Statistics -2

9709

Sampling and Estimation

Exercise (June series 2019 – 2002)

With marking scheme

Manjula Balaji

CAIE math facilitator

Chennai Public Global School

Chennai,

India

SAMPLING AND ESTIMATION

1) S/2002/7/Q1

The result of a fitness trial is a random variable X which is normally distributed with mean μ and standard deviation 2.4. A researcher uses the results from a random sample of 90 trials to calculate a 98% confidence interval for μ. What is the width of this interval? [4]

2) S/2003/7/Q3

A consumer group, interested in the mean fat content of a particular type of sausage, takes a random sample of 20 sausages and sends them away to be analysed. The percentage of fat in each sausage is as follows.

26 27 28 28 28 29 29 30 30 31 32 32 32 33 33 34 34 34 35 35

Assume that the percentage of fat is normally distributed with mean μ, and that the standard deviation is known to be 3.

(i) Calculate a 98% confidence interval for the population mean percentage of fat. [4]

(ii) The manufacturer claims that the mean percentage of fat in sausages of this type is 30. Use your answer to part **(i)** to determine whether the consumer group should accept this claim. [2]

3) S/2004/7/Q4

Packets of cat food are filled by a machine.

(i) In a random sample of 10 packets, the weights, in grams, of the packets were as follows.

374.6 377.4 376.1 379.2 371.2 375.0 372.4 378.6 377.1 371.5

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

(ii) In a random sample of 200 packets, 38 were found to be underweight. Calculate a 96% confidence interval for the population proportion of underweight packets. [4]

4) S/2005/7/Q2

Jenny has to do a statistics project at school on how much pocket money, in dollars, is received by students in her year group. She plans to take a sample of 7 students from her year group, which contains 122 students.

(i) Give a suitable method of taking this sample. [1]

Her sample gives the following results.

13.40 10.60 26.50 20.00 14.50 15.00 16.50

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

(iii) Is the estimated population variance more than, less than or the same as the sample variance? [1]

(iv) Describe what you understand by 'population' in this question. [1]

5) S/2005/7/Q3

A survey of a random sample of n people found that 61 of them read *The Reporter* newspaper. A symmetric confidence interval for the true population proportion, p, who read *The Reporter* is $0.1993 < p < 0.2887$.

(i) Find the mid-point of this confidence interval and use this to find the value of n. [3]

(ii) Find the confidence level of this confidence interval. [4]

6) S/2006/7/Q1

Packets of fish food have weights that are distributed with standard deviation 2.3 g. A random sample of 200 packets is taken. The mean weight of this sample is found to be 99.2 g. Calculate a 99% confidence interval for the population mean weight. [3]

7) S/2007/7/Q6

The daily takings, $\$x$, for a shop were noted on 30 randomly chosen days. The takings are summarised by $\Sigma x = 31\,500$, $\Sigma x^2 = 33\,141\,816$.

(i) Calculate unbiased estimates of the population mean and variance of the shop's daily takings. [3]

(ii) Calculate a 98% confidence interval for the mean daily takings. [3]

The mean daily takings for a random sample of n days is found.

(iii) Estimate the value of n for which it is approximately 95% certain that the sample mean does not differ from the population mean by more than \$6. [3]

8) S/2008/7/Q1

A magazine conducted a survey about the sleeping time of adults. A random sample of 12 adults was chosen from the adults travelling to work on a train.

(i) Give a reason why this is an unsatisfactory sample for the purposes of the survey. [1]

(ii) State a population for which this sample would be satisfactory. [1]

A satisfactory sample of 12 adults gave numbers of hours of sleep as shown below.

4.6 6.8 5.2 6.2 5.7 7.1 6.3 5.6 7.0 5.8 6.5 7.2

(iii) Calculate unbiased estimates of the mean and variance of the sleeping times of adults. [3]

9) S/2009/7/Q2

The weights in grams of oranges grown in a certain area are normally distributed with mean μ and standard deviation σ. A random sample of 50 of these oranges was taken, and a 97% confidence interval for μ based on this sample was (222.1, 232.1).

(i) Calculate unbiased estimates of μ and σ^2. [4]

(ii) Estimate the sample size that would be required in order for a 97% confidence interval for μ to have width 8. [3]

10) S/2010/71/Q2

A random sample of n people were questioned about their internet use. 87 of them had a high-speed internet connection. A confidence interval for the population proportion having a high-speed internet connection is $0.1129 < p < 0.1771$.

(i) Write down the mid-point of this confidence interval and hence find the value of n. [3]

(ii) This interval is an $\alpha\%$ confidence interval. Find α. [4]

11) S/2010/71/Q3

Metal bolts are produced in large numbers and have lengths which are normally distributed with mean 2.62 cm and standard deviation 0.30 cm.

(i) Find the probability that a random sample of 45 bolts will have a mean length of more than 2.55 cm. [3]

(ii) The machine making these bolts is given an annual service. This may change the mean length of bolts produced but does not change the standard deviation. To test whether the mean has changed, a random sample of 30 bolts is taken and their lengths noted. The sample mean length is m cm. Find the set of values of m which result in rejection at the 10% significance level of the hypothesis that no change in the mean length has occurred. [4]

12) S/2010/73/Q3

The weight, in grams, of a certain type of apple is modelled by the random variable X with mean 62 and standard deviation 8.2. A random sample of 50 apples is selected, and the mean weight in grams, $\overline{X}$, is found.

(i) Describe fully the distribution of $\overline{X}$. [3]

(ii) Find $P(\overline{X} > 64)$. [3]

13) S/2011/71/Q2

(a) The time taken by a worker to complete a task was recorded for a random sample of 50 workers. The sample mean was 41.2 minutes and an unbiased estimate of the population variance was 32.6 minutes^2. Find a 95% confidence interval for the mean time taken to complete the task. [3]

(b) The probability that an $\alpha\%$ confidence interval includes only values that are lower than the population mean is $\frac{1}{16}$. Find the value of α. [2]

14) S/2011/71/Q3

Past experience has shown that the heights of a certain variety of rose bush have been normally distributed with mean 85.0 cm. A new fertiliser is used and it is hoped that this will increase the heights. In order to test whether this is the case, a botanist records the heights, x cm, of a large random sample of n rose bushes and calculates that $\bar{x} = 85.7$ and $s = 4.8$, where $\bar{x}$ is the sample mean and s^2 is an unbiased estimate of the population variance. The botanist then carries out an appropriate hypothesis test.

(i) The test statistic, z, has a value of 1.786 correct to 3 decimal places. Calculate the value of n. [3]

(ii) Using this value of the test statistic, carry out the test at the 5% significance level. [3]

15) S/2011/72/Q4

A doctor wishes to investigate the mean fat content in low-fat burgers. He takes a random sample of 15 burgers and sends them to a laboratory where the mass, in grams, of fat in each burger is determined. The results are as follows.

9 7 8 9 6 11 7 9 8 9 8 10 7 9 9

Assume that the mass, in grams, of fat in low-fat burgers is normally distributed with mean μ and that the population standard deviation is 1.3.

(i) Calculate a 99% confidence interval for μ. [4]

(ii) Explain whether it was necessary to use the Central Limit theorem in the calculation in part **(i)**. [2]

(iii) The manufacturer claims that the mean mass of fat in burgers of this type is 8 g. Use your answer to part **(i)** to comment on this claim. [2]

16) S/2011/73/Q2

In a random sample of 70 bars of Luxcleanse soap, 18 were found to be undersized.

(i) Calculate an approximate 90% confidence interval for the proportion of all bars of Luxcleanse soap that are undersized. [4]

(ii) Give a reason why your interval is only approximate. [1]

17) S/2012/71/Q1

The weights, in grams, of packets of sugar are distributed with mean μ and standard deviation 23. A random sample of 150 packets is taken. The mean weight of this sample is found to be 494 g. Calculate a 98% confidence interval for μ. [3]

18) S/2012/71/Q6

A survey taken last year showed that the mean number of computers per household in Branley was 1.66. This year a random sample of 50 households in Branley answered a questionnaire with the following results.

Number of computers	0	1	2	3	4	> 4
Number of households	5	12	18	10	5	0

(i) Calculate unbiased estimates for the population mean and variance of the number of computers per household in Branley this year. [3]

(ii) Test at the 5% significance level whether the mean number of computers per household has changed since last year. [5]

(iii) Explain whether it is possible that a Type I error may have been made in the test in part **(ii)**. [1]

(iv) State what is meant by a Type II error in the context of the test in part **(ii)**, and give the set of values of the test statistic that could lead to a Type II error being made. [2]

19) S/2012/72/Q1

The number of new enquiries per day at an office has a Poisson distribution. In the past the mean has been 3. Following a change of staff, the manager wishes to test, at the 5% significance level, whether the mean has increased.

(i) State the null and alternative hypotheses for this test. [1]

The manager notes the number, N, of new enquiries during a certain 6-day period. She finds that $N = 25$ and then, assuming that the null hypothesis is true, she calculates that $P(N \geqslant 25) = 0.0683$.

(ii) What conclusion should she draw? [2]

20) S/2012/72/Q2

A population has mean 7 and standard deviation 3. A random sample of size n is chosen from this population.

(i) Write down the mean and standard deviation of the distribution of the sample mean. [2]

(ii) Under what circumstances does the sample mean have

(a) a normal distribution, [1]

(b) an approximately normal distribution? [1]

21) S/2012/72/Q3

In a sample of 50 students at Batlin college, 18 support the football club Real Madrid.

(i) Calculate an approximate 98% confidence interval for the proportion of students at Batlin college who support Real Madrid. [4]

(ii) Give one condition for this to be a reliable result. [1]

22) S/2012/73/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.

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

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

23) S/2012/73/Q3

The lengths, x mm, of a random sample of 150 insects of a certain kind were found. The results are summarised by $\Sigma x = 7520$ and $\Sigma x^2 = 413\,540$.

(i) Calculate unbiased estimates of the population mean and variance of the lengths of insects of this kind. [3]

(ii) Using the values found in part (i), calculate an estimate of the probability that the mean length of a further random sample of 80 insects of this kind is greater than 53 mm. [3]

24) S/2012/73/Q5

(i) Deng wishes to test whether a certain coin is biased so that it is more likely to show Heads than Tails. He throws it 12 times. If it shows Heads more than 9 times, he will conclude that the coin is biased. Calculate the significance level of the test. [3]

(ii) Deng throws another coin 100 times in order to test, at the 5% significance level, whether it is biased towards Heads. Find the rejection region for this test. [5]

25) S/2013/71/Q4

The lengths, x m, of a random sample of 200 balls of string are found and the results are summarised by $\Sigma x = 2005$ and $\Sigma x^2 = 20\,175$.

(i) Calculate unbiased estimates of the population mean and variance of the lengths. [3]

(ii) Use the values from part **(i)** to estimate the probability that the mean length of a random sample of 50 balls of string is less than 10 m. [3]

(iii) Explain whether or not it was necessary to use the Central Limit theorem in your calculation in part **(ii)**. [2]

26) S/2013/71/Q7

Leila suspects that a particular six-sided die is biased so that the probability, p, that it will show a six is greater than $\frac{1}{6}$. She tests the die by throwing it 5 times. If it shows a six on 3 or more throws she will conclude that it is biased.

(i) State what is meant by a Type I error in this situation and calculate the probability of a Type I error. [3]

(ii) Assuming that the value of p is actually $\frac{2}{3}$, calculate the probability of a Type II error. [3]

Leila now throws the die 80 times and it shows a six on 50 throws.

(iii) Calculate an approximate 96% confidence interval for p. [4]

27) S/2013/72/Q4

The masses, in grams, of a certain type of plum are normally distributed with mean μ and variance σ^2. The masses, m grams, of a random sample of 150 plums of this type were found and the results are summarised by $\Sigma m = 9750$ and $\Sigma m^2 = 647\,500$.

(i) Calculate unbiased estimates of μ and σ^2. [3]

(ii) Calculate a 98% confidence interval for μ. [3]

Two more random samples of plums of this type are taken and a 98% confidence interval for μ is calculated from each sample.

(iii) Find the probability that neither of these two intervals contains μ. [2]

28) S/2013/72/Q6

The number of cases of asthma per month at a clinic has a Poisson distribution. In the past the mean has been 5.3 cases per month. A new treatment is introduced. In order to test at the 5% significance level whether the mean has decreased, the number of cases in a randomly chosen month is noted.

(i) Find the critical region for the test and, given that the number of cases is 2, carry out the test. [5]

(ii) Explain the meaning of a Type I error in this context and state the probability of a Type I error. [2]

(iii) At another clinic the mean number of cases of asthma per month has the independent distribution Po(13.1). Assuming that the mean for the first clinic is still 5.3, use a suitable approximating distribution to estimate the probability that the total number of cases in the two clinics in a particular month is more than 20. [5]

29) S/2013/73/Q3

Each of a random sample of 15 students was asked how long they spent revising for an exam. The results, in minutes, were as follows.

50 70 80 60 65 110 10 70 75 60 65 45 50 70 50

Assume that the times for all students are normally distributed with mean μ minutes and standard deviation 12 minutes.

(i) Calculate a 92% confidence interval for μ. [4]

(ii) Explain what is meant by a 92% confidence interval for μ. [1]

(iii) Explain what is meant by saying that a sample is 'random'. [1]

30) S/2013/73/Q7

In the past the weekly profit at a store had mean \$34 600 and standard deviation \$4500. Following a change of ownership, the mean weekly profit for 90 randomly chosen weeks was \$35 400.

(i) Stating a necessary assumption, test at the 5% significance level whether the mean weekly profit has increased. [6]

(ii) State, with a reason, whether it was necessary to use the Central Limit theorem in part **(i)**. [2]

The mean weekly profit for another random sample of 90 weeks is found and the same test is carried out at the 5% significance level.

(iii) State the probability of a Type I error. [1]

(iv) Given that the population mean weekly profit is now \$36 500, calculate the probability of a Type II error. [5]

31) S/2014/71/Q1

The masses, in grams, of apples of a certain type are normally distributed with mean 60.4 and standard deviation 8.2. The apples are packed in bags, with each bag containing 8 randomly chosen apples. The bags are checked by Quality Control and any bag containing apples with a total mass of less than 436 g is rejected. Find the proportion of bags that are rejected. [4]

32) S/2014/71/Q2

A die is biased. The mean and variance of a random sample of 70 scores on this die are found to be 3.61 and 2.70 respectively. Calculate a 95% confidence interval for the population mean score. [5]

33) S/2014/71/Q5

The score on one throw of a 4-sided die is denoted by the random variable X with probability distribution as shown in the table.

x	0	1	2	3
$P(X = x)$	0.25	0.25	0.25	0.25

(i) Show that $\text{Var}(X) = 1.25$. [1]

The die is thrown 300 times. The score on each throw is noted and the mean, $\overline{X}$, of the 300 scores is found.

(ii) Use a normal distribution to find $P(\overline{X} < 1.4)$. [3]

(iii) Justify the use of the normal distribution in part **(ii)**. [1]

34) S/2014/72/Q1

The weights, in grams, of a random sample of 8 packets of cereal are as follows.

250 248 255 244 259 250 242 258

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

35) S/ 2014/72/Q7

A researcher is investigating the actual lengths of time that patients spend with the doctor at their appointments. He plans to choose a sample of 12 appointments on a particular day.

(i) Which of the following methods is preferable, and why?

- Choose the first 12 appointments of the day.
- Choose 12 appointments evenly spaced throughout the day. [2]

Appointments are scheduled to last 10 minutes. The actual lengths of time, in minutes, that patients spend with the doctor may be assumed to have a normal distribution with mean μ and standard deviation 3.4. The researcher suspects that the actual time spent is more than 10 minutes on average. To test this suspicion, he recorded the actual times spent for a random sample of 12 appointments and carried out a hypothesis test at the 1% significance level.

(ii) State the probability of making a Type I error and explain what is meant by a Type I error in this context. [2]

(iii) Given that the total length of time spent for the 12 appointments was 147 minutes, carry out the test. [5]

(iv) Give a reason why the Central Limit theorem was not needed in part **(iii)**. [1]

36) S/2014/73/Q3

A die is thrown 100 times and shows an odd number on 56 throws. Calculate an approximate 97% confidence interval for the probability that the die shows an odd number on one throw. [4]

37) S/2014/73/Q4

The weights, X kilograms, of rabbits in a certain area have population mean μ kg. A random sample of 100 rabbits from this area was taken and the weights are summarised by

$$\Sigma x = 165, \qquad \Sigma x^2 = 276.25.$$

Test at the 5% significance level the null hypothesis $H_0 : \mu = 1.6$ against the alternative hypothesis $H_1 : \mu \neq 1.6$. [6]

38) S/2015/71/Q5

The masses, m grams, of a random sample of 80 strawberries of a certain type were measured and summarised as follows.

$$n = 80 \qquad \Sigma m = 4200 \qquad \Sigma m^2 = 229\,000$$

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

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

50 random samples of size 80 were taken and a 98% confidence interval for the population mean, μ, was found from each sample.

(iii) Find the number of these 50 confidence intervals that would be expected to include the true value of μ. [1]

39) S/2015/72/Q5

The volumes, v millilitres, of juice in a random sample of 50 bottles of Cooljoos are measured and summarised as follows.

$$n = 50 \qquad \Sigma v = 14\,800 \qquad \Sigma v^2 = 4\,390\,000$$

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

(ii) An α% confidence interval for the population mean, based on this sample, is found to have a width of 5.45 millilitres. Find α. [4]

Four random samples of size 10 are taken and a 96% confidence interval for the population mean is found from each sample.

(iii) Find the probability that these 4 confidence intervals all include the true value of the population mean. [2]

40) S/2015/73/Q3

A die is biased so that the probability that it shows a six on any throw is p.

(i) In an experiment, the die shows a six on 22 out of 100 throws. Find an approximate 97% confidence interval for p. [4]

(ii) The experiment is repeated and another 97% confidence interval is found. Find the probability that exactly one of the two confidence intervals includes the true value of p. [2]

41) S/2015/73/Q4

The marks, x, of a random sample of 50 students in a test were summarised as follows.

$$n = 50 \qquad \Sigma x = 1508 \qquad \Sigma x^2 = 51\,825$$

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

(ii) Each student's mark is scaled using the formula $y = 1.5x + 10$. Find estimates of the population mean and variance of the scaled marks, y. [3]

42) S/2015/73/Q5

The mean breaking strength of cables made at a certain factory is supposed to be 5 tonnes. The quality control department wishes to test whether the mean breaking strength of cables made by a particular machine is actually less than it should be. They take a random sample of 60 cables. For each cable they find the breaking strength by gradually increasing the tension in the cable and noting the tension when the cable breaks.

(i) Give a reason why it is necessary to take a sample rather then testing all the cables produced by the machine. [1]

(ii) The mean breaking strength of the 60 cables in the sample is found to be 4.95 tonnes. Given that the population standard deviation of breaking strengths is 0.15 tonnes, test at the 1% significance level whether the population mean breaking strength is less than it should be. [4]

(iii) Explain whether it was necessary to use the Central Limit theorem in the solution to part **(ii)**. [2]

43) S/2016/71/Q2

A researcher is investigating the lengths, in kilometres, of the journeys to work of the employees at a certain firm. She takes a random sample of 10 employees.

(i) State what is meant by 'random' in this context. [1]

The results of her sample are as follows.

1.5 2.0 3.6 5.9 4.8 8.7 3.5 2.9 4.1 3.0

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

(iii) State what is meant by 'population' in this context. [1]

44) S/2016/71/Q3

Based on a random sample of 700 people living in a certain area, a confidence interval for the proportion, p, of all people living in that area who had travelled abroad was found to be $0.5672 < p < 0.6528$.

(i) Find the proportion of people in the sample who had travelled abroad. [1]

(ii) Find the confidence level of this confidence interval. Give your answer correct to the nearest integer. [4]

45) S/2016/72/Q3

(i) Give a reason for using a sample rather than the whole population in carrying out a statistical investigation. [1]

(ii) Tennis balls of a certain brand are known to have a mean height of bounce of 64.7 cm, when dropped from a height of 100 cm. A change is made in the manufacturing process and it is required to test whether this change has affected the mean height of bounce. 100 new tennis balls are tested and it is found that their mean height of bounce when dropped from a height of 100 cm is 65.7 cm and the unbiased estimate of the population variance is 15 cm^2.

(a) Calculate a 95% confidence interval for the population mean. [3]

(b) Use your answer to part **(ii)(a)** to explain what conclusion can be drawn about whether the change has affected the mean height of bounce. [1]

46) S/ 2016/72/Q1

The length of time, in minutes, taken by people to complete a task has mean 53.0 and standard deviation 6.2. Find the probability that the mean time taken to complete the task by a random sample of 50 people is more than 51 minutes. [4]

47) S/2016/73/Q1

The time taken for a particular type of paint to dry was measured for a sample of 150 randomly chosen points on a wall. The sample mean was 192.4 minutes and an unbiased estimate of the population variance was 43.6 $minutes^2$. Find a 98% confidence interval for the mean drying time. [3]

48) S/2016/73/Q3

1% of adults in a certain country own a yellow car.

(i) Use a suitable approximating distribution to find the probability that a random sample of 240 adults includes more than 2 who own a yellow car. [4]

(ii) Justify your approximation. [2]

49) S /2017/71/Q3

(a) The waiting time at a certain bus stop has variance 2.6 $minutes^2$. For a random sample of 75 people, the mean waiting time was 7.1 minutes. Calculate a 92% confidence interval for the population mean waiting time. [3]

(b) A researcher used 3 random samples to calculate 3 independent 92% confidence intervals. Find the probability that all 3 of these confidence intervals contain only values that are greater than the actual population mean. [2]

50) S/2017/72/Q1

In a survey of 2000 randomly chosen adults, 1602 said that they owned a smartphone. Calculate an approximate 95% confidence interval for the proportion of adults in the whole population who own a smartphone. [4]

51) S/2017/72/Q3

Household incomes, in thousands of dollars, in a certain country are represented by the random variable X with mean μ and standard deviation σ. The incomes of a random sample of 400 households are found and the results are summarised below.

$$n = 400 \qquad \Sigma x = 923 \qquad \Sigma x^2 = 3170$$

(i) Calculate unbiased estimates of μ and σ^2. [3]

(ii) A random sample of 50 households in one particular region of the country is taken and the sample mean income, in thousands of dollars, is found to be 2.6. Using your values from part **(i)**, test at the 5% significance level whether household incomes in this region are greater, on average, than in the country as a whole. [5]

52) S/2017/73/Q2

In a random sample of 200 shareholders of a company, 103 said that they wanted a change in the management.

(i) Find an approximate 92% confidence interval for the proportion, p, of all shareholders who want a change in the management. [3]

(ii) State the probability that a 92% confidence interval does not contain p. [1]

53) S/2017/73/Q1

A residents' association has 654 members, numbered from 1 to 654. The secretary wishes to send a questionnaire to a random sample of members. In order to choose the members for the sample she uses a table of random numbers. The first line in the table is as follows.

1096 4357 3765 0431 0928 9264

The numbers of the first two members in the sample are 109 and 643. Find the numbers of the next three members in the sample. [3]

54) S/2017/73/Q3

The mass, in tonnes, of iron ore produced per day at a mine is normally distributed with mean 7.0 and standard deviation 0.46. Find the probability that the total amount of iron ore produced in 10 randomly chosen days is more than 71 tonnes. [5]

55) S/2017/73/Q4

Last year the mean level of a certain pollutant in a river was found to be 0.034 grams per millilitre. This year the levels of pollutant, X grams per millilitre, were measured at a random sample of 200 locations in the river. The results are summarised below.

$$n = 200 \qquad \Sigma x = 6.7 \qquad \Sigma x^2 = 0.2312$$

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

(ii) Test, at the 10% significance level, whether the mean level of pollutant has changed. [5]

56) S/2018/71/Q1

A random sample of 75 values of a variable X gave the following results.

$$n = 75 \qquad \Sigma x = 153.2 \qquad \Sigma x^2 = 340.24$$

Find unbiased estimates for the population mean and variance of X. [3]

57) S/2018/71/Q2

A six-sided die is suspected of bias. The die is thrown 100 times and it is found that the score is 2 on 20 throws. It is given that the probability of obtaining a score of 2 on any throw is p.

(i) Find an approximate 94% confidence interval for p. [3]

(ii) Use your answer to part **(i)** to comment on whether the die may be biased. [1]

58) S/2018/71/Q5

The mass, in kilograms, of rocks in a certain area has mean 14.2 and standard deviation 3.1.

(i) Find the probability that the mean mass of a random sample of 50 of these rocks is less than 14.0 kg. [3]

(ii) Explain whether it was necessary to assume that the population of the masses of these rocks is normally distributed. [1]

59) S/2018/72/Q3

The management of a factory wished to find a range within which the time taken to complete a particular task generally lies. It is given that the times, in minutes, have a normal distribution with mean μ and standard deviation 6.5. A random sample of 15 employees was chosen and the mean time taken by these employees was found to be 52 minutes.

(i) Calculate a 95% confidence interval for μ. [3]

Later another 95% confidence interval for μ was found, based on a random sample of 30 employees.

(ii) State, with a reason, whether the width of this confidence interval was less than, equal to or greater than the width of the previous interval. [1]

60) S/2018/72/Q4

The mean mass of packets of sugar is supposed to be 505 g. A random sample of 10 packets filled by a certain machine was taken and the masses, in grams, were found to be as follows.

500 499 496 495 498 490 492 501 494 494

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

The mean mass of packets produced by this machine was found to be less than 505 g, so the machine was adjusted. Following the adjustment, the masses of a random sample of 150 packets from the machine were measured and the total mass was found to be 75 660 g.

(ii) Given that the population standard deviation is 3.6 g, test at the 2% significance level whether the machine is still producing packets with mean mass less than 505 g. [5]

61) S/2018/73/Q2

Amy has to choose a random sample from the 265 students in her year at college. She numbers the students from 1 to 265 and then uses random numbers generated by her calculator. The first two random numbers produced by her calculator are 0.213 165 448 and 0.073 165 196.

(i) Use these figures to find the numbers of the first four students in her sample. [2]

There were 25 students in Amy's sample. She asked each of them how much money, $\$x$, they earned in a week, on average. Her results are summarised below.

$$n = 25 \qquad \Sigma x = 510 \qquad \Sigma x^2 = 13\,225$$

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

(iii) Explain briefly what is meant by 'population' in this question. [1]

62) S/2018/73/Q3

A researcher wishes to estimate the proportion, p, of houses in London Road that have only one occupant. He takes a random sample of 64 houses in London Road and finds that 8 houses in the sample have only one occupant. Using this sample, he calculates that an approximate α% confidence interval for p has width 0.130. Find α correct to the nearest integer. [5]

63) S/2019/71/Q2

The time, in minutes, that John takes to travel to work has a normal distribution. Last year the mean and standard deviation were 26.5 and 4.8 respectively. This year John uses a different route and he finds that the mean time for his first 150 journeys is 27.5 minutes.

(i) Stating a necessary assumption, test at the 1% significance level whether the mean time for his journey to work has increased. [6]

(ii) State, with a reason, whether it was necessary to use the Central Limit theorem in your answer to part **(i)**. [1]

64) S/2019/71/Q6

Ramesh plans to carry out a survey in order to find out what adults in his town think about local sports facilities. He chooses a random sample from the adult members of a tennis club and gives each of them a questionnaire.

(i) Give a reason why this will not result in Ramesh having a random sample of adults who live in the town. [1]

(ii) Describe briefly a valid method that Ramesh could use to choose a random sample of adults in the town. [2]

Ramesh now uses a valid method to choose a random sample of 350 adults from the town. He finds that 47 adults think that the local sports facilities are good.

(iii) Calculate an approximate 90% confidence interval for the proportion of all adults in the town who think that the local sports facilities are good. [4]

(iv) Ramesh calculates a confidence interval whose width is 1.25 times the width of this 90% confidence interval. Ramesh's new interval is an x% confidence interval. Find the value of x. [3]

65) S/2019/72/Q3

It is claimed that, on average, a particular train journey takes less than 1.9 hours. The times, t hours, taken for this journey on a random sample of 50 days were recorded. The results are summarised below.

$$n = 50 \qquad \Sigma t = 92.5 \qquad \Sigma t^2 = 175.25$$

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

(ii) Test the claim at the 5% significance level. [5]

66) S/2019/73/Q1

A coin is thrown 100 times and it shows heads 60 times. Calculate an approximate 98% confidence interval for the probability, p, that the coin shows heads on any throw. [3]

67) S/2019/73/Q2

The length of worms is denoted by X cm. The lengths of a random sample of 50 worms were measured. Some of the results were lost, but the following results are available.

- $\Sigma x^2 = 4361$
- An unbiased estimate of the population variance of X is 9.62.

Calculate the mean length of the 50 worms. [3]

68) S/2019/73/Q8

The four sides of a spinner are A, B, C, D. The spinner is supposed to be fair, but Sonam suspects that the spinner is biased so that the probability, p, that it will land on side A is greater than $\frac{1}{4}$. He spins the spinner 10 times and finds that it lands on side A 6 times.

(i) Test Sonam's suspicion using a 1% significance level. [5]

Later Sonam carries out a similar test at the 1% significance level, using another 10 spins of the spinner.

(ii) Calculate the probability of a Type I error. [2]

(iii) Assuming that the value of p is actually $\frac{3}{5}$, calculate the probability of a Type II error. [3]

Answers

1)

$\bar{x} \pm 2.326 \times \frac{2.4}{\sqrt{90}}$

Width $2.326 \times \frac{2.4}{\sqrt{90}} \times 2$

$= 1.18$

2)

(i) $31 \pm 2.326 \times \frac{3}{\sqrt{20}}$

$= (29.4, 32.6)$

(ii) 30% is inside interval
Accept claim (at 2% level)

3)

(i) $\bar{x} = 375.3$
$\sigma^2_{n-1} = 8.29$

(ii) $p = 0.19$ or equiv.

$0.19 \pm 2.055 \times \sqrt{\frac{0.19 \times 0.81}{200}}$

$0.133 < p < 0.247$

4)

(i) Put names in a hat and draw out, or assign a number to each person in year and generate 7 random numbers by calculator.

(ii) est pop mean 116.5/7 (= 16.6)

est pop var = 27.1

(iii) more

(iv) (pocket money of) all pupils in Jenny's year at school

5)

(i) (0.1993 + 0.2887)/2 (= 0.244)
= $61/n$
$n = 250$

(ii) $0.0447 = z \times \sqrt{\frac{0.244(1-0.244)}{250}}$

(or equiv. equ. leading to this)

$z = 1.646$

90% confidence interval

6)

$99.2 \pm 2.576 \times \frac{2.3}{\sqrt{200}}$
$= (98.8, 99.6)$

7)

(i) $\bar{x} = 1050$

$s^2 = \frac{1}{29}\left(33141816 - \frac{31500^2}{30}\right)$

$= 2304$

(ii) $1050 \pm 2.326 \times \frac{48}{\sqrt{30}}$

$= (1030, 1070)$

(iii) $1.96 \times \frac{48}{\sqrt{n}} = 6$

$n = 246$

8)

(i) commuters are not representative of the whole population

(ii) people who travel to work on (this) train

(iii) mean = 6.17 o.e.

$\text{variance} = \frac{1}{11}\left(463.56 - \frac{74^2}{12}\right)$

$= 0.657$

9)

(i) $\hat{\mu} = 227.(1)$

$5 = 2.17 \times \sqrt{\frac{\hat{\sigma}^2}{50}}$

$\hat{\sigma}^2 = 265$ or 266

(ii) $4 = 2.17 \times \frac{16.3}{\sqrt{n}}$

$n = 78$

10)

(i) 0.145
$= 87 / n$
$n = 600$

(ii) $0.0321 = z \times \sqrt{\frac{0.145(1-0.145)}{600}}$

$z = 2.233 \quad \Phi(z) = 0.9872$

width of CI is $1 - 2 \times (1 - 0.9872)$

$\alpha = 97.4\%$

11)

(i) $z = \frac{2.55 - 2.62}{0.3/\sqrt{45}} = -1.565$

$P(z > -1.565) = 0.941$

(ii) rejection region is $m < a_1$ and $m > a_2$

where $\frac{a_1 - 2.62}{0.3/\sqrt{30}} = -1.645$

and $\frac{a_2 - 2.62}{0.3/\sqrt{30}} = 1.645$

$m < 2.53$ and $m > 2.71$

12)

(i) (Approx) normal
mean 62
sd $= \frac{8.2}{\sqrt{50}} = 1.16$ (3 sfs)

(ii) $\frac{64-62}{\text{"}1.16\text{"}}$ (= 1.725 or 1.724)
$1 - \Phi(\text{“}1.725\text{”})$
$= (1 - 0.9577)$
$= 0.0423$ (3 sfs)

13)

(a) $41.2 \pm z \times \sqrt{\frac{32.6}{50}}$
$z = 1.96$
[39.6, 42.8] (3 sfs)

(b) $2 \times \frac{1}{16}$ or $\frac{1}{8}$ or 0.125 or 12.5%
$\alpha = 87.5\%$

14)

(i) $\frac{85.7-85}{\frac{4.8}{\sqrt{n}}}$ (= 1.786)

$n = \left(\frac{1.786 \times 4.8}{0.7}\right)^2$
$= 150$

(ii) $H_0: \mu = 85.0 \quad H_1: \mu > 85.0$

$z = 1.645$

Evidence that μ increased

15)

(i) $\bar{x} = 8.4$
$8.4 \pm z \frac{1.3}{\sqrt{15}}$
$z = 2.576$
[7.54, 9.26]

(ii) No because pop normal
so $\bar{X}$ normally distr

(iii) 8 within CI
Claim justified

16)

(i) $\frac{\frac{18}{70}\times(1-\frac{18}{70})}{70}$ (= 0.00272886..)

$z = 1.645$

$\frac{18}{70} \pm z \times \sqrt{\text{"0.00272886"}}$

0.171 to 0.343

(ii) Var (or sd) estimated or N ~ B used

17)

$z = 2.326$

$494 \pm z \times \frac{23}{\sqrt{150}}$

= 490 to 498 (3 sfs)

18)

(i) $\bar{x} = 1.96$

$(\Sigma x^2 f = 254)$

$S^2 = \frac{50}{49} x\left(\frac{254}{50} - 1.96^2\right)$

$= \frac{1548}{1225}$ or 1.2637

(ii) H_0: Pop mean = 1.66

H_1: Pop mean ≠ 1.66

$\frac{1.96-1.66}{\sqrt{\frac{1.2637}{50}}}$

= 1.887

$z = 1.96 \quad 1.887 < 1.96$

No evidence that mean has changed

(iii) No because H_0 not rejected

(iv) State mean not changed when it has

$-1.96 <$ test stat < 1.96

19)

(i) H_0: Pop mean = 3 H_1: Pop mean > 3

(ii) 0.0683 > 0.05

No evidence that pop mean increased

20)

(i) 7, $^3/_{\sqrt{n}}$

(ii) (a) Pop is normal

(b) Large sample

21)

(i) $p = {}^{18}/_{50}$ or 0.36 oe

$z = 2.326$

$0.36 \pm z\sqrt{\frac{0.36\times(1-0.36)}{50}}$

= 0.202 to 0.518 (3 sfs)

(ii) Sample random

22)

(i) $z = 2.574$ to 2.576

$12.5 \pm z\frac{3.2}{\sqrt{250}}$

12.0 to 13.0 (3 sfs)

(ii) 0.005 or 0.5%

23)

(i) $\bar{x} = \frac{7520}{150} = (50.1)$ (3 sfs)

$s^2 = \frac{150}{149}\left(\frac{413540}{150} - \left(\frac{7520}{150}\right)^2\right)$

= 245 or 246 (3 sfs)

(ii) $\frac{53-\frac{7520}{150}}{\sqrt{\frac{\text{'245.217'}}{80}}}$ (= 1.637 to 1.638)

$1 - \Phi(\text{'1.637'})$

= 0.0488 to 0.0509

24)

(i) P(> 9 Heads | unbiased) = $^{12}C_{10} \times 0.5^{10} \times 0.5^2 + 12 \times 0.5^{11} \times 0.5 + 0.5^{12}$

= 0.0193

Level is 1.93% or 1.9%

24)

(ii) B(100, 0.5) ≈ N(50, 25)

$\frac{x - 0.5 - '50'}{\sqrt{'25'}} = z$

$z = 1.645$

$x = 58.7$

Rejection region is > 59

25)

(i) est(μ) = 2005/200 = (10.025)

est(σ^2) = $\frac{1}{99}20175 - \frac{2005^2}{200}$)

= 0.376 (3 sf)

(ii) $\frac{10 - '10.025'}{\sqrt{\frac{'0.376256'}{50}}}$

$1 - \Phi('0.288')$

= 0.387 (3 sf)

(iii) Yes; (assumed distr of $\overline{X}$ normal) although distr of X unknown

26)

(i) Conclude die is biased when it isn't oe

${}^5C_3\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 + 5\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right) + \left(\frac{1}{6}\right)^5 + 5$

$= \frac{23}{648}$ or 0.0355 (3 sf)

(ii) State or attempt P(0, 1, 2) with $p = \frac{2}{3}$

${}^5C_2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^3 + 5\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^4 + \left(\frac{1}{3}\right)^5$

$= \frac{17}{81}$ or 0.210 (3 sf)

(iii) Est Var(P_s) = $\frac{0.625 \times (1 - 0.625)}{80}$

$(= \frac{3}{1024})$

$z = 2.054$ (or 2.055)

$0.625 \pm z \times \sqrt{\frac{'3'}{1024}}$

= 0.514 to 0.736 (3 sf)

27)

(i) est(μ) = 9750/150 = (65)

est(σ^2) = $\frac{1}{149}(647500 - \frac{9750^2}{150})$

= 92.3 (3 s.f.)

(ii) $z = 2.326$

$'65' \pm z \times \frac{\sqrt{'92.28188'}}{\sqrt{150}}$

= 63.2 to 66.8 (3 s.f.)

(iii) 0.02^2

= 0.0004 o.e.

28)

(i) H_0: Pop mean (or λ or μ) is 5.3

H_1: Pop mean (or λ or μ) is less than 5.3

$P(X \le 1) = e^{-5.3}(1 + 5.3)$

$P(X \le 2) = e^{-5.3}(1 + 5.3 + \frac{5.3^2}{2})$/P(X=2)

$P(X \le 1) = 0.0314$ or 0.0315

& $P(X \le 2) = 0.102$/ P(X=2)=0.7071

CR is 0 or 1 cases

No evidence mean has decreased

(ii) Concluding mean has decreased when it hasn't

'0.0314 or 0.0315'

(iii) (Po(18.4))

N(18.4, 18.4)

$\frac{20.5 - 18.4}{\sqrt{18.4}}$ (= 0.490)

$1 - \Phi('0.490')$

= 0.312 (3 s.f.)

29)

(i) $\bar{x} = 930/15 = (62)$
$z = 1.751$
$'62' \pm z \times \frac{12}{\sqrt{15}}$
$= 56.6$ to 67.4 (3 sf)

(ii) 92% of such intervals will contain μ

(iii) Each possible sample of this size is equally likely

30)

(i) Assume sd unchanged or 4500

H_0: Pop mean = 34600
H_1: Pop mean > 34600

$\frac{35400-34600}{\frac{4500}{\sqrt{90}}}$
$= 1.687/1.686\ (1.69)$
cf $1.645 < 1.686$
Evidence that mean wkly profit has increased

(ii) Distr'n of X unknown.

Yes

(iii) 0.05 or 5%

(iv) $\frac{a-34600}{\frac{4500}{\sqrt{90}}} = 1.645$
$a = 35380$
$\frac{35380-36500}{\frac{4500}{\sqrt{90}}}\ (= -2.361)$
$1 - \Phi('2.361')$
$= 0.0091$

31)

$N(483.2, 537.92)$ or $N(483.2, 23.2^2)$

$\frac{436-483.2}{\sqrt{537.92}}$ or $\frac{436-483.2}{23.2}$ $(= -2.035)$

$\Phi(\text{“}-2.035\text{”}) = 1 - \Phi(\text{“}2.035\text{”})$
$= 0.021$ or 2.1%

32)

$\frac{70}{69} \times 2.70 \qquad = 2.73913$

$3.61 \pm z\sqrt{\frac{\text{"2.73913"}}{70}}$

$z = 1.96$
3.22 to 4.00 (3 sf)

33)

(i) $0.25(1 + 4 + 9) - 1.5^2$
$(= 1.25$ **AG**$)$

(ii) $\frac{1.4-1.5}{\sqrt{\frac{5}{4} \div 300}} \qquad (= -1.549)$

$\Phi(\text{“}-1.549\text{”}) = 1 - \Phi(\text{“}1.549\text{”})$
$= 0.0607$ (3 sf)

(iii) Large sample or large n
($\bar{X}$ (approx) normally distr)
or
Central Limit Theorem

34)

$\frac{\Sigma x}{8} = \frac{2006}{8} = 250.75$ or 251 (3 s.f.)

$(\Sigma x^2 = 503274)$

$\frac{8}{7}\left(\frac{\text{"503274"}}{8} - \text{"250.75"}^2\right)$

$= 38.5$ o.e. (accept 6.204^2)

35)

(i) 2nd

More representative of all appointments
or Lengths may vary during the day
or 1st does not include later appts so not representative

(ii) 0.01 o.e.

Concluding that times spent are too long when they are not.

35)

(iii) H_0: Pop mean appt time (or μ) = 10
H_1: Pop mean appt time (or μ) > 10

$$\frac{\frac{147}{12}-10}{\frac{3.4}{\sqrt{12}}}\ (\pm)$$

= (±)2.292 or (0.0109 if area comparison done)

"2.292" < 2.326 o.e.

(No evidence to reject H_0.)
No reason to believe appts are too long

(iv) Normal population

36)

$p = 0.56$

$$\text{'}0.56\text{'} \pm z \times \sqrt{\frac{0.56\times0.44}{100}}$$

$z = 2.17$, or 2.169 or 2.171

0.452 to 0.668 (3 s.f.)

37)

$\bar{x} = 1.65$

$$\text{est}(\sigma^2) = \frac{100}{99}\left(\frac{276.25}{100} - 1.65^2\right)$$

= 0.040404… = 4/99

$$(\pm)\frac{1.65-1.6}{\sqrt{\frac{\text{"}0.040404\text{"}}{100}}}$$

= (±) 2.487/2.488 accept 2.49 Or 0.0065/0.0064 if area comparison done

comp with 1.96

There is evidence that μ is not 1.6

38)

(i) 4200/80 (=52.5)

$$= \frac{80}{79}\left(\frac{229\,000}{80} - \text{'}52.5\text{'}^2\right) \quad (= 107.595)$$

= 108 (3 sf)

(ii) $\text{'}52.5\text{'} \pm z\sqrt{\frac{\text{'}107.595\text{'}}{80}}$

$z = 2.326$
49.8 to 55.2

(iii) 49

39)

(i) 14800/50 or 296

$$\frac{50}{49}\left(\frac{4390000}{50} - \text{'}296\text{'}^2\right) \quad (= 187.755)$$

= 188 (3 sf)

(ii) $2 \times z \times \sqrt{\frac{\text{'}187.755\text{'}}{50}} = 5.45$ oe

$z = 1.406$ or 1.405
$\Phi(\text{'}1.406\text{'})$ (= 0.92 or 0.9199)

$\alpha = 84$ (2 sf) allow 83.98

(iii) 0.96^4
= 0.849 (3 sf)

40)

(i) $$\text{Var}(p_s) = \frac{0.22\times(1-0.22)}{100}$$

$$\left(= \frac{429}{250\,000} \text{ or } 0.001716\right)$$

$$0.22 \pm z\sqrt{\text{'}\frac{429}{250\,000}\text{'}}$$

$z = 2.17$ or 2.168/9 or 2.171

0.13(0) to 0.31(0) (2 sf)

(ii) '2' × (1 – 0.97) × 0.97
= 0.0582

41)

(i) $\left(\frac{1508}{50}\right) = 30.16\ (30.2)$

$\frac{50}{49}\left(\frac{51825}{50} - ('30.16'^2)\right)$

$= 129$ (3 sf) Or 130

(ii) $(1.5 \times \text{'}30.16\text{'} + 10)$
$= 55.24$

$(1.5^2 \times \text{'}129\ldots\text{'})$

$= 291$ (3 sf)

42)

(i) Cables broken
or not all cables can be accessed oe
or Too many cables oe
or too time consuming oe

(ii) H_0: Pop mean brk str (or μ) = 5
H_1: Pop mean brk str (or μ) < 5

$(\pm)\frac{4.95 - 5}{\frac{0.15}{\sqrt{60}}}$

$(= \pm 2.582)$

comp ± 2.326
There is evidence that mean breaking strength is less than it should be
Or reject H_0 (H_0 correctly defined)

(iii) Population not necessarily normal
so yes

43)

(i) Each employee has an equal chance of being chosen

(ii) Est (μ) = 4
Est (σ^2) = $\frac{10}{9}\left(\frac{199.22}{10} - \text{'}4\text{'}^2\right)$

$= 4.36$ (3 sf)

(iii) Distances travelled by all employees at the firm

44)

(i) $((0.5672 + 0.6528) \div 2)$
$= 0.61$

(ii) $\text{'}0.61\text{'} + z\sqrt{\frac{\text{'}0.61\text{'}\times(1-\text{'}0.61\text{'})}{350}} = 0.6528$

$z = 0.0428 \times \sqrt{\frac{700}{\text{'}0.61\text{'}\times(1-\text{'}0.61\text{'})}}$ oe

$= 2.321$
98% confidence

45)

(i) Pop too big or takes too long oe
or testing destroys articles oe

(ii) **(a)** $z = 1.96$
$65.7 \pm z \times \frac{\sqrt{15}}{10}$
$= 64.9$ to 66.5 (3 sf)

(b) CI does not include 64.7
Probably has affected (or increased) mean bounce ht.

47)

$192.4 \pm z\sqrt{\frac{43.6}{150}}$

$z = 2.326$ to 2.329
191 to 194 (3 sf)

48)

(i) Use of Poisson
Mean = 2.4
$1 - e^{-2.4}(1 + 2.4 + \frac{2.4^2}{2})$
$= 0.43(0)$ (3 sf)

(ii) $240 > 50$ or $n>50$
$240 \times 0.01 = 2.4 < 5$ or $np<5$ or $p<0.1$

49)

(a) $7.1 \pm z \times \sqrt{\frac{2.6}{75}}$

$z = 1.751$

6.77 to 7.43 (3 sfs)

(b) 0.04^3

(c) e.g. Particular day or time of day

50)

$\dfrac{0.801\times(1-0.801)}{2000}$ (

$0.801 \pm z \times \sqrt{"0.0000797"}$

$z = 1.96$

0.784 to 0.818 (3 sf)

51)

(i) Est (μ) = 923/400 or 2.3075 or 2.31 (3 sf)

$\text{Est}(\sigma^2) = \dfrac{400}{399}\left(\dfrac{3170}{400} - "2.3075"^2\right)$ OE

= 2.60696 or 2.61 (3 sf)

(ii) H_0: Pop mean (or μ) = "2.31" or "2310"
H_1: Pop mean (or μ) > "2.31" or "2310"

$\pm \dfrac{2.6-"2.310"}{\sqrt{2.60696 \div 50}} = 1.27$

Comp 1.645 (OE)

No evidence that incomes in the region greater

52)

(i) $z = 1.751$

$\dfrac{103}{200} \pm z\sqrt{\dfrac{\frac{103}{200}\times(1-\frac{103}{200})}{200}}$ oe

= 0.453 to 0.577 (3 sf) as final answer

(ii) 0.08 oe 8%, 8/100

53) 573, 43 (or 043), 289

54)

10×0.46^2 (= 2.116) or $\dfrac{0.46}{\sqrt{10}}$

Total mass of ore ~ N(70, 2.116) or $\sim N\left(7, \left(\dfrac{0.46}{\sqrt{10}}\right)^2\right)$

$\pm\dfrac{71-"70"}{\sqrt{"2.116"}}$ or $\pm\dfrac{7.1-"7.0"}{0.46/\sqrt{10}}$ (= 0.687)

$1 - \phi("0.687")$

= 0.246 (3 sf)

55)

(i) $\bar{x} = 6.7/200$ (= 67/2000 = 0.0335)

$s^2 = \dfrac{200}{199}\times\left(\dfrac{0.2312}{200} - "0.0335"^2\right)$

= 0.0000339(2) = 27/796000

(ii) H_0: Pop mean level = 0.034
H_1: Pop mean level ≠ 0.034

$\dfrac{"030335" - 0.034}{\dfrac{\sqrt{"0.00003392"}}{\sqrt{200}}}$

= −1.21(4) (3 sfs) (−1.22 ↔ −1.21)

Comp with $z = -1.645$ (or 0.1124>0.05)

No evidence that (mean) pollutant level has changed, accept H_0

56)

est(μ) (= 153.2 ÷ 75) = 2.04 (3 sf)

$\text{est}(\sigma^2) = \frac{75}{74}\left(\frac{340.24}{75} - "2.04267"^2\right)$ oe

= 0.369 (3 sf)

57)

(i) $\dfrac{20}{100} \pm z \times \sqrt{\dfrac{0.2\times(1-0.2)}{100}}$

$z = 1.881$ or 1.882

= 0.125 to 0.275

(ii) $\frac{1}{6}$ is within this range
No evidence of bias concerning 2

58)

(i) $\dfrac{14-14.2}{\frac{3.1}{\sqrt{50}}}$

$1 - \Phi(\text{"0.456"})$

= 0.324 (3 sfs)

(ii) No because n large

58)

(iii) $H_0: \mu = 14.2$
$H_1: \mu < 14.2$

$\dfrac{13.5 - 14.2}{\frac{3.1}{\sqrt{100}}}$

$= -2.258$

comp -2.054 (or -2.055)

There is evidence (at 2% level) that mean mass in this area < 14.2

59)

(i) $52 \pm z \times \frac{6.5}{\sqrt{15}}$

$z = 1.96$

48.7 to 55.3 (3 sf)

(ii) Narrower
because more information or
because $\frac{\sigma}{\sqrt{n}}$ smaller

60)

(i) $\text{Est}(\mu) = 495.9$

$\text{Est}(\sigma^2) = \frac{10}{9}\left(\frac{2459283}{10} - \text{"}495.9\text{"}^2\right)$

$= 12.8$ (3 sf) or 383/30

(ii) $H_0: \mu = 505$
$H_1: \mu < 505$
$\dfrac{\frac{75660}{150} - 505}{3.6 \div \sqrt{150}}$

$= -2.04$

comp $z = -2.054$

No evidence (at 2%) that machine pkts mean mass < 505

61)

(i) 213, 165, 73, 196

61)

(ii) $\dfrac{510}{25} = \dfrac{102}{5}$ or 20.4

$\dfrac{25}{24}\left[\dfrac{13225}{25} - \left(\dfrac{102}{5}\right)^2\right]$

118 (3 sf) or $\dfrac{2821}{24}$

(iii) (Average) weekly earnings of all students in Amy's year

62)

$\dfrac{\frac{8}{64} \times (1 - \frac{8}{64})}{64}$ $\left(= \dfrac{7}{4096}\right.$ or $\left.0.00171\right)$

$2 \times z\sqrt{\text{"}\dfrac{7}{4096}\text{"}} = 0.130$

$z = 1.572$

$\phi(\text{"}1.572\text{"})$ $(= 0.942)$
$(0.942 - (1 - 0.942) = 0.884)$

$\alpha = 88$

63)

(i) Assume sd still 4.8 or is unchanged

H_0: Pop mean = 26.5
H_1: Pop mean > 26.5

$\dfrac{27.5 - 26.5}{\frac{4.8}{\sqrt{150}}}$

$= 2.552$

Comp with z-value
'2.552' > 2.326

There is evidence time has increased

(ii) No because pop is normal so distr of $\bar{X}$ is normal

64)

(i) Biased towards people who like tennis
Excludes people who don't like tennis

(ii) Obtain a list of all people in the town

Use random numbers

64)

(iii)

$\text{Var}(p) = \dfrac{\frac{47}{350}(1-\frac{47}{350})}{350}$ (= 0.000332152)

$z = 1.645$

$\dfrac{47}{350} \pm z\sqrt{\dfrac{\frac{47}{350}(1-\frac{47}{350})}{350}}$

0.104 to 0.164 (3 sf)

(iv)

1.25×1.645

$\Phi(\text{'}2.056\text{'})$

$x = 96$ (2 sf)

65)

i)

$\text{Est}(\mu) = 1.85$

$\text{Est}(\sigma^2) = \dfrac{50}{49}\left(\dfrac{175.25}{50} - \text{'}1.85\text{'}^2\right)$

$= 0.0842$ (3 sf) or $\dfrac{33}{392}$

(ii)

H_0: Pop mean time = 1.9 (h)
H_1: Pop mean time < 1.9 (h)

$\pm\dfrac{1.85-1.9}{\sqrt{\dfrac{\text{'}0.0842\text{'}}{50}}}$

$= -1.22$

comp $z = -1.645$

No evidence that mean time < 1.9 h

66)

$0.6 \pm z\sqrt{\dfrac{0.4\times0.6}{100}}$

$z = 2.326$

0.486 to 0.714 (3 sf)

67)

$\frac{50}{49}\left(\frac{4361}{50} - \bar{x}^2\right) = 9.62$

$\bar{x}^2 = \frac{4361}{50} - 9.62\times\frac{49}{50} = 77.7924$

$\bar{x} = 8.82$ (3 sf)

68)

(i)

$H_0: p = \frac{1}{4}$
$H_1: p > \frac{1}{4}$

${}^{10}C_6(\frac{1}{4})^6(\frac{3}{4})^4 + {}^{10}C_7(\frac{1}{4})^7(\frac{3}{4})^3 + {}^{10}C_8(\frac{1}{4})^8(\frac{3}{4})^2 + 10(\frac{1}{4})^9(\frac{3}{4}) + (\frac{1}{4})^{10}$

$= 0.0197$

comp '0.0197' with 0.01

No evidence to conclude $p > \frac{1}{4}$

(ii)

${}^{10}C_7(\frac{1}{4})^7(\frac{3}{4})^3 + {}^{10}C_8(\frac{1}{4})^8(\frac{3}{4})^2 + 10(\frac{1}{4})^9(\frac{3}{4}) + (\frac{1}{4})^{10}$

P(Type I) = 0.00351 (3 sf)

(iii)

C.R is $X \geqslant 7$
P(Type II) = $1 - P(X \geqslant 7 \mid p = \frac{3}{5}) =$

$1 - ({}^{10}C_7(\frac{3}{5})^7(\frac{2}{5})^3 + {}^{10}C_8(\frac{3}{5})^8(\frac{2}{5})^2 + 10(\frac{3}{5})^9(\frac{2}{5}) + (\frac{3}{5})^{10})$

$= 0.618$