}=\frac{34.7}{7}=4.9571$ or 4.96 (3 sf)<br>($\Sigma x^2 = 175.15$) | **B1** |
| | $\frac{7}{6}\left(\frac{"175.15"}{7}-"4.9571"^2\right)$ | **M1** |
| | 0.523 (3 sf) | **A1** |
| | | **3** |
| (b) | $'4.96' \pm z \times \sqrt{\frac{0.55}{7}}$<br>(**FT** *their* mean) | **M1** |
| | $z = 1.96$ | **B1** |
| | 4.41 to 5.51 (3 sf) | **A1** |

**9) .**

| | | |
|---|---|---|
| (a) | $E(X) = 2$ | **B1** |
| | $0.2 \times 1 + 0.4 \times 2^2 + 0.2 \times 3^2 + 0.1 \times 4^2 - 2^2 (= 1.2)$ **AG** | **B1** |
| | | **2** |
| (b) | $\frac{a-2}{\sqrt{1.2 \div 200}} = \phi^{-1}(0.9)$ | **M1** |
| | (**M1** for LHS, **M1** for RHS) | **M1** |
| | $a = 2 + \sqrt{1.2 \div 200} \times 1.282$ | **M1** |
| | 2.10 (3 sf) | **A1** |
| (c) | Yes, because $X$ is not normally distributed. | **B1** |
| | | **1** |
| (d) | $H_0$: pop mean = 2<br>$H_1$: pop mean < 2 | **B1** |
| | $\frac{1.86-2}{\sqrt{1.2 \div 200}}$ | **M1** |
| | 1.807 | **A1** |
| | comp $z = 1.645$ | **M1** |
| | There is evidence that the spinner is biased so that mean is less than 2 | **A1** |

**10)** .

| |
|---|
| Est $\mu = 15.56$ |
| Est $\sigma^2 = \frac{100}{99}\left(\frac{29004}{100} - "15.56"^2\right)$<br>or $= \frac{1}{99}\left(29004 - \frac{1556^2}{100}\right)$ |
| $48.4105 = 48.4$ (3 sf) |

**11)** .

| | | |
|---|---|---|
| (a) | $p = \frac{70}{500}$ or 0.14 | **B1** |
| | $z = 2.576$ | **B1** |
| | $"0.14" \pm z \times \sqrt{\frac{"0.14"(1-"0.14")}{500}}$ | **M1** |
| | 0.100 to 0.180 | **A1** |

| | |
|---|---|
| (b) | 0.1666... is within confidence interval<br>Belief supported or justified |
| | |

| | |
|---|---|
| (c) | $z \times \sqrt{\frac{"0.14"(1-"0.14")}{500}} = 0.02$ |
| | $z = 1.289$ |
| | $\Phi('1.289') = 0.9013$ |
| | $\alpha = '0.9013' - (1 - '0.9013')$ |
| | 80.3% (3 sf) |

**12)** .

| | | |
|---|---|---|
| (a) | $\pm\frac{123-125}{\frac{6}{\sqrt{40}}}$ [= –2.108...] | M1 |
| | P($z <$ '–2.108') = 1 – $\Phi$('2.108') | M1 |
| | = 0.0175 or 0.0176 (3 sf) | A1 |
| | | **3** |
| (b) | No, population is normal | B1 |

**13) .**

| | |
|---|---|
| $\frac{3820}{100}$ [= 38.2] | B1 |
| $\frac{100}{99}\left(\frac{182200}{100} - '38.2'^2\right)$ or $\frac{1}{99}\left(182200 - \frac{3820^2}{100}\right)$ | M1 |
| $= \frac{12092}{33}$ or 366.424 or 366 (3 sf) | A1 |
| $'38.2' \pm z \times \sqrt{\frac{'366.424'}{100}}$ | M1 |
| $z = 1.881$ or 1.882 | B1 |
| 34.6 to 41.8 (3 sf) | A1 |

**14) .**

| | | |
|---|---|---|
| (a) | est $(\mu) = \frac{7570}{100}$ (= 75.7) | B1 |
| | est$(\sigma^2) = \frac{100}{99}\left(\frac{\sum h^2}{100} - '75.7'^2\right)$ or $\frac{1}{99}\left(588050 - \frac{7570^2}{100}\right)$<br>$= \frac{100}{99}\left(\frac{588050}{100} - '75.7'^2\right)$ [= 151.525] | M1 |
| | = 152 (3 sf) | A1 |
| (b) | $'75.7' \pm z\sqrt{\frac{'151.525'}{100}}$ | M1 |
| | $z = 2.576$ | B1 |
| | 72.5 to 78.9 | A1 FT |
| | | **3** |
| (c) | $0.99^4$ | B1 |
| | 0.961 (3 sf) | B1 |

**15) .**

| | | |
|---|---|---|
| (a) | $\frac{4509}{90}$ [= 50.1] | B1 |
| | $\frac{90}{89}(\frac{225950}{90} - '50.1'^2)$ or $\frac{1}{89}\left(225950 - \frac{4509^2}{90}\right)$ | M1 |
| | $\frac{491}{890}$ or 0.552 (3 sf) | A1 |
| (b) | '50.1' $\pm z\sqrt{\frac{'\frac{491}{890}'}{90}}$ | M1 |
| | $z = 2.326$ | B1 |
| | 49.9 to 50.3 (3 sf) | A1 |
| (c) | Population of masses is unknown | B1 |
| | | 1 |
| (d) | 1 – 0.98 | M1 |
| | 0.02 ÷ 2 = 0.01 | A1 |

**16) .**

| | | |
|---|---|---|
| (a) | Est($\mu$) = $\frac{2520}{200}$ [= 12.6] | B1 |
| | Est($\sigma^2$) = $\frac{200}{199}\left(\frac{31582}{200} - '12.6'^2\right)$ or $\frac{1}{199}\left(31852 - \frac{2520^2}{200}\right)$ | M1 |
| | = 0.5025 or 0.503 or $\frac{100}{199}$ | A1 |
| | $z = 1.96$ | B1 |
| | '12.6' $\pm z \times \sqrt{'0.5025' \div 200}$ | M1 |
| | CI = 12.5 to 12.7 (3 sf) | A1 |
| (b) | 0.95 × 40 [= 38] | B1 |

**17) .**

| | |
|---|---|
| $\bar{X} \sim N(2.9, \frac{2.9}{100})$ OR Totals method N(290, 290) | **B1** |
| | **B1** |
| $\frac{2.88-2.90}{\sqrt{\frac{2.9}{100}}}$ [= –0.1174] OR $\frac{288-290}{\sqrt{290}}$ | **M1** |
| 1– Φ('0.1174') | **M1** |
| 0.453 (3 sf) | **A1** |

**Alternative method for question 7**

| | |
|---|---|
| $\bar{X} \sim N(2.9, \frac{2.9}{100})$ OR Totals method N(290, 290) | **B1** |
| | **B1** |
| $\frac{(2.88-\frac{1}{200})-2.90}{\sqrt{\frac{2.9}{100}}}$ [= –0.1468] OR (287.5 – 290)/√290 | **M1** |
| 1 – ϕ('0.1468') | **M1** |
| 0.442 (3 sf) | **A1** |

**18) .**

| | |
|---|---|
| $\frac{1300+\frac{1}{200}-1250}{\frac{480}{10}}$ or $\frac{1300-1250}{\frac{480}{10}}$ [= 1.042] | **M1** |
| 1 – Φ('1.042') | **M1** |
| 0.149 ( 3 s.f.) | **A1** |

**19) .**

| | | |
|---|---|---|
| (a) | $\frac{13+a}{5}$ | **B1** |
| (b) | $\frac{5}{4}\left(\frac{47+a^2}{5}-\left(\frac{13+a}{5}\right)^2\right)=4$ or $\frac{1}{4}\left(47+a^2-\frac{(13+a)^2}{5}\right)=4$ | **M1** |
| | $2a^2-13a-7=0$ | **A1** |
| | $a=7$ | **A1** |

**20) .**

| | | |
|---|---|---|
| (a) | $z\times\sqrt{\frac{11.2}{n}}=1.4076\div 2$ | **M1** |
| | $z=1.881$ or $1.882$ | **B1** |
| | $[n=\left(\frac{1.881}{0.7038}\right)^2\times 11.2]$<br>$n=80$ | **A1** |
| (b) | Jan, Feb and March not typical of whole year. | **B1** |
| | | **1** |
| (c) | $0.94^3\times 0.06\times 4$ | **M1** |
| | $=0.199$ (3 sf) | **A1** |

**21) .**

| | |
|---|---|
| $\frac{5}{4}\left(\frac{1+2^2+6^2+1+a^2}{5}-\left(\frac{1+2+6+1+a}{5}\right)^2\right)=\frac{11}{2}$<br>or $\frac{1}{4}\left((42+a^2)-\frac{(10+a^2)}{5}\right)=\frac{11}{2}$ | **M1*** |
| $4a^2-20a+0=0$ or $a^2-5a+0=0$ | **DM1** |
| $a=5$ | **A1** |

**22) .**

| | |
|---|---|
| $0.23 \pm z \times \sqrt{\frac{0.23\times(1-0.23)}{200}}$ | **M1** |
| $z = 1.811$ or $1.812$ | **B1** |
| 0.176 to 0.284 (3 sf) | **A1** |

**23) .**

| | | |
|---|---|---|
| (a) | 180, 227 | **B1** |
| | | **B1** |
| | | **2** |
| (b) | These numbers are not independent of the previous numbers<br>OR<br>Only a finite number of digits used | **B1** |

**24) .**

| | | |
|---|---|---|
| (a) | $z = 1.645$ | **B1** |
| | $z \times \frac{\sqrt{\frac{x}{100}\times(1-\frac{x}{100})}}{100} = 0.07896$ | **M1** |
| | $[x(100-x) = 100^3 \times 0.07896^2 \div 1.645^2]$<br>$x^2 - 100x + 2304 = 0$ | **A1** |
| | $x = 36$ or $64$ | **A1** |
| | | **4** |
| (b) | $0.1^2 = 0.01$ | **B1** |

**25) .**

| | | |
|---|---|---|
| (a) | $\frac{\frac{102}{250}\times\frac{250-102}{250}}{250}$ (= 0.000966144)<br>$\frac{102}{250}\pm z\sqrt{'0.00096614'}$ | **M1** |
| | $z = 1.645$ | **B1** |
| | Confident Interval is 0.357 to 0.459 (3 sf) | **A1** |
| | | **3** |
| (b) | Estimate of mean $\left(\frac{50460}{250}\right)$ = \$201.84 | **B1** |
| | $\frac{250}{249}\left(\frac{19854200}{250}-\left(\frac{50460}{250}\right)^2\right)$ or $\frac{1}{249}\left(19854200-\frac{50460^2}{250}\right)$ | **M1** |
| | Estimate of variance = 38 832.75 dollars$^2$ or 38 800 (3 sf) | **A1** |
| | | **3** |
| (c) | e.g. Every house doesn't have an equal chance of being selected or most houses have no chance of being selected. | **B1** |
| | | **1** |

**26) .**

| | | |
|---|---|---|
| 2(a) | $\frac{56}{300} \pm z \times \sqrt{\frac{\frac{56}{300} \times \frac{244}{300}}{300}}$ | **M1** |
| | $z = 2.054$ or $2.055$ | **B1** |
| | 0.14(0) to 0.233 (3sf) or 0.141 to 0.233 (3sf) | **A1** |
| | | **3** |
| 2(b) | $\frac{1}{6}$ $(= 0.167)$<br>This is within confidence interval, so no reason to believe die is biased. | **B1 FT** |

**27) .**

| | | |
|---|---|---|
| (a) | N(12.5, ….) | **B1** |
| | Variance = 0.4096 | **B1** |
| (b) | $\frac{13 - '12.5'}{\sqrt{'0.4096'}}$ [= 0.781] | **M1** |
| | ϕ(‘0.781’) – (1 – Φ(‘0.781’)) | **M1** |
| | 0.565 (3sf) | **A1** |

**28) .**

| | |
|---|---|
| est(p) = 0.2 accept $\frac{15}{75}$ | **B1** |
| $2 \times z \times \sqrt{\frac{0.2 \times 0.8}{75}} = 0.162$ | **M1** |
| $z \left[ = 0.081 \times \sqrt{\frac{75}{0.2 \times 0.8}} \right] = 1.754$ | **A1** |
| Φ(‘1.754’) = 0.96[03]<br>'0.96' – (1 – '0.96') | **M1** |
| $\alpha = 92$ | **A1** |

**29) .**

| | | |
|---|---|---|
| (a) | est $\mu = 14$ accept $\frac{560}{40}$ | **B1** |
| | est $\sigma^2 = \frac{40}{39}\left(\frac{7850}{40} - 14^2\right)$ or $\frac{1}{39}\left(7850 - \frac{560^2}{40}\right)$ | **M1** |
| | 0.25641 or 0.256 (3sf) | **A1** |
| (b) | E($S - T$) = 14.2 – '14' [= 0.2 ] | **B1 FT** |
| | Var($S - T$) = 0.3 + ` 0.256` [= 0.55641] | **B1 FT** |
| | $\frac{0.1 - '0.2'}{\sqrt{'0.55641'}}$ [= –0.134] | **M1** |
| | P($S$ – T > 0.1) = 1 – Φ('–0.134') = Φ('0.134') | **M1** |
| | 0.553 (3sf) | **A1** |

**30) .**

| | | |
|---|---|---|
| (a) | $\frac{20.5}{40} = 0.5125$ | **B1** |
| | $\frac{40}{39}\left(\frac{10.728}{40} - \left('0.5125'^2\right)\right)$ or $\frac{1}{39}\left(10.728 - \frac{20.50^2}{40}\right)$ | **M1** |
| | 0.0056859 or 0.00569 (3 sf) or $\frac{887}{156\,000}$ | **A1** |
| (b) | [11 × '0.5125' + 0.5]) = 6.1375 or $\frac{491}{80}$ or 6.14 (3sf) | **B1 FT** |
| | $11^2$ × '0.0056859' | **M1** |
| | 0.688 (3sf) | **A1** |

**31) .**

| | | |
|---|---|---|
| (a) | E.g. Bias towards students who play instruments or only music students<br>or e.g. the six will possibly be friends/have similar music preferences | **B1** |
| (b) | 28, 119, 207 | **B1** |
| | | **B1** |

**32) .**

| | | |
|---|---|---|
| (a) | $0.25 \pm z\sqrt{\frac{0.25\times0.75}{140}}$ | **M1** |
| | $z = 2.054$ or $2.055$ | **B1** |
| | 0.175 to 0.325 (3sf) | **A1** |
| (b) | $0.90 \times 0.95 \times 0.01$<br>$+ 0.90 \times 0.05 \times 0.99$<br>$+ 0.10 \times 0.95 \times 0.99$ | **M1 M1** |
| | 0.147 | **A1** |

**33) .**

| | |
|---|---|
| $\frac{62.1}{10} = 6.21$ | **B1** |
| $[\Sigma x^2 = 387.05]$ | **M1** |
| $\frac{10}{9}\left(\frac{\textit{their '387.05'}}{10} - \left(\textit{their '6.21'}\right)^2\right)$<br>**or** $\frac{1}{9}\left(\frac{\textit{their '387.05'}}{10} - \frac{\left(\textit{their '6.21'}\right)^2}{10}\right)$ | |
| $= 0.157$ (3 sf) or $\frac{1409}{9000}$ | **A1** |

**34) ,**

| | | |
|---|---|---|
| (a) | $45 \pm z \times \frac{6}{\sqrt{200}}$ | **M1** |
| | $z = 1.96$ | **B1** |
| | 44.2 to 45.8 (3 sf) | **A1** |
| (b) | $z \times \frac{11}{\sqrt{200}} = 2$ | **M1** |
| | $z = 2.571$ | **A1** |
| | $\phi$(*their* '2.571') = 0.9949<br>**and** *their* '0.9949' – (1 – *their* '0.9949') [= 0.9898] | **M1** |
| | $\alpha = 99.0$ (3 sf) | **A1** |

**35) .**

| | | |
|---|---|---|
| (a) | Est $\mu = 25.6$ or $\frac{2048}{80}$ or $\frac{128}{5}$ | **B1** |
| | Est $\sigma^2 = \frac{80}{79}\left(\frac{52760}{80} - \left(\frac{2048}{80}\right)^2\right)$ or $\frac{1}{79}\left(52760 - \frac{2048^2}{80}\right)$ | **M1** |
| | $= 4.19$ (3 sf) or $\frac{1656}{395}$ | **A1** |
| (b) | '25.6' $+ z\sqrt{\frac{'4.19'}{80}} = 26.0$ | **M1** |
| | $z = 1.748$ or 1.747 | **A1** |
| | ($\phi$('1.748') = 0.960) '0.960' – (1 – '0.960') | **M1** |
| | $\alpha = 92.0$ or 91.9 | **A1** |

**36) .**

| | | |
|---|---|---|
| (a) | For $X$, $\mu = 2$ $\sigma^2 = 1.6$ | |
| | Mean = 2 | **B1** |
| | Variance $= \frac{1.6}{160}$ or $\frac{1}{100}$ or 0.01 | **B1** |
| | Normal | **B1** |
| (b) | $\pm\frac{1.8-\frac{1}{320}-'2'}{\sqrt{'0.01'}}$ or $\pm\frac{1.8-'2'}{\sqrt{'0.01'}}$ [= –2.03 or –2]<br>or $\pm$ (287.5 – '320') / $\sqrt{'256'}$ or $\pm$ (288 – '320') / $\sqrt{'256'}$ [= –2.03 or –2] | **M1** |
| | $\phi$('–2.03') = 1 – $\phi$('2.03') | **M1** |
| | = 0.0212 or 0.0228 (3 sf) | **A1** |